

Acoustic detection range of right whale upcalls identified in near-real time from a moored buoy and a Slocum glider

Hansen D. Johnson, Christopher T. Taggart, Arthur E. Newhall, et al.

Citation: [The Journal of the Acoustical Society of America](#) **151**, 2558 (2022); doi: 10.1121/10.0010124

View online: <https://doi.org/10.1121/10.0010124>

View Table of Contents: <https://asa.scitation.org/toc/jas/151/4>

Published by the [Acoustical Society of America](#)

ARTICLES YOU MAY BE INTERESTED IN

[Acoustic vector sensor analysis of the Monterey Bay region soundscape and the impact of COVID-19](#)

[The Journal of the Acoustical Society of America](#) **151**, 2507 (2022); <https://doi.org/10.1121/10.0010162>

[An overview of array invariant for source-range estimation in shallow water](#)

[The Journal of the Acoustical Society of America](#) **151**, 2336 (2022); <https://doi.org/10.1121/10.0009828>

[Wave equations for porous media described by the Biot model](#)

[The Journal of the Acoustical Society of America](#) **151**, 2576 (2022); <https://doi.org/10.1121/10.0010164>

[Oceanography by ear](#)

[The Journal of the Acoustical Society of America](#) **151**, R7 (2022); <https://doi.org/10.1121/10.0009957>

[The combined importance of finite dimensions, anisotropy, and pre-stress in acoustoelastography](#)

[The Journal of the Acoustical Society of America](#) **151**, 2403 (2022); <https://doi.org/10.1121/10.0010110>

[Auralization of aircraft flyovers with turbulence-induced coherence loss in ground effect](#)

[The Journal of the Acoustical Society of America](#) **151**, 2453 (2022); <https://doi.org/10.1121/10.0010121>






**Advance your science and career
as a member of the**

ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE



Acoustic detection range of right whale upcalls identified in near-real time from a moored buoy and a Slocum glider

Hansen D. Johnson,^{1,a)}  Christopher T. Taggart,¹ Arthur E. Newhall,² Ying-Tsong Lin,² 
and Mark F. Baumgartner³ 

¹Oceanography Department, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4R2, Canada

²Applied Ocean Physics and Engineering Department, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, Massachusetts 02543, USA

³Biology Department, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, Massachusetts 02543, USA

ABSTRACT:

The goal of this study was to characterize the detection range of a near real-time baleen whale detection system, the digital acoustic monitoring instrument/low-frequency detection and classification system (DMON/LFDCS), equipped on a Slocum glider and a moored buoy. As a reference, a hydrophone array was deployed alongside the glider and buoy at a shallow-water site southwest of Martha's Vineyard (Massachusetts, USA) over a four-week period in spring 2017. A call-by-call comparison between North Atlantic right whale upcalls localized with the array ($n = 541$) and those detected by the glider or buoy was used to estimate the detection function for each DMON/LFDCS platform. The probability of detection was influenced by range, ambient noise level, platform depth, detection process, review protocol, and calling rate. The conservative analysis of near real-time pitch tracks suggested that, under typical conditions, a 0.33 probability of detection of a single call occurred at 6.2 km for the buoy and 8.6–13.4 km for the glider (depending on glider depth), while a 0.10 probability of detection of a single call occurred at 14.4 m for the buoy and 22.6–27.5 km for the glider. Probability of detection is predicted to increase substantially at all ranges if more than one call is available for detection.

© 2022 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1121/10.0010124>

(Received 24 December 2020; revised 6 March 2022; accepted 21 March 2022; published online 13 April 2022)

[Editor: Aaron M. Thode]

Pages: 2558–2575

I. INTRODUCTION

Mitigation of anthropogenic impacts on North Atlantic right whales (*Eubalaena glacialis*; hereafter “right whales”) and other at-risk species is critical for effective conservation but challenging given limited survey resources and the cryptic nature of whale behavior. Nearly all risk mitigation and management strategies rely on knowledge of whale distribution collected by monitoring surveys (e.g., Vanderlaan *et al.*, 2008). Conventional visual survey methods provide important information for population and health assessment, but they alone cannot cover the time and space scales required to resolve range-wide distribution patterns. Passive acoustic monitoring (PAM) can complement visual survey methods by offering the ability to autonomously monitor remote areas persistently for months to years at a time (e.g., Davis *et al.*, 2017).

Numerous efforts have demonstrated the efficacy of PAM for right whales. Clark *et al.* (2010) conducted an extensive comparison between aerial and acoustic surveys for right whales in Cape Cod Bay and demonstrated that visual surveys detected right whales on two-thirds of the

days for which they were detected acoustically. The same authors concluded that PAM is more reliable than visual methods for determining right whale presence over daily time scales in Cape Cod Bay and strongly recommended that PAM be used to inform management decisions. In a similar comparison on the southwestern Scotian Shelf, Durette-Morin *et al.* (2019) reached similar conclusions and highlighted the capacity of PAM to extend monitoring beyond visual surveys constrained by limited resources and poor sighting conditions. Davis *et al.* (2017) collated and analyzed an acoustic dataset spanning 35 600 days over 2004–2014 on 324 recorders located in the western North Atlantic from the Caribbean to the Davis Strait and Iceland. Their analyses documented shifts in the range-wide distribution pattern of right whales since 2010 as well as persistent wintertime presence in most regions; observations that would not have been possible if reliant on sporadic visual surveys expended over the last decade.

Archival PAM data are rich in information but typically not available on time scales required to inform risk-mitigation strategies and dynamic management of activities that affect whales. The Woods Hole Oceanographic Institution (WHOI) developed a PAM system comprised of the low-power digital acoustic monitoring instrument (DMON) (Johnson and Hurst, 2007) and an on-board detection algorithm [low-frequency detection and classification system (LFDCS)] (Baumgartner

^{a)}Also at: Biology Department, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, MA 02543, USA. Electronic mail: hansen.johnson@dal.ca

and Mussoline, 2011) that detects, classifies, and reports the sounds of baleen whales (right, fin, sei, blue, and humpback) in near real-time from autonomous platforms (Baumgartner *et al.*, 2013; Baumgartner *et al.*, 2019; Baumgartner *et al.*, 2020). The LFDCS algorithm produces spectrograms of the audio data, removes spurious broadband noise and continuous tonal noise and then uses a contour-following algorithm to create pitch tracks of tonal sounds from the spectrogram. Each pitch track is classified by comparing attributes of the pitch track to a library of call types using quadratic discriminate function analysis. The DMON/LFDCS then sends a subset of the pitch tracks and classifications to a land station via Iridium satellite every 2 h where they are divided into 15 min analysis (tally) periods that are manually reviewed by a trained analyst for the acoustic presence of several species, including right whales (Baumgartner *et al.*, 2013; Baumgartner *et al.*, 2019; Baumgartner *et al.*, 2020).

The DMON/LFDCS is fully operational on Slocum gliders (Baumgartner *et al.*, 2013; Baumgartner *et al.*, 2020) and moored buoys (Baumgartner *et al.*, 2019). These platforms are particularly useful for management applications as they can monitor persistently for days to years at a time, regardless of weather conditions, at no risk to human operators, and at a relatively low cost compared to conventional visual surveys. From 2013 to 2021, the DMON/LFDCS system has been deployed on at least 50 Slocum gliders and 10 moored buoy missions in the Northwest Atlantic, amassing over 4500 days at sea, and recording over 1500 validated right whale detections. All of these data are made available in near real-time in a variety of ways, including email and text messages, websites (robots4whales.who.edu;¹ whalesafe.com;² see also Johnson *et al.*, 2021), and a mobile app (Whale Alert³). The system has demonstrated its effectiveness in several monitoring initiatives with the National Oceanic and Atmospheric Administration (NOAA) Northeast Fisheries Science Center (NEFSC), the U.S. Navy, U.S. Coast Guard, Fisheries and Oceans Canada (DFO), Transport Canada (TC), and the Department of National Defense Canada (DND).

The LFDCS detector and validation protocol have been extensively used and quantitatively evaluated for right whales. Davis *et al.* (2017) used the LFDCS for their analysis of archival recordings, and Baumgartner *et al.* (2019) and Baumgartner *et al.* (2020) recently evaluated the accuracy of the LFDCS on the DMON for near real-time detections from moored buoys and Slocum gliders. Baumgartner *et al.* (2019) found that the false-positive rate for moored buoys was 0%, meaning that right whales were never detected in near real-time when they were not acoustically present, and that the system missed right whale occurrence 27% of the time on daily time scales. Using the same near real-time review protocol, Baumgartner *et al.* (2020) found a 0% false-positive rate and an 18% missed occurrence rate for Slocum gliders on daily time scales. The near real-time review protocol was designed to be conservative in recognition of the high operational costs of a false detection, but it can be adjusted depending on the application (Baumgartner *et al.*, 2019).

As with visual surveys, PAM detection performance depends on a variety of species-, site-, and platform-specific factors. Sound source level, background noise, propagation conditions, receiver characteristics, and detection processes all influence the probability of detecting a call. A challenge when using detection information from many PAM systems (including the DMON/LFDCS) for science, conservation, and mitigation applications is the uncertainty in the relationship between the probability of detecting a whale call and the range to a calling animal, which can lead to misinterpretation of PAM results (e.g., Helble *et al.*, 2013b). The platforms on which the DMON/LFDCS has been integrated currently relay only the position of the platform when a sound is detected, not the position of the sound source. Determining whether positional uncertainty is tolerable for a particular application depends on the acoustic detection range for a species of concern; for short detection ranges (e.g., hundreds of meters), the position of the platform may be an acceptable proxy for the position of the animal, but for large detection ranges (e.g., tens of kilometers), lack of location specificity may limit mitigation options over short response time scales (see Johnson *et al.*, 2020b).

The acoustic detection range of a PAM system is best described by a detection function, which refers to the continuous relationship between the probability of detection and the horizontal distance between a sound source and the acoustic receiver. Estimating the site- and species-specific detection function is necessary to properly interpret and compare PAM results (e.g., Helble *et al.*, 2013b) and is a prerequisite of acoustic density estimation using distance sampling methods (e.g., Buckland *et al.*, 2004). A detection function can be estimated empirically using measured distances to both detected and undetected calls (as in this study), statistically by fitting a function to the distribution of distances to detected calls (e.g., Marques *et al.*, 2011; Harris *et al.*, 2013), or computationally based on simulations of some or all of the call production, propagation, and detection processes (e.g., Küsel *et al.*, 2011; Helble *et al.*, 2013a; Harris *et al.*, 2018). Each method has distinct advantages and disadvantages (Marques *et al.*, 2013), but if applied correctly with valid assumptions, they can provide reliable estimates of the detection function.

The empirical approach involves measuring distances to both detected and undetected calls, then estimating the detection function based on the proportion of calls detected at each range using logistic regression (e.g., Buckland *et al.*, 2006; this study) or a generalized additive model (GAM) (e.g., Marques *et al.*, 2009). This approach is desirable because it requires relatively few assumptions, but it is often impeded by the challenge of measuring the distances to undetected calls. Previous studies have overcome this difficulty using a variety of methods, including deploying animal-borne acoustic tags (e.g., Marques *et al.*, 2009), combining visual and acoustic observations (e.g., Kyhn *et al.*, 2012), or conducting playback experiments (e.g., Nuuttila *et al.*, 2018). A common approach for visual surveys is to have observers with high-power binoculars set up

trials for observers that are using the naked eye (e.g., Buckland and Turnock, 1992). Here, we employ an analogous study design by using a multi-channel hydrophone array to set up trials for single sensor DMON/LFDCS platforms with the goal of assessing the range-dependent accuracy of the DMON/LFDCS for detecting right whale upcalls using a mobile (glider) and a fixed (buoy) platform.

II. METHODS

A. Site description

From 28 February to 24 March 2017, we deployed horizontal and vertical line arrays of hydrophones forming an L-shaped array (hereafter “array” unless otherwise specified) as well as a DMON/LFDCS Slocum glider adjacent to an extant DMON/LFDCS buoy at a nominal position of $41^{\circ}8.8' \text{ N}$, $70^{\circ}56.7' \text{ W}$, $\sim 15 \text{ m}$ southwest of Nomans Land, a small island southwest of Martha’s Vineyard, MA, USA. The water depth was $\sim 30 \text{ m}$ at the buoy. The bathymetry (from the ETOPO1 Global Relief model) (Amante and Eakins, 2009) is relatively flat and featureless to a range of $\sim 15 \text{ km}$ with the notable exception of a steep shoal near Nomans Land beginning $\sim 8 \text{ km}$ northeast of the deployment site [Fig. 1(A)]. The glider held station within $\pm 2 \text{ km}$ of the array for the first 2 weeks of the study period before making longer (up to 10 km), roughly circular forays along a

predefined course away from the array for the remainder of the mission [Fig. 1(B)]. The array was positioned within $\sim 150 \text{ m}$ of the DMON/LFDCS buoy, with horizontal and vertical components separated by $\sim 120 \text{ m}$ [Fig. 1(C)]. The water column was uniformly mixed for the duration of the study. Sea state during the study period was assessed using hourly observations of significant wave height recorded at the Block Island meteorological buoy (Station 44 097) $\sim 10 \text{ km}$ SW of the study site (see footnote 4).

The study site was chosen because the DMON/LFDCS buoy located there was originally deployed to monitor right whale presence in near real-time close to several U. S. Coast Guard gunnery ranges. This area is also of particular interest because it is targeted for wind energy development in the near future. We chose to deploy the glider and array in the early Spring based on historical right whale presence in the region at that time of year (Davis *et al.*, 2017).

B. System specifications

The glider and buoy were each equipped with DMON/LFDCS real-time PAM systems. In addition to generating pitch tracks in real-time, the glider DMON/LFDCS recorded audio at 2 kHz continuously while the buoy DMON/LFDCS recorded audio at 2 kHz on a 50% duty cycle (0.5 h on, 0.5 h off) due to memory constraints imposed by the year-long deployment. The audio recorded while the glider was at the surface,

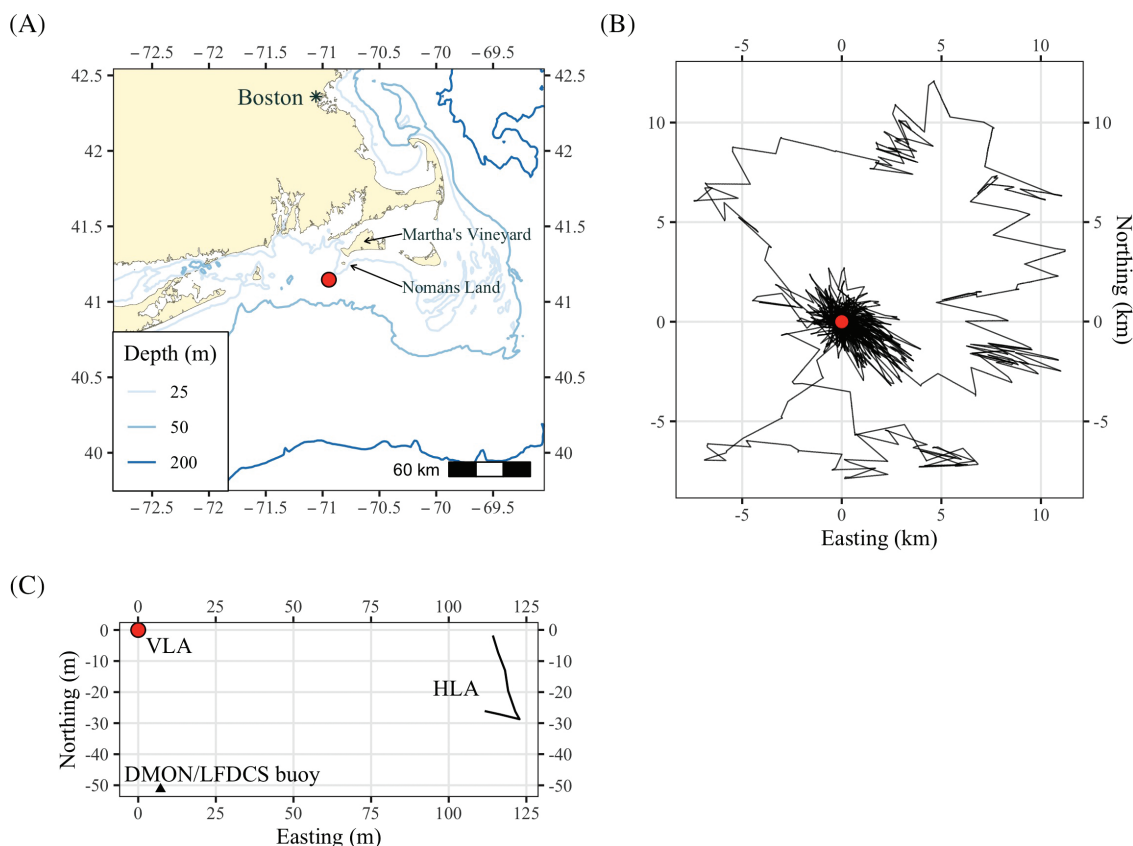


FIG. 1. The position of the vertical line array (VLA) (red circle) at (A) the study site in $\sim 30 \text{ m}$ water depth $\sim 15 \text{ km}$ southwest of Nomans Land, MA, USA, (B) relative to the trajectory of the glider (black line) from 28 February through 24 March, and (C) relative to the positions of the horizontal line array (HLA) and DMON/LFDCS buoy.

approximately 12% of the deployment, was contaminated by platform noise and not included in the analysis. Each system generated and classified pitch tracks of tonal signals over the full 2 kHz bandwidth and transmitted them back to shore every 2 h. Pitch tracking was continuous on the buoy but was suspended while the glider was at the surface. The DMON hydrophone system had a sensitivity of -203 dB re 1 V μPa^{-1} , gain of 33.2 dB, zero-to-peak voltage of 1.5 V, and flat frequency response between approximately 50 and 1000 Hz. Additional details on the specifications of the PAM system on the glider and buoy are available in Baumgartner *et al.* (2013); Baumgartner *et al.* (2020); and Baumgartner *et al.* (2019), respectively.

The vertical component of the array (referred to as the vertical line array or VLA) consisted of a several hydrophone receiving unit (SHRU), four hydrophones, multiple environmental sensors, and a number of additional mooring components. The SHRU was suspended several meters above the anchor and acoustic release system and sampled the hydrophones continuously at a rate of 9765.625 Hz for the full deployment period. The hydrophone sensitivity was -170 dB re 1 V μPa^{-1} and recorder gain was 26 dB. The hydrophones and environmental sensors were secured to a 15 m wire rope that extended from the top of the SHRU to a steel sphere suspended ~ 8 m below the surface. Hydrophones were positioned at approximately 27.4, 23, 18, and 13.4 m depth (nominal spacing of 5 m). The environmental sensors included two temperature loggers and a temperature-pressure logger positioned at intervals along the extent of the array to measure the temperature profile, depth, and array tilt at 0.5 Hz throughout the deployment. All environmental sensors operated without any detectable acoustic signature. The horizontal component of the array (referred to as the horizontal line array or HLA) was comprised of 8 hydrophones positioned at 7.5 m intervals along a 60 m cable coated with hairy fairing. The hydrophone sensitivity was -173 dB re 1 V μPa^{-1} and recorder gain was 23 dB. The hydrophones were sampled continuously at 4 kHz using a multichannel recorder built by Webb Research Corporation (East Falmouth, MA). The HLA also had a single temperature-pressure instrument to record bottom water properties for the full deployment. Additional details on array specifications and configuration are provided by Johnson *et al.* (2020a).

C. Call detection and localization

1. Call detection using array

Right whale upcalls (hereafter “upcalls”) were chosen as the focal call for this analysis because they are used by the LFDCS to determine right whale presence and were amenable to localization (see Sec. II C 4). Upcalls are frequency modulated upsweeps from approximately 100–200 Hz with a duration of ~ 0.75 s (see Parks *et al.*, 2009, for a detailed description and discussion of upcall acoustic parameters). The full 12-channel acoustic record from the array was decimated to 2 kHz, displayed as spectrograms and visually/aurally reviewed for upcalls by an experienced analyst (HDJ) using Raven Pro 2.0

(Bioacoustics Research Program, 2011) and consistent spectrogram parameters (512 sample DFT, 50% overlap, Hann window) which yielded a time resolution of 6.25 ms and a frequency resolution of 3.9 Hz. Only upcalls that were present on one or more channels and could be confidently scored as “detected” were included in the analysis. We assumed that the performance advantage from the simultaneous review of multiple channels located at different depths allowed the array to serve as a suitable reference to determine the probability of detection and the detection range of the DMON/LFDCS single-hydrophone platforms.

2. Call detection using near real-time pitch track data

Pitch tracks and automated detector output for the buoy and glider were displayed chronologically using custom-written software designed to mimic the interface used to validate near real-time detection results on the DMON/LFDCS website (robots4whales.whoi.edu). The full pitch track datasets were visually reviewed independently (i.e., without access to the archival audio data or any other detection or localization results) by the same experienced analyst performing call detection in the array data (HDJ) who was also well-versed in the review of pitch track data. Pitch tracks of upcalls were scored as “detected” or “possibly detected” depending on the confidence of the analyst following a similar protocol as described by Baumgartner *et al.* (2019); also available at robots4whales.whoi.edu.¹ In brief, upcalls scored as “detected” convincingly adhered to the general time/frequency characteristics of upcalls (see Sec. II C 1) and were isolated from competing noise processes, while those scored as “possibly detected” only partially satisfied these criteria. Classification by the LFDCS was not required for a score of “detected” or “possibly detected.”

3. Call detection using archival audio data

The complete archival audio records from the glider and buoy were displayed as spectrograms and visually/aurally reviewed chronologically for upcalls by the experienced analyst (HDJ) also using Raven Pro 2.0 (Bioacoustics Research Program, 2011). The spectrogram parameters were the same as those used for the analysis of the array audio (512 sample DFT, 50% overlap, Hann window). Upcalls were given a score of “detected” or “possibly detected” depending on the confidence of the analyst. As with the pitch track analysis, the analyses of the archival audio from the glider and buoy were done independently without access to detection or localization results from any other platform.

4. Call localization

A normal mode backpropagation method (Lin *et al.*, 2012) using the array data were utilized to estimate range and bearing to each detected call. The method allows localization of low-frequency signals from a single array station,

as opposed to the distributed arrays required for conventional arrival time difference methods (e.g., Cato, 1998). The technique exploits the modal dispersion of a shallow water waveguide that is well-represented by normal mode theory (Frisk, 1994). The vertical component of the array can be used to spatially filter modal arrivals, the arrival time differences between which can be used to make inferences about signal propagation (Fig. 2). The general steps of the localization workflow were to (1) isolate an upcall in time and frequency space using spectrograms of the array data (see Sec. II C 1), (2) use a normal mode model (KRAKEN) (Porter, 1992) and pseudo-inverse mode filtering to isolate the modal arrivals of the call (Fig. 2), (3) use the estimated group velocities of each modal arrival to beamform to determine the arrival bearing of the call, (4) use the same mode model to estimate mode structures along the arrival path, and (5) back-propagate the received signal along the arrival path until the back-propagated modes converged (Fig. 3). The range with the best convergence was used as the estimated range to the call. With this estimated bearing and

range, the position of the calling whale was calculated. For more details on the localization methods, see Lin *et al.* (2012), and for an application to sei whale call localization, see Newhall *et al.* (2012).

The uncertainty in the range estimates was qualitatively assessed from the produced ambiguity surfaces [e.g., Fig. 3(A)] and was estimated to be ~ 1 km. The normal mode backpropagation method for localizing long distant sound sources requires the excitation of two or more propagating modes. The cut-off frequency for propagating mode 2 at the study site (~ 30 m water depth) was approximately 80 Hz, which prevented localization of any distant calls with substantial energy at lower frequencies. The cut-off frequency for propagating mode 3 was approximately 300 Hz. Thus, mode 3 was not reliably present in all distant upcalls and therefore, not used for localization.

During the recovery of the horizontal component of the array (the HLA), it was immediately evident that it had moved from its initial deployment position. There were several storm events (2 and 15 March), characterized by high

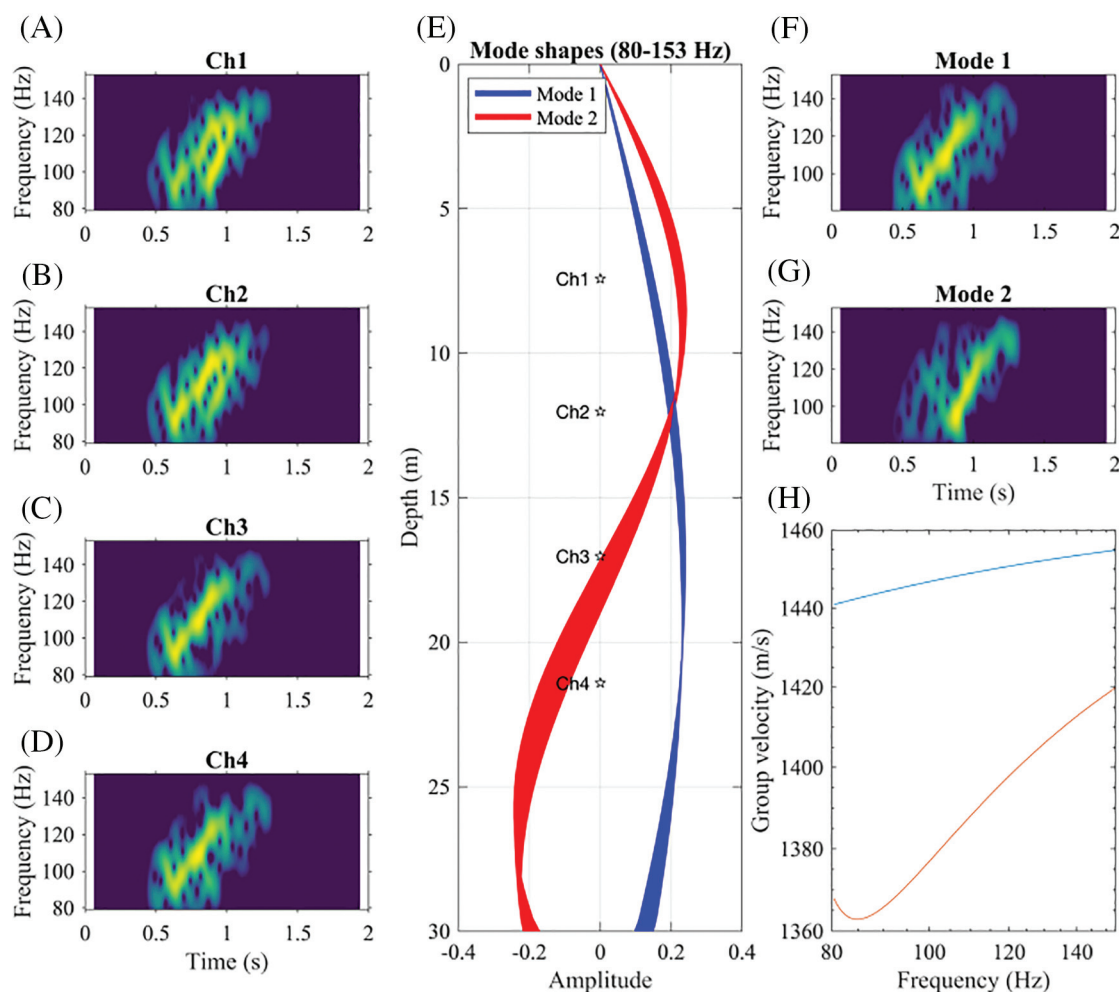


FIG. 2. Overview of the mode filtering procedure. Panels A–D show spectrograms of a right whale upcall received on each channel of the vertical line array. Panel E shows the theoretical shapes of mode 1 (blue) and mode 2 (red) at 146 discrete frequencies within the 80–153 Hz band. These were generated using the KRAKEN normal mode model parameterized using site-specific conditions at the time of the right whale upcall. The labeled stars indicate the depth of each channel of the array. The same model produced the group velocity estimates for each mode shown in panel H. Panels F and G show spectrograms of modes 1 and 2, respectively, after application of the pseudo-inverse mode filter.

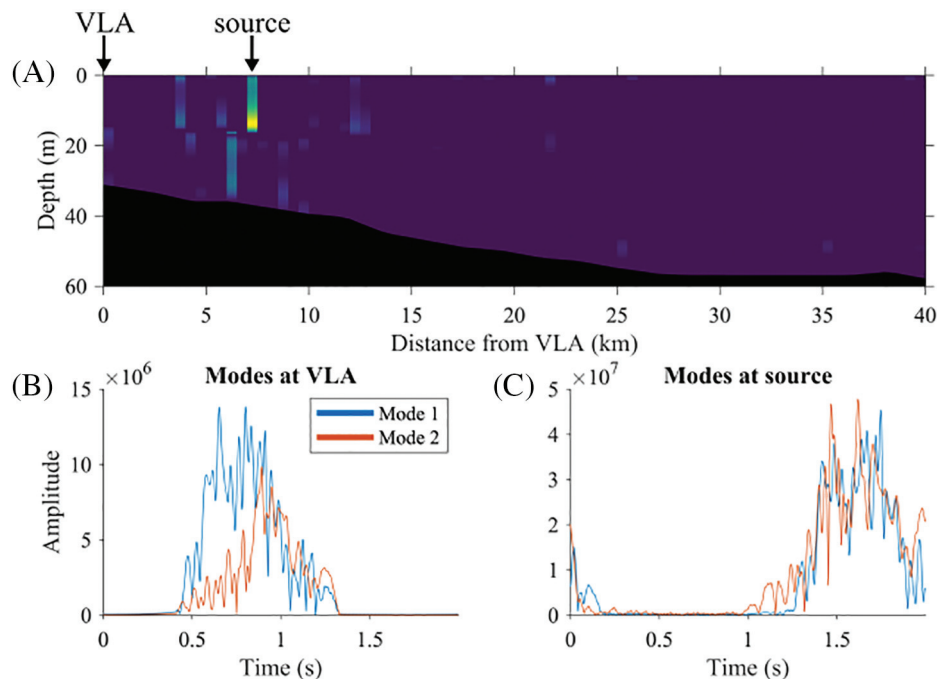


FIG. 3. Overview of the backpropagation and ranging procedure. Panel A shows a normalized probability map of the backpropagation results with an arrow indicating the most likely range to the calling whale. Panel B shows the timing and amplitude of mode 1 (blue) and mode 2 (red) as received at the vertical line array (VLA), and panel C shows the same modes at the source after backpropagation.

ambient noise levels and wave heights [Figs. 4(A) and 4(B)], during which the HLA may have moved [Johnson *et al.*, 2020(a)]. Precise estimates of the location of each HLA element were critical to our localization methodology, as errors in HLA element location prevent accurate beamforming for call bearing estimation. To correct for storm-induced movement, the HLA elements were re-localized several times using known vessel noise emitted from the WHOI coastal research vessel *Tioga* during cruises in the area (after Morley *et al.*, 2009; details in Johnson *et al.*, 2020a). These analyses provided evidence that several of the array elements moved negligible distances (< 2 m) in the storm event on 2 March, and substantial distances (up to ~ 15 m) in the storm event on 15 March. We did not correct for movement during the first storm. For the second storm on 15 March, we assumed that the array movement occurred at the beginning of the day, such that bearings to calls localized from 15 through 23 March ($n = 368$) were computed using the post-storm array position [Fig. 4(C)]. Since we were able to update the array element locations to account for storm-induced movement, the HLA beamforming results were considered reliable throughout the deployment.

5. Signal and noise level estimation

Acoustic data were calibrated using technical specifications of the recording systems as described by Merchant *et al.* (2015). Signal level, noise level, and signal-to-noise ratio (SNR) were estimated for each call on the glider and buoy. The signal level was defined as the median power spectral density (PSD) (dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$) within the time-frequency “bounding box” of the call assigned during manual review of the audio data. The noise level was determined by calculating the median PSD within the same frequency and duration of the call at each time step within a 30 s audio

snippet centered on each call, and then selecting the lowest median PSD of this 30 s period. This was done to avoid including transient impulsive signals in the noise level estimate. The median PSD for both the signal and noise levels was calculated by computing a spectrogram (2000 sample DFT, 50% overlap, Hann window), collating all the time-frequency cells within the bounding box, and extracting the median from the distribution. The SNR (in dB) was defined as the difference between the signal level and noise level. Signal levels that were contaminated by transient impulsive sounds were rejected and not used to calculate SNR.

D. Platform detection probability

The array was used as the reference for comparison between the DMON/LFDCS single hydrophone platforms. For each call detected and localized on the array, a score of zero was assigned if the call was not detected and one if the call was detected in the pitch track data generated by the DMON/LFDCS on the buoy. The same scoring protocol was applied to the DMON/LFDCS on the glider. This scoring protocol was used for both the pitch tracks that were available in near real-time and the archival audio that was available after platform recovery. Two separate analyses were conducted for the pitch track data to inform how the review protocol affects the probability of detection. These protocols differed in their treatment of calls scored as “possibly detected.” The first used a conservative protocol in which the “possibly detected” calls were treated as if they were scored as “not detected.” This protocol was designed to minimize false detections at the expense of increased missed detections. It has been extensively employed on deployments in the NW Atlantic (e.g., Baumgartner *et al.*, 2019) and is therefore the primary focus of this study. The second was a precautionary protocol in which the “possibly detected” pitch tracks were treated as if they were “detected.”

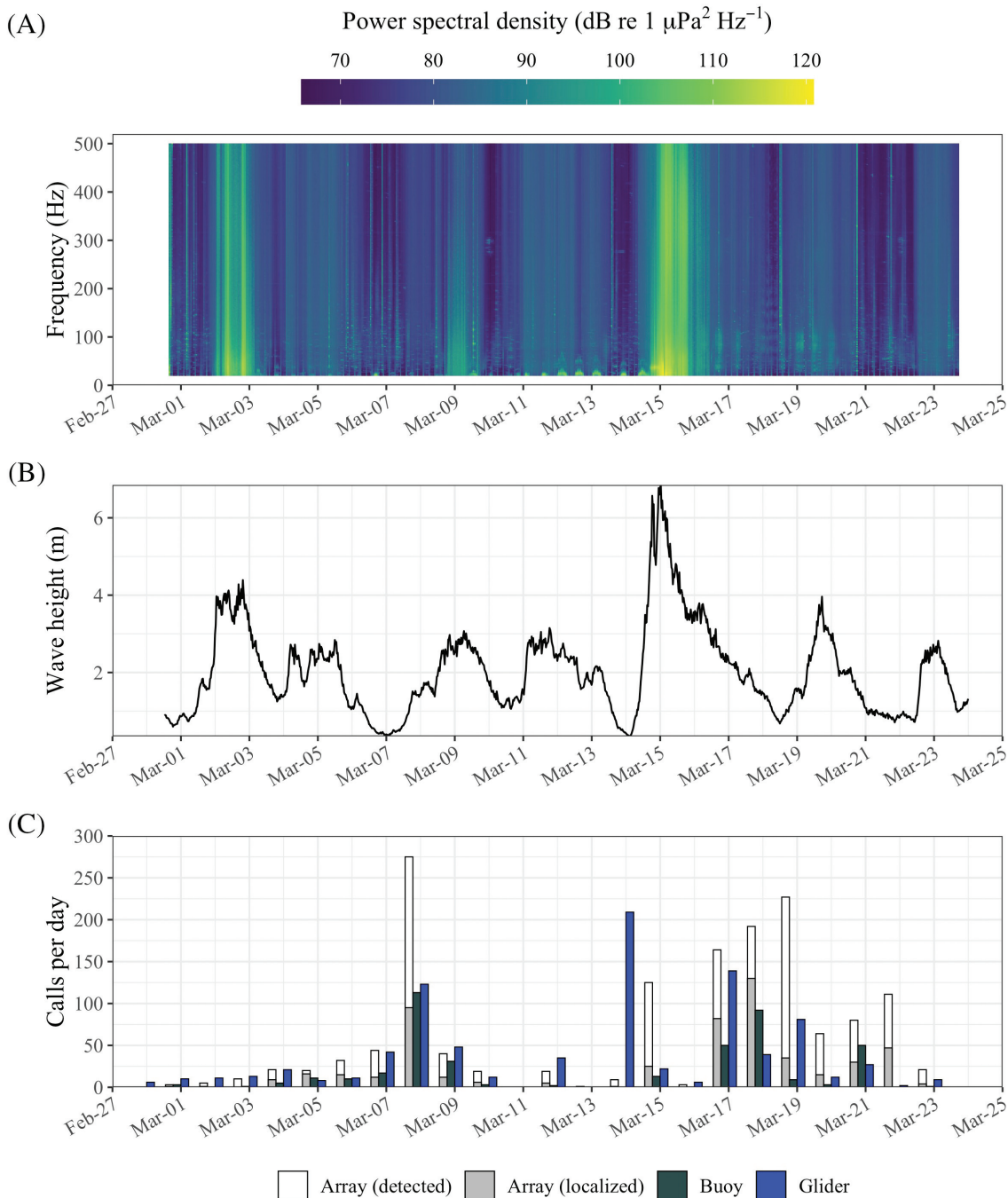


FIG. 4. (A) Power spectral density ($\text{dB re } 1 \mu\text{Pa}^2 \text{Hz}^{-1}$) of channel 1 of the horizontal line array computed from 1 s time segments averaged to 1 h resolution via the Welch method. (B) Hourly observations of significant wave height from the Block Island meteorological buoy $\sim 10 \text{ km}$ SW of the study site. (C) Daily counts (calls per day) of right whale upcalls detected in the array audio (white bars; $n = 1485$), the buoy pitch tracks (black bars; $n = 414$), and the glider pitch tracks (blue bars; $n = 886$), as well as numbers of calls that were successfully localized (gray bars; $n = 541$).

This protocol was designed to minimize missed detections at the expense of increased false detections. The archival audio data were only scored using the conservative protocol; there were too few calls scored as “possibly detected” in the review of the archival audio data ($n = 5$ for the buoy; $n = 3$ for the glider) to justify a protocol comparison.

The detection probability of the DMON/LFDCS single hydrophone platforms, $P_S(R)$, at range R was defined as follows:

$$P_S(R) = N_S(R) / N_A(R), \quad (1)$$

where $N_S(R)$ is the number of localized calls detected by the single-hydrophone platform (i.e., the buoy or the glider) at range R and $N_A(R)$ is the total number of localized calls at range R . Critically, the detection probability of the DMON/LFDCS platforms was evaluated using only calls that were first detected and then localized by the array. This approach requires that each system be analyzed independently and

that the single-hydrophone platforms are not used in the localization process, both of which are met here. It also assumes that the detectability of localized calls is representative of the detectability of all calls on the array. This assumption may be violated, as localization typically requires higher SNR than detection (e.g., Thode *et al.*, 2012). The median SNR of detected calls that were not localized by the array (2.3 dB; IQR: 2.7 dB) was 0.3 dB lower than the median SNR of localized calls (2.6 dB; IQR 2.3 dB) but results of a Mann–Whitney U test failed to reject the null hypothesis that both distributions are equal ($p = 0.112$) (see the supplementary material⁵). Though there is little direct evidence that this assumption is violated, the most conservative approach would be to interpret the results presented here as an estimate of the upper bound of the detection function of each single hydrophone system in this environment. The true detection function will likely be somewhat reduced depending on variations in source level and depth distributions of the underlying calls.

Detection functions for each DMON/LFDCS platform were quantified using logistic regression analysis. The series of detected/not-detected scores was used as the dependent variable. Candidate models were constructed using various combinations of detection range, noise level, and glider depth as independent variables. The glider depth term was used in the glider analysis only and was expressed in a parabolic form, based on the observed relationship between glider depth and proportion of calls detected [Fig. 6(F)] (see supplementary material⁵). The influence of autocorrelation in detected calls was deemed minimal, based on a preliminary analysis using generalized estimating equations with a first order autoregressive covariance structure implemented with the *geepack* package in R (Halekoh *et al.*, 2006). SNR was not used as a model covariate because it was correlated with both range and noise level. The most parsimonious model was selected using Akaike Information Criterion (AIC). Wald tests were used to evaluate the contribution of each independent variable to the overall model. Drop-in-deviance tests were used to compare among models. Separate logistic regressions were conducted for the buoy and the glider using scores from (1) pitch tracks scored using the conservative protocol, (2) pitch tracks scored using the precautionary protocol, and (3) archived audio (i.e., six logistic regressions were conducted). The fitted logistic regressions were used to estimate the probability of detecting a localized call at a given range, noise level, and glider depth, and also used to compute the effective detection radius (EDR) as described by Buckland *et al.* (2004). For all undetected calls, we also examined the buoy and glider audio and pitch track records to determine why they were not detected by the analyst and DMON/LFDCS, respectively.

All analyses were conducted using MATLAB (The Mathworks Inc., Natick, MA) and R (R Core Team, 2019) programming languages. Analyses in R were conducted using the *oce* (Kelley and Richards, 2019), *shiny* (Chang *et al.*, 2019), and *tidyverse* (Wickham *et al.*, 2019) packages, and visualizations in R were created using the *ggplot2* package (Wickham, 2016).

III. RESULTS

A. Call detection and localization

A total of 1485 right whale upcalls were detected by the array between 28 February and 24 March. The DMON/LFDCS on the glider and buoy pitch tracked (i.e., detected) 886 and 414 right whale upcalls, respectively, during the same period. Calls occurred throughout the monitoring period but were especially abundant on 8 March and from 17 through 19 March [Fig. 4(C)]. Of the calls detected on the array, 36% (541 of 1485) could be confidently localized such that the back-propagated modes converged at a single range. There were several calls with potential broad-side bearing ambiguity, but range estimates using either bearing were consistent, likely owing to the relatively uniform bathymetry at the site, so these calls were retained in the analysis. The spatial distribution of localized calls was not uniform; most calls originated from the area south of the DMON/LFDCS buoy and the array (Fig. 5). The distances to localized calls from each platform ranged from 0.4–30.1 km on the glider (median = 5.3 km), and from 0.3–29.7 km on the buoy (median = 6.2 km). Noise levels associated with calls ranged from 83.9–108 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$ (median = 99.4 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$) on the glider and from 85.2–110 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$ on the buoy (median = 98.9 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$). The depth of the glider at the time of call reception ranged from 0.62–32.0 m with a median of 13.6 m (Fig. 6).

B. Platform detection probability

For the buoy, the proportion of localized calls detected using pitch tracks and the conservative protocol generally decreased with range [Fig. 6(A)]; 55.0% of localized calls within 5 km (111/202) were detected while 21.1% of localized calls between 15 and 40 km (4/19) were detected. The proportion of localized calls detected also decreased with noise [Fig. 6(C)]; 46.0% of localized calls received in noise levels below 100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$ (64/139) were detected, while 24.0% of localized calls received in louder noise conditions (24/100) were detected. A SNR of more than 3 dB was required to detect at least 50% of localized calls [Fig. 6(G)]. Calls were missed for a variety of reasons: of the 541 localized calls, 46.6% were missed due to absent or poor pitch tracks, 4.6% were missed due to interfering biological sounds (i.e., humpback whale song), 5.7% were missed due to interfering non-biological sounds (e.g., platform noise, ship noise), and 3.9% were missed due to analyst error in scoring the pitch tracks (Table I) [Fig. 7(A)].

The logistic regressions for the buoy were conducted using the 239 localized calls for which archival audio were available and noise levels could be calculated. The most parsimonious logistic regression model and subsequent significance testing provided evidence that the probability of detecting localized calls was negatively related to both range and noise level for all analyses (Table II). In average noise conditions (100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$), the fitted regression suggested that a probability of detection of 0.5 [95%

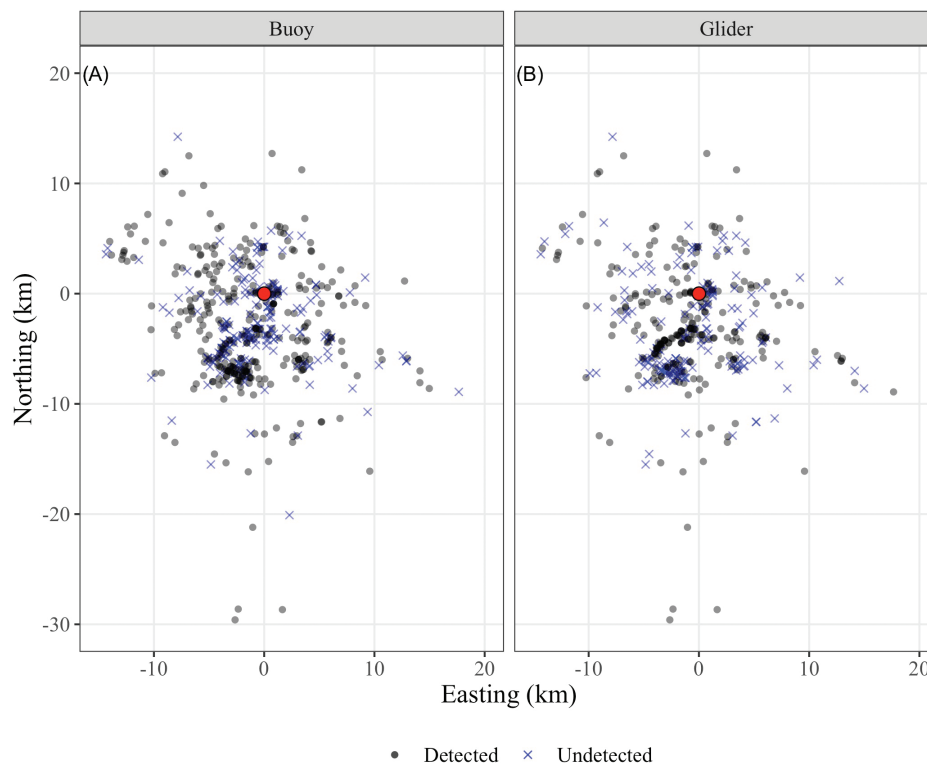


FIG. 5. The spatial distribution of localized right whale upcalls that were either detected (gray circles) or not detected (blue crosses) by the buoy (panel A; $n = 541$) or the glider (panel B; $n = 426$) in the near real-time pitch track record using a conservative protocol. The red circle at the origin indicates the location of the array. [See supplementary material for analogous results using a precautionary protocol or archival audio (footnote 5).]

confidence interval (CI): 0.385–0.613] occurred at 2.3 km and the effective detection radius was 8.3 km (Fig. 8) (Table III). In low noise conditions (95 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$), the range to a probability of detection of 0.5 increased to 7.6 km. A probability of detection of 0.5 was not achieved at any range in high noise conditions (Fig. 8).

For the glider, many localized calls ($n = 115$) occurred during periods when the glider produced acoustic noise during activation of the buoyancy pump during profiling inflections (typically 30 s every 3.5 min in 30–35 m depths for the glider used in this study) or electrical noise during satellite communications at the sea surface (typically 10–15 min every 2–2.25 h, or $\sim 12\%$ of the deployment). The LFDSCS recognizes these periods of noise and terminates pitch tracking during them. As calls during these periods were not available for detection and therefore not useful in determining our assessment of the effect of range on the accuracy of the DMON/LFDSCS, they were excluded from the analysis. The distribution of these excluded calls was uniform with respect to range. The proportion of the remaining 426 localized calls detected using the pitch tracks and the conservative protocol decreased with range [Fig. 6(B)]; 51.6% of localized calls within 5 km (95/184) were detected while 18.8% of localized calls between 15 and 40 km (3/16) were detected. There was a parabolic relationship between the proportion of calls detected and glider depth, with the greatest proportion of calls detected at mid depths and lower proportion detected near the surface and bottom [Fig. 6(F)] (see supplementary material for figures⁵). This was not explained by the proportion of time spent at depth, which was relatively uniform for 0–25 m. The influence of noise on the empirical proportion of detected calls was not obvious for the pitch track analyses [Fig. 6(D)], but increasing

noise was associated with a decrease in the proportion of detected calls for the archival audio analysis (see supplementary material for figures⁵) and the proportion of detected calls increased with SNR in all analyses [Fig. 6(H)]. An SNR of more than 5 dB was required to detect at least 50% of localized calls [Fig. 6(G)]. Calls were missed for a variety of reasons: of the 426 available calls, 36.6% were missed due to absent or poor pitch tracks, 9.6% were missed due to interfering biological sounds, 11.2% were missed due to interfering non-biological sounds, and 0.9% were missed due to analyst error in scoring the pitch tracks (Table I) [Fig. 7(B)].

The most parsimonious logistic regression models for the glider provided evidence that the probability of detecting localized calls was related to range, noise level, and glider depth for all analyses (Table II). At an average noise level (100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$) and glider depth (15 m), the fitted regression suggested that a probability of detection of 0.5 (95% CI: 0.430–0.569) occurred at 6.8 km and the effective detection radius was 15.2 km (Fig. 8; Table III). In low noise conditions (95 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$) the range to a probability of detection of 0.5 increased to 9.0 km, while in high noise conditions (100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$), the range decreased to 4.5 km (Fig. 8). (See supplementary material for results for analogous analyses of the pitch tracks using the precautionary protocol and of the archived audio recordings from the glider and buoy.⁵)

IV. DISCUSSION

A. Estimating and reporting detection range

Several previous efforts have succeeded in ranging and localizing baleen whale calls for purposes, such as density estimation (e.g., Harris *et al.*, 2013), call attribution

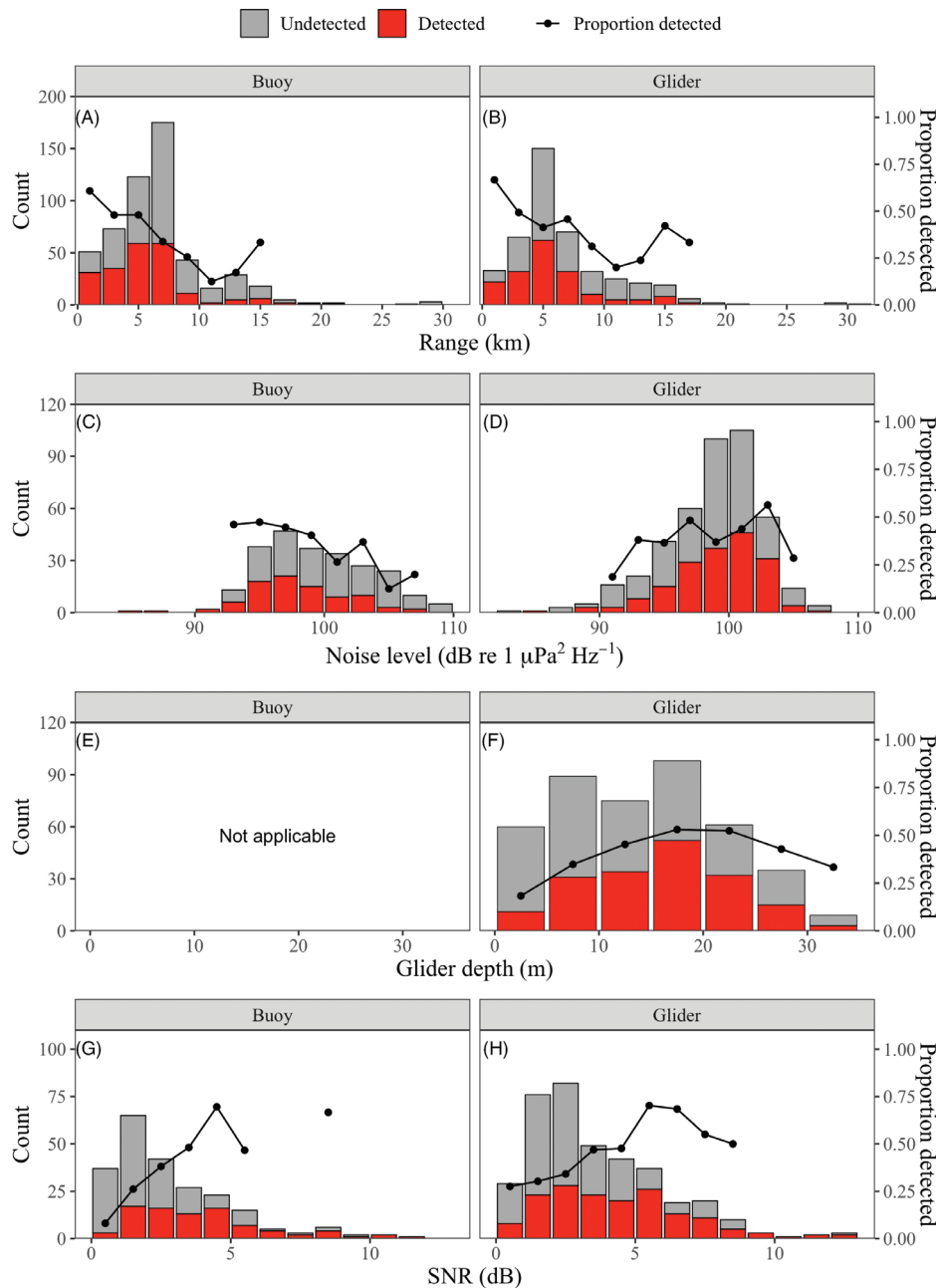


FIG. 6. Distribution of ranges (A) and (B), noise levels (C) and (D), glider depths (E) and (F), and signal-to-noise ratios (SNR) (G) and (H) from the buoy (left column) and glider (right column) of right whale upcalls localized by the array ($n=541$ buoy, $n=426$ glider) and detected via near real-time pitch track analysis using the conservative protocol. Total number of localized calls are shown in gray and localized calls detected in near real-time are shown in red in each bin. The black line shows the proportion of localized calls detected in bins with more than five calls. [See supplementary material for analogous results using the precautionary protocol or archival audio (footnote 5).]

TABLE I. Results from manual scoring of glider and buoy pitch track records of calls localized by the array using a conservative protocol (total number of calls = 541). Here, n refers to the number of calls, while % is the percentage of total localized calls available for detection (i.e., does not consider excluded calls). See supplementary materials for analogous results using the precautionary protocol or archival audio (footnote 1).

Score	Definition	Glider		Buoy	
		n	%	n	%
Absent	Calls were not pitch tracked at all because of low amplitude	72	16.9	102	18.9
Poor	Calls were not pitch tracked accurately/completely because of low amplitude or poor shape	84	19.7	150	27.7
Song	Uncertainty due to interfering species calls	41	9.6	25	4.6
Noise	Calls were not pitch tracked accurately/completely because of interfering sound	48	11.2	31	5.7
Missed	Human error (analyst chose wrong score erroneously)	4	0.9	21	3.9
Detected	Calls were pitch tracked and scored as detected by analyst	178	41.7	212	39.2
Exclude	Calls were not available for pitch tracking because the platform was not monitoring	114	N/A	0	N/A

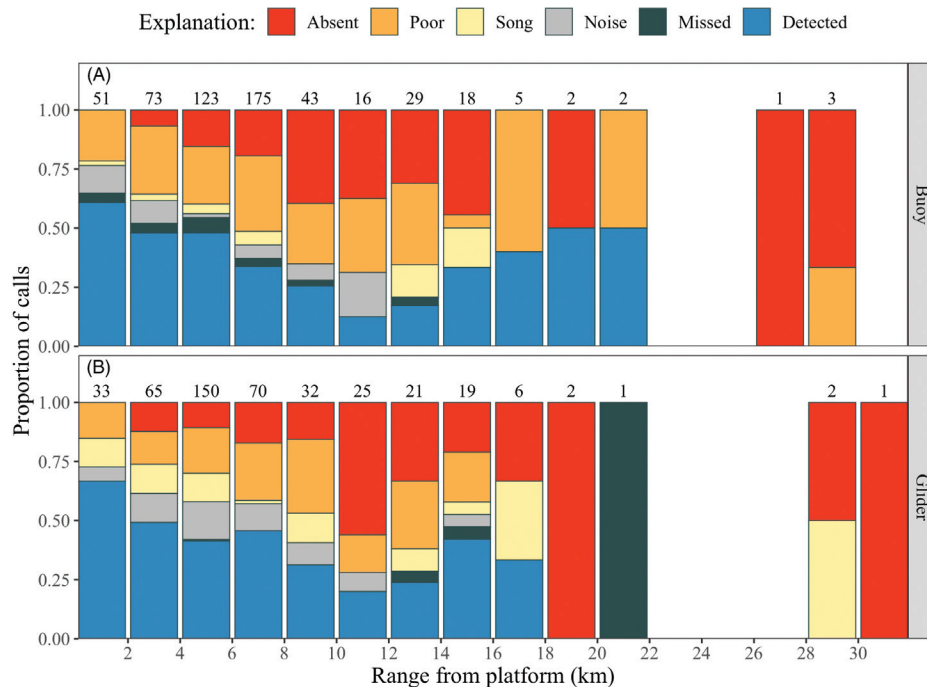


FIG. 7. Proportion of localized calls assigned to each score category based on a conservative review of the near real-time pitch track data as a function of range from the buoy (panel A) and glider (panel B). Colors indicate the proportion of calls of a given score in 2 km range bins, while the number of calls in each bin is shown above each bar. Definitions of each category are provided in Table I. [See supplementary materials for analogous results using the precautionary protocol or archival audio (footnote 5).]

(e.g., Baumgartner *et al.*, 2008), or measuring noise impacts (e.g., Thode *et al.*, 2020). Few studies have attempted to quantify the probability of detecting localized calls. To our knowledge, this is the first study to empirically derive a detection function for right whales, and as such, there is a paucity of other observations available for comparison. Laurinoli *et al.* (2003) localized tonal right whale calls to a maximum range of approximately 29 km in the Bay of Fundy, and Thode *et al.* (2017) observed a maximum range to a North Pacific right whale (*Eubalaena japonica*) upcall of approximately 30 km in the Bering Sea. These observations are similar to the maximum range of a localized call from our study (~ 30 km), but greater than the observed maximum detection range of the DMON/LFDCS on the glider or buoy (~ 20 km). The estimated probability of detection at 30 km is low, but non-zero, so it is possible that we did not have a large enough sample size to detect a call at 30 km. Tennessen and Parks (2016) used a modeling approach to estimate a maximum propagation distance of approximately 16 km for a right whale upcall in optimal noise conditions ($85 \text{ dB re } 1 \mu\text{Pa}^2 \text{ Hz}^{-1}$) in the Bay of Fundy. This is lower than the maximum detection ranges observed in our study and in the Laurinoli *et al.* (2003) study. The reasons for such discrepancies are unclear, which highlights the challenges in simulating detection range.

Clark *et al.* (2010) conducted an excellent study of right whale upcalls in Cape Cod Bay, a shallow habitat similar to our study area, and they stated that the “acoustic detection area was reliably found to be within a range of approximately 9 km (~ 5 nmi) from a recorder” (Clark *et al.*, 2010). The comparability of our observations to this acoustic detection range estimate depends on the definition of “reliably.” If the definition of the acoustic detection range is the range at which the probability of detecting a calling whale is 0.5

(i.e., “reliable” is defined as a 1 in 2 chance of detection), then our range estimates are shorter than those of Clark *et al.* (2010). However, if we define the detection range as the range at which the probability of detecting a calling whale is 0.33 (i.e., “reliable” is defined as a 1 in 3 chance of detection), then our detection range estimates are similar to those of Clark *et al.* (2010). Finally, if Clark *et al.* (2010) were reporting a maximum detection range, then our maximum estimated detection range of more than 30 km exceeds the detection range that they reported. We present this comparison to highlight something that is likely obvious, but perhaps underappreciated: the use of a single number for detection range is an incomplete description of the area that is effectively monitored by a passive acoustic system. From our study, we estimated that whales calling at >20 km can be detected in near real time by the DMON/LFDCS system carried aboard either a glider or a buoy, but the chances of a calling whale being detected at those distances are low. The detection function, or the curve describing the range-dependent probability of detection (Fig. 8), is a near-complete description of the site-, environment-, and species-specific detection range of a PAM system. Efforts should be made to estimate and report the detection function whenever possible, as it provides a vastly more accurate and appropriate description of a system’s detection range than a single number.

B. Factors influencing detection range

Our results indicate that other covariates in addition to range, such as noise level and platform depth, play an important role in the probability of detecting a call. The performance of both DMON/LFDCS platforms was significantly reduced in louder noise environments. The buoy was especially sensitive to noise, where an increase in

TABLE II. Selection and statistical evaluation of candidate logistic regression models describing the probability of detection of the glider and buoy. The logistic regressions used scores as the dependent variable derived from pitch track analysis with a conservative protocol, pitch track analysis with a precautionary protocol, and archival audio analysis. The full models (highlighted in gray) were the most parsimonious for all analyses. Formulas for each candidate model are as follows. G1: score \sim range. G2: score \sim range + noise. G3: score \sim range + glider_depth + glider_depth². G4: score \sim range + noise + glider_depth + glider_depth². B1: score \sim range. B2: score \sim noise. B3: score \sim range + noise.

Platform	Analysis	Model*	AIC	Wald tests			Drop-in-deviance tests	
				Variable	Coefficient	P-value	Test	P-value
Glider (n = 426)	Real-time pitch tracks (conservative protocol)	G1	272.5	range	−0.19	<0.001*	NA	NA
		G2	274	range	−0.19	<0.001*	G2 v G1	0.454
				noise	−0.03	0.48		
		G3	248.1	range	−0.2	<0.001*	G3 v G1	<0.001*
				glider_depth	0.35	<0.001*		
				glider_depth ²	−0.01	<0.001*		
		G4	235.7	range	−0.23	<0.001*	G4 v G2	<0.001*
				noise	−0.18	<0.001*		
				glider_depth	0.5	<0.001*	G4 v G3	<0.001*
				glider_depth ²	−0.01	<0.001*		
	Real-time pitch tracks (precautionary protocol)	G1	306.6	range	−0.17	<0.001*	NA	NA
		G2	307.4	range	−0.17	<0.001*	G2 v G1	0.22
				noise	−0.04	0.29		
		G3	281.7	range	−0.19	<0.001*	G3 v G1	<0.001*
				glider_depth	0.35	<0.001*		
				glider_depth ²	−0.01	<0.001*		
		G4	268.4	range	−0.21	<0.001*	G4 v G2	<0.001*
				noise	−0.17	<0.001*		
				glider_depth	0.48	<0.001*	G4 v G3	<0.001*
				glider_depth ²	−0.01	<0.001*		
	Archival audio	G1	278.1	range	−0.15	<0.001*	NA	NA
		G2	259.7	range	−0.19	<0.001*	G2 v G1	<0.001*
				noise	−0.22	<0.001*		
		G3	262.6	range	−0.17	<0.001*	G3 v G1	<0.001*
				glider_depth	0.31	<0.001*		
				glider_depth ²	−0.01	<0.001*		
		G4	225.2	range	−0.23	<0.001*	G4 v G2	<0.001*
				noise	−0.33	<0.001*		
				glider_depth	0.45	<0.001*	G4 v G3	<0.001*
				glider_depth ²	−0.01	<0.001*		
Buoy (n = 239)	Real-time pitch tracks (conservative protocol)	B1	143.7	range	−0.24	<0.001*	NA	NA
		B2	125.2	noise	−0.32	<0.001*	NA	NA
		B3	91.7	range	−0.42	<0.001*	B3 v B1	<0.001*
				noise	−0.48	<0.001*	B3 v B2	<0.001*
	Real-time pitch tracks (precautionary protocol)	B1	164.7	range	−0.21	<0.001*	NA	NA
		B2	144	noise	−0.31	<0.001*	NA	NA
		B3	110.9	range	−0.37	<0.001*	B3 v B1	<0.001*
				noise	−0.43	<0.001*	B3 v B2	<0.001*
	Archival audio	B1	33.6	range	−0.16	0.23	NA	NA
		B2	21.6	noise	−0.88	0.02*	NA	NA
		B3	9.8	range	−9.25	0.46	B3 v B1	<0.001*
				noise	−10.5	0.44	B3 v B2	<0.001*

10 dB re 1 $\mu\text{Pa}^2 \text{Hz}^{-1}$ translated to a nearly 50% reduction in the probability of detecting a call at 5 km, compared to a 20% reduction on the glider. These results emphasize the importance of considering the impact of noise on the interpretation of PAM results, which, if left unaccounted for, can introduce artificial trends in detection results (e.g., Helble *et al.*, 2013b; Fregosi *et al.*, 2020).

The improved probability of detection and reduced sensitivity to noise of the glider relative to the buoy is likely

due in large part to differences in platform depth, and, by extension, transmission loss. Transmission loss refers to the decrease in acoustic intensity due to spreading or attenuation as a sound propagates. During March (i.e., prior to the onset of stratification), transmission loss generally varies parabolically with depth in this environment, with the lowest values in the middle of the water column and the highest values near the top and bottom boundaries (Fig. 9). This agrees well with the distribution of detections versus glider depth,

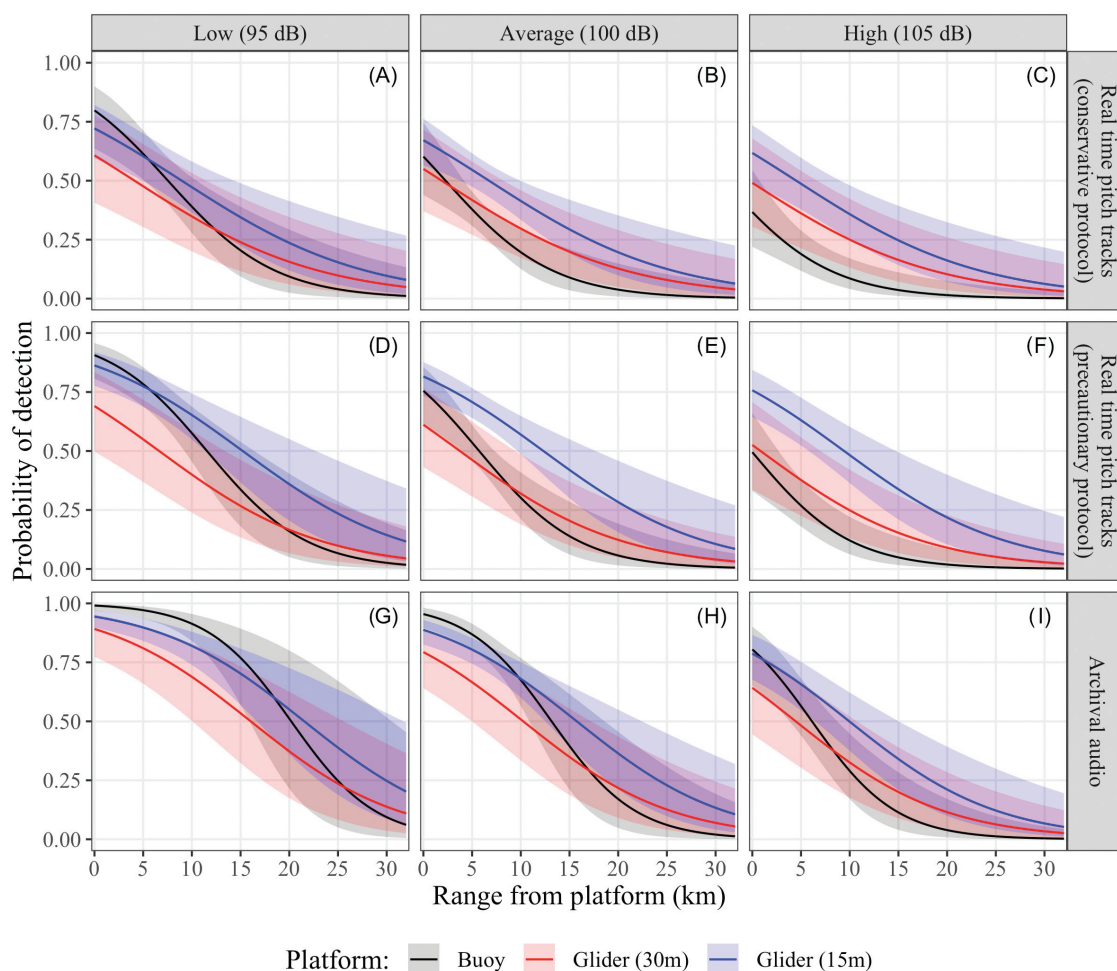


FIG. 8. Estimated probability of detection of localized right whale upcalls as a function of range to the buoy (black lines) and glider at a depth of either 15 m (blue lines) or 30 m (red lines) at low (95 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$; left column), average (100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$; middle column), and high (105 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$; right column) noise levels based on the conservative (top row) or the precautionary (middle row) analyses of near real-time pitch track data, or based on the manual review of archival audio (bottom row). The fitted regression models are shown as solid lines, while the 95% confidence intervals are shown as shaded regions. [See supplementary material for an alternate representation of these data showing the influence of noise levels by platform (footnote 5).]

where a higher proportion of calls were detected when the glider was located in the middle of the water column rather than near the surface or the bottom [Fig. 6(F)] (see supplementary material for photograph⁵). When the glider depth covariate was fixed at 15 m, the depth stratum with minimum transmission loss (Fig. 9), the logistic regressions

suggested the performance of the glider was consistently better than that of the buoy. In contrast, when the glider depth covariate was fixed to a value of 30 m, the same depth as the hydrophone on the buoy, the logistic regressions suggested that the performance of the platforms was nearly identical (Fig. 8). Thus, the regular vertical profiling of the

TABLE III. The ranges (in km) for a given probability of detecting a right whale upcall from the glider or the buoy estimated from the most parsimonious logistic regression models (Table II). Noise was fixed at an intermediate value of 100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$. Glider depths were set to either 15 or 30 m. The bottom row shows the effective detection radius (EDR) computed using a 40 km truncation range.

Probability	Range (km)								
	Real-time pitch tracks (conservative protocol)			Real-time pitch tracks (precautionary protocol)			Archival audio		
	Glider (15 m)	Glider (30 m)	Buoy	Glider (15 m)	Glider (30 m)	Buoy	Glider (15 m)	Glider (30 m)	Buoy
0.5	6.8	1.9	2.3	12.3	3.8	5.8	15.8	10.2	13.2
0.33	13.4	8.6	6.2	18.2	9.6	9.3	21.2	15.7	16.2
0.25	17.1	12.2	8.3	21.4	12.9	11.3	24.1	18.6	17.9
0.1	27.5	22.6	14.4	30.6	22	16.9	32.5	27	22.7
EDR	15.2	12.5	8.3	17.9	12.4	10.1	19.9	15.9	15.2

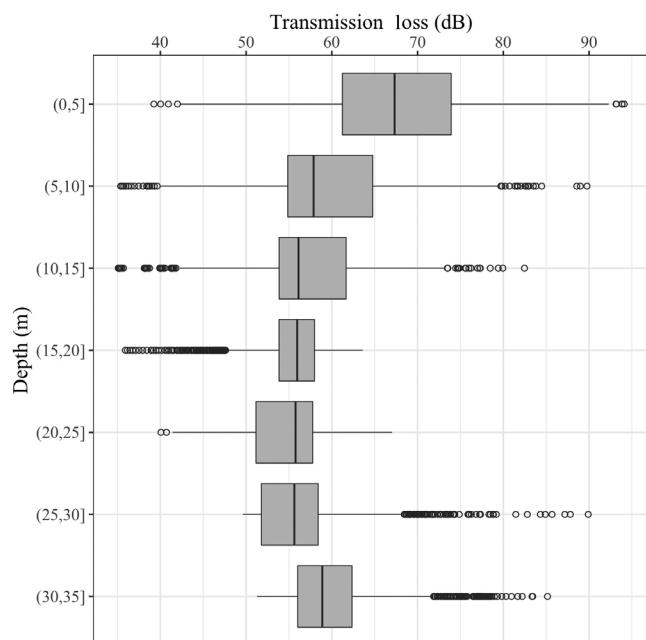


FIG. 9. Depth distribution of transmission loss estimated using the KRAKEN normal mode model parameterized with a 100 Hz frequency source at 15 m depth, two propagating modes, and environmental conditions consistent with those of the study site. Transmission loss estimates were computed at each point in a grid with 250 m range and 0.5 m depth resolution to a maximum range and depth of 35 km and 35 m, respectively, and then aggregated into 5 m depth bins.

glider through regions of low transmission loss at intermediate depths confers a performance improvement relative to the buoy.

The probability of detection also varied depending on the data source (real-time pitch tracks versus archival audio), platform, and detection protocol. The probability of detecting a call was lower using near real-time pitch track data compared to archival audio data. This is not surprising given that pitch tracks are an abstraction of the audio and do not possess all of the cues that an analyst would use to confidently detect and classify a call. Inspection of spectrograms and aural review of audio by an experienced analyst is considered the gold standard for detecting marine mammal calls in passive acoustic monitoring data (e.g., Baumgartner and Mussoline, 2011). The difference between the audio and pitch track results indicates the cost of relying on an automated detector and not directly reviewing the audio; improvements could potentially be made by either upgrading the existing detection system or using a different detector (e.g., Simard *et al.*, 2019; Kirsebom *et al.*, 2020). These differences are most pronounced at close ranges. The majority of the undetected calls within 5 km of both platforms were not detected because of poor pitch tracks, meaning that the pitch tracks were present but could not be confidently classified by the analyst because of poor shape or low amplitude. Poorly formed pitch tracks from close range calls could be due to competing biological- and platform-related noise processes, receiver depth, or variation in source levels (e.g., Parks and Tyack, 2005). A higher proportion of

close-range calls was missed by the glider due to the presence of competing biological signals (humpback whale song). As with right whale calls, these signals may have been more detectable on the glider owing to its vertical profiling through depth strata characterized by low transmission loss.

Our results demonstrate that the detection function of a given platform changes depending on the analysis protocol. The near real-time detection protocol currently used with DMON/LFDCS gliders and buoys is conservative by design in recognition of the high costs associated with triggering a management measure based on a false-positive detection (Baumgartner *et al.*, 2019). This protocol can be relaxed to reduce missed detections at the expense of allowing false positives. This would represent a more precautionary management approach. We examined the use of a precautionary protocol where right whale calls that were scored as “possibly detected” were considered “detected” in the estimation of the detection function. We found that employing a precautionary protocol increased the real-time probability of detection by approximately 15% at close ranges (≤ 10 km) (see supplementary material⁵) (Fig. 8). In other words, the call detection protocol directly influences the detection range of a platform and thus must be designed with care and implemented with consistency.

The real-time detection rates reported here are much lower than those previously presented for the glider (Baumgartner *et al.*, 2013; Baumgartner *et al.*, 2020) or buoy (Baumgartner *et al.*, 2020) (i.e., the missed detection rates in this study are higher than previously reported). In this study, pitch tracks were reviewed on a call-by-call basis to determine if individual localized calls were detected or not detected by the glider or buoy. In contrast, the near real-time validation protocol operates on nominal 15 min “tally periods” to determine whale occurrence (Baumgartner *et al.*, 2013; Baumgartner *et al.*, 2019; Baumgartner *et al.*, 2020). The tally period approach is more robust to missing occasional calls when individual whales are calling repeatedly within a tally period. For example, if the probability of missing a single call is 0.5, the probability of missing three calls in a row is reduced to 0.125 ($0.5 \times 0.5 \times 0.5 = 0.125$) if the calls are independent. The probability of detecting at least one of those three calls is 0.875 ($1 - 0.5 \times 0.5 \times 0.5 = 0.875$). As a simple thought experiment, we can apply this logic to the per-call detection functions determined in this study to illustrate how the probability of detection may change when considering multiple calls within a given time period. The assumption of independence of calls is almost certainly violated due to correlation in a number of factors (e.g., calling behavior, background noise levels, interference from other species), so the probabilities of detection in this thought experiment are likely overestimated. However, the key concept is that the detection function of a given platform changes based on the number of calls available for detection in a given period. If the assumption of independence was not violated, for example, attempting to detect one of two available calls in average noise conditions on the buoy increases the range to the

0.5 probability of detection from 2.3–7.1 km, while attempting to detect one of 5 or 10 available calls increases this range to 12.8 or 16.8 km, respectively (Fig. 10). The probability of detection of a single call in a tally period is dependent on the number of calls that are available; however, it is nearly impossible to know the number of calls that are available for detection in a tally period, particularly when calling rates likely vary widely depending on location, time of year, whale density and whale behavior. Consequently, the probability of detection results reported here for single calls (i.e., attempting to detect one of one available call) (Fig. 8) should be considered a minimum estimate that is likely improved substantially, but by an unknown amount, by using a tally period approach when the goal of monitoring is to assess right whale occurrence over time scales that are longer than instantaneous (i.e., the goal is not to detect every call at all times, but to detect occurrence over, say, daily time scales).

C. Assumptions and caveats

We chose to calculate the detection function empirically to avoid making assumptions about the source (e.g., source depth, level, frequency), environmental (e.g., ambient or platform noise level), or detector characteristics (e.g., detection threshold). A potential source of bias in our approach is that each call used to estimate the detection functions of the buoy or glider first had to be detected and localized on the array. Results from a simple simulation suggest that the methodology employed here is robust to implicit bias introduced by imperfect array detection and localization (see supplementary material⁵). We also assumed that the detection functions were logistically monotonic and thus well represented by the logistic regression model. There was a slight increase in the proportion of calls detected beyond 10 km, but this was driven by a very small number of calls, so we do not consider this assumption to be violated [Figs. 6(A) and 6(B)] (see supplementary material⁵). We also used generalized additive models (GAMs) to estimate the probability of detection (not shown) and found

that the shapes of the resulting functions were well represented with logistic functions. The detection function may take on non-monotonic shapes in more complex propagation environments (e.g., Helble *et al.*, 2013a), but this is unlikely in the relatively range-independent conditions of this study site.

Our analysis only considers calls that were available for detection by the glider or buoy, meaning they were received when each platform was actively monitoring. The buoy produced pitch tracks continuously but only recorded audio 50% of the time. In contrast, pitch track and audio data were not available approximately 12% of the time for the glider due to noise associated with surfacing or inflecting. We configured the glider to surface and inflect at this rate to facilitate shallow water navigation and the reporting of detection results every ~ 2 h, but these parameters can be adjusted depending on the environment and monitoring objectives. We do not make an effort to correct for differences in duty cycling between platforms. Scientists or regulators seeking to employ the DMON/LFDCS on one of these platforms for a particular application should consider the relative differences in monitoring effort between platforms in mission planning.

The single-station ranging method we employed does require some assumptions to be made about signal transmission and the propagation environment. The assumptions include range-independent sound speed and bottom type, and the assumption that the propagation of the calls was well approximated by normal mode theory. These assumptions are likely justified, as numerous studies have demonstrated the efficacy of normal mode ranging of low-frequency signals in shallow water environments (e.g., D'Spain *et al.*, 1997; Thode *et al.*, 2000; Wiggins *et al.*, 2004; Thode *et al.*, 2006; Munger *et al.*, 2011; Newhall *et al.*, 2012; Abadi *et al.*, 2014; Bonnel *et al.*, 2014; Thode *et al.*, 2017). We made efforts to account for variation in bathymetry by using a range- and bearing-dependent back-propagation method. The array- and glider-based environmental sensors revealed that the water column was entirely

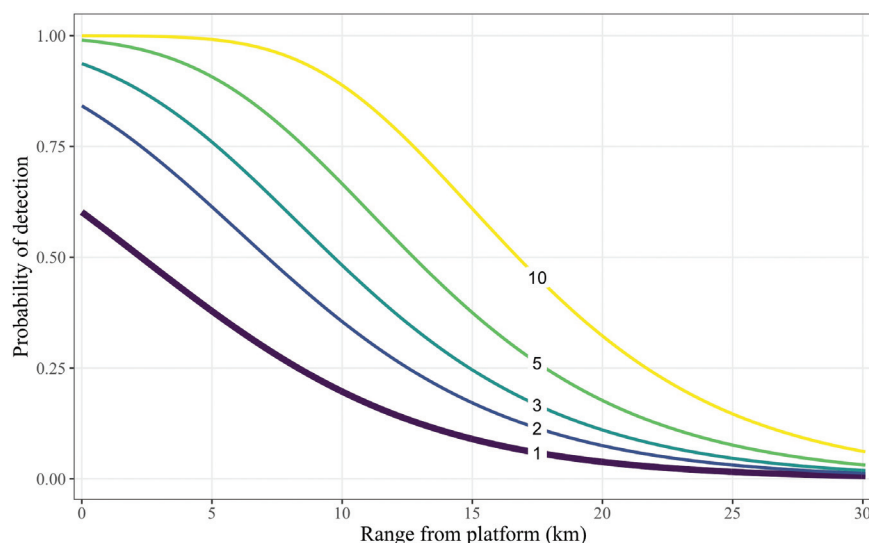


FIG. 10. Results of a thought experiment showing the probability of detecting one of a given number of available right whale upcalls as a function of range to the buoy based on the conservative analysis of near real-time pitch track data and a fixed noise level of 100 dB re $1 \mu\text{Pa}^2 \text{Hz}^{-1}$. Each line shows the probability of detecting one call out of 1, 2, 3, 5, or 10 available calls during a fixed period of time. This analysis relies upon the unlikely assumption that calls are detected independently, so the probabilities of detection are likely overestimated (see Sec. IV B in the text). [See supplementary material for results from all combinations of platform, data source and protocol (footnote 5).]

mixed throughout the study, so depth-varying sound speed was unlikely to contribute to ranging error. Conventional long-baseline array localization methods would require similar simplifying assumptions about the propagation environment (e.g., uniform bathymetry, constant sound speed).

Our analysis made no attempt to quantify the likelihood that a right whale will produce a call; the probability of detection examined here assumes a call is already available to be detected. Call types, rates, depths, and spectral characteristics (e.g., frequency, amplitude) vary depending on the time of day, season, location, environment, behavior, and individual. Some of this variability has been characterized for right whales (e.g., Parks *et al.*, 2011a, 2011b) but small sample sizes have often precluded range-wide characterization.

Future efforts should be made to improve array detection and localization to increase sample size. We did not attempt to quantify the probability of detection for the array, but the success rate for localizing calls of 36% (541 of 1485) was similar to the success rate of Laurinolli *et al.* (2003) for loud tonal sounds in the Bay of Fundy using traditional cross correlation and time difference of arrival methods and was substantially higher than in other studies (e.g., Cummings and Holliday, 1986). Many detected calls could not be localized owing to noise on one or more VLA or HLA channel(s) that prevented accurate mode filtering or beamforming, respectively. Improvements in array design and mooring configuration to reduce platform noise, as well as noise-adaptive filtering and beamforming algorithms, could be pursued to increase localization success rates.

The results presented here are specific to the conditions in the area and at the time of our study. They provide an indication of how these PAM systems might perform in similar conditions, but caution is warranted when applying our results to other areas or times. Many efforts must be made to characterize variability in the source, background noise, and transmission conditions before detection probability estimates can be generalized. That said, these results have already been applied to inform dynamic management of right whales; preliminary versions of the detection functions derived here were used to parameterize a modeling effort that suggests whale movement causes visual and acoustic detections to provide equally uncertain estimates of whale location on dynamic management timescales (Johnson *et al.*, 2020b). We anticipate and hope that the results presented here will also prove useful for refining management policies in the United States and Canada as both countries have recently committed to using near real-time PAM to trigger management measures.

V. CONCLUSIONS

Our primary motivation for this work was to improve conservation outcomes for right whales by using an effective and reliable near real-time passive acoustic monitoring system. One such system, the DMON/LFDCS, has been operational for several years but has only recently been used

to inform dynamic management measures owing partly to the uncertainty in the acoustic detection range. We were able to successfully address this uncertainty by conducting a dedicated study using a multi-channel reference hydrophone array to empirically quantify the probability of detecting localized right whale upcalls from autonomous DMON/LFDCS platforms in different noise conditions. Our results provide a near-complete description of both near real-time and archival performance of both monitoring platforms for a shallow water site. We quantify the impact of noise conditions and platform depth on performance and provide evidence that the profiling glider gains an advantage over the buoy by occupying depth strata characterized by low transmission loss. We also demonstrate how the detection range is influenced by the review protocol where a more conservative protocol effectively reduces the detection range of the system. Our analysis was conducted on a call-by-call basis and therefore provides a minimum estimate of the platform detection range that can be increased by considering multiple calls. All the results presented here are specific to the conditions in the area and at the time of our study, and caution is required to apply them more broadly.

Given its economy and performance, we anticipate near real-time PAM will become even more widely used in the future. We recommend that new systems quantify and report their performance (e.g., Baumgartner *et al.*, 2019; Baumgartner *et al.*, 2020) before being used operationally for management, and that the detection function should be characterized (this study) to inform mitigation applications. Furthermore, we encourage visual survey teams to conduct and report similar analyses, as they are subject to many analogous detection challenges. Such analyses are difficult but illuminating; they can aid in the proper interpretation of the survey results, allow the standardization and inter-comparison of survey methodologies, and identify issues or sources of bias. More thorough evaluation of both acoustic and visual survey performance will help us determine which survey methodology is optimal for a particular application, how they can better complement one another, and how to best consolidate and compare these data sources for management and research purposes.

ACKNOWLEDGMENTS

We are indebted to Ken Houtler and Ian Hanley for their skilled operation of the R/V *Tioga*; John Kemp, Don Peters, Meghan Donohue, Jeff Pietro, Kris Newhall, Jim Dunn, and Nico Llanos of the WHOI Mooring Operations and Engineering Group for development, deployment, and recovery of the mooring systems; Phil Alatalo for assistance at sea; Ben Hodges for preparation, deployment, and recovery of the Slocum glider; Ed O'Brien and Giorgio Caramanna of the WHOI Dive Group for assisting in mooring recovery; Peter Koski, Julien Bonnel, and Dan Zitterbart of WHOI Applied Ocean Physics & Engineering Department for guidance and advice; and Kim Davies, Delphine Durette-Morin, Meg Carr, Marcia Pearson, and

Kimberly Franklin at Dalhousie University for providing helpful feedback.

Support for this study was provided by the Massachusetts Clean Energy Center (MassCEC), Bureau of Ocean and Energy Management (BOEM), and the Nova Scotia Offshore Energy Research Association (OERA). Support for H.D.J. was provided by the Marine Environmental Prediction and Response Network (MEOPAR) Whales Habitat and Listening Experiment (WHALE), the Killam Foundation, Vanier Canada Graduate Scholarship program, Dalhousie University, the Nova Scotia Graduate Scholarship program, and the Canada Graduate Scholarships–Michael Smith Foreign Study Supplements (CGS-MSFSS) program.

¹See <http://robots4whales.whoi.edu/#protocol> for a detailed description of the near real-time analysis protocol for DMON/LFDCS platforms (Last viewed April 4, 2022).

²See <https://www.whalesafe.com> for more information on the Whale Safe system for reducing whale-ship collisions off the West Coast of the United States (Last viewed April 4, 2022).

³See <https://whalealert.org/> for more information on the Whale Alert App and its role in reducing ship strikes to large whales (Last viewed April 4, 2022).

⁴See <https://www.ndbc.noaa.gov/> for information on data available from the National Data Buoy Center (Last viewed April 4, 2022).

⁵See supplementary material at <https://www.scitation.org/doi/suppl/10.1121/10.0010124> for more information on the precautionary analysis of near real-time pitch tracks, the analysis of archival audio, and results of a simulation to evaluate the methodology.

Abadi, S. H., Thode, A. M., Blackwell, S. B., and Dowling, D. R. (2014). “Ranging bowhead whale calls in a shallow-water dispersive waveguide,” *J. Acoust. Soc. Am.* **136**, 130–144.

Amante, C., and Eakins, B. W. (2009). “ETOPO1 1 arc-minute global relief model: Procedures, data sources and analysis,” *NOAA Technical Memorandum NESDIS NGDC-24*, National Geophysical Data Center, NOAA.

Baumgartner, M. F., Bonnell, J., Corkeron, P. J., Van Parijs, S. M., Hotchkiss, C., Hodges, B. A., Bort Thornton, J., Mensi, B. L., and Bruner, S. M. (2020). “Slocum gliders provide accurate near real-time estimates of baleen whale presence from human-reviewed passive acoustic detection information,” *Front. Mar. Sci.* **7**, 100.

Baumgartner, M. F., Bonnell, J., Van Parijs, S. M., Corkeron, P. J., Hotchkiss, C., Ball, K., Pelletier, L.-P., Partan, J., Peters, D., Kemp, J., Pietro, J., Newhall, K., Stokes, A., Cole, T. V. N., Quintana, E., and Kraus, S. D. (2019). “Persistent near real-time passive acoustic monitoring for baleen whales from a moored buoy: System description and evaluation,” *Meth. Ecol. Evol.* **10**, 1476–1489.

Baumgartner, M. F., Fratantoni, D. M., Hurst, T. P., Brown, M. W., Cole, T. V. N., Van Parijs, S. M., and Johnson, M. P. (2013). “Real-time reporting of baleen whale passive acoustic detections from ocean gliders,” *J. Acoust. Soc. Am.* **134**, 1814–1823.

Baumgartner, M. F., and Mussoline, S. E. (2011). “A generalized baleen whale call detection and classification system,” *J. Acoust. Soc. Am.* **129**, 2889–2902.

Baumgartner, M. F., Van Parijs, S. M., Wenzel, F. W., Tremblay, C. J., Carter Esch, H., and Warde, A. M. (2008). “Low frequency vocalizations attributed to sei whales (*Balaenoptera borealis*),” *J. Acoust. Soc. Am.* **124**, 1339–1349.

Bioacoustics Research Program (2011). *Raven Pro: Interactive Sound Analysis Software (Version 2.0)* (The Cornell Lab of Ornithology, Ithaca, NY) (<http://www.birds.cornell.edu/raven>) (Last viewed April 4, 2022).

Bonnel, J., Thode, A. M., Blackwell, S. B., Kim, K., Macrander, A. M., and Macrander, A. M. (2014). “Range estimation of bowhead whale (*Balaena mysticetus*) calls in the Arctic using a single hydrophone,” *J. Acoust. Soc. Am.* **136**, 145–155.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2004). *Advanced Distance Sampling* (Oxford University Press, Oxford, UK), Vol. 2.

Buckland, S. T., Summers, R. W., Borchers, D. L., and Thomas, L. (2006). “Point transect sampling with traps or lures: Point transect sampling,” *J. Appl. Ecol.* **43**, 377–384.

Buckland, S. T., and Turnock, B. J. (1992). “A robust line transect method,” *Biometrics* **48**, 901.

Cato, D. H. (1998). “Simple methods of estimating source levels and locations of marine animal sounds,” *J. Acoust. Soc. Am.* **104**, 1667–1678.

Chang, W., Cheng, J., Allaire, J. J., Xie, Y., and McPherson, J. (2019). “shiny: Web application framework for R,” R package version 1.4.0, <https://CRAN.R-project.org/package=shiny> (Last viewed April 4, 2022).

Clark, C. W., Brown, M. W., and Corkeron, P. J. (2010). “Visual and acoustic surveys for North Atlantic right whales, *Eubalaena glacialis*, in Cape Cod Bay, Massachusetts, 2001–2005: Management implications,” *Mar. Mamm. Sci.* **26**, 837–854.

Cummings, W. C., and Holliday, D. V. (1986). “Sounds and source levels from bowhead whales off Pt. Barrow, Alaska,” *J. Acoust. Soc. Am.* **82**, 814–821.

Davis, G. E., Baumgartner, M. F., Bonnell, J. M., Bell, J., Berchok, C. L., Bort Thornton, J., Brault, S., Brault, S., Buchanan, G., Charif, R. A., Cholewiak, D., Clark, C. W., Corkeron, P., Delarue, J., Dudzinski, K., Hatch, L., Hildebrand, J., Hodge, L., Klinck, H., Kraus, S., Martin, B., Mellinger, D. K., Moors-Murphy, H., Nieukirk, S., Nowacek, D. P., Parks, S., Read, A. J., Rice, A. N., Risch, D., Širović, A., Soldevilla, M., Stafford, K., Stanistreet, J. E., Summers, E., Todd, S., Warde, A., and Van Parijs, S. M. (2017). “Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014,” *Sci. Rep.* **7**, 13460.

D’Spain, G. L., Thode, A. M., and Kuperman, W. A. (1997). “Three-dimensional localizations of vocalizing blue whales,” *J. Acoust. Soc. Am.* **102**, 3212.

Durette-Morin, D., Davies, K. T. A., Johnson, H. D., Brown, M. W., Moors-Murphy, H., Martin, B., and Taggart, C. T. (2019). “Passive acoustic monitoring predicts daily variation in North Atlantic right whale presence and relative abundance in Roseway Basin, Canada,” *Mar. Mamm. Sci.* **35**, 1280–1303.

Fregosi, S., Harris, D. V., Matsumoto, H., Mellinger, D. K., Negretti, C., Moretti, D. J., Martin, S. W., Matsuyama, B., Dugan, P. J., and Klinck, H. (2020). “Comparison of fin whale 20 Hz call detections by deep-water mobile autonomous and stationary recorders,” *J. Acoust. Soc. Am.* **147**, 961–977.

Frisk, G. V. (1994). *Ocean and Seabed Acoustic: A Theory of Wave Propagation* (Prentice-Hall, Englewood Cliffs, NJ).

Halekoh, U., Højsgaard, S., and Yan, J. (2006). “The R Package geepack for generalized estimating equations,” *J. Stat. Softw.* **15**(2), 1–11.

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,” *J. Acoust. Soc. Am.* **134**, 3522–3535.

Harris, D. V., Miksis-Olds, J. L., Vernon, J. A., and Thomas, L. (2018). “Fin whale density and distribution estimation using acoustic bearings derived from sparse arrays,” *J. Acoust. Soc. Am.* **143**, 2980–2993.

Helble, T. A., D’Spain, G. L., Campbell, G. S., and Hildebrand, J. A. (2013b). “Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics,” *J. Acoust. Soc. Am.* **134**, EL400–EL406.

Helble, T. A., D’Spain, G. L., Hildebrand, J. A., Campbell, G. S., Campbell, R. L., and Heaney, K. D. (2013a). “Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones,” *J. Acoust. Soc. Am.* **134**, 2556–2570.

Johnson, H. D., Baumgartner, M. F., and Taggart, C. T. (2020b). “Estimating North Atlantic right whale (*Eubalaena glacialis*) location uncertainty following visual or acoustic detection to inform dynamic management,” *Conservation Sci. Practice* **2**(10), e267.

Johnson, H. D., Newhall, A., Lin, Y.-T., and Baumgartner, M. F. (2020a). “Acoustic and oceanographic observations collected during the Nomans Island experiment in spring 2017,” Technical report WHOI-2020-04 (Woods Hole Oceanographic Institution, Woods Hole, MA).

Johnson, M., and Hurst, T. (2007). “The DMON: An open-hardware/open-software passive acoustic detector,” in *3rd International Workshop on the*

- Detection and Classification of Marine Mammals Using Passive Acoustics*, Boston, MA (July 24–26, 2007).
- Johnson, H., Morrison, D., and Taggart, C. (2021). “WhaleMap: A tool to collate and display whale survey results in near real-time,” *J. Open Source Software* **6**, 3094.
- Kelley, D., Richards, C. (2019). “oce: Analysis of Oceanographic Data,” R package version 1.1-1. <https://CRAN.R-project.org/package=oce> (Last viewed April 4, 2022).
- Kirsebom, O. S., Frazao, F., Simard, Y., Roy, N., Matwin, S., and Giard, S. (2020). “Performance of a deep neural network at detecting North Atlantic right whale upcalls,” *J. Acoust. Soc. Am.* **147**, 2636–2646.
- Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D. J., and Ward, J. (2011). “Cetacean population density estimation from single fixed sensors using passive acoustics,” *J. Acoust. Soc. Am.* **129**, 3610–3622.
- Kyhn, L. A., Tougaard, J., Thomas, L., Duve, L. R., Stenback, J., Amundin, M., Desportes, G., and Tellman, J. (2012). “From echolocation clicks to animal density—Acoustic sampling of harbor porpoises with static data-loggers,” *J. Acoust. Soc. Am.* **131**, 550–560.
- Laurinoli, M., Hay, A. E., Desharnais, F., and Taggart, C. T. (2003). “Localization of North Atlantic right whale sounds in the Bay of Fundy using a sonobuoy array,” *Mar. Mamm. Sci.* **19**, 708–723.
- 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.
- Marques, T. A., Munger, L. M., Thomas, L., Wiggins, S. M., and Hildebrand, J. A. (2011). “Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting,” *Endanger. Spec. Res.* **13**, 163–172.
- Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). “Estimating animal population density using passive acoustics,” *Biol. Rev.* **88**, 287–309.
- Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (2009). “Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales,” *J. Acoust. Soc. Am.* **125**, 1982–1994.
- Merchant, N. D., Frstrup, K. M., Johnson, M. P., Tyack, P. L., Witt, M. J., Blondel, P., and Parks, S. E. (2015). “Measuring acoustic habitats,” *Meth. Ecol. Evol.* **6**, 257–265.
- Morley, M. G., Dosso, S. E., and Chapman, N. R. (2009). “Array element localization using ship noise,” *J. Acoust. Soc. Am.* **125**, 1403–1409.
- 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.
- Newhall, A. E., Lin, Y.-T., Lynch, J. F., Baumgartner, M. F., and Gawarkiewicz, G. G. (2012). “Long distance passive localization of vocalizing sei whales using an acoustic normal mode approach,” *J. Acoust. Soc. Am.* **131**, 1814–1825.
- Nuuttila, H. K., Brundiers, K., Dähne, M., Koblit, J. C., Thomas, L., Courtene-Jones, W., Evans, P. G. H., Turner, J. R., Bennell, J. D., and Hiddink, J. G. (2018). “Estimating effective detection area of static passive acoustic data loggers from playback experiments with cetacean vocalisations,” *Meth. Ecol. Evol.* **9**, 2362–2371.
- Parks, S. E., Johnson, M. P., Nowacek, D. P., and Tyack, P. L. (2011a). “Individual right whales call louder in increased environmental noise,” *Biol. Lett.* **7**, 33–35.
- Parks, S. E., Searby, A., Célrier, A., Johnson, M. P., Nowacek, D. P., and Tyack, P. L. (2011b). “Sound production behavior of individual North Atlantic right whales: Implications for passive acoustic monitoring,” *Endanger. Spec. Res.* **15**, 63–76.
- Parks, S. E., and Tyack, P. L. (2005). “Sound production by North Atlantic right whales (*Eubalaena glacialis*) in surface active groups,” *J. Acoust. Soc. Am.* **117**, 3297–3306.
- Parks, S. E., Urazghildiev, I. R., and Clark, C. W. (2009). “Variability in ambient noise levels and call parameters of North Atlantic right whales in three habitat areas,” *J. Acoust. Soc. Am.* **125**, 1230–1239.
- Porter, M. B. (1992). “The KRAKEN Normal Mode Program,” Naval Res. Lab. 1–198.
- R Core Team. (2019). “R: A language and environment for statistical computing,” in *R Foundation for Statistical Computing* (Vienna, Austria).
- Simard, Y., Roy, N., Giard, S., and Aulancier, F. (2019). “North Atlantic right whale shift to the Gulf of St. Lawrence in 2015, revealed by long-term passive acoustics,” *Endanger. Spec. Res.* **40**, 271–284.
- Tennessen, J., and Parks, S. E. (2016). “Acoustic propagation modeling indicates vocal compensation in noise improves communication range for North Atlantic right whales,” *Endanger. Spec. Res.* **30**, 225–237.
- Thode, A., Blackwell, S., Conrad, A., Kim, K., Marques, T., Thomas, L., Oedekoven, C., Harris, D., and Bröker, K. (2020). “Roaring and repetition: How bowhead whales adjust their call density and source level (Lombard effect) in the presence of natural and seismic airgun survey noise,” *J. Acoust. Soc. Am.* **147**, 2061–2080.
- Thode, A., Bonnel, J., Thieury, M., Fagan, A., Verlinden, C., Wright, D., Berchok, C., and Crance, J. (2017). “Using nonlinear time warping to estimate North Pacific right whale calling depths in the Bering Sea,” *J. Acoust. Soc. Am.* **141**, 3059–3069.
- Thode, A. M., D’Spain, G. L., and Kuperman, W. A. (2000). “Matched-field processing, geoacoustic inversion, and source signature recovery 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, M. D., 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., Nations, C. S., McDonald, T. L., and Macrander, A. M. (2012). “Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys,” *J. Acoust. Soc. Am.* **131**, 3726–3747.
- Vanderlaan, A. S. M., Taggart, C. T., Serdynska, A. R., Kenney, R. D., and Brown, M. W. (2008). “Reducing the risk of lethal encounters: Vessels and right whales in the Bay of Fundy and on the Scotian shelf,” *Endanger. Species Res.* **4**, 283–297.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, York).
- Wickham, H., Averick, M., Bryan, J., Chang, W., D’Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). “Welcome to the tidyverse,” *J. Open Source Softw.* **4**(43), 1686.
- Wiggins, S. M., McDonald, M. A., 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.