

# Coastal Acoustic Buoy for Offshore Wind: Project Synthesis

Prepared for DOE EERE & MD Department of Natural  
Resources

[July 2022]

---

SMRU Consulting

815D Spring St, Unit 1  
Friday Harbor, WA 98250  
USA

604 – 55 Water Street  
Vancouver, BC V6B 1A1  
Canada

---

# Coastal Acoustic Buoy for Offshore Wind Project Synthesis

28 July 2022

Prepared by SMRU Consulting

Authors:

Paul King  
Director of Engineering

Kaitlin Palmer, PhD  
Senior Research Scientist

Sam Tabbutt  
Field Engineer

Jason Wood, PhD  
Managing Director

For its part, the Buyer acknowledges that Reports supplied by the Seller as part of the Services may be misleading if not read in their entirety, and can misrepresent the position if presented in selectively edited form. Accordingly, the Buyer undertakes that it will make use of Reports only in unedited form, and will use reasonable endeavours to procure that its client under the Main Contract does likewise. As a minimum, a full copy of our Report must be appended to the broader Report to the client.

## Cover Page

**Agencies to which Report is submitted:**

DOE EERE –Wind Energy Technologies Office  
Maryland Department of Natural Resources

**Recipient:** SMRU, LLC (DUNS# 968300074)

**Award Number:** DE-EE0008732

**MD DNR CCS Contract Number:** 14-20-2731 MEA

**Project Title:** Coastal Acoustic Buoy for Offshore Wind (CABOW)

**Project Period:** 8/1/2019-7/31/2022

**Principle Investigator:** Jason Wood, [jw@smruconsulting.com](mailto:jw@smruconsulting.com), 360-370-5493

**Report Submitted by:** Jason Wood, [jw@smruconsulting.com](mailto:jw@smruconsulting.com), 360-370-5493

**Date of Report:** July 28, 2022

**Working Partners:**

Kaitlin Palmer, [kjp@smruconsulting.com](mailto:kjp@smruconsulting.com)

Paul King, [pk@smruconsulting.com](mailto:pk@smruconsulting.com)

Sam Tabbutt, [st@smruconsulting.com](mailto:st@smruconsulting.com)

Douglas Gillespie, [dg50@st-andrews.ac.uk](mailto:dg50@st-andrews.ac.uk)

**Cost-Sharing Partners:** SMRU, LLC

**DOE Project Team:** DOE HQ Technology Manager – Joy Page  
DOE HQ Technical Advisor – Naomi Lewandowski  
DOE Field Contract Officer – Laura Merrick  
DOE Field Grants Management Specialist – Stephanie Sites  
DOE Field Project Officer – Michael Carella  
DOE/AST Project Monitor – Patrick Linden  
DOE Technology Closeout Coordinator – Corrin MacLuckie

**MD DNR Project Team:** Director, Office of Coastal and Ocean Management –  
Catherine McCall

## Funding Acknowledgement

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 (WETO) Award Number DE-EE0008732. The Maryland Department of Natural Resources and the Maryland Offshore Wind Development Fund at the Maryland Energy Administration co-sponsored this work.

## Disclaimer

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 favoring 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.

## Executive Summary

The population of North Atlantic right whales is critically endangered and their habitat overlaps with offshore windfarm leases. It is therefore imperative that effective mitigation strategies be used to avoid impacts on right whales during the construction of offshore windfarms. The Department of Energy issued FOA Number DE-FOA-0001924 to encourage the development of technology that could monitor large exclusion zones for right whales in order to mitigate potential impact of construction noise on right whales.

This report summarizes the past two years of the development and evaluation of the Coastal Acoustic Buoy for Offshore Wind (CABOW) project which aimed to develop technology to monitor large exclusion zones for North Atlantic right whales. Over the course of the project SMRU Consulting have implemented a rigorous design process including comparison of different approaches (e.g., single sensor vs multiple sensors), as well as consideration of placement and timing of acoustic monitoring.

We have evaluated critical components of the CABOW system including, reliability, detection range, and bearing accuracy in areas adjacent to offshore windfarm leases in Maryland by conducting 3,536 playbacks of simulated right whale upcalls. The maximum call detection range was 7.5 km when noise was 99 dB re 1 $\mu$ Pa rms (50-225 Hz), but this reduced to < 1 km when ambient noise levels were high. Our detection probability in the field was measured as a function of range as well as the source-to-noise level ratio allowing us to build a model to predict the probability of detection under various scenarios (sample size: 3,536 calls x 5 buoys = 17,680). The median bearing error was -0.25° but this is likely an underestimate of error due to experimental design.

Using the published recall and precision of the two detectors we implemented in the CABOW system, we estimate that at a recall of 80%, our precision was > 80%, within the range of what we were aiming for in this project. To estimate our exclusion zone false negative and false positive rates, we built a simulation model using the empirical data from our field trial. We modelled three to nine CABOW units placed on the 10 km exclusion zone and estimated our false negative rate to be 1% or less (which was our project goal) and our false positive rate to be between seven and nine percent, slightly above our goal of 5%. However, we also modelled an equivalent PAM system that does not have bearing capabilities and found the false positive rates for that system to be six to eight times higher than the CABOW rate. This higher false positive rate of PAM systems with low spatial information could have significant cost repercussions for offshore wind developers by adding work shutdowns or delays without providing additional protection for right whales. The model we built allows us to explore the placement of PAM systems under various scenarios and will thus help facilitate planning of PAM mitigation systems to meet NOAA Incidental Harassment Authorizations for specific windfarms.

We achieved an average system uptime of 98.3%, just below our goal of 99%. The issues that caused these short losses of data have been identified and fixed. Right whale detections and audio clips were typically transferred via radio from the buoys to the base station in two to four seconds. We therefore believe we have developed a highly robust real-time PAM system.

Based on the above, we feel we have achieved the stated funding goal of developing a cost-effective and robust real-time PAM system that enables the monitoring of large exclusion zones for right whales during the constructions of offshore windfarms. This should lead to decreased costs and risks for the offshore wind sector while providing robust mitigation for right whales. It is important to state that PAM mitigation will need to be implemented with other mitigation strategies (e.g., visual observers) to provide a complete mitigation strategy to ensure that any effects on right whales from offshore windfarm construction is minimized.

## Table of Contents

Cover Page .....	i
Funding Acknowledgement .....	ii
Disclaimer .....	ii
Executive Summary .....	iii
<b>1 Introduction .....</b>	<b>1</b>
<b>1.1 Project Description .....</b>	<b>1</b>
1.1.1 Project Goals .....	1
1.1.2 Project Scope .....	1
1.1.3 Performance Metrics .....	1
<b>2 Design Phase .....</b>	<b>2</b>
<b>2.1 PAM Approach .....</b>	<b>2</b>
<b>2.2 Hardware Design .....</b>	<b>5</b>
<b>2.3 Software Design .....</b>	<b>8</b>
<b>3 East Coast Field Trials .....</b>	<b>10</b>
<b>3.1 Trial Methods .....</b>	<b>10</b>
<b>3.2 Trial Results .....</b>	<b>13</b>
3.2.1 Detection Probability .....	13
3.2.2 Bearing Error .....	15
3.2.3 System Effectiveness .....	16
3.2.4 System Comparison (Modelled CABOW vs Single Sensor) .....	18
3.2.5 System Uptime .....	20
3.2.6 Communications .....	20
<b>4 Discussion .....</b>	<b>22</b>
<b>5 Acknowledgements .....</b>	<b>24</b>
<b>6 References .....</b>	<b>25</b>

## List of Figures

Figure 1. Existing Coastal Acoustic Buoy (CAB) which was used as a design starting point.....	2
Figure 2. Mooring options considered in the design phase. The left option was considered the most effective PAM solution and most cost effective.....	4
Figure 3. SolidWorks model of the lander design.....	6
Figure 4. Schematic of the CABOW mooring components.....	7
Figure 5. Clip display view of the CABOW base station graphical user interface.....	8
Figure 6. Main display of the CABOW base station graphical user interface, including LTSA display (top left), temporal detection display for each buoy (mid section), alarm history (bottom left), and map (top right) containing bearing estimation to the detected animal. ....	9
Figure 7. Deployment locations of CABOWs (yellow dots) and playback locations (black dots). ....	11
Figure 8. Ambient noise levels at the five CABOW locations. Horizontal lines indicate median, 25th and	

75th percentiles. .... 13

Figure 9. Proportion of upcalls detected as a function of range from a sensor and source-to-noise level ratio (SLNR). The color indicates the probability of detection. .... 14

Figure 10. Proportion of upcalls detected as a function of SNR in the original field data (black circles) and reprocessed recordings with lower gain (gray triangles). .... 15

Figure 11. Bearing error (degrees) at each CABOW location. Graph truncated at 10 degrees..... 16

Figure 12. Simulation comparison of 5 CABOW units on the 10 km exclusion border and 5 single acoustic sensors with the same detection capability. .... 19

Figure 13. Communications latency for audio clips (left) and acoustic detections (right)..... 21

**List of Tables**

Table 1. Pros and cons of several localization methods used on acoustic buoys. .... 3

Table 2. CABOW deployment summary for the field trial. .... 10

Table 3. Upcall arrival ‘confusion’ table indicating the number of correct and incorrectly localized calls (columns) and the total number of calls in the dataset played from inside or outside of the exclusion zone (rows). Peach shading indicates correct location. .... 17

Table 4. Confusion matrix for the individual calls indicating whether each call would have instigated an operational shutdown. Peach shading indicates correct ‘conservation action’ ..... 17

Table 5. Performance metrics between hypothetical CABOW setups and single acoustic sensors. .... 19

Table 6. CABOW system uptime during the field trial. .... 20

## 1 Introduction

The Department of Energy issued the Funding Opportunity Announcement Number DE-FOA-0001924 on 17 July 2018. Topic Area 3 within this FAO focused on the development and validation of offshore wind monitoring and mitigation technologies. Of particular interest to this topic area was ‘the development of tools to address noise impacts on marine species, including systems designed to either mitigate noise or monitor exclusion zones around construction activities, with a particular emphasis on systems for detection and possibly tracking of North Atlantic right whales in or approaching exclusion zones’. SMRU Consulting was successful in our bid to develop and test the Coastal Acoustic Buoy for Offshore Wind (CABOW) under this FOA.

### 1.1 Project Description

#### 1.1.1 Project Goals

The goals of the CABOW project were to develop technology that will monitor exclusion zones for North Atlantic right whales (NARW) during offshore wind construction (with an emphasis on pile driving). In achieving these objectives, we will develop a cost-effective and robust system for monitoring and mitigating U.S. offshore wind construction and thus help minimize potential underwater noise effects on NARW.

#### 1.1.2 Project Scope

The technical scope was to build on existing SMRU Consulting Coastal Acoustic Buoy (CAB) technology by tailoring real-time communications for East Coast offshore wind areas and developing methods to monitor exclusion zones for NARW. We felt this would result in increased performance over current Passive Acoustic Monitoring (PAM) systems, as well as reduce costs and risks for the U.S. offshore wind sector. Specific to increased performance, this project sought to improve NARW detection rates, allow for determination of NARW presence inside or outside exclusion zones, increase the reliability of real-time communication and overall system up time and decrease costs for PAM monitoring of NARW.

#### 1.1.3 Performance Metrics

The following end of project goals were set.

- Recall > 90%,
- Precision > 70%,
- Exclusion zone False Negative < 1%,
- Exclusion zone False Positive < 5%, and
- System up time > 99%.

## 2 Design Phase

The early part of this project focused on developing an effective PAM approach to meet the project goals stated above. During this process we focused on keeping designs as simple as possible so that costs could be kept down, and reliability kept high. Once the PAM approach was agreed we went on to design each necessary subsystem. We used an iterative design, bench test, and field test approach to efficiently progress through the design phase and build of sub-systems. This initial field testing was conducted in Washington State waters for efficient logistics.

### 2.1 PAM Approach

The existing SMRU Consulting CAB system consists of a surface buoy housing an acoustic processing computer, a sound board, batteries, and a communications system (radio or cell modem). The acoustic processor runs marine mammal acoustic detectors and measures ambient noise levels using a variant of PAMGuard software, which is open source and an industry standard for marine mammal mitigation in the oil and gas sector. The surface buoy is attached to an anchor via a rope to which a hydrophone is attached (Figure 1). The CAB is designed for short term (~2 week) deployments aimed at mitigation work in shallow waters. These short-term deployments are limited by battery life. These can be recharged, or a different unit swapped out to maintain coverage. As such it can be deployed by hand from a small work boat. We started our design process from our CAB technology to avoid the inefficiencies of a complete system redesign.

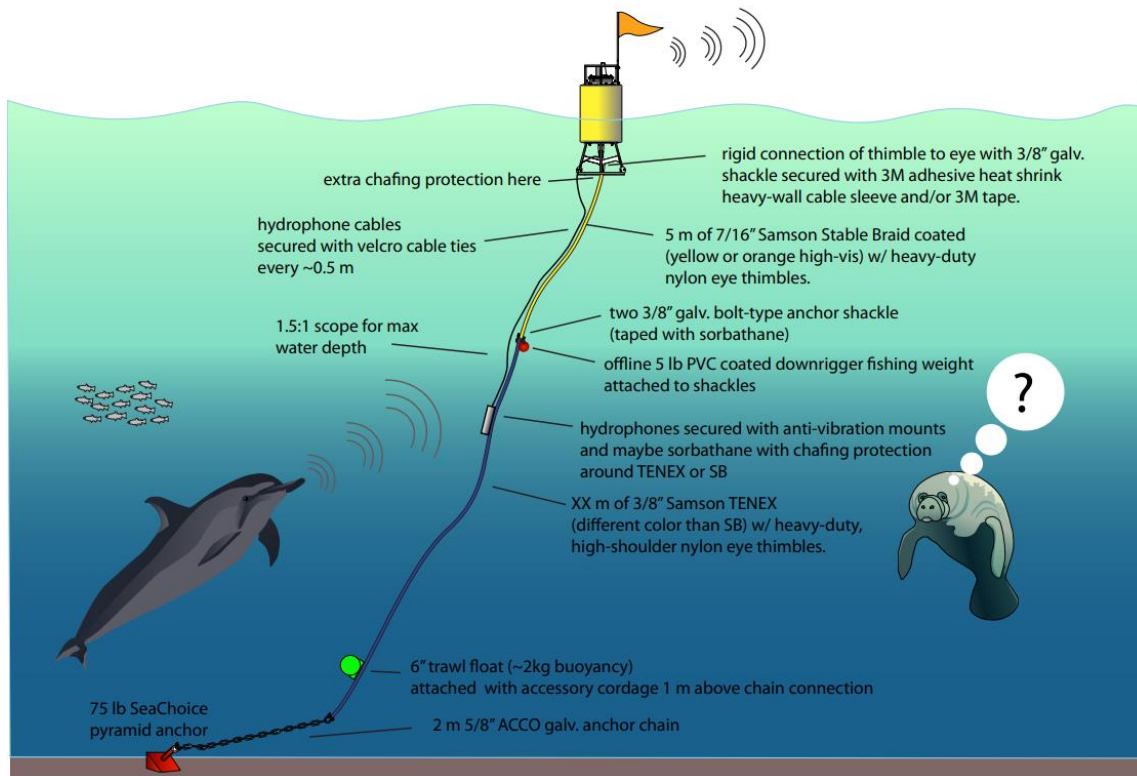


Figure 1. Existing Coastal Acoustic Buoy (CAB) which was used as a design starting point.

One of the key factors affecting a PAM system’s ability to detect NARW or other marine mammal calls is the signal to noise ratio (SNR). In lay terms, a signal of interest that is higher in amplitude than ‘ambient’ noise the system is also recording. Pile driving and other construction noise (e.g., support vessels) will significantly reduce a PAM system’s detection ability. This argues for placing the PAM system away from these noise sources. In addition, how and where you mount your hydrophones can affect noise levels recorded on the hydrophone. This includes flow noise from ocean currents and knocking or banging from the mounting platform itself. This argues for mounting hydrophones where they will be in minimal current and on a stable, quiet platform.

Another key factor in the PAM system’s ability to determine if a calling NARW is inside the exclusion zone is whether some form of localization is being used. Acoustic localization can be achieved with clusters of hydrophones that use time-difference-of-arrival of the signal to estimate a bearing to the calling animal, or a sparse array that uses time-difference-of-arrival of the signal to estimate a location using hyperbolic estimation. The later approach works best if the calling animal is located inside the sparse array. Table 1 lays out the pros and cons of using an acoustic buoy with a single hydrophone (and therefore no localization), an acoustic buoy with a single hydrophone that forms part of a sparse array that is used for hyperbolic localization, and an acoustic buoy with three hydrophones in a cluster to estimate bearings to the calling animal. Based on these pros and cons, it was decided to push forward with a PAM design that uses three hydrophones to estimate bearings.

**Table 1. Pros and cons of several localization methods used on acoustic buoys.**

Monitoring Method	Pros	Cons
1 sensor/buoy: no localization	<ul style="list-style-type: none"> <li>• Least complicated approach</li> <li>• Least expensive for hardware &amp; subsequent software development</li> <li>• Only need 1 sensor</li> </ul>	<ul style="list-style-type: none"> <li>• No directional capabilities</li> <li>• Cannot determine whether signal is within or outside of exclusion zone</li> </ul>
1 sensor/buoy: localization	<ul style="list-style-type: none"> <li>• Only need 1 sensor</li> <li>• Time-difference-of arrival can provide most precise localization</li> </ul>	<ul style="list-style-type: none"> <li>• Calls must be detected by 3+ sensors on 3+ buoys</li> <li>• Under high ambient noise conditions, no localization possible and same problem as single sensor approach</li> <li>• Where multiple animals are calling associating calls to correct time on the hydrophones becomes extremely challenging</li> </ul>
3 sensors/buoy: bearing	<ul style="list-style-type: none"> <li>• Allows for bearing estimates</li> <li>• Miss-association from multiple calling animals not an issue</li> <li>• Bearing angle in combination with known maximum detection range sufficient to determine whether source was within or outside of the exclusion area</li> <li>• Onboard bearing processing capability through PAMGuard</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple sensors increase cost</li> <li>• Ascertaining exact location of calling animal not feasible</li> </ul>

Once the decision was made to cluster three hydrophones on each buoy and estimate bearings, we considered different options for the placement of the key subsystems (communications, electronics, battery, hydrophones). Figure 2 illustrates the options we considered. The option on the left moves all subsystems to the lander except for communications, which are at the surface. The option in the middle moves all the subsystems midwater, except for communications, which are at the surface. The option on the right keeps all subsystems at the surface, except for the hydrophones, which are placed midwater. The option on the right was assessed to drive costs up by requiring a large surface float which drives the need for a larger mooring system, which in turn drives up the cost of deploying and recovering the buoy. The option in the middle places the hydrophones midwater where they are more exposed to ocean currents, and the ability of the hydrophone cluster to rotate creates complexity in estimating the bearing to calling animals. The option on the left was considered the most effective acoustic option and most cost effective and therefore was selected for use in the CABOW system.

In summary, the final PAM approach that we selected moved all subsystems, except communications, to the lander, which also doubles as a mooring system. To minimize construction related noise and take advantage of the bearing capability of this design, the CABOW units would be deployed on or near the edge of the exclusion zone.

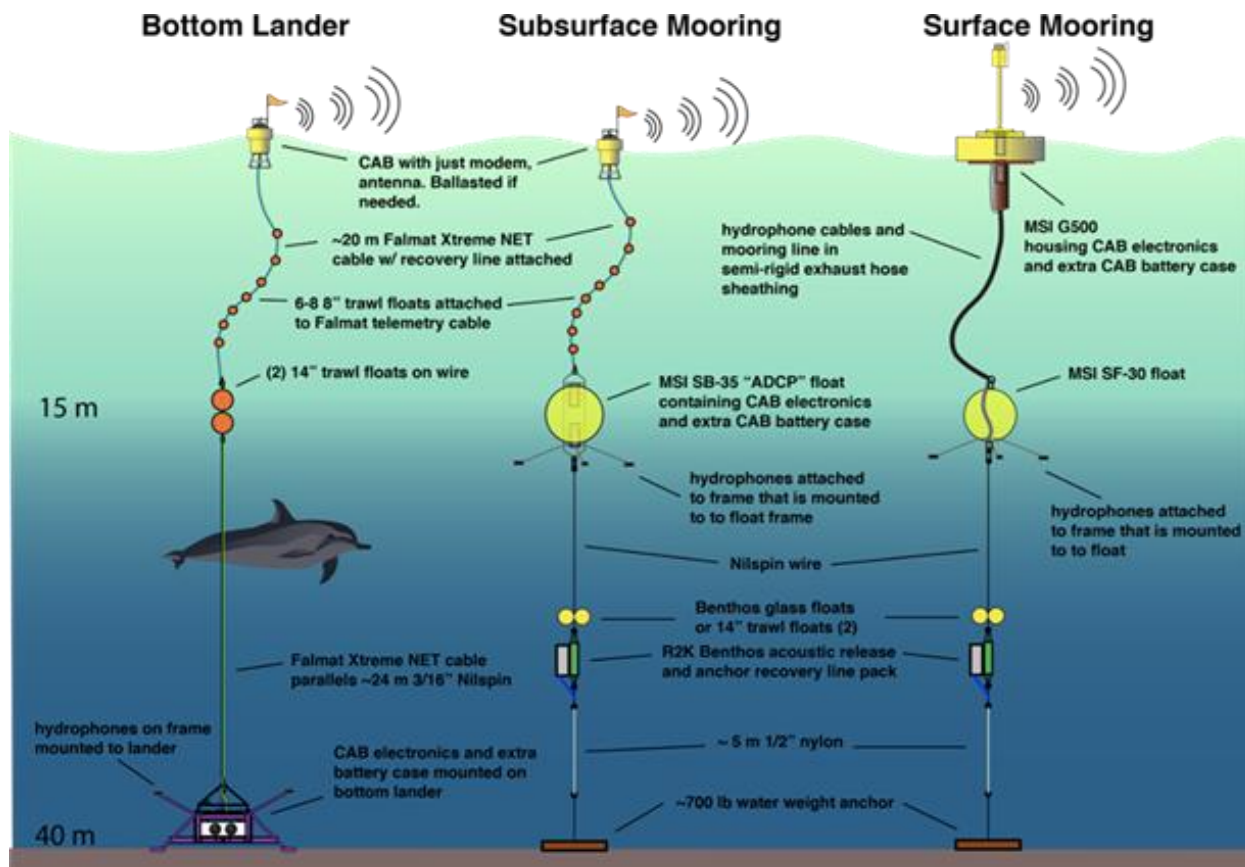


Figure 2. Mooring options considered in the design phase. The left option was considered the most effective PAM solution and most cost effective.

## 2.2 Hardware Design

With the PAM approach selected, we moved onto hardware design. The lander design is shown in Figure 3. The metal frame (yellow components) provides stability and a way to attach components. The outer arms provide hydrophone attachment locations to keep the hydrophones 2 m apart, which was determined to be the appropriate separation distance based on the wavelength of NARW upcalls. HTI 96min hydrophones were selected based on their frequency sensitivity (2 Hz to 30 kHz) and cost. The three housings on the lander are eight-inch schedule 80 PVC which provides cost effective watertight housings with operational depths of 60 m. The outer two housings contain lithium-ion rechargeable batteries with 385-amp hour capacity at 16.8V which provide deployment durations up to 7 weeks. The central housing holds electronics. These include a four-channel sound board capable of sampling at 250 kHz and 16-bit depth, as well as a Linux processor that runs the real-time acoustic analyses. For this project we sampled at 2 kHz. The analyzed acoustic data are stored onboard the buoy and sent via the communication system to the base station. In addition, raw audio data are stored onboard the buoy for later auditing or evaluation. A fiber grate (not shown) prevents the unit from sinking into the substrate and allows for attachment of additional ballast weight if conditions require.

The lander is connected to the surface float (tele-buoy) with a Falmat cable which provides data and power to the communications subsystem (Figure 4). This cable was selected for its strength to lift the lander (~1,000 lbs) and to minimize any chance of loops forming in the cable which might cause an entanglement risk for NARW. The trawl floats attached to the Falmat cable decouple the surface motion of the tele-buoy from lander and therefore minimize potential acoustic noise in the system. The tele-buoy houses a Zumlink PE9 spread spectrum radio (900 MHz). This was selected after bench and field testing several other radio systems. The tele-buoy also houses a Wi-Fi module that allows for wireless communication with the buoy when in proximity (<100 m).

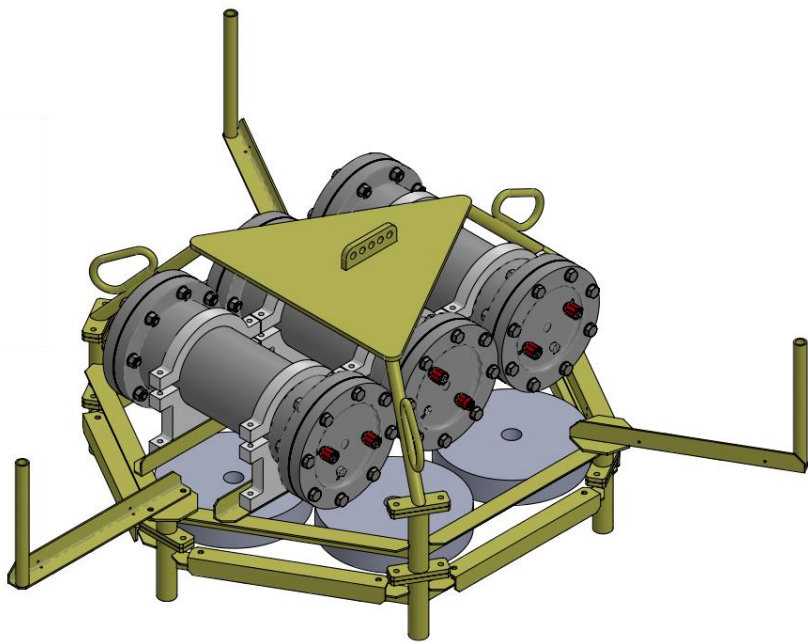


Figure 3. SolidWorks model of the lander design.

### SMRU CABOW Lander 65 m

Version 1.2: Sept 8, 2020

Event Log / Notes:

---



---

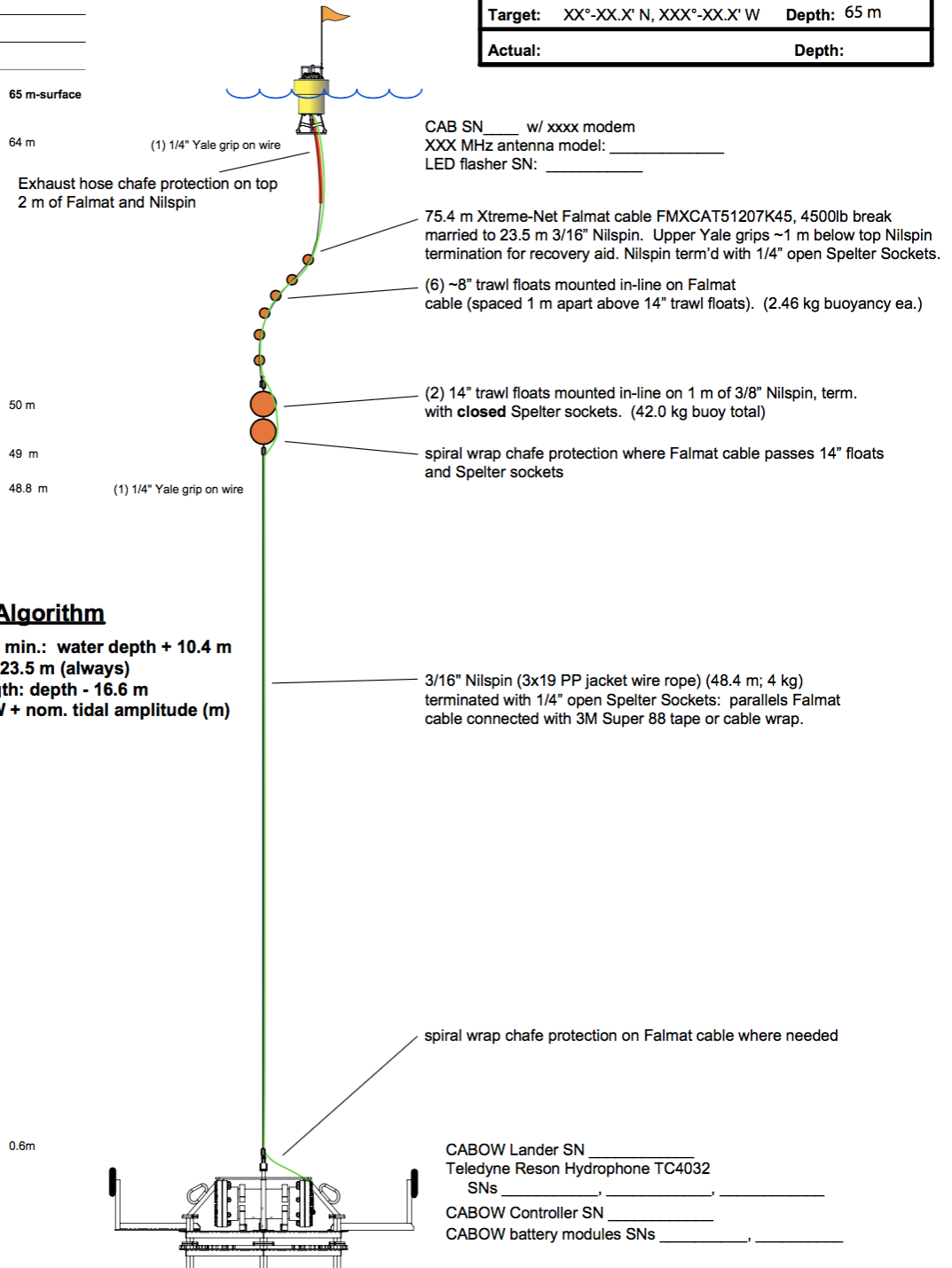


---



---

<b>Target:</b> XX°-XX.X' N, XXX°-XX.X' W	<b>Depth:</b> 65 m
<b>Actual:</b>	<b>Depth:</b>



#### Cable Length Algorithm

Falmat cable length min.: water depth + 10.4 m  
 Top Nilspin length: 23.5 m (always)  
 Bottom Nilspin length: depth - 16.6 m  
 Water depth= MLLW + nom. tidal amplitude (m)

Figure 4. Schematic of the CABOW mooring components.

## 2.3 Software Design

The monitoring software was designed to provide efficient and reliable collection, processing, transmission, filtration, and presentation of data to inform the presence or absence of NARW in the designated exclusion zone.

The Linux processor embedded in each CABOW unit actively collects acoustic data. It processes the acoustic data to generate a Long Term Spectral Average (L TSA) which allows initial acoustic quality assurance. The system also monitors for potential NARW upsweeps in the acoustic signal (detections). In the event of a detection, an estimated bearing to the sound source, the contour of the detection, and a two-second audio clip surrounding the call are transmitted to a centralized computer (base station) for filtration and presentation. LTSA measurements are transmitted to the base station at 30 second intervals. Each CABOW unit works and sends its data to the base station independently.

The base station runs PAMGuard software, with additional processing and filtration modules. We used a trained Convolutional Neural Network (CNN) built on publically available NARW data to act as the second stage filter for detections (Shiu et al., 2020). If the CNN rates a detection with a score greater than 0.5, then the detection passes through the filter to be presented on a graphical user interface (GUI) for a PAM operator to review.

The PAM operator is presented with a spectrogram of the detection and a two-second clip to listen to the detection (Figure 5). The PAM operator is then able to determine whether the detection is a true positive or false positive and annotate the detection as appropriate. This information is stored in a database on the base station.

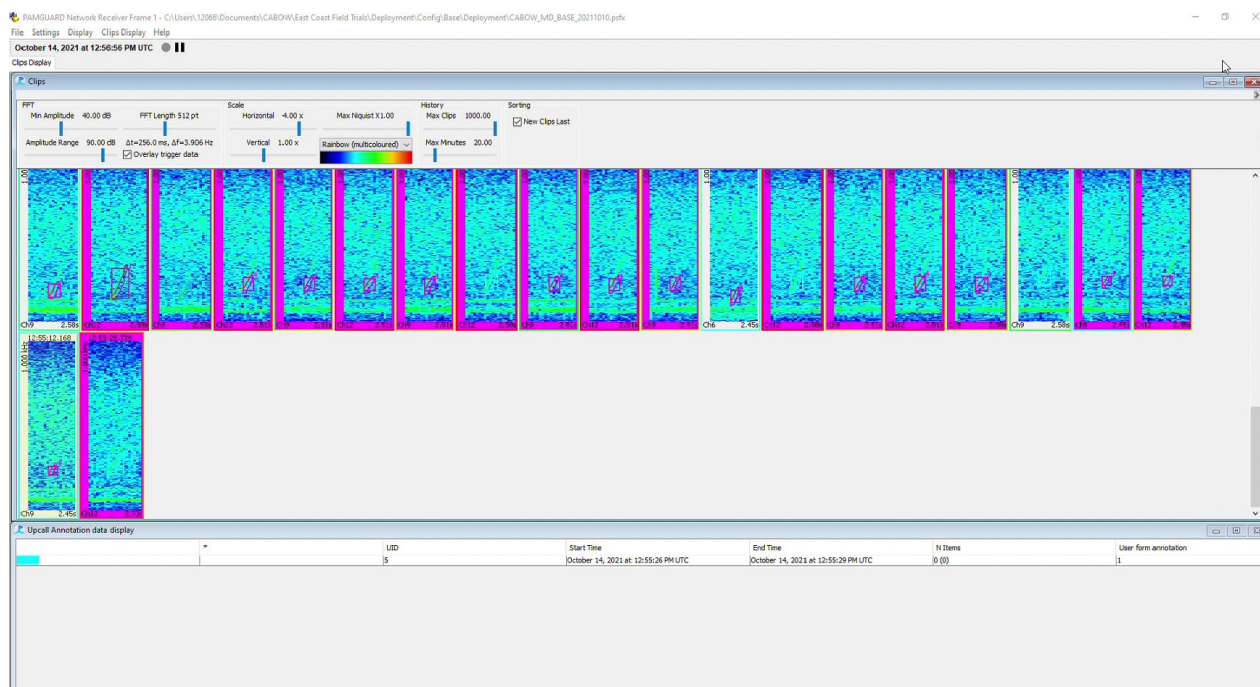


Figure 5. Clip display view of the CABOW base station graphical user interface.

If the PAM operator determines a detection is a true positive, then the calculated bearing to the sound source is displayed on a map view. At this point the PAM operator can determine if the calling animal is inside or outside the designated exclusion zone, and annotate the detection appropriately (Figure 6). If the PAM operator determines the calling animal to be inside the exclusion zone, then an alarm is triggered, which can be configured to play an audible alarm, send an email, or send the detecton information onto a protected species observer application so that it can be more widely shared on or off site. The history of each alarm trigger is tracked (Figure 6).

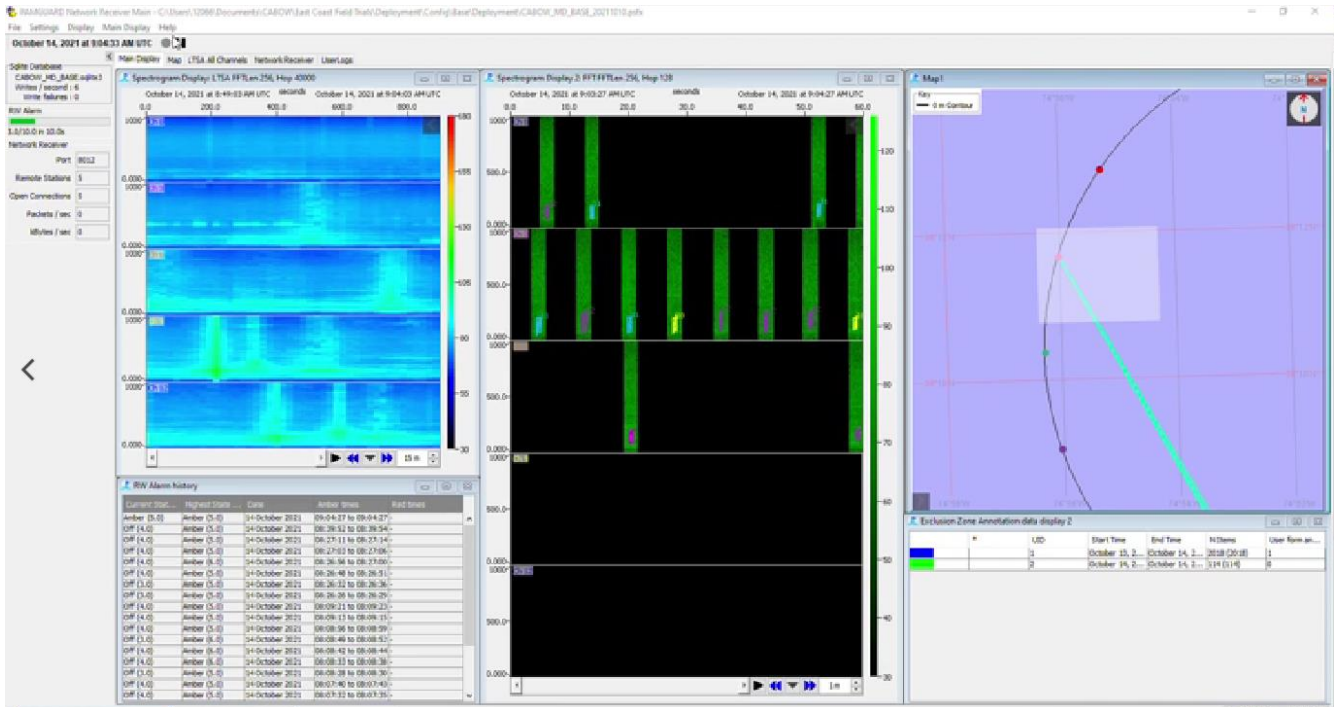


Figure 6. Main display of the CABOW base station graphical user interface, including LTSA display (top left), temporal detection display for each buoy (mid section), alarm history (bottom left), and map (top right) containing bearing estimation to the detected animal.

Each step of the process, including raw audio recordings, is recorded in multiple formats either on the CABOW embedded computer, the base station, or, in most cases, on both systems. These data can be reviewed after NARW detections in many formats for auditing purposes. Raw, continuous audio can also be accessed after recovery of the buoy to evaluate the detector or the presence of other species.

Throughout the project we encountered software bugs, challenges, and a need for restructuring the software components on the CABOW system. We were able to fix these issues along the way to provide a simple and clear operating interface by the end of the project. The software issues were mostly related to visualization, data flow, and data computations. These updates were made to PAMGuard software and are incorporated into the current release of this software.

### 3 East Coast Field Trials

Final testing to determine if we met project performance targets was conducted offshore Maryland in an area adjacent to an offshore windfarm lease.

#### 3.1 Trial Methods

The goals of the field trial were to

- 1) Evaluate total system performance
- 2) Evaluate the range at which calls could be detected
- 3) Evaluate the bearing accuracy

With these goals we could then estimate total system performance in a hypothetical scenario. For this we created a simulation in which any number of CABOW units are deployed at user-defined ranges from an exclusion zone at set source and noise level. This simulation was based on empirical data collected during these trials.

The field trial using simulated right whale up-sweeps was undertaken in the offshore waters of Maryland. The deployment area was chosen based on its proximity to established offshore wind lease areas and relative ease of access. 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 2). CABOW units were spaced approximately 2.5 km apart to ensure that the detection function at short ranges was well characterized. We refer to multiple deployed CABOW units as an ‘array’. 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.

Playback locations were chosen to maximize characterization of the detection function at small ranges and under varying signal and source to noise level ratios and to ensure that the base station, situated on the vessel, could maintain constant connection with all CABOW units while undertaking the playbacks.

Table 2. CABOW deployment summary for the field trial.

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

The playback signal consisted of 30 right whale-like upcall signals separated by seven seconds of silence. Upcall source levels across the 50-225 Hz band were 150-162 dB re  $1\mu\text{Pa}_{\text{rms}}$  at one meter. Calls were played from a Sound Devices 702 recorder and through a Lubell VC2C transducer.

A total of 3,536 simulated upcalls were played in the vicinity (< 10 km) of the CABOW units. After each deployment, one playback set consisting of 30 simulated upcalls was played to orient the CABOW unit. Following orientation, the remainder of the playback locations largely followed one of three defined transects (Figure 7).

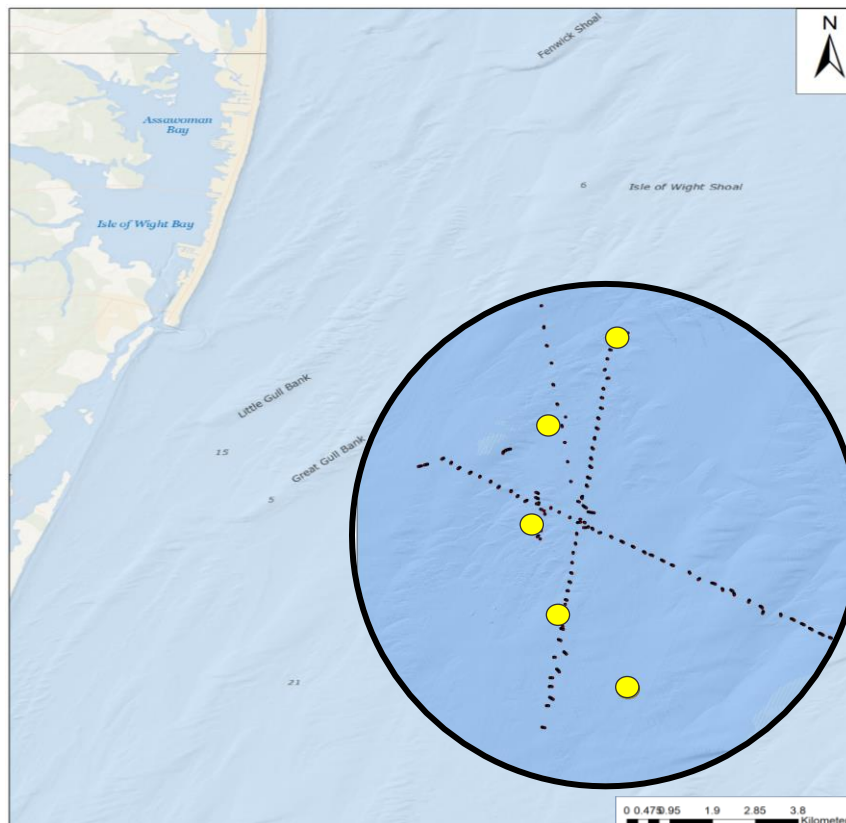


Figure 7. Deployment locations of CABOWs (yellow dots) and playback locations (black dots).

We report the probability of detection as a function of range ( $r$ ) and the source-to-noise level ratio (SLNR) as presented by Thode et al. (2020). Source to noise level is the difference between the signal level at the source (here the underwater speaker) and the received noise level at the recorder (hear each CABOW unit) in dB. 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 moving value based on the ambient noise conditions.

For detected upcalls, ambient noise levels were measured by taking the minimum rms spectra averaged over 0.1 sec, taken from one second before and one second after the call measured over the 50-225 Hz band ( $\text{dB}_{\text{rms}}$ ). The minimum value was chosen over the mean or median to avoid transient signals biasing the ambient noise level and because in bench testing this measurement approach resulted in the most consistent SNR values under fixed signal levels and a variety of different types of background noise. For the non-detected calls, the expected arrival time at the CABOW was estimated assuming a speed of sound of 1462m/s and the ambient noise levels were processed according to the procedure above. The speed of sound value was derived from a CTD cast taken at the beginning of the cruise that indicated a well-mixed layer with no evidence of a thermocline.

A SoundTrap ST300 was used to calculate source level measurements for each of the playbacks in the study. The SoundTrap was affixed to the deployment line 1 m from the playback transducer. Source levels were defined as the rms of the 0.8 sec of data measured over the 50-225 Hz band ( $\text{dB}_{\text{rms}}$ ).

During the playback trials it was observed that the high source level-to-noise level ratios ( $> 60 \text{ dB}$ ) were not being detected at the nearest CABOW units likely due to saturation of the signal, causing spectral leakage obscuring the harmonic structure. To address this and determine appropriate SLNR values that would result in perfect detection at small ranges from the system, the source level was subsequently reduced by 10 dB. The issue was also clear when the proportion of calls detected was plotted against the SNR; very high SNR calls were detected at lower rates than calls with SNR below 15 dB. In post processing we re-processed all data with lower input gain. This was a software problem that was identified and fixed and will not be an issue if future deployments.

Precision is defined by the total number of true detections (e.g. upcalls) divided by total number of detections (upcalls and false positive detections). Recall is the number of upcalls detected divided by the total number of upcalls produced. Precision and recall are typically evaluated for new detectors using annotated data from real (e.g., biological) sources. The evaluation process has previously been done for both detectors used in this project. It would not be sensible to reevaluate the system using playbacks for several reasons. First and foremost, precision and recall have already been established using standardized methods for these two detectors. Because this is a compound system with two detectors an 'edge detector' (Gillespie 2004) and neural network detector (Shiu et al., 2020) the maximum recall is limited by the ability of the edge detector (the first phase) to detect faint or non-standard calls and the precision is limited by the second phase in which the neural network is used to reduce false positive detections. Second, the design of the field trials focused on determining the range at which calls could and could not be detected rather than collecting an ecologically reasonable dataset. The calls used here consisted of simulated upsweeps with sufficient similarity to real right whale calls to trigger both detectors but should not be presumed to match the call diversity of animals. Based on the published values for these two detectors, we expect that for recall values of 0.8 (the maximum reported for the edge detector) we obtained precision values exceeding 0.8.

To evaluate the effectiveness of the system we compared several performance metrics in a simulated

scenario between CABOW and identical single sensor units. We report false shutdown rate as a proportion of detected calls as defined as when a call originating from outside the exclusion zone erroneously triggers a shutdown. We also report a false negative rate as defined by the proportion of detections erroneously placed outside the detection radius that should have triggered a call. In all scenarios the exclusion zone was set to be 10 km with a constant noise field. We report on scenarios with three to nine units to monitor the exclusion zone. For calls detected by multiple instruments, a conservation action was triggered if any bearing estimates crossed the exclusion zone.

## 3.2 Trial Results

### 3.2.1 Detection Probability

Ambient sound levels during the playback period ranged from 91 to 131 dB<sub>rms</sub> (50-255 Hz) re 1 $\mu$ Pa and seemed to be mostly driven by vessel noise. There was very little self-noise from the mooring evident in the recordings. Sound levels across the deployment area were consistent with a maximum of 3 dB difference in minimum, and median noise levels between CABOW units (Figure 8).

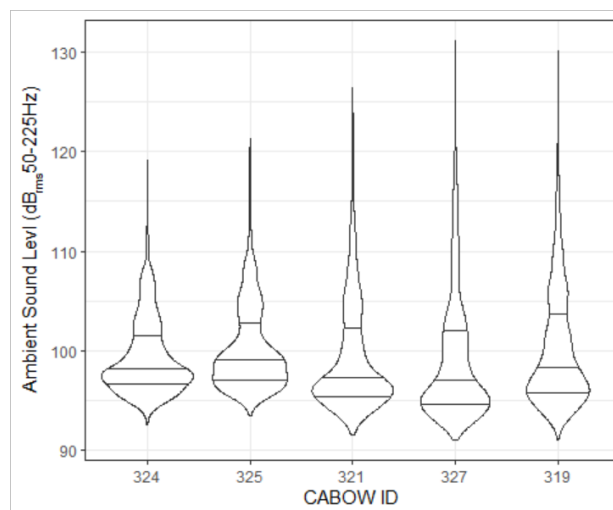


Figure 8. Ambient noise levels at the five CABOW locations. Horizontal lines indicate median, 25th and 75th percentiles.

The observed relationship between the proportion of calls detected, SLNR, and range as determined by the field trials is shown in (Figure 9). Under low SLNR conditions, the maximum detection range was less than 1 km. With high SLNR (>60 dB<sub>50-225Hz</sub>) the maximum detection range exceeded 5km, with one detection at 7.5 km when noise was 99 dB re 1 $\mu$ Pa rms (50-225 Hz). However, at high signal SLNR and low range the probability of detecting an upcall at near distances also decreased. Inspection of the data for low range and high signal excess 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.

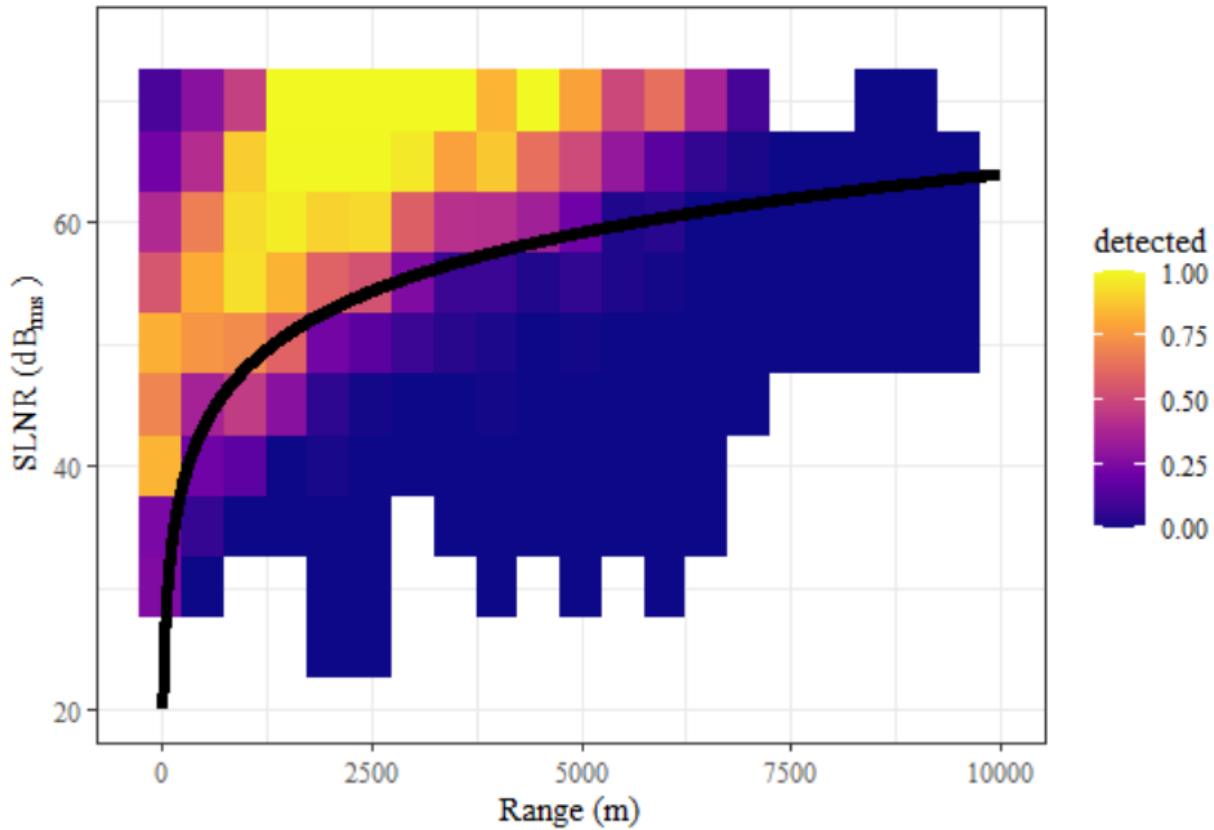


Figure 9. Proportion of upcalls detected as a function of range from a sensor and source-to-noise level ratio (SLNR). The color indicates the probability of detection.

This can also be seen in Figure 10 where the probability of detecting very high SNR calls dropped unexpectedly. We addressed this issue by lowering the gain on the input sensors then reprocessing the data. This resulted in vast majority of high SNR calls being detected (as expected).

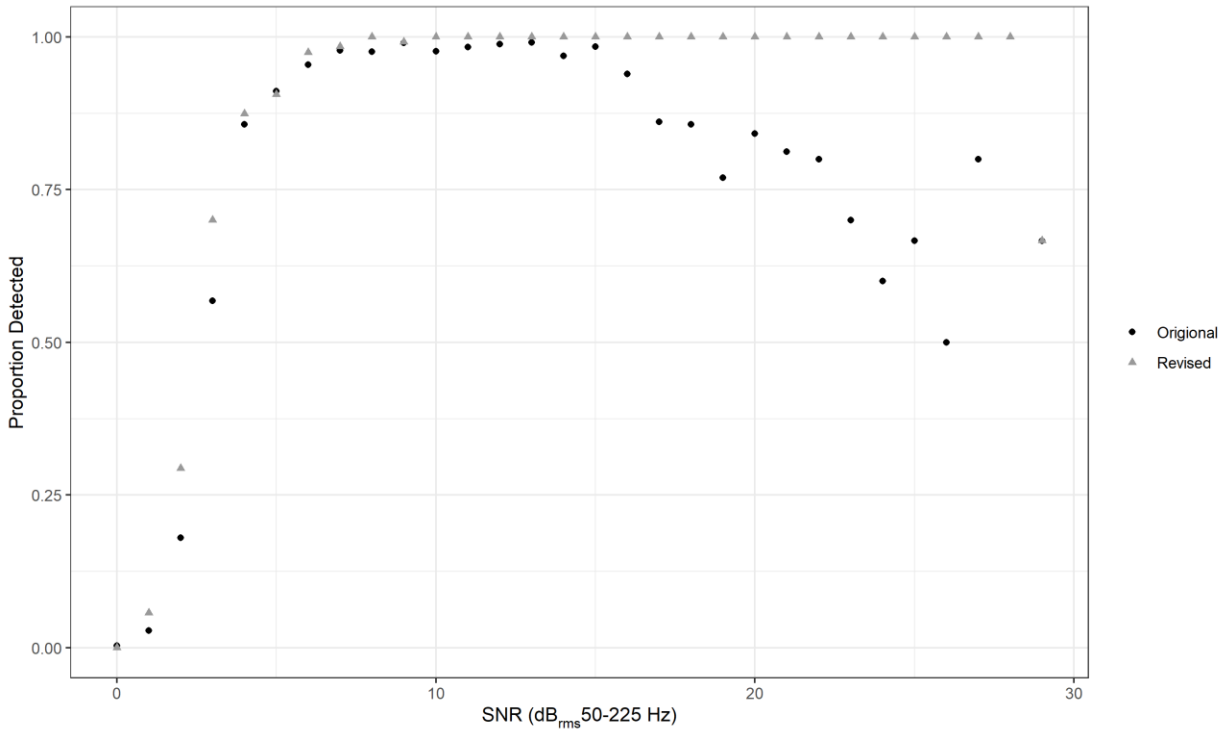


Figure 10. Proportion of upcalls detected as a function of SNR in the original field data (black circles) and reprocessed recordings with lower gain (gray triangles).

### 3.2.2 Bearing Error

The observed bearing error averaged across all units was small. The median, 25<sup>th</sup> and 75<sup>th</sup> quantiles of bearing error were -0.25°, -1.6°, 0.93° respectively (Figure 11). Some amount of bias in bearing error for each CABOW unit 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 error as the playback vessel could not be completely shut down during field trials. Thus, an ambient noise field coherent with the upcalls resulted in better bearing estimates than would otherwise be expected in a real-world situation where the direction of the dominant ambient sound field and the upcalls are independent.

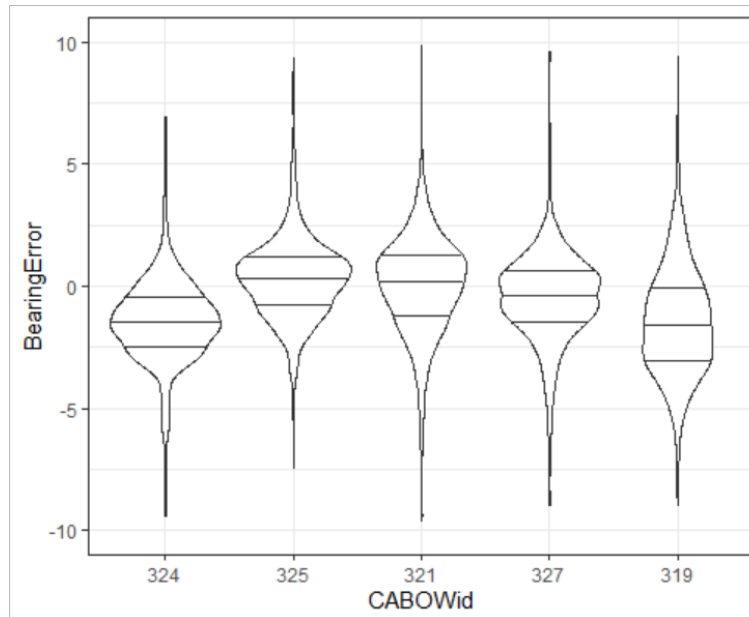


Figure 11. Bearing error (degrees) at each CABOW location. Graph truncated at 10 degrees.

### 3.2.3 System Effectiveness

PAM operators provided by Smultea Sciences were not privy to the location of the vessel or the timing of upcall playbacks. The PAM operators successfully used the PAMGuard interface on the base station (Figure 6) to detect upcalls and determine if they were inside or outside the exclusion zone. Bearing information is provided for every call that is classified as a right whale.

Inspection of the playback data resulted in the discovery of limited and well understood issues in the PAMGuard user interface. These issues resulted in the removal of annotations from the final evaluation dataset. While this had no impact on the evaluation of bearing error and understanding of the detection function, it did result in confusion with regards to the user decision concerning whether a call was in or outside of the exclusion boundary. To avoid any misrepresentation of the improperly annotated data, a subset of data from the field trial was selected which included only detections where we could be confident about the in versus out of exclusion zone annotations were consistent with the original analyst’s intent. We also restricted the data to call arrivals that were detected by the PAM operator, as non-detected calls would not have been localized by the analyst. This resulted in a subset of data consisting of 1,577 arrivals. Of those arrivals, 1,201 originated from inside the exclusion zone and 376 originated from outside of the exclusion zone. A total of 1,420 (90%) upcall arrivals were correctly localized as either in or outside of the exclusion zone and 152 (10%) were incorrectly localized (Table 3).

Table 3. Upcall arrival ‘confusion’ table indicating the number of correct and incorrectly localized calls (columns) and the total number of calls in the dataset played from inside or outside of the exclusion zone (rows). Peach shading indicates correct location.

	Correctly Localized	Incorrectly Localized	Total
Call in Zone	1,187	14	1,201
Call out of Zone	233	143	376
Correct ID	1,420	157	

Table 3 represents the total number of possible arrivals for each call and individual playback calls could be detected by multiple instruments. In the simulation scenario we specify that any call arrival localized within the exclusion zone would result in a shutdown regardless of whether another arrival or arrivals of the same call were localized outside of the exclusion zone. To estimate the overall performance, we looked at individual playback calls and determined whether a call at each location would ultimately trigger a shutdown.

The dataset above represented 694 individual calls played and detected by at least one CABOW unit. We define a ‘correct conservation action’ as either 1) triggering a shutdown if an animal is localized within the exclusion zone or 2) not triggering a shutdown if an animal is localized outside of the exclusion zone and no shutdown is triggered. Of the 694 unique calls included in the analysis, a total of 617 or 88% represented a correct conservation action and 78 calls (12%) represented an incorrect conservation action (Table 4).

Table 4. Confusion matrix for the individual calls indicating whether each call would have instigated an operational shutdown. Peach shading indicates correct ‘conservation action’.

	Triggered Exclusion	Did Not Trigger Exclusion	Total
Call in Zone	523	73	596
Call out of Zone	5	94	99
	528	167	

When interpreting these numbers, it is important to consider that the majority of calls (76%) originated inside the exclusion zone and within detection range of multiple units. With our stringent methodology this likely biased the potential unnecessary shutdowns. In a less contrived deployment where ambient noise levels inside the exclusion zone exceed those outside, we expect lower rates of unnecessary shutdowns due to the geometry of the detection radii and the most likely location of animal detections. Likewise, because many of the playbacks were conducted near the exclusion zone perimeter (Figure 7) the 12% of playbacks inside the exclusion zone that did not trigger an operational shutdown are likely an overestimate. With right whales more randomly spread, this estimate would be lower.

### 3.2.4 System Comparison (Modelled CABOW vs Single Sensor)

With the field results we created a model that compares the monitoring performance of CABOW systems to equivalent single hydrophone PAM systems. In this analysis we sought to quantify the difference in the probability of instituting an unnecessary shutdown between a CABOW system and one where bearing is not known. We did this by first creating a model of the probability of detecting a whale as a function of range and SLNR (i.e., Figure 9). We then simulated all potential right whale locations on a grid with a hypothetical exclusion zone. Noise and whale call source levels were set to biologically reasonable levels and for each grid square we simulated whether or not a call would have been detected based on a model of our empirical data. A Bernoulli random distribution determined the final outcome (correct or incorrect conservation action). For the CABOW bearing approach, we simulated bearing error for each detection by adding random error drawn from a normal distribution with a standard deviation of 15 degrees. This is more than we observed in our empirical data so therefore underestimates the bearing accuracy (and is thus a conservative assumption).

From the simulation we estimated a false positive rate associated with unnecessary shutdowns. An unnecessary shutdown occurs when calls originating from outside the exclusion zone trigger an alarm, either because the bearing estimate was wrong or, for the single sensor, no bearing included (Figure 12). Assuming 165 dB call source level, 100 dB noise level, and 5 units, the likelihood of instigating an unnecessary shutdown was eight times lower with a CABOW system than an equivalent system that does not include bearing, while the probability of missing a right whale inside the exclusion zone was 1% (Table 5). In several additional scenarios derived from the simulation, including CABOW bearing capabilities resulted in a six to eight fold improvement in the probability of instituting an unnecessary shutdown (Table 5). The false negative rate of the CABOW system with a 15° bearing error remained at or below 1% in all scenarios and decreased where calls were detected by multiple instruments (high source level-to-noise ratio or number of units).

### Detection Probability Simulation

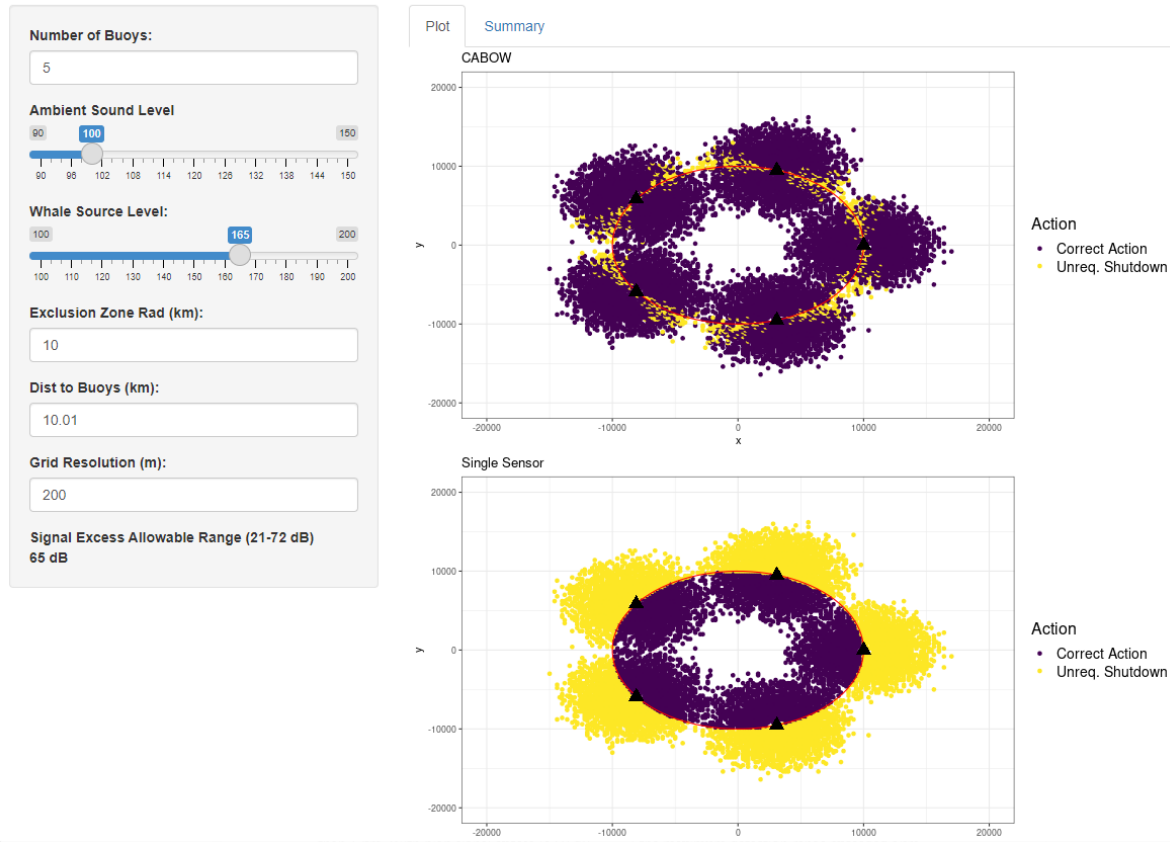


Figure 12. Simulation comparison of 5 CABOW units on the 10 km exclusion border and 5 single acoustic sensors with the same detection capability.

Table 5. Performance metrics between hypothetical CABOW setups and single acoustic sensors.

Exclusion Zone Radius (km)	Units	NL (dB)	SL (dB)	False Positive Rate (CABOW)	False Positive Rate (Detection Only)	False Negative Rate (CABOW)	Percent Improvement	Exclusion Zone Precision (CABOW)	Exclusion Zone Precision (Detection only)
10	3	95	160	0.07	0.57	0.01	662	0.93	0.43
	5	95	160	0.07	0.57	0.01	690	0.93	0.43
	7	95	160	0.07	0.61	0.01	836	0.93	0.39
	9	95	160	0.09	0.62	<0.01	604	0.91	0.38

### 3.2.5 System Uptime

System uptime is defined as the time the system is actively recording acoustic data and analyzing those data for detections. During the three-day field trial, two of the buoys were up 100% of the time, two for more than 97% of the time, and one for just under 97% of the time (Table 6). This resulted in an average uptime of 98.30%. The systems which were not available 100% of the time were down for a maximum of under 2 hours during the field trial, due to a series of system reboots, the cause of which have been identified and corrected.

Table 6. CABOW system uptime during the field trial.

Buoy	Deployed	Recovered	Number of Data Gaps	Total Down Time (Hours)	System Availability (%)
Pb324	10/13/21 19:35	10/15/21 15:33	6	1.86	96.50%
Pb325	10/13/21 20:11	10/15/21 14:46	6	1.29	97.01%
Pb321	10/13/21 20:42	10/15/21 13:41	0	0.00	100%
Pb327	10/13/21 21:09	10/15/21 12:53	0	0.00	100%
Pb319	10/13/21 21:56	10/15/21 11:48	4	0.78	97.98%

### 3.2.6 Communications

Radio communications were reliable and provided the required range and data throughput to meet system operating objectives. The communication system performance was consistent with previous field trials which found reliable communication over 10 km ranges.

Any time that communications were lost between the base station and any of the buoys, the system continued to capture and analyze acoustic data. When communications were re-established, the base station processed the buffered detections (if any) so that a complete record was maintained.

Communications latency was measured for both acoustic clips and for detections. Clip latency had a median duration of 4.25 seconds, and detection latency had a median duration of 2 seconds (Figure 13).

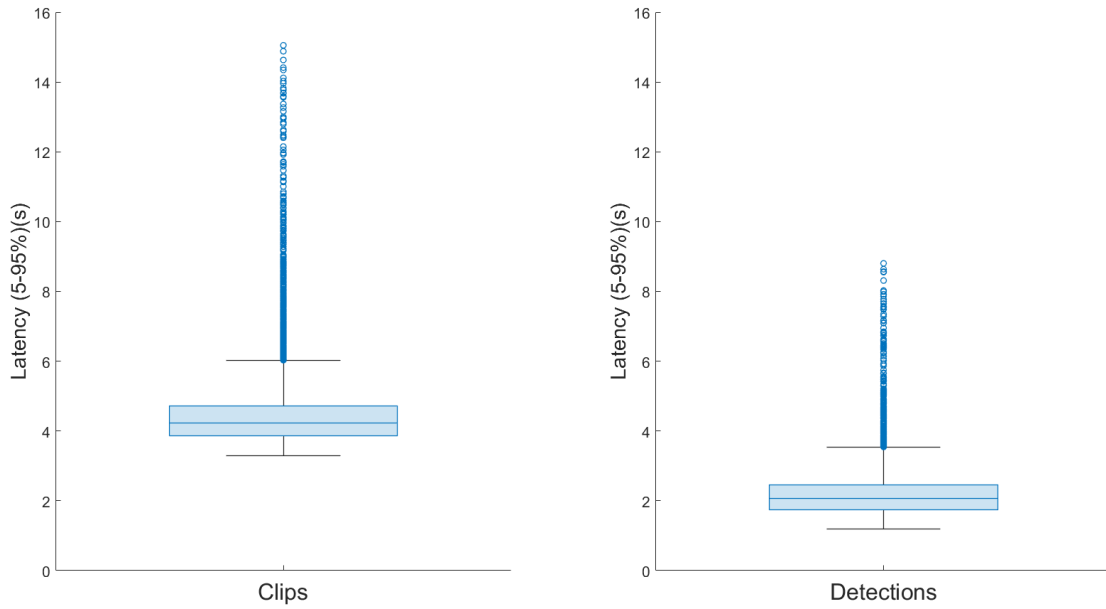


Figure 13. Communications latency for audio clips (left) and acoustic detections (right).

## 4 Discussion

Over the course of this project SMRU Consulting has designed, built and evaluated new PAM technology for monitoring right whales that we believe will help reduce cost and risk of offshore construction in the US offshore wind sector. Throughout the design process we sought to maximize our ability to detect whales and provide as much situational information as possible to PAM operators in the field. We have done this by building robust equipment that provides spatial information on calling NARWs and by using the most up to date acoustic detectors available. We have also striven to keep costs down and maintain transparency by using off-the shelf hardware and open-source software on the user end.

The field evaluation of the CABOW represents, to our knowledge, the best characterized detection system for monitoring NARW available. Our choice to use simulated calls allowed us to avoid relying on very few detections of real NARW calls of unknown location and amplitude, which subsequently resulted in more datapoints with which to estimate the detection range and bearing error than any other system.

We set challenging goals for this project. We aimed for detection recall > 90% and precision > 70%. Because we needed to focus our field trials on estimating a detection function and bearing errors, we could not estimate precision and recall directly. However, these have been done extensively for the two detectors we implemented in the CABOW system. The maximum recall value reported for the PAMGuard edge detector is 80% (Gillespie 2004). At this level of recall, we estimate our precision using the neural network (Shiu et al., 2020) to be > 80%. Increasing our recall will decrease our precision rate, but these numbers suggest we are close to our targets for this project.

Our exclusion zone goals required combining the results from our field trials (i.e., the detection probabilities and bearing errors) into a simulation model that could predict exclusion zone false negative (proportion of detected calls inside the exclusion zone that did not trigger an alarm) and false positive (proportion of detected calls outside the exclusion zone that did trigger an alarm) rates. Based on a few model assumptions of ambient noise levels and NARW call source levels as well as a 10 km exclusion zone, we estimated our false negative rate at 1% or less when three to nine CABOW units are placed on the periphery of the exclusion zone. The number of units will depend on the propagation conditions, ambient noise level, expected source level of the animals in the region and tolerance for false positive and false negative errors. For instance, on calving grounds where animals call less frequently and at lower amplitude and the cost of failing to detect a call is high, we expect regulators to place stricter limits on the false negative rate. In other regions, such as the Gulf of Maine, where animals call more frequently there is a greater likelihood of detecting each animal and as such the tolerance for false negatives may be (slightly) higher, assuming similar noise regimes and propagation loss. This meets our target of <1% exclusion zone false negative rate. In these same model scenarios, our estimated false positive rate ranged from seven to nine percent which is slightly above our goal of < 5%. However, when compared to PAM systems with the same capabilities and locations as the CABOW units, but without bearing or localization capacities, we estimate the CABOW bearing approach to have six to eight times lower false positive rates. This could translate into significant cost savings for the offshore wind sector by minimizing work delays or shutdowns, while

still providing protection for NARW. In addition, the model we have built allows us to quantify the efficacy of different CABOW deployment scenarios with different exclusion zone sizes, ambient noise and NARW call source levels. This could be used as a tool in planning the PAM mitigation required for specific offshore windfarms to meet the Incidental Harassment Authorization requirements issued by NOAA. Factors not included in our simulation that will change the efficacy rate are biased noise fields caused by stationary ships, the presence of acoustically similar species including humpbacks which will necessarily increase the false positive rate and aberrant calling behaviour of the animals such that the edge detector is incapable of picking up the calls.

We achieved an average uptime across the five deployed CABOW units of 98.30%, tantalizingly close to our goal of > 99%. The few data gaps we had occurred for reasons that have been identified and since fixed. Therefore, we are confident that we have designed and built a highly reliable system for monitoring large exclusion zones for NARW. Based on the efficacy and reliability of the CABOW system as well as its six to eight time reduction in unnecessary and costly shutdowns, we feel CABOW provides a cost-effective solution for monitoring large exclusion zones for NARW.

We hope the CABOW system will help lower costs and risks for the US offshore wind sector while providing robust protection for NARW during offshore wind construction, and therefore support a rapid transition to renewable energy. The CABOW system provides a cost-effective approach to monitoring very large NARW exclusion zones by letting PAM operators know whether an acoustic detection is inside of the exclusion zone. In doing so, this system could reduce unnecessary construction shutdowns by factors of six to eight, when compared to PAM systems that can't localize these calls. Based on the above, we feel that we have contributed significantly to the DOE goals laid out in the FOA from 2018.

## 5 Acknowledgements

We are grateful to for the guidance of Dr. Sophie Van Parjis, the support of vessel crews of the RV Sharpe, Jon Swallow, and the Kittiwake, Jakob Bueche. 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.

## 6 References

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