Received: 26 April 2022 | Accepted: 17 June 2022

DOI: 10.1111/2041-210X.13973

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Evaluation of a coastal acoustic buoy for cetacean detections, bearing accuracy and exclusion zone monitoring

Kaitlin J. Palmer[1] | Sam Tabbutt[1] | Douglas Gillespie[2] | Jesse Turner[1] | Paul King[1] | Dominic Tollit[1] | Jessica Thompson[1] | Jason Wood[1]

[1]SMRU Consulting, Friday Harbor, Washington, USA

[2]Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, UK

Correspondence
Kaitlin J. Palmer
Email: kjp@smruconsulting.com

Funding information
The Maryland Department of Natural Resources, Offshore Wind Development Fund; US Department of Energy, Grant/Award Number: DE-EE00087

Handling Editor: Chloe Robinson

Abstract

1. There is strong socio-political support for offshore wind development in US territorial waters and construction is planned off several east coast states. Some of the planned development sites coincide with important habitat for critically endangered North Atlantic right whales. Both exclusion zones and passive acoustic monitoring are important tools for managing interactions between marine mammals and human activities. Understanding where animals are with respect to exclusion zones is important to avoid costly construction delays while minimizing the potential for negative impacts. Impact piling from construction of hundreds of offshore wind turbines likely require exclusion zones as large as 10 km.
2. We have developed a three-hydrophone passive acoustic monitoring system that provides bearing information along with marine mammal detections to allow for informed management decisions in real-time. Multiple units form a monitoring system designed to determine whether marine mammal calls originate from inside or outside of an exclusion zone. In October 2021, we undertook a full system validation, with a focus on evaluating the detection range and bearing accuracy of the system with respect to right whale upcalls. Five units were deployed in Mid-Atlantic waters and we played more than 3500 simulated right whale upcalls at known locations to characterize the detection function and bearing accuracy of each unit. The modelled results of the detection function error were then used to compare the effectiveness of a bearing-based system to a single sensor that can only detect a signal but not ascertain directivity.
3. Field trials indicated maximum detection ranges from 4–7.3 km depending on source and ambient noise levels. Simulations showed that incorporating bearing detections provide a substantial improvement in false alarm rates (6 to 12 times depending on number of units, placement and signal to noise conditions) for a small increase in the risk of missed detections inside of an exclusion zone (1%–3%).
4. We show that the system can be used for monitoring exclusion zones and clearly highlight the value of including bearing estimation into exclusion zone

monitoring plans while noting that placement and configuration of units should reflect anticipated ambient noise conditions.

KEYWORDS
acoustics, autonomous, conservation, exclusion zone, localization, mitigation, real-time, whale

1 | INTRODUCTION

Exclusion or closure zones are areas within which potentially harmful human activities are managed to reduce the likelihood of injury, or disturbance, to protected species. Exclusion zones are important conservation tools and have been used to mitigate diverse negative impacts ranging from sound exposure and light exposure (Pendoley & Kamrowski, 2016; Weir & Dolman, 2007), overfishing (Okey et al., 2004), to domestic cat predation (Metsers et al., 2010). Exclusion zones can be either static, protecting a stationary habitat from harm (Davies & Brillant, 2019; Spaulding et al., 2009), or movable. Seismic surveys, for example, maintain exclusion zones centred on the primary vessel as it completes its survey (Bröker et al., 2015; Racca et al., 2015).

North Atlantic right whales (NARWs), *Euablaena galcialis* are an endangered species inhabiting the coastal waters of the U.S. and Canadian eastern seaboard (Davis et al., 2017). Current mortality rates are mainly due to entanglement and ship strikes (Davies & Brillant, 2019), much of which have resulted from the climate-induced abrupt change in habitat use. Presently, a variety of mitigation efforts including seasonal and dynamic mitigation zones are used with the intention of lowering mortality rates (Baumgartner et al., 2020; Cole et al., 2021; Van Parijs et al., 2009). A substantial number of offshore wind farms are planned for the east coast of the United States, which will necessitate extensive offshore construction involving both pile-driving and regionally increased ship activity. This results in the potential for disturbance throughout the construction, operation and decommissioning periods. In addition to increased risk of ship strikes, pile driving noise, particularly, has the potential to cause harm and harassment to marine mammals (Bailey et al., 2010, 2014; Madsen et al., 2006; Southall et al., 2019; Teilmann & Carstensen, 2012; Tougaard et al., 2009). An incidental harassment authorization for a recently proposed offshore wind farm requires a 10 km radius mitigation zone to be monitored with visual and passive acoustics for cetaceans, including right whales, before and during pile driving activity (Offshore Wind Energy Development in New England/Mid-Atlantic Waters). The confirmed presence of right whales within the resulting 314 km^2 exclusion zone will result in mandatory delays in pile driving.

Monitoring such a large area, often during times of low visibility, is particularly challenging. While some low-visibility monitoring technologies are available (Verfuss et al., 2018), many have limited range and area coverage. If animals are actively calling, passive acoustic monitoring (PAM) can be used to detect and potentially localize animals. Common approaches include single sensors that can detect animals but provide no localisation, thereby providing only presence/absence information (Spaulding et al., 2009). Large arrays can similarly use time-difference of arrival (TDOA) to estimate animal location, but they are often cost prohibitive and must cover the entire study area. Such approaches are typically used in long term studies of archival data. Finally, clusters of hydrophones spaced centimetres to a few metres apart can provide intermediate location information by estimating the bearings to detected sounds.

If no localisation information is available, then a precautionary approach would have to be adopted, whereby economically important activities are curtailed, even though the animal may be outside the mitigation zone. Localisation is therefore important, to avoid unnecessary shutdowns. Deploying large, spatially distributed arrays is particularly challenging as accurate clock synchronization between nodes is required (Marchetto et al., 2012; Palmer, Wu, et al., 2022). To be used in mitigation, data are needed in real or near real-time. Retrieving data from multiple nodes in real-time poses additional challenges but is currently achievable at scales currently required for NARW.

In this paper, we present an approach that uses relatively small clusters of hydrophones, which can therefore easily share a common clock, and measure accurate bearings to detected sources. By placing multiple such clusters close to the mitigation zone boundary, crossed bearing localisation, which does not require accurate clock synchronization, can be used to localize calls, and is particularly good for answering the key question as to whether a source is inside or outside an exclusion zone (Swartz et al., 2000). Having the sensors at the boundary of the mitigation zone has the additional advantages of being as far possible from confounding noise sources at the centre of the zone and can also help to detect approaching animals.

We have developed one such approach to monitoring large exclusion zones at wind farm construction sites in the US, with particular emphasis on NARWs. The coastal acoustic buoy for offshore wind (CABOW) system consists of multiple remote PAM units (CABOWs) and a remote processing node (base station). Each CABOW unit is capable of real-time automated acoustic right whale detection and transmits the detection data to the base station for manual review. PAM operators, mandated to be present for all offshore wind pile driving activities, are then able to (1) confirm whether each detection was a right whale and (2) determine if the call was likely produced inside or outside of the predefined exclusion zone, by mapping the directions of the bearing produced by the CABOW unit with respect to the exclusion zone boundary.

In the fall of 2021, we conducted extensive field trials off the east coast of the United States to evaluate the full CABOW system including multiple units. We report the detection range of the system under varying signal and noise conditions and estimate the detection function and bearing estimation accuracy. We then used simulations to compare the effectiveness of a CABOW system (with bearing information) in monitoring a large exclusion zone with a PAM system using only a single sensor (detection only) approach.

2 | MATERIALS AND METHODS

2.1 | CABOW

2.1.1 | Hardware

Each CABOW unit consists of the following hardware components: a weighted bottom lander designed by the University of Washington Applied Physics Laboratory with three attached hydrophones, one or more battery housings, and an instrument housing; a combined network and power cable; a large (40 cm dimeter) mooring float to keep the data cable vertically oriented in the water column, six small trawl floats to reduce the risk of entanglement causing loops, and a surface floating communications buoy. Mounted on the bottom lander are three High-Tech Inc 96 min hydrophones separated by 2 m in a triangular pattern; two lithium-ion battery pressure housings each with 385-amp hour capacity at 16.8 V, which provide deployment durations up to 7 weeks; and a pressure housing containing the acoustic instrumentation. The instrument housing contains a Decimus (single board embedded computer running Linux with an integrated digital signal processor), a custom data acquisition board and power management circuitry. Falmat Xtreme-Cat-5, Kevlar reinforced, Ethernet/power cabling is used to transmit data and power from the instrument housing to the surface buoy. This cabling also serves as the mooring for the surface float, and as the strength member for the deployment and recovery of the system. The surface buoy contains a Zumlink PE-9900 MHz broadband frequency hopping spread spectrum radio transmitter and receiver as an endpoint in a star radio network (Figure 1).

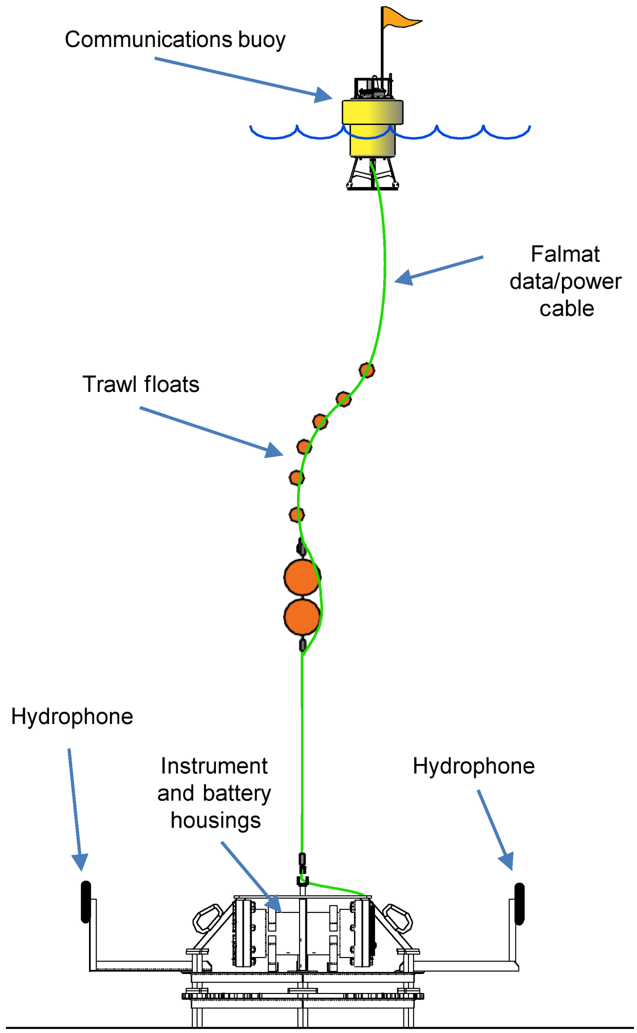

FIGURE 1 Schematic of coastal acoustic buoy for offshore wind lander, data cable, and communications buoy.

2.1.2 | Software

All right whales produce stereotyped ‘upcalls’ throughout their range and life stage (Clark & Gagnon, 2002; Parks et al., 2007). Upcalls are frequency modulated signals between ~50 and 350 Hz and 0.25–1.25 s in duration. Because of their ubiquity, upcalls are used as indicators of right whale presence in nearly all passive acoustic mitigation projects (Baumgartner & Mussoline, 2011; Davis et al., 2017; Fladung et al., 2011; Palmer, Wu, et al., 2022). Thus, as with other real or near real-time systems, the right whale detection algorithms within the CABOW systems target upcalls.

Each CABOW system runs an embedded version of the right whale detection algorithm described in Gillespie (2004) and integrated into PAMGuard software (www.pamguard.org). PAMGuard is an open-source software package commonly used for real-time monitoring of cetaceans and is familar to most PAM operators (Barlow et al., 2018; Gillespie et al., 2022; Romagosa et al., 2015). TDOA measurements using the cross correlation and interpolation methods described in Gillespie and Macaulay (2019) measure the time delays between each hydrophone channel on the CABOW. These measurements and a 2 s clip of a single channel of audio data for each detection are transmitted using transmission control protocol via radio to a base station computer in near real-time. The remote base station computer runs a modified version of the PAMGuard software (Gillespie & Caillat, 2008), which can receive data packets sent from multiple CABOW communication buoys and then store, display and further process those data, just as it would treat detections created within PAMguard during continuous operation.

The detection process incorporates two-level classification to accurately and precisely detect right whale upcalls. The first level of classifier (Gillespie, 2004) runs onboard each CABOW lander and is set to detect sounds with a low specificity (minimum classification score of 4). This provides a high recall (proportion of true calls detected), but also a high number of false detections (i.e., low precision). At the base station, the 2 s clip around each of these detections is then re-classified using the algorithm described in Shiu et al., (2020) to increase accuracy. This convolutional neural network was developed using publicly available data and evaluated on competition data collected from Cape Cod Bay, MA (Gillespie, 2019) as well as sampled recordings throughout much of the right whales known extent (see Shiu et al., 2020 for details on recordings). Two-second spectrograms (fft 0.256 s, 0.125 s overlap) of each detection are presented to an operator for visual and aural validation. The estimated bearing lines to each detection is also shown on the PAMGuard map display. CABOW units also measure 10 s average power spectra and transmit these data to the PAM operator, to allow them to easily determine whether the system is running properly when no detections are present. During these trials, the system was configured to continuously store the raw audio data on a solid-state drive for post-processing validation of the detection function. All sound pressure measurements are referenced to 1 μPa unless otherwise specified.

2.2 | Field evaluation

2.2.1 | Study area

Field trials using simulated right whale upcalls were undertaken in the offshore waters of Maryland, USA to

1. Measure the bearing error and
2. Determine a site-specific detection function, the probability of detecting a call at a given range given the source level of the call and the ambient noise level at the sensor.

The deployment area was chosen based on its proximity to established offshore wind lease areas and relative ease of access (Figure 2). The sediment at the deployment site was characterized as sand/silt/clay (USGS Woods Hole Field Center, 2000). A conductivity, temperature and depth cast was taken at the beginning of the cruise to ensure that the underwater speaker depth (8 m) exceed the thermocline. The cast indicated that the water column was well mixed and there was no indication of a thermocline. Wind conditions during cruise were generally calm with swells from distant storm systems ranging 1–1.5 m.

Five calibrated CABOW units were evenly deployed along the arc of a 10 km radius circle to simulate an exclusion zone of the same size (Table 1). CABOW units were spaced approximately 2.5 km apart, in order to ensure that the detection function at short ranges was well characterized. We refer to multiple deployed CABOW units as an 'array'. All CABOW units were successfully deployed and recovered on 13 October and 15 October 2021, respectively. Playback locations were chosen to maximize characterization of the detection function at small ranges $g(0)$ and to ensure that the base

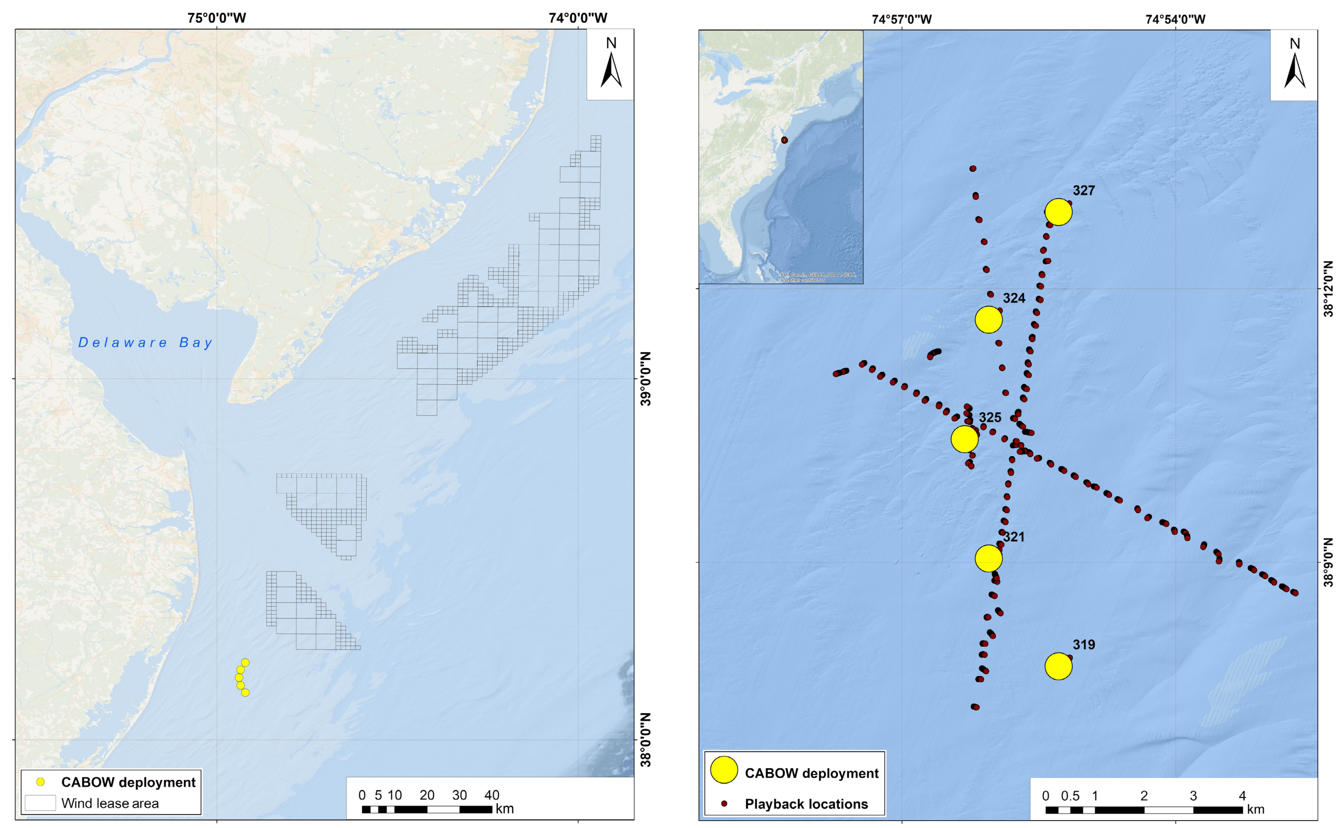

FIGURE 2 Left panel, location of coastal acoustic buoy for offshore wind (CABOW) deployment (yellow circles) with respect to renewable energy leases and planning areas (black lines, BOEM). Right panel, detailed deployment map showing CABOW unit locations and playback locations (red dots). Point in inset represents deployment location relative to the entirety of the eastern US seaboard.

station, situated on the vessel, could maintain constant connection with all CABOW units while undertaking the playbacks. Data for the detection function were processed iteratively throughout the deployment to ensure sufficient replicates in each range and signal excess bin. For example, noise levels were lower than initially predicted resulting in good characterization of the detection function under high source-to-noise level conditions but few playbacks under low source-to-noise conditions. As the passive sonar equations (Urick, 1983) dictate that low signal–to-noise ratios (SNR) or source-to-noise level can be produced by either increasing the ambient noise, or decreasing the source level, the amplitude of the calls were decreased partway through the deployment. This resulted in more low SNR replicates that would otherwise been possible given the short deployment duration.

The playback signal consisted of 30 right whale-like upcall signals (henceforth 'upcalls') with 7 s of silence separating each upcall. Signals were created using open-source software, Audacity 3.0 (https://www.audacityteam.org/). Upcall start frequencies ranged between 25 and 80 Hz; end frequencies between 250 and 300 Hz; and duration between 0.4 and 0.8 s. Upcall source levels across the 50–225 Hz band were 150–162 dB re: 1 μPa$_{rms}$ at 1 m. These upcalls aimed to replicate the most frequently detected calls made by right whales in the Atlantic. Calls were played from a Sound Devices 702 recorder and through a Lubel 9162 underwater speaker. The University of Delaware's R/V Hugh R. Sharp was used for both deploying the CABOW units and as the playback vessel. The Sharp was designed with ICES 209 sound emission standards and was run in 'quiet' mode throughout the playbacks to reduce masking and potential bias in bearing error estimation. A radio receiver was temporarily mounted to the mast of the RV Sharp approximately 20 m above the water surface resulting in a line-of-sight communication range of ~15 km. Mooring permit is a Nation Wide Permit issued by the US Army Corps of Engineers: 2021-61257.

2.2.2 | Detection function

We report the results of two methods for evaluating the detection performance. First, we measure the probability of detection as a function of the SNR of the received call. This approach is site independent and, where the propagation loss is known, allows for detection range estimation at future deployment locations. The second approach considers the probability of detection (g) as a function of range (r) and the source-to-noise level ratio (SLNL) as presented by Thode et al., (2020). Both approaches represent slightly different interpretations of the passive sonar equations with the g(SNR) approach considering the detector performance exclusively and the g(SLNL|r) approach incorporating explicitly measured range. While this approach was originally developed to evaluate the Lombard effect in marine mammals, it allows for detailed evaluation of the probability of detection within the context of a fixed environment. In doing so, we seek to incorporate a level of portability into the system to answer the fundamental question of 'How far can it hear a right whale?'. Using this approach, if the local propagation conditions can be estimated, the probability of detecting a call at a given range can be provided as a function of the ambient noise Detection Function.

3 | RESULTS

3.1 | Detection function

Ambient noise levels during the playback period ranged from 91 to 131 dB$_{rms}$. Sound levels across the deployment area were consistent with a maximum 3 dB difference in minimum, and median source levels between the lowest and highest amplitude deployment area (Figure 3).

The observed relationship between the proportion of calls detected, source-to-noise level and range are shown in (Figure 4). Also shown is the simplified propagation model using 16.1 $\log_{10}(r)$ as estimated for the region by Bailey et al. (2019). The amplitude of the simulated calls must be greater than the transmission loss and the ambient noise level at the sensor to be detected. As the proportion of calls detected at high SLNL decreases with range more rapidly than the simplified TL model, it indicates that this model likely underestimates transmission loss at ranges greater than 6 km. Under low SLNL conditions, the maximum detection range was less than 1 km. With high SLNL (>60 dB), the maximum detection range exceeded 7 km. However, at very high SLNL, the probability of detecting an upcall at short ranges (<1.25 km) also decreased indicating a limitation of the system. Inspection of the data for low range and high SLNL showed merging of some of the harmonic structure thereby reducing the ability of the edge detection system to detect the calls, despite the low initial threshold.

The same effect was observable in the SNR performance curves (Figure 5). The system performed well under low and moderate SNR conditions. However, under high SNR conditions consistent with low range, high source level, or low ambient noise level, the system performance dropped off unexpectedly. Detailed analysis of the recovered data indicated that two issues were present resulting in the failure of the statistical assumption of perfect detection at low ranges. First, the fixed-point precision within the DSP chip on the Decimus failed to properly compute high amplitude received levels, due to a fixed-point overflow on the processor. Second, the edge detection algorithm was not initially trained on (rarely available) high SNR calls containing harmonics. This resulted in harmonic merging, and ultimately causing the edge detector to consider the bandwidth of the call too high. These issues were addressed in post processing. Field recordings were re-processed using the new lowered system gain and revised version of the edge detection algorithm. The two revisions resulted in improvement of the system performance at high SNR without decreasing the probability of detecting low SNR signals (Figure 5). As such the low $g(0)$ at high SLNL/SNR observed in the field recordings (Figure 4) should not be considered indicative of future CABOW deployments.

TABLE 1 Deployment location and playback summary for each coastal acoustic buoy for offshore wind (CABOW) unit

CABOW unit #	End to end calibration (dB re 1 volt)	Latitude	Longitude	Depth (m)
324	−159.6	38.214	−74.921	23
325	−160.7	38.194	−74.934	22
321	−159.8	38.173	−74.939	21
327	−160.4	38.151	−74.934	21
319	−160.1	38.131	−74.921	25

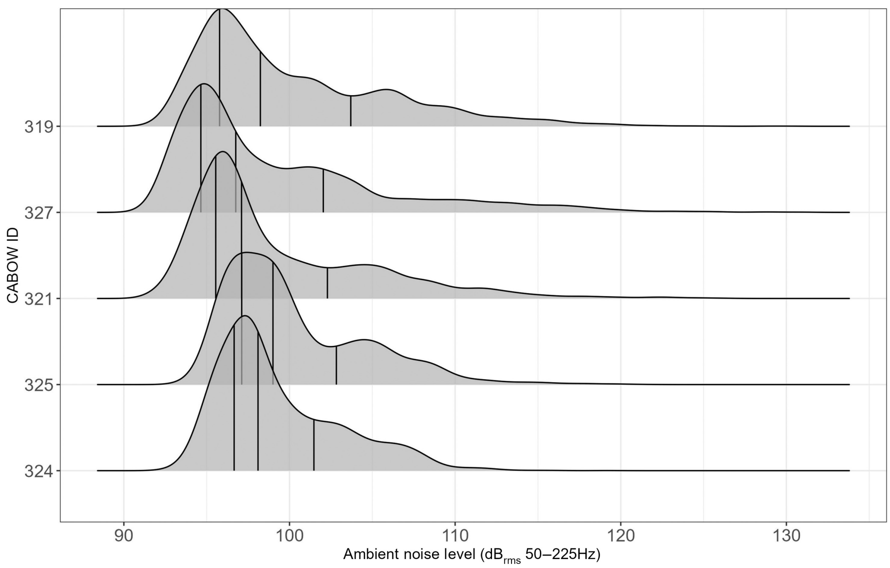

FIGURE 3 Density distributions of the ambient noise levels recorded at the five coastal acoustic buoy for offshore wind (CABOW) locations during the field trials. Vertical lines indicate 25th, 50th and 75th percentiles.

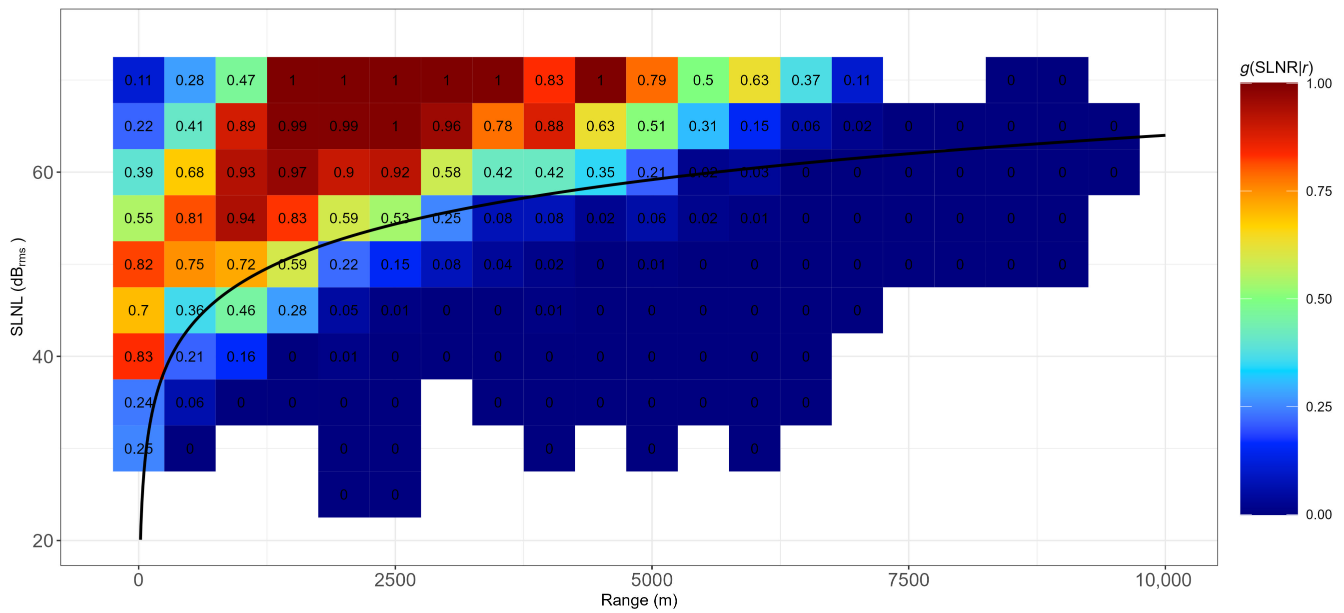

FIGURE 4 Proportion of upcalls detected as a function of range from a sensor and source-to-noise level ratio. Grid cells represent the proportion of calls detected at each source-to-noise level ratio (SLNL) and detection range. Black line represents simplified estimated transmission loss in the region. Lower than expected detection rates at low ranges and high SLNL values (upper left) was principally due to a internal gain error that was fixed upon recovery.

3.2 | Bearing error

The observed bearing error averaged across all units was small. The median, 25th and 75th quantiles of bearing error were −0.25°, −1.6°, 0.93°, respectively (Figure 6). Some amount of bias in bearing error for each CABOW was attributed to uncertainty during the orientation playback. Less than 1% of the overall bearing errors exceeded ±16.5°. However, this is likely an underestimation of the true bearing

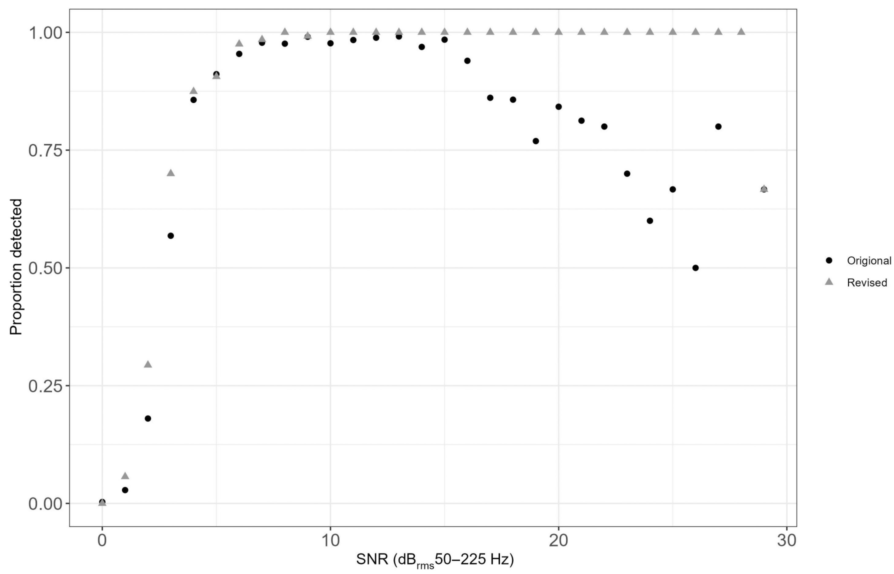

FIGURE 5 Proportion of calls detected in the field (black circles) as a function of signal–to-noise ratio (SNR) indicating a decrease in the proportion of high SNR calls (>12 dB) detected by the system. Grey triangles indicate the reprocessed-field data with the system gain reduced and the bug in the edge detection addressed.

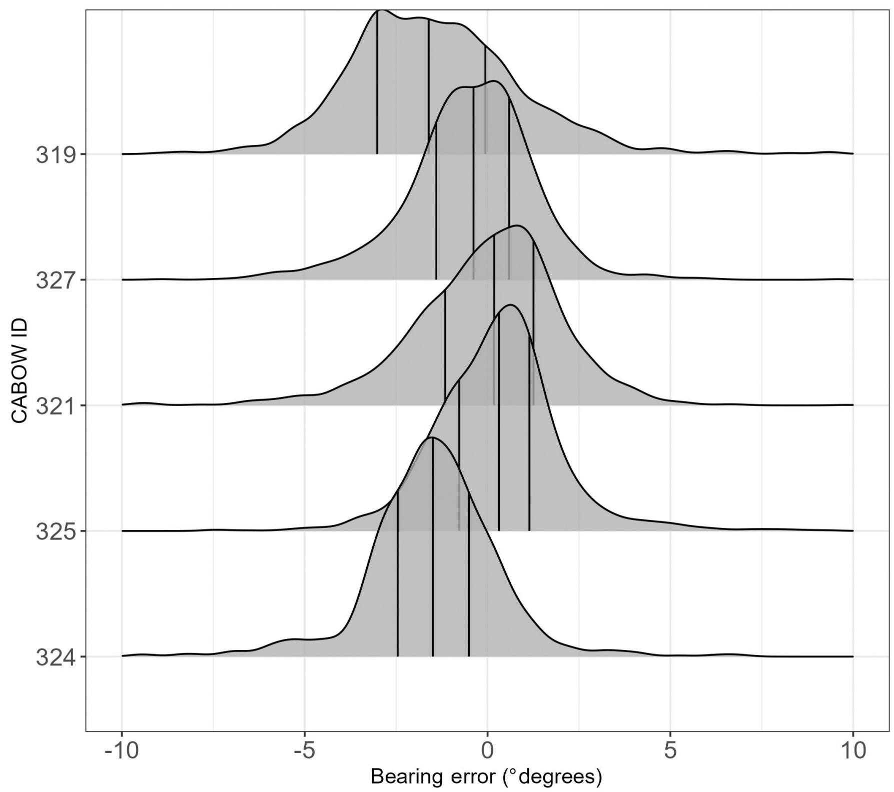

FIGURE 6 Density distribution of measured bearing errors from each coastal acoustic buoy for offshore wind (CABOW). Vertical lines indicate 25th, 50th and 75th percentiles. Not shown for scale, 179 detections (1% of total) with errors >±15°.

error as the playback vessel could not be completely shut down during field trials. Thus, an ambient noise field coherent with the upcalls may have resulted in better bearing estimates than would otherwise be expected in a real-world situation where the direction of the dominant ambient noise field and the upcalls would be independent.

3.3 | Single sensor comparison

In all scenarios, including bearing resulted in a 6–12 times reduction in the probability of instituting an unnecessary shutdown following a detection (Table 2). The false negative rate of the CABOW system with a 15° bearing error remained at or below 3% in all scenarios and decreased where calls were detected by multiple instruments (high source-to-noise ratio or number of units). The later was related to the same increase in the false positive rate in the CABOW as any bearing pointing inside the exclusion zone triggered a shutdown.

4 | DISCUSSION

Exclusion zones can aid in limiting the negative impacts of anthropogenic activity by ensuring that animals in the area are not exposed to unnecessary danger (Lebon & Kelly, 2019; Pendoley &

TABLE 2 Performance metrics between hypothetical system setups. Number of units indicates the number of hypothetical units equally spaced along the exclusion zone radius. Source and noise level are simulated rms values in the 50–225 Hz range. False positive rate is the number of unrequired shutdowns triggered by the system given the coastal acoustic buoy for offshore wind (CABOW) approach and single sensor detection only approach divided by the total number of detections. False negative rate is the proportion of detections that were inside the exclusion zone and should have triggered a shutdown, but for which the bearing error indicated they were outside of the exclusion zone.

Unit range from exclusion zone centre (km)	Number of units	Ambient noise level (dB)	Source level (dB)	False positive rate (CABOW)	False positive rate (detection only)	False negative rate (CABOW)
10	3	95	160	0.07	0.57	0.01
	5	95	160	0.07	0.57	0.01
	7	95	160	0.07	0.61	0.01
	9	95	160	0.09	0.62	<0.01
11	3	95	160	0.05	0.66	0.03
	5	95	160	0.05	0.66	0.03
	7	95	160	0.05	0.68	0.01
	9	95	160	0.06	0.69	0.01
10	3	105	160	0.04	0.53	0.02
	5	105	160	0.05	0.54	0.02
	7	105	160	0.05	0.54	0.02
	9	105	160	0.05	0.53	0.02

Kamrowski, 2016). However, mitigating underwater noise potentially requires large exclusion zones, which are technically challenging and costly to monitor. Current options include either single PAM sensors which provide no location information (Baumgartner et al., 2020; Fladung et al., 2011; Klinck et al., 2012) or larger interconnected arrays which are costly and difficult to deploy and maintain (Guazzo et al., 2020). The PAM system presented here addresses this issue by incorporating a small aperture array on each CABOW unit to estimate bearings to calling animals. Bearing-only approaches such as those implemented in the CABOW system provide an intermediate option between full localization of calling animals and single sensor detections. Such systems can be strategically placed in order to achieve optimal performance, in terms of precision, recall, false positive rate or a weighted average of the three within the exclusion zone. Having a well characterized detection function is key in the ability of any system to generalize and plan effective mitigation strategies.

Regardless of the choice of monitoring system, a thorough understanding of the site and noise-specific detection probability is critical. Underestimating the detection range could result in missed detections and increases the risk of acoustic trauma whereas overestimating the detection range would result in unnecessary and costly shutdowns. Here, we opted to use many simulated upcalls, rather than relying on unpredictable real right whale calls, to produce a well parameterized end-to-end detection function that incorporates range as well as SNR or SLNL. Thus far, all right whale detectors (Baumgartner & Mussoline, 2011; Gillespie, 2004; Shiu et al., 2020; Smirnov, 2013) have been built, trained, and evaluated using passive acoustic data from real animals. These training data contain few, if any, very high SNR calls and as such do not afford detectors the ability to learn the harmonic structure present at short ranges. This information should be included when deciding on instrument placement around proposed exclusion zones to avoid potential hazards around $g(0)$ shown here. In real-world applications, the probability of such high SNR calls is quite low as the area monitored by an acoustic sensor (generally) increases with range and therefore animals are more likely to be detected at greater ranges. Even so, we recommend that systems be validated with high SNR calls containing harmonic structure to identify and account for any acoustic blind spots.

The source-to-noise ratio approach used here provides a site-specific but robust methodology for estimating the area monitored by a sensor under a variety of acoustic conditions. The deployment duration was short due to weather and cost constraints thereby limiting the variability in natural ambient noise. However, by varying the amplitude of the playback calls, we were able to investigate system performance at both the low and high end of the SNR/SLNL spectrum. In this approach, the effective detection range can be adjusted in real-time as the ambient noise levels fluctuate. In doing so, users can determine times during which the 'listening space' of the sensor is likely to be reduced and subsequently shift priority to visual monitoring. Right whales are also known to change their call amplitude according to behaviour state, region, and ambient noise levels (Parks, 2003; Parks et al., 2011, 2019). The source-to-noise function estimated here allows users to modify the expected detection range based on the expected source levels in the area as well as ambient noise. It will thus be useful to managers in optimizing the number and placement of PAM units to achieve a desired level of mitigation in a specific area.

It is well established that that the probability of detecting a call depends greatly on the environment and this is especially true for low frequency sounds with large propagation ranges (Helble et al., 2013; Širović et al., 2007; Thode et al., 2016; Van Parijs et al., 2021).

Accounting for propagation loss can either be done in situ, as was here, or by modelling propagation prior to deployment (Farcas et al., 2016). However, by not explicitly evaluating the propagation conditions, the approach is not location invariant. The bathymetry in the survey region is relatively consistent and there was no indication of a strong thermocline. This allowed us to generalize the detection function across bearing angles. If the instruments are deployed in a more complex location, proper characterization of both the bearing error and detection probability either through empirical studies such as this or propagation modelling must include angular dependence. This principle holds for all PAM studies and especially those used in real-time monitoring and or mitigation.

The bearing error observed here is likely an underestimate of what would be expected when deployed in a real-world noisy underwater construction site as it was not possible to completely shut down the R.V. Sharp during playbacks. This resulted in coupling between the ambient noise field and the playback source. We accounted for this bias in the evaluation of the system by modelling the bearing error distribution with the average error (15°) rather than the median (1.25°) or 75th percentile (2.26°). The grid-based model clearly suggested that even with a wider distribution of bearing errors, significant improvements in unnecessary shutdowns are likely at the cost of a relatively small (1%–3%) increase in missed detection rate.

The practical application of the CABOW systems is limited by the deployment logistics, battery life, operational depth, and communications. To that end, the CABOW system has been designed to be easy to deploy; work at depths meeting the offshore wind industry's current operational needs (typically <50 m); and provide options of radio, cellular, or satellite communications to a central base station. Current mitigation-driven applications for CABOW have resulted in the current configuration having a battery life of 7 weeks which can be extended to 10 weeks with additional battery packs. For these reasons the CABOW system is not ideal for some long-term large-scale baseline monitoring applications. For applications that prioritize mobility, gliders provide a better solution (Baumgartner et al., 2019, 2020; Klinck et al., 2012), though for real-time mitigation, pump and flow noise may result in less reliable performance during the surface return portion of the flight. Additionally, the use of the neural network classification step necessitates that audio samples be returned to the base station via the communication buoys. At remote study sites satellite communications are required and the amount of data needed for the final classification step may be cost prohibitive in some cases.

Like all real or near real-time detection systems, the ultimate decision concerning whether or not a member target species is present is given to an experienced analysist (Baumgartner et al., 2019, 2020; Gillespie et al., 2013; Johnson et al., 2022; Klinck et al., 2012). With this system, users are provided with a 2 s clip with which to make their final decision. This is consistent with the Woods Hole Oceanographic Institute/Cornell real-time buoys that have been in operation in the Boston shipping channel for more than two decades (Spaulding et al., 2009). However, the small amount of data presented to the analysists can be limiting when acoustically similar species such as humpback whales, *Megaptera novaeanglea* are present. In this situation the false positive rate may be higher than is otherwise achievable. In continuous recordings, skilled analysts often look at the pattern of calls in order to discriminate between two species (Gillespie, 2019). This approach requires larger contextual data to be sent to the user. Sending longer clips is possible for the CABOW systems depending on the ultimate goals of the deployment (e.g., longevity or false positive rate).

As the population of NARWs continues to fall the need to reduce stress, potential hearing trauma and mortality has become increasingly urgent. However, in the drive for clean renewable energy, anthropogenic activities in right whale habitat are and will continue to be an inescapable reality. Acoustic monitoring allows for sanctioned activities to occur while limiting the likelihood of negative interactions between animals and industry. Bearing-only approaches are a middle ground between single sensors that are incapable of determining source location and large-scale arrays that provide localization but are immovable and prohibitively expensive to deploy and maintain. The CABOW system was specifically designed and represents an effective and well characterized method for monitoring large acoustic exclusion zones in real-time for mitigation purposes. Through field evaluation we have thoroughly parameterized the performance metrics of this system including detection probability and bearing error. In building the CABOW system, we have also included the most thoroughly characterized upcall detector to date thereby limiting false positive detections while maintaining a high recall. The approach outlined here is one of several used for effective mitigation monitoring and also does lay down and/or reinforce fundamental principles of bioacoustics monitoring that will be needed for mandated acoustic mitigation zones in east coast waters in the near future. This includes thoroughly characterizing the detection function with both prerecorded calls (Gillespie, 2004, 2019; Shiu et al., 2020) as well as in-situ playbacks where site-specific propagation conditions can be directly accounted for. In doing so, this comprehensive field test has highlighted how bearing information can aid reducing unnecessary and costly shutdowns, while largely maintaining robust conservation outcomes.

It is worth noting that there are a variety of acoustic systems now available for real-time monitoring of right whales, each with its own features and drawbacks. The CABOW system provides continuous noise monitoring, bearing estimation and benefits from a two-stage classification system. However, moving them requires a bigger boat than glider deployment. Other systems including gliders (Johnson et al., 2022) benefit from having longer deployment durations and the ability to move out of high noise areas but are largely acoustically blind during the ascent period of their dives and while data are transmitting. The onboard detector also transmits lower quality, but longer, detection information to the analyst(s) making the final regulatory decision. Cornell University and Woods Hole Oceanographic Institute have also operated a series of near real-time right whale buoys that provide users with audio files of each detection and the systems require very infrequent servicing due to their large on-board battery capacity and additional solar power. However, they do not provide bearing information and are extremely costly to maintain.

Each of these systems can be appropriately deployed for effective right whale management so long as the regulatory structure takes into consideration the benefits and limitations of the systems as well as the environment in which they are placed. If acoustic monitoring platforms are placed near pile driving operations, they should not be used once piling starts as animal sounds will likely be masked. Similarly, all acoustic systems need to alert users to times when high noise effectively blinds the system such that they cannot be relied upon to detect animals. Right whales also change their calling behaviour throughout their range. Lactating females call less frequently and softer than other age and sex classes thereby limiting their detection range (Parks et al., 2019). In such cases, failing to detect a single call may result in failing to detect the animal. In other context where animals are chattier, such as in Cape Cod Bay, the probability of detecting an animal is still quite high even if a small proportion of calls are missed. Finally, no acoustic system can monitor whales when they are not calling, which can be a significant portion of the time (Matthews & Parks, 2021). Users and regulators must therefore consider the entirety of each system within the context it's being deployed as well as rely on other detection modalities including visual observations, in order to ensure maximum protection for the species while supporting the transition to renewable energy.

AUTHOR CONTRIBUTIONS

Kaitlin J. Palmer, Sam Tabbutt, Douglas Gillespie, Jesse Turner and Jason Wood conceived the ideas and designed the methodology. All authors participated in the data collection; Kaitlin J. Palmer, Sam Tabbutt and Douglas Gillespie analysed the data; Kaitlin J. Palmer and Jason Wood led the writing of the manuscript. Jason Wood and Dominic Tollit acquired initial funding. All authors contributed critically to the drafts and gave the final approval for publication.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Wind Energy Technologies Office Award Number DE-EE0008732. The Maryland Department of Natural Resources and the Maryland Offshore Wind Development Fund at the Maryland Energy Administration cosponsored this work. We are grateful to for the guidance of Dr. Sofie Van Parjis and the crews of the RV Sharpe and John Swallow. We are grateful to Geoff Cram, John Mickett, and Derek Martin for their support on hardware design. We thank Smultea Sciences and their PAM operators (Shannon Cremeans and Victoria Schaefer) for help in improving the PAMGuard interface and field trial validations. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer or otherwise does not necessarily constitute or imply its endorsement, recommendation or favouring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

CONFLICT OF INTEREST

The authors do not declare any conflict of interests.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13973.

DATA AVAILABILITY STATEMENT

All data in this study are freely accessible to the public. MATLAB and r scripts are available on Zenodo (https://doi.org/10.5281/zenodo.6830862; Palmer, 2022). Compressed audiofiles and extraction software files are downloadable from Dryad https://doi.org/10.5061/dryad.98sf7m0mn (Palmer, Gillespie, et al., 2022).

ORCID

Kaitlin J. Palmer https://orcid.org/0000-0001-7669-9469
Sam Tabbutt https://orcid.org/0000-0003-2221-4555
Douglas Gillespie https://orcid.org/0000-0001-9628-157X
Jesse Turner https://orcid.org/0000-0002-6951-4101
Paul King https://orcid.org/0000-0002-1413-5817
Dominic Tollit https://orcid.org/0000-0001-5822-0194
Jason Wood https://orcid.org/0000-0003-3086-075X

REFERENCES

Bailey, H., Brookes, K. L., & Thompson, P. M. (2014). Assessing environmental impacts of offshore wind farms: Lessons learned and recommendations for the future. *Aquatic Biosystems*, *10*, 1–13.

Bailey, H., Rice, A., Wingfield, J., Hodge, K., Estabrook, B., Hawthorne, D., Garrod, A., Fandel, A., Fouda, L., & McDonald, E. (2019). *Determining habitat use by marine mammals and ambient noise levels using passive acoustic monitoring offshore of Maryland*. OCS study, BOEM, 1-232. US Department of the Interior, Bureau of Ocean Energy Management.

Bailey, H., Senior, B., Simmons, D., Rusin, J., Picken, G., & Thompson, P. M. (2010). Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Marine Pollution Bulletin*,*60*, 888–897.

Barlow, J., Griffiths, E. T., Klinck, H., & Harris, D. V. (2018). Diving behavior of Cuvier's beaked whales inferred from three-dimensional acoustic localization and tracking using a nested array of drifting hydrophone recorders. *The Journal of the Acoustical Society of America*, *144*, 2030–2041.

Baumgartner, M. F., Bonnell, J., Corkeron, P. J., Van Parijs, S. M., Hotchkin, C., Hodges, B. A., Bort Thornton, J., Mensi, B. L., & Bruner, S. M. (2020). Slocum gliders provide accurate near real-time estimates of baleen whale presence from human-reviewed passive acoustic detection information. *Frontiers in Marine Science*, *7*, 100.

Baumgartner, M. F., Bonnell, J., Van Parijs, S. M., Corkeron, P. J., Hotchkin, C., Ball, K., Pelletier, L. P., Partan, J., Peters, D., & Kemp, J. (2019).

Persistent near real-time passive acoustic monitoring for baleen whales from a moored buoy: System description and evaluation. *Methods in Ecology and Evolution*, *10*, 1476–1489.

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

Bröker, K., Gailey, G., Muir, J., & Racca, R. (2015). Monitoring and impact mitigation during a 4D seismic survey near a population of grey whales off Sakhalin Island, Russia. *Endangered Species Research*, *28*, 187–208.

Clark, C. W., & Gagnon, G. C. (2002). Low-frequency vocal behaviors of baleen whales in the North Atlantic: Insights from IUSS detections, locations and tracking from 1992 to 1996. *Journal of Underwater Acoustics*, *52*, 609–640.

Cole, A. K., Brillant, S. W., & Boudreau, S. A. (2021). Effects of time-area closures on the distribution of snow crab fishing effort with respect to entanglement threat to North Atlantic right whales. *ICES Journal of Marine Science*, *78*(6), 2109–2119.

Davies, K. T., & Brillant, S. W. (2019). Mass human-caused mortality spurs federal action to protect endangered North Atlantic right whales in Canada. *Marine Policy*, *104*, 157–162.

Davis, G. E., Baumgartner, M. F., Bonnell, J. M., Bell, J., Berchok, C., Thornton, J. B., Brault, S., Buchanan, G., Charif, R. A., & Cholewiak, D. (2017). Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014. *Scientific Reports*, *7*, 13460.

Farcas, A., Thompson, P. M., & Merchant, N. D. (2016). Underwater noise modelling for environmental impact assessment. *Environmental Impact Assessment Review*, *57*, 114–122.

Fladung, S., Robbins, M., Spaulding, E., & Clark, C. W. (2011). *Flexible infrastructure for near-real-time acoustic monitoring of right whales and other marine species*. *OCEANS 2011*. IEEE.

Gillespie, D. (2004). Detection and classification of right whale calls using an 'edge' detector operating on a smoothed spectrogram. *Canadian Acoustics*, *32*, 39–47.

Gillespie, D. (2019). *DCLDE 2013 workshop dataset*. University of St Andrews Research Portal.

Gillespie, D., & Caillat, M. (2008). Statistical classification of odontocete clicks. *Canadian Acoustics*, *36*, 20–26.

Gillespie, D., Castellote, M., Maginnis, A., Garner, C. & Small, R. (2013). *Real-time monitoring for beluga whales* (Delphinapterus leucas*) in the Eagle River, Alaska using a PAMBuoy® detection system*. Final Contract Report by SMRU, LLC and Marine Instruments Ltd., 19 p.

Gillespie, D., & Macaulay, J. (2019). Time of arrival difference estimation for narrow band high frequency echolocation clicks. *The Journal of the Acoustical Society of America*, *146*, EL387–EL392.

Gillespie, D., Oswald, M., Hastie, G., & Sparling, C. (2022). Marine mammal HiCUP: A high current underwater platform for the long-term monitoring of fine-scale marine mammal behavior around tidal turbines. *Frontiers in Marine Science*, *283*. https://doi.org/10.3389/fmars.2022.850446

Guazzo, R. A., Helble, T. A., Alongi, G. C., Durbach, I. N., Martin, C. R., Martin, S. W., & Henderson, E. E. (2020). The Lombard effect in singing humpback whales: Source levels increase as ambient ocean noise levels increase. *The Journal of the Acoustical Society of America*, *148*, 542–555.

Helble, T. A., D'Spain, G. L., Hildebrand, J. A., Campbell, G. S., Campbell, R. L., & Heaney, K. D. (2013). Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones. *The Journal of the Acoustical Society of America*, *134*, 2556–2570.

Johnson, H., Taggart, C. T., Newhall, A. E., Lin, Y.-S., & Baumgartner, M. F. (2022). Acoustic detection range of right whale upcalls identified in near real time from a moored buoy and a Slocum glider. *Journal of the Acoustical Society of America*, *151*, 2558–2575.

Klinck, H., Mellinger, D. K., Klinck, K., Bogue, N. M., Luby, J. C., Jump, W. A., Shilling, G. B., Litchendorf, T., Wood, A. S., & Schorr, G. S. (2012). Near-real-time acoustic monitoring of beaked whales and other cetaceans using a Seaglider™. *PLoS One*, *7*, e36128.

Lebon, K. M., & Kelly, R. P. (2019). Evaluating alternatives to reduce whale entanglements in commercial Dungeness crab fishing gear. *Global Ecology and Conservation*, *18*, e00608.

Madsen, P. T., Wahlberg, M., Tougaard, J., Lucke, K., & Tyack, A. P. (2006). Wind turbine underwater noise and marine mammals: Implications of current knowledge and data needs. *Marine Ecology Progress Series*, *309*, 279–295.

Marchetto, P., Strickhart, A., Mack, R. & Cheyne, H. (2012). *Temperature compensation of a quartz tuning-fork clock crystal via post-processing*. IEEE International Frequency Control Symposium (FCS), pp. 1–4.

Matthews, L. P., & Parks, S. E. (2021). An overview of North Atlantic right whale acoustic behavior, hearing capabilities, and responses to sound. *Marine Pollution Bulletin*, *173*, 113043.

Metsers, E. M., Seddon, P. J., & van Heezik, Y. M. (2010). Cat-exclusion zones in rural and urban-fringe landscapes: How large would they have to be? *Wildlife Research*, *37*, 47–56.

Okey, T. A., Banks, S., Born, A. F., Bustamante, R. H., Calvopiña, M., Edgar, G. J., Espinoza, E., Fariña, J. M., Garske, L. E., & Reck, G. K. (2004). A trophic model of a Galápagos subtidal rocky reef for evaluating fisheries and conservation strategies. *Ecological Modelling*, *172*, 383–401.

Palmer, K., Wu, G.-M., Clark, C., & Klinck, H. (2022). Accounting for the Lombard effect in estimating the probability of detection in passive acoustic surveys: Applications for single sensor mitigation and monitoring. *The Journal of the Acoustical Society of America*, *151*, 67–79.

Palmer, K.J. (2022). *CABOW processing scripts*. https://doi.org/10.5281/zenodo.6830862

Palmer, K. J., Gillespie, D., Turner, J., Tabbutt, S., King, P., Tollit, D., Thompson, J., & Wood, J. (2022). Recordings from the evaluation of a coastal acoustic buoy for cetacean detections, bearing accuracy, and exclusion zone monitoring. *Dryad*, *Dataset*. https://doi.org/10.5061/dryad.98sf7m0mn

Parks, S., Searby, A., Celerier, A., Johnson, M., Nowacek, D., & Tyack, P. (2011). Sound production behavior of individual North Atlantic right whales: Implications for passive acoustic monitoring. *Endangered Species Research*, *15*, 63–76.

Parks, S. E. (2003). Response of North Atlantic right whales (*Eubalaena glacialis*) to playback of calls recorded from surface active groups in both the north and South Atlantic. *Marine Mammal Science*, *19*, 563–580.

Parks, S. E., Clark, C. W., & Tyack, P. L. (2007). Short- and long-term changes in right whale calling behavior: The potential effects of noise on acoustic communication. *The Journal of the Acoustical Society of America*, *122*, 3725–3731.

Parks, S. E., Cusano, D. A., Van Parijs, S. M., & Nowacek, D. P. (2019). North Atlantic right whale (*Eubalaena glacialis*) acoustic behavior on the calving grounds. *The Journal of the Acoustical Society of America*, *146*, EL15–EL21.

Pendoley, K., & Kamrowski, R. L. (2016). Sea-finding in marine turtle hatchlings: What is an appropriate exclusion zone to limit disruptive impacts of industrial light at night? *Journal for Nature Conservation*, *30*, 1–11.

Racca, R., Austin, M., Rutenko, A., & Bröker, K. (2015). Monitoring the grey whale sound exposure mitigation zone and estimating acoustic transmission during a 4-D seismic survey, Sakhalin Island, Russia. *Endangered Species Research*, *29*, 131–146.

Romagosa, M., Boisseau, O., Cucknell, A.-C., Moscrop, A., & McLanaghan, R. (2015). Source level estimates for sei whale (*Balaenoptera borealis*) vocalizations off the Azores. *The Journal of the Acoustical Society of America*, *138*, 2367–2372.

Shiu, Y., Palmer, K., Roch, M. A., Fleishman, E., Liu, X., Nosal, E.-M., Helble, T., Cholewiak, D., Gillespie, D., & Klinck, H. (2020). Deep neural networks for automated detection of marine mammal species. *Scientific Reports*, *10*, 1–12.

Širović, A., Hildebrand, J. A., & 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.

Smirnov, E. (2013). *North Atlantic right whale call detection with convolutional neural networks*. Proc. int. conf. on machine learning, Atlanta, USA, Citeseer, pp. 78–79.

Southall, B. L., Finneran, J. J., Reichmuth, C., Nachtigall, P. E., Ketten, D. R., Bowles, A. E., Ellison, W. T., Nowacek, D. P., & Tyack, P. L. (2019). Marine mammal noise exposure criteria: Updated scientific recommendations for residual hearing effects. *Aquatic Mammals*, *45*, 125–232.

Spaulding, E., Robbins, M., Calupca, T., Clark, C.W., Tremblay, C., Waack, A., Warde, A., Kemp, J. & Newhall, K. (2009). *An autonomous, near-real-time buoy system for automatic detection of North Atlantic right whale calls*. Proc. meet. Acoust., Acoustical Society of America, pp. 010001.

Swartz, S., Clapham, P., Cole, T., Barlow, J., McDonald, M., Oleson, E., & Hildebrand, J. (2000). Locating and enumerating endangered humpback whales in the eastern Caribbean with directional (DIFAR) sonobuoys. *The Journal of the Acoustical Society of America*, *108*, 2540.

Teilmann, J., & Carstensen, J. (2012). Negative long term effects on harbour porpoises from a large scale offshore wind farm in the Baltic—Evidence of slow recovery. *Environmental Research Letters*, *7*, 045101.

Thode, A. M., Blackwell, S. B., Conrad, A. S., Kim, K. H., Marques, T., Thomas, L., Oedekoven, C. S., Harris, D., & 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. *The Journal of the Acoustical Society of America*, *147*, 2061–2080.

Thode, A. M., Blackwell, S. B., Seger, K. D., Conrad, A. S., Kim, K. H., & Michael Macrander, A. (2016). Source level and calling depth distributions of migrating bowhead whale calls in the shallow Beaufort Sea. *The Journal of the Acoustical Society of America*, *140*, 4288–4297.

Tougaard, J., Carstensen, J., Teilmann, J., Skov, H., & Rasmussen, P. (2009). Pile driving zone of responsiveness extends beyond 20 km for harbour porpoises (*Phocoena phocoena* [L.]). *The Journal of the Acoustical Society of America*, *126*, 11–14.

Urick, R. J. (1983). *Principles of underwater sound* (3rd ed.). Tata McGraw-Hill Education.

USGS Woods Hole Field Center. (2000). *USGS east-coast sediment analysis and georeferenced displays*. Woods Hole.

Van Parijs, S. M., Baker, K., Carduner, J., Daly, J., Davis, G. E., Esch, C., Guan, S., Scholik-Schlomer, A., Sisson, N. B., & Staaterman, E. (2021). NOAA and BOEM minimum recommendations for use of passive acoustic listening Systems in Offshore Wind Energy Development Monitoring and Mitigation Programs. *Frontiers in Marine Science*, *8*, 760840.

Van Parijs, S. M., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., & van Opzeeland, I. (2009). Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Marine Ecology Progress Series*, *395*, 21–36.

Verfuss, U. K., Gillespie, D., Gordon, J., Marques, T. A., Miller, B., Plunkett, R., Theriault, J. A., Tollit, D. J., Zitterbart, D. P., & Hubert, P. (2018). Comparing methods suitable for monitoring marine mammals in low visibility conditions during seismic surveys. *Marine Pollution Bulletin*, *126*, 1–18.

Weir, C. R., & Dolman, S. J. (2007). Comparative review of the regional marine mammal mitigation guidelines implemented during industrial seismic surveys, and guidance towards a worldwide standard. *Journal of International Wildlife Law and Policy*, *10*, 1–27.

How to cite this article: Palmer, K. J., Tabbutt, S., Gillespie, D., Turner, J., King, P., Tollit, D., Thompson, J., & Wood, J. (2022). Evaluation of a coastal acoustic buoy for cetacean detections, bearing accuracy and exclusion zone monitoring. *Methods in Ecology and Evolution*, *13*, 2491–2502. https://doi.org/10.1111/2041-210X.13973