Collaborative Whale Detection Technology Evaluation Virtual Workshop Series

SESSION 1 | APRIL 18, 2024

RWSC
Regional Wildlife Science Collaborative for Offshore Wind

MTS
marine technology society
Opportunity runs deep™

Collaborative Technology Workshop Series

Workshop Session 1: Overview of Whale Detection Technologies and Ways to Evaluate Efficacy

Table of Contents

RWSC
Regional Wildlife Science Collaborative for Offshore Wind

MTS marine technology society
Opportunity runs deep™

Executive Summary

NOAA Fisheries, the Bureau of Ocean Energy Management (BOEM), and U.S. Department of Energy (DOE) are working with each other through the Regional Wildlife Science Collaborative for Offshore Wind (RWSC) to identify models, technologies, and information from other sectors, research areas, and potential partners in support of whale conservation and responsible offshore wind development. To support this effort, a collaborative virtual workshop series is being hosted by the RWSC, the Marine Technology Society (MTS), and the Consensus Building Institute (CBI), in partnership with the Pacific Northwest National Laboratory (PNNL) and National Renewable Energy Laboratory (NREL), with support from DOE and contributions from NOAA, BOEM, and Turn Forward.

The objective of the series is to assess the state of the science and metrics for evaluating technologies, tools, and methods for monitoring baleen whales during sound-producing offshore wind construction activities. PNNL and NREL are leading the development of the technical workshop materials and outputs, with oversight by DOE, NOAA, and BOEM. RWSC, MTS, and CBI are providing the forum, workshop facilitation, and are developing workshop proceedings for each session that capture participant input and discussion.

The first session of the Collaborative Technology Workshop Series was held virtually on April 18, 2024. This report summarizes discussions and key takeaways from this first session. The objectives for this session were to:

- Develop a shared understanding of the applicability of existing technologies or classes of technologies to monitor baleen whales around sound-producing offshore wind construction activities.
- Discuss the types of measures that can be used to evaluate technology efficacy.

Materials from this session including an agenda and draft products produced by PNNL and NREL can be found here: https://bit.ly/3y20gs5.

Workshop participants heard introductory presentations from DOE, RWSC, PNNL and NREL, and an expert in whale behavior. Participants requested clarification from the organizers about the goals of the workshop discussion and the PNNL/NREL technical report. The workshop and the PNNL/NREL technical report are focused on how to evaluate technologies that can be used for real-time baleen whale detection during offshore wind pile driving and construction activities. Some participants felt that the missed detection rate (false negative rate) was a critical metric by which systems could be evaluated and, if an acceptable threshold is identified, systems could be designed to meet that threshold. Organizers asked participants to consider ways to evaluate the quality, quantity, and types of information that a technology can provide in real-time about whale presence and behavior. This information could then be used to inform decisions about systems to deploy and expectations around their performance. There was also general agreement that whale behavior directly influences cue availability, necessitating its consideration in technology evaluation, and that complementary and combined detection technologies could outperform singular methods.

Participants selected a breakout group focused on a particular category of technology (passive acoustic monitoring, thermal imaging, visible light cameras, protected species observers, and others). They discussed attributes of the technology that make it useful or limit its applicability to detect baleen whales during pile driving and what measures or metrics could be used to characterize the technology's performance.

In breakout discussions and in the final plenary session, participants identified the following key themes and next steps for technology evaluation:

- Several categories of technology can inform or contribute to the accuracy of whale detection data, but not all are suitable for real-time monitoring during offshore wind pile driving. The development and evaluation of these other, less suitable technologies should continue to be advanced and flagged as methods that might provide context for or verification of real-time baleen whale detections.
- The integration of multiple technologies will ensure comprehensive monitoring of baleen whale presence during offshore wind construction activities.
- The functionality of technology in different environmental conditions needs to be measured to best evaluate technology effectiveness.
- More discussion of technology performance metrics is needed. Potential metrics to evaluate readiness and effectiveness include:
    - Timeliness of data availability
    - Missed detections, or false negative rate
    - Ability to detect at the species level
    - Detection range and accuracy across levels of environmental factors (sea state, visibility, and others) and across various deployment platform types
    - Level of human intervention needed to interpret results
    - Participants flagged price as a consideration, especially for technologies where multiple devices may be needed at a single offshore wind construction site, but also noted that price is not a traditional performance metric
- There are knowledge gaps in standardization of data produced by various technologies and around how decisions are made with the data collected.

Participants also shared feedback about workshop logistics and structure:

- It is essential to include subject matter experts in PNNL's and NREL's work as it moves forward, including maximizing their participation in future workshops, particularly ensuring the balanced distribution of those experts among any breakout discussions.
- Providing workshop materials ahead of the next workshop session will ensure all participants are prepared for productive discussion.

Acknowledgements

The Collaborative Technology Workshop Series was conceived through a collaboration between the Regional Wildlife Science Collaborative (RWSC), the Marine Technology Society (MTS), American Clean Power (ACP), Turn Forward, several environmental nonprofits, the U.S. Department of Energy (DOE), the Pacific Northwest National Laboratory (PNNL), and the National Renewable Energy Laboratory (NREL). The organizers are grateful for the facilitation provided by Patrick Field from the Consensus Building Institute (CBI).

Opening and Welcome

Presenters: Joy Page, Department of Energy, and Emily Shumchenia, Ph.D, Regional Wildlife Science Collaborative for Offshore Wind

The workshop opened with a welcome from Joy Page, DOE, and Emily Shumchenia, Ph.D, RWSC, who provided the background, context, and purpose of the Collaborative Technology Workshop Series. Bringing together various stakeholders including agencies, eNGOs, and industry players, the series will aim to gather technical insights on monitoring baleen whales during sound-producing offshore wind construction activities.

Key objectives include developing a shared understanding of relevant technologies, discussing their applicability and limitations, and informing the production of a technical report by PNNL and NREL.

Through breakout group discussions in session 1, participants will explore technology categories, understand whale behavior's impact on detection rate, and define performance measurement criteria.

Following each workshop session, RWSC-MTS will produce proceedings summarizing key themes and feedback. PNNL and NREL will develop a technical report outlining technology profiles and opportunities for using combinations of technologies. Key milestones in this process include:

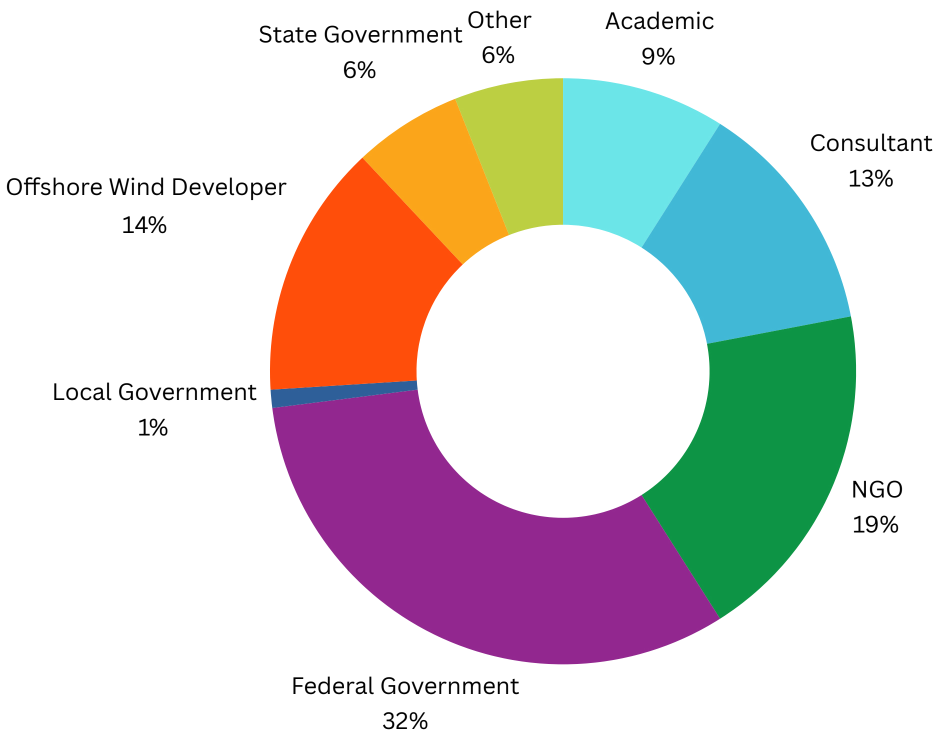
- June: PNNL and NREL finalize a draft report characterizing technologies for monitoring baleen whales around sound-producing offshore wind construction activities. Second workshop session is held.
- June – July: PNNL and NREL conduct peer review on the draft report and revise the draft according to expert input.
- August: PNNL and NREL add to the draft report to include considerations/opportunities for using combinations of technologies. Third workshop session is held.
- September: PNNL and NREL finalize technical report. A final public webinar to present the final report is held.

The workshop attracted significant interest, with nearly 400 registrations and over 200 attendees spanning various sectors and locations.

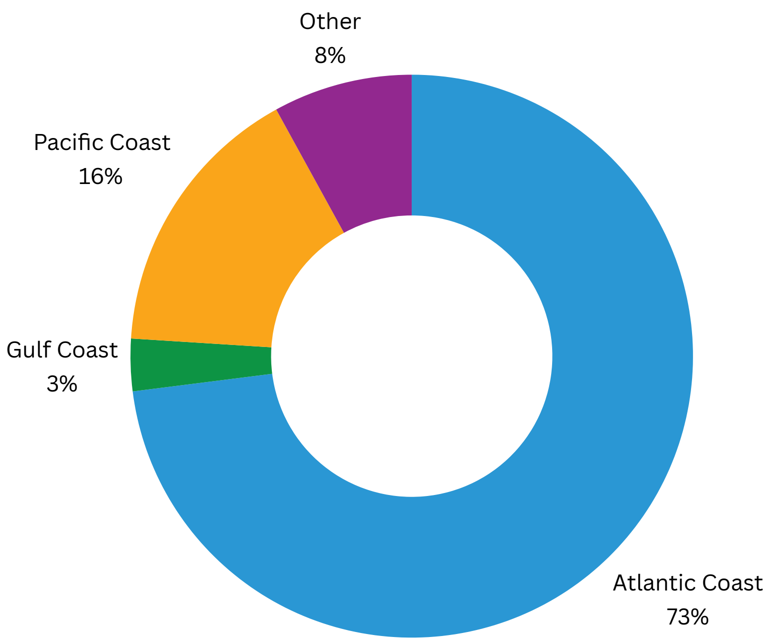
Participant Overview

399 Persons Registered

225 Persons Attended



Participants by Sector



Location of Participants

Overview of Real-Time Whale Detection Technology Categories and Potential Means to Evaluate Efficacy

Presenters: Jeff Clerc, Ph.D, National Renewable Energy Laboratory, and Angela Szesciorka, Ph.D., Pacific Northwest National Laboratory and National Renewable Energy Laboratory

Key Takeaways

A team from PNNL and NREL developed two draft products to guide the evaluation of existing technologies to detect and monitor baleen whales in real-time during offshore wind turbine construction (pile driving).

- Product 1: The Technology Matrix provides an initial high-level evaluation of existing detection technologies and uses draft inclusion criteria to determine if technologies have practical applications for monitoring during pile driving.
- Product 2: The Technical Report Containing Technology Profiles presents a draft performance evaluation framework for applicable technologies (identified in Product 1) by analyzing performance drivers, assessing approaches to validation, and identifying research and development needs.

Jeff Clerc, Ph.D. (NREL), and Angela Szesciorka, Ph.D. (PNNL, NREL) presented an overview of the two draft products, developed by PNNL and NREL, aimed at identifying and describing a set of technologies that may be useful for detecting baleen whale presence during pile driving associated with offshore wind turbine construction. This presentation included a walkthrough of the products and their development process to inform workshop participants' discussion during the breakout session.

Product 1: Technology Matrix

This product was developed as a high-level characterization of existing baleen whale detection technologies. The matrix compiles relevant specifications of monitoring technology such as detection cue, estimated detection ranges of the cue, and automation capabilities. The matrix is being used by PNNL and NREL to identify technologies applicable to real-time baleen whale detection during offshore wind construction activities by using the following draft inclusion criteria:

1. Ability to record the time and location of an animal?
2. Can the detection cue be detected and delivered within the decision-making time window?
3. Practical ability to detect any individual whale that enters the zone of perception in ideal conditions and displays/produces the cue?

Technologies that meet all three of these criteria will be included in Product 2.

Product 2: Technical Report Containing Technology Profiles

This product proposes a performance evaluation framework for evaluating real-time baleen whale detection technologies to understand how compounding variables influence system performance and detection probability. The final version of this report will include detailed profiles of each applicable technology, discuss drivers of performance, and address approaches for validation with research and development needs. The report will not propose discrete criteria or thresholds to initiate action during offshore wind construction. The draft report discussed in this workshop proposes to include 3 technologies that meet the draft inclusion criteria developed in Product 1: Passive Acoustic Monitoring, Thermal Imaging, and Visible Light Camera.

Summary of Q&A Session

Questions and discussions from workshop participants included:

Clarification on Scope and Inclusion Criteria: The presenters clarified that the scope of these products is limited to technologies that can be used to detect baleen whales real-time during offshore wind construction, primarily pile driving. They clarified that technologies that do not meet the initial high-level criteria outlined in the Technology Matrix may still be critical in understanding whale behavior and validating technologies.

Detection Probability and Grouping of Animals: Workshop participants encouraged PNNL and NREL to be more specific in the detection probability category and consider factors such as detection range and animal behavior. Participants also encouraged PNNL and NREL to consider whether technologies simply detect or can enumerate animals as the current categories do not include the detection of individuals vs. groups of baleen whales.

Considering Covariates in Analysis: PNNL and NREL clarified that they are looking at covariates such as animal behavior and sensor and platform configurations for their technology evaluation. Participants encouraged the presenters to also consider how environmental conditions such as sea surface conditions and temperatures may impact detection probability. Participants also recommended that the team consider animal physiology as a covariate for detection.

Inclusion of Telemetry and Tagging: Participants noted that telemetry and tagging did not meet the initial high-level criteria applied to the Technology Matrix. Participants added that tagging advancements have increased the utility of this method for characterizing species' cue rates and behavior, which provide important context for assessing the efficacy of other real-time systems, like passive acoustics. Participants were encouraged to further this discussion in the "Other Technologies" breakout group.

Inclusion of Night Vision: Participants indicated that night vision was not discussed in the presentation. The team from PNNL and NREL explained this would fall under visible light and the team will clarify the wavelengths devices are tuned to capture.

Outreach to Technology Development Industry: Participants recommended that the workshop hosts reach out to technology development industry members and include them in future workshop sessions.

Baleen Whale Behavior and the Use of Whale Detection Technologies

Presenters: Julia Dombroski, Ph.D., National Offshore Wind Research and Development Consortium

Key Takeaways
• Behavior directly influences cue availability, necessitating its consideration in technology evaluation. • Behavior is dynamic, leading to changing cue availability over time. • Complementary detection technologies outperform singular methods.

Julia Dombroski from the National Offshore Wind Research and Development Consortium (NOWRDC) provided examples of how baleen whale behavior can impact the efficacy of whale detection technologies. The presentation highlighted ways that behavior directly influences the availability of cues needed for detection. No technology type can detect all types of cues in all scenarios, and whales do not produce cues constantly (behavior is complex). Julia advocated for the use of multiple complementary technologies instead of solely relying on one method. A complimentary approach significantly enhances the ability to detect whales and make well-informed decisions.

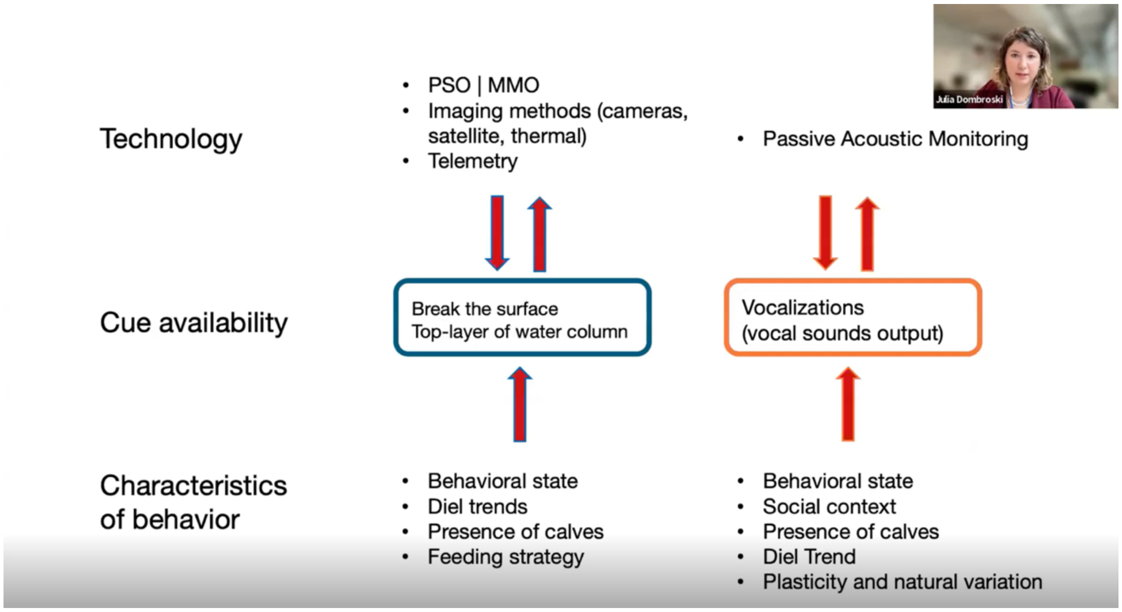


Figure 1. Examples were provided to show how whale surface cues and vocal cues needed for detection are affected by behavior.

Surface Cues

Surface cues refer primarily to visual detections at the water's surface or in the top layer of the water column. Technologies that detect surface cues include protected species observers and imaging methods (such as cameras, satellites, and thermal imaging). Satellite telemetry devices also require the animal to surface in order to transmit signals via satellite. Julia provided examples for how whale surface cues are affected by behaviors including feeding and traveling, and by diel factors such as light availability.

Vocal Cues

Vocal cues involve the vocal outputs of whales such as calls and are detected with passive acoustic monitoring technologies. Julia provided examples for how whale vocal cues differ in frequency and type during different behaviors such as traveling, milling, foraging, and resting.

Julia described a study (Dombroski et. al., 2021) that revealed that baleen whales exhibit intermittent calling patterns with periods of high vocal activity followed by extended periods of silence, posing challenges for continuous acoustic tracking. Group size also influences call rates with solitary whales producing fewer calls. North Atlantic right whale females accompanied by calves produce high rates of low-intensity calls which are less detectable by passive acoustic monitoring. Furthermore, whales exhibit plasticity in call patterns in response to noise, sometimes ceasing calling altogether to cope with noisy environments, which further reduces the availability of cues.

Summary of Q&A

Following Julia's presentation, there were a few clarifying points made:

- It was noted that tagging data is an important part of the toolkit for understanding cue rates and relationships between cues and whale behavior.
- The discussion also touched upon the significance of group size variation among different whale species. While some species consistently exhibit similar group sizes, others display variability. This variability in group size is crucial to consider as it directly impacts detectability, offering another factor to account for in research and monitoring efforts.
- With new suppliers entering the market, there's a need for standards to evaluate the efficacy of various equipment. This is important for developers and regulators alike to understand capabilities in detection. We need to integrate this efficacy assessment into the broader context of creating appropriate mitigation systems.

Conclusion

In conclusion, understanding baleen whale behavior is essential for optimizing detection technologies. By considering behavior-induced variations in cue availability, researchers can develop more effective detection strategies.

Pre-Breakout Group Discussion

Key Takeaways

- The workshop and the PNNL/NREL technical report are focused on how to evaluate technologies that can be used for real-time baleen whale detection during offshore wind pile driving and construction activities.
- Some participants felt that the missed detection rate (false negative rate) was a critical metric by which systems could be evaluated and, if an acceptable threshold is identified, systems could be designed to meet this threshold.
- Organizers asked participants to consider ways to evaluate the quality, quantity, and types of information that a technology can provide, about whale presence and behavior in real-time. This information could be used to inform decisions about systems to deploy and expectations around their performance.

Before moving to breakout groups, participants discussed the scope of the workshop and asked about workshop outcomes. The organizers provided clarification:

The decision to narrow the workshop's focus to real-time baleen whale monitoring during construction activities was made to ensure meaningful and focused outcomes. It was acknowledged that there is potential crossover in the relevance of technologies applied to other monitoring activities such as for vessel strike risk reduction. One workshop series outcome could be to provide guidance to industry for how to demonstrate the effectiveness of their activities and technologies, particularly during construction, but also more generally, with respect to metrics that wildlife biologists and decision-makers consider important.

Participants requested clarification from the organizers about the goals of the workshop discussion and the PNNL/NREL technical report. The central issue driving this request was the notion that once a cue has been made, experts felt that the performance of various technologies in detecting those cues are generally well-understood. However, the realistic scenario is one where we are unsure whether a cue has been made and we therefore cannot be sure if we have successfully detected it. One expert distilled the need as "identify the acceptable failed detection rate, for example, 5% or 10%" because they felt they could easily design a multi-modal detection system that could achieve that rate (given some understanding or prediction of whale presence and cue rates during a given time window).

Workshop organizers agreed that the central issue was one of assessing system(s) where the successful detection rate and missed detection rate are known and optimized to minimize the missed detection rate. An example provided included technology systems that could provide more context about whales' behavior versus simply confirming whale presence. Such systems could help decision-makers understand whether whales are traveling through the area of interest or likely staying for several days to forage.

Breakout Groups

Key Takeaways from All Breakout Sessions

• Several categories of technology can inform or contribute to the accuracy of whale detection data but not all are suitable for real-time monitoring during offshore wind pile driving. The development and evaluation of these other technologies should continue to be advanced and perhaps flagged as methods that can provide near real-time or slower verification of or context around real-time whale detections.
• The integration of multiple technologies ensures comprehensive monitoring of marine mammal presence during offshore wind construction activities.
• The functionality of technology in different environmental conditions needs to be measured to best evaluate technology effectiveness.
• Standard performance metrics, including methods and assumptions, across all technologies are needed. Potential examples to evaluate readiness and effectiveness include:
  ○ Timeliness of data availability
  ○ Missed detections, or false negative rate
  ○ Detection range and accuracy across levels of sea state, visibility, and other environmental factors, and across various deployment platform types
  ○ Level of human intervention needed to interpret results
  ○ Participants flagged price as a consideration, especially for technologies where multiple devices may be needed at a single offshore wind construction site, but also noted that price is not a traditional performance metric

The breakout sessions served as a platform for participants to discuss the metrics needed to evaluate the readiness and effectiveness of technology deployments before and during pile-driving. Attendees were encouraged to select a breakout group based on their specific interests, as discussions centered around distinct technology categories. Six breakout groups were formed, each focusing on a particular aspect of technology deployment:

- Passive Acoustic Monitoring - Mobile
- Passive Acoustic Monitoring - Stationary
- Thermal Imaging
- Visible Light Cameras
- Protected Species Observers
- Other Technologies (such as satellite imagery and eDNA sampling).

Each breakout session intended to address key questions (below) with representatives from CBI, RWSC, DOE, MTS, PNNL, and NREL providing facilitation and notetaking. The insights generated from these discussions are being used to inform the reference document developed by the PNNL and NREL, which will continue to evolve beyond the workshop. In groups where participants provided specific references to the literature, they were included below.

- What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?
- What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?
- What measurements should be used to characterize the performance of this technology?
- What information gaps limit our ability to use this technology for offshore wind construction monitoring and what are the priority R&D needs to fill those gaps?

Passive Acoustic Monitoring – Stationary Breakout Group

Facilitator: Julia Dombroski, Ph.D., National Offshore Wind Research & Development Consortium

Subcategories within this Technology Category

Breakout room participants identified the following subcategories for stationary PAM systems:

- Bottom-Mounted Real-Time Buoys
- Bottom-mounted Arrays
- Fiber Optic Cables

What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?

- Noise Levels: Noise levels during construction may constrain the effectiveness of stationary PAM systems. Breakout room participants suggested considering a large range of ambient noises in evaluations of this technology.

What measurements should be used to characterize the performance of this technology?

- Detection of Cues: Breakout room participants identified detection function and detection range as critical measures for stationary PAM systems. Understanding how missed detection rates are impacted by different variables such as range is important for the evaluation of stationary PAM systems.
- Speed of Data Availability: Breakout room participants flagged the speed of data availability as a key measure for technology evaluation. Participants noted that it is important to know when data will be available to decision-makers to ensure timeliness of mitigation efforts.

- Engineering Metrics: Measuring engineering metrics such as uptime, radio communication time, etc. was identified as critical for characterizing the performance of stationary PAM. This measure should assess the design and packages of systems to understand compatibility with different platforms.
- Other: Other key measures identified by participants include characterizing localization ability, missed alerts, and false detection rates.

What information gaps limit our ability to use this technology for offshore wind construction monitoring and what are the priority R&D needs to fill those gaps?

- Standardization of Validation Methods: Breakout room participants advocated for the development of standardized validation methods of stationary PAM systems, including how data are collected and benchmarks for practitioners. Standardization should focus on metrics such as methods, assumptions, source levels, and recognition potential. This standardization should include an evaluation of both the software (detection functions) and hardware (recording performance). This standardization can apply to mobile PAM systems as well. Participants suggested the development of a working group to establish this standardization. These validation discussions should happen before pile-driving begins and continue temporally throughout monitoring programs. Participants identified standardization of validation methods as an R&D funding priority.

Passive Acoustic Monitoring – Mobile Breakout Group

Facilitator: Patrick Field, Consensus Building Institute

Subcategories within this Technology Category

Breakout room participants identified the following overarching categories for mobile PAM systems: subsurface autonomous, surface autonomous, buoyancy driven, and drifters.

What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?

- Movement: The greatest advantage of mobile PAM systems is their ability to be moved. This allows operators to move systems to different sites and tune to conditions users need to explore in 3D. This also offers advantages in deployment as some mobile PAM systems can be deployed from an onshore location.
- Real-Time Vocalization: Data from mobile PAM systems can provide useful information on the positions, bearing, and range of whales.
- Long Recording Duration: Several mobile PAM systems can record data from weeks to months and monitor an area at different points in time.
- Flexibility: The diversity of mobile systems allows different tactics to be used in a survey area. This flexibility can fill gaps between stationary areas, provide more statistically robust

data, and provide more flexibility in design. This also allows for understanding the probability of detection and the probability of the equipment shutting down.
- Cost: Autonomous systems reduce costs as these systems are unmanned and require less equipment. Having fewer vessels for equipment also reduces vessel strikes.

What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?

- Deployment and Sea States: Although some mobile PAM systems have advantages in deployment, not every mobile PAM system can drive itself. This causes safety concerns for operators deploying equipment offshore. Furthermore, there are limitations as not all mobile PAM systems have the ability to operate in certain sea states. Stationary platforms can operate and detect better owing to their stability.
- Endurance: Participants noted that endurance of systems is a challenge and the need for continuous monitoring may create issues with biofouling.
- System Noise: System noises from locomotion, pumps, propulsion systems, etc. may impact timing of data collection.
- Depth Restrictions: Some mobile PAM systems have depth restrictions and cannot operate in shallow waters. This applies especially to gliders.
- Platform Compatibility: There may be constraints on compatibility with technologies that are being widely used.

What measurements should be used to characterize the performance of this technology?

- Environmental Efficacy: There is a need to ensure mobile PAM systems evaluated have undergone appropriate testing and are feasible in different environmental conditions. Understanding the tradeoff of different platforms is also important especially when thinking about coverage of systems and requiring integrated solutions.
- Environmental Impacts: It is important to measure the environmental impacts of active acoustic modems as they may cause compliance issues for developers.
- Cost: Participants flagged cost as an important measure for mobile PAM systems. Participants noted that more detailed data can be collected at a higher price point.
- Timing: Participants identified the need to understand the way and how frequently mobile PAM systems are sending information. The frequency at which information is transmitted can impact efficacy of mitigation.
- Deployment Capabilities: With the diversity of mobile PAM systems, it is important to consider the deployment capabilities of these platforms. Breakout room participants suggested that deployment capabilities such as where a device can be deployed, self-righting functionality, and operation in different sea states be considered.

- Defining Localization Systems: Breakout room participants suggested that systems that can localize be treated as a separate system. There will be cost differences between these systems and other mobile PAM systems that users need to consider. It is important to specify that not all mobile PAM platforms are capable of localization.

Breakout room participants suggested the hosts consult with companies to understand the flexibility and functionality of mobile PAM platforms.

Thermal Imaging Breakout Group

Facilitator: Genevra Harker-Klimes, Ph.D., U.S. Department of Energy

Subcategories within this Technology Category

Breakout room participants identified the following subcategories for thermal imaging: vessel handheld, vessel mounted, stationary/fixed offshore, stationary/fixed land-based, and drone, aircraft, autonomous surface vehicle.

What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?

- Day and Night Detection: Thermal imaging systems are able to detect animals during the day and night. This is beneficial as some developers have requested authorization to pile drive at night. Furthermore, thermal imaging devices can supplement PAM data if animals are not actively vocalizing during the day or at night.
- Flexibility: Thermal imaging systems are flexible in different environmental conditions. Higher quality systems offer advantages such as stabilization, rotating camera, and high resolution.

What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?

- Limited Detection Capabilities: Thermal imaging devices record surface-only detections and have limited ability to detect to the species level.
- Cost: Cost of thermal imaging devices is a prohibitive factor to developers and may limit deployments to one camera system per monitoring area.
- Battery power: Limited battery power on thermal imaging devices is a constrain on performance. This constrains data transfer, data storage, and detection power.
- Camera Resolution and Glare: A major limiting factor is camera resolution as it impacts detection ability. Glare on the cameras is also a concern.
- Stabilization: Lower quality thermal image devices may not be stable in certain sea states.

- Need for Staff: There is a need for human verification of data and/or to train an identification algorithm. This limits developer's use as cost and staffing needs increase. Furthermore, the data need to be transferred to humans in a timely manner.
- Range: Some thermal systems do not meet range criteria and have a narrow field of view. Sea states can further impair range detections. Participants suggested developing standards for range.

What measurements should be used to characterize the performance of this technology?

- Range: Range is an important aspect to characterize by system. Range in distance and field of view needs to be considered to evaluate the number of devices needed.
- Resolution: Image resolution needs to be measured to characterize the performance of thermal imaging devices. There is also a need to measure the stabilization, camera movement, and camera angle capabilities of thermal imaging devices.
- Functionality: Participants noted the need to understand the usability of thermal imaging devices. For example, can systems be used by protected species observers. Furthermore, there is a need to measure if thermal imaging devices can be connected to third party devices such as auto-detection systems.

What information gaps limit our ability to use this technology for offshore wind construction monitoring and what are the priority R&D needs to fill those gaps?

- Automation and Machine Learning: Information gaps on automation and machine learning need to be prioritized. This includes developing algorithms for false positives, studies of stabilization, assessing 2D vs. 3D imaging, and rotating vs. forward-looking cameras.
- Standardization: Participants suggested the development of standards for thermal imaging for regulators. These standards should include probability detection curves, curves for different environmental conditions, and validation standards. This will allow for greater investments in systems regulators will use.
- Device Testing and Compatibility: Breakout room participants identified the need to test thermal imaging devices in actual pile-driving situations as noise may impact animal behavior. Furthermore, these devices need to be tested to confirm compatibility with protected species observer work and autonomous platforms.

Visible Light Camera Breakout Group

Facilitator: Caisey Hoffman, Marine Technology Society

Subcategories within this Technology Category

Breakout room participants identified the following subcategories for visible light camera: vessel-mounted, fixed/stationary platform offshore, fixed/stationary platform land-based, aerial (drone, aircraft), and autonomous surface vehicle (ASV). They also stated the need to include subcategories based on light wavelength, specifically differentiating between visual light and night-vision imagery.

What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?

- Automated Identification: Visual imagery can be processed through models for automated identification, leveraging AI to detect whales promptly.
- Historical Data: Visual imagery provides a record, allowing for the maintenance of historical data for future re-analysis.
- Presence: While visual light cameras don't operate in the dark, they can still detect large objects like whales at night if bioluminescence is present. Any image that can detect a signature of whale movement is valuable.
- Targeted Search Area: Pile-driving delineates a definite and limited area of search, which may be covered easily by aerial platforms equipped with visual light cameras.

What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?

- Processing Images: Data processing of images is time consuming and presents challenges for use in real-time decision-making. Humans are needed to verify AI detections. Although a lab has shown real-time identification for beluga whales, this capability isn't available commercially.
- Night-time: Limitations exist in capturing images in low-light conditions, but night-vision technology could be a solution.
- Environmental Factors and Platforms: Functionality is impacted by environmental factors such as lighting, sea state, weather, cloud coverage, ice, and glare. Glare has been documented to affect false detections. The recommended weather conditions and sea state for conducting surveys depends on the type of platform being used (i.e. drones are unsuitable during poor weather, with fixed platforms offering more stability).
- Limited Spatial Coverage: Visual imagery covers only a small area. It is noted that the limited search area for pile-driving may mitigate concerns.
- Trust: Lack of transparency in the development of technology affects trust in AI systems. Human involvement in the development stage is essential, with scientists required to evaluate models. This process can be time-consuming.
- Individual Identification: Individual identification of animals is not easily achievable. It is noted that the primary objective is to identify any relevant species in the potential shutdown area, making identification of specific individuals less necessary.

What measurements should be used to characterize the performance of this technology?

- Speed of Decision-Making: Evaluate the technology's ability to facilitate accurate decisions quickly.
- Co-variates: Consider site conditions necessary for performance characterization, accounting for factors like inter-observer variability and false positives.
- Utilize Published Metrics: Rely on published metrics for evaluating technology, prioritizing peer-reviewed literature over internal company reports.
- Literature Availability: "Scaling whale monitoring using deep learning: A human in the loop solution for analyzing aerial data sets" Boulent, et al. 2023.

What information gaps limit our ability to use this technology for offshore wind construction monitoring and what are the priority R&D needs to fill those gaps?

- Detectability Co-variates: Conduct studies to understand how detectability varies under different weather conditions and for different species.
- Speed of Decision-Making: Evaluate whether real-time decision-making is necessary and feasible with this technology, particularly during the construction phase where immediate actions are crucial. Develop technology capable of providing timely decisions to enhance monitoring effectiveness.
- Guidance: Standardize data collection methods and protocols to streamline monitoring efforts across organizations. Establish guidelines for the use of drones, cameras, and transect procedures to ensure consistency and reliability in data collection.
- Satellites: Explore potential of low earth orbit satellites equipped with advanced infrared components for daily coverage of turbine construction areas. Assess the feasibility and cost-effectiveness of using satellites as a primary detection system or with other technologies for real-time or next-day decision-making.
- Models: Explore the modeling of probabilistic projections of whale distribution (i.e. Fathom Science is working on this).
- Laws Regarding Drones: Investigate regulations concerning the operation of drones, including requirements for visual range of the operator and coverage limitations within the area of pile driving.

Protected Species Observers Breakout Group

Facilitator: Abbey Wakely, Marine Technology Society

Subcategories within this Technology Category

Breakout room participants identified three subcategories for Protected Species Observers (PSOs) systems and outlined their respective pros and cons for monitoring for baleen whales:

- Vessel-based:
    - Pros: Flexibility to adjust proximity, can move closer for clarification.

- Cons: Limited by vessel type and space, potential bias due to social pressure from vessel operator(s).
- Fixed/Stationary:
  - Pros: Individual observer may be personally less affected by weather conditions, potentially less social pressure.
  - Cons: Limited perspective compared to moving platforms, constraints on mobility.
- Aerial:
  - Pros: Greater field of view, deeper visibility into water, platform mobility.
  - Cons: Safety concerns, higher cost, limited airtime, weather dependency, limited time in monitoring zone, additional training requirements.

What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?

- Reliable Visual Cue: Capable of detecting baleen whale respiration, which serves as a consistent visual cue as animals always have to breathe, unlike vocalizations. There also might not always be a sufficient thermal difference between the air and the water for detection on a thermal camera.
- Environmental Conditions: Night-time operation is possible, but only with technology assistance (like night vision).
- Human Decision-making Ability: PSOs possess the capability to analyze multiple data streams in real-time and incorporate external information to make informed decisions, a capacity not fully replicated by automated technologies.
- Behavioral Data Collection: PSOs record behavioral characteristics of baleen whales, providing valuable insights.
- Real-time Communication: Facilitates real-time communication of sightings to stakeholders.
- Ancillary Data Collection: Allows for the collection of supplementary information such as species identification, behavioral observations, group size, and mother-calf relationships.

What are the attributes that constrain or potentially limit this technology for the purpose of decision-making around piledriving?

- Variability Due to Human Factors: The level of training, expertise, and individual ability of observers may lead to inconsistent results that are challenging to standardize or quantify, unlike equipment calibration.
- Observer Fatigue: Prolonged shifts and operational fatigue, particularly during periods of low baleen whale density, can diminish detection capabilities.
- Spatial Limitations: The spatial detection range is limited by environmental conditions, potentially impacting its effectiveness in certain conditions.
- Environmental Conditions: Environmental conditions such as light, sea state, visibility, and cloud cover can significantly impact detection capabilities.

- Human Bias: Human observers may introduce bias, especially during periods of low whale density, potentially influencing decision-making processes.
- Behavioral Complexity: The diverse behavioral characteristics of baleen whales pose challenges for accurate monitoring and interpretation.

What measurements should be used to characterize the performance of this technology?

- Observer Calibration: Attendees inquired if it was feasible to standardize observer performance similar to equipment calibration and if there are avenues to provide experience outside of field work.
- Environmental Logging and Log Evaluation: PSOs possess the capability to record observational and environmental data in a manner not achievable through technology. Evaluation of PSO logs allows for real-time and continuous assessment of attentiveness, objectivity, and effectiveness.
- Simulators: Simulators offer a means to evaluate performance. Employing multiple observers can aid in identifying perception and search image biases.

What information gaps limit our ability to use this technology for offshore wind construction monitoring and what are the priority R&D needs to fill those gaps?

- Literature Available: Protocols and evaluations of PSOs have been developed by the National Marine Fisheries Service (NMFS) and are continuously refined.
- Multi-Modal Assessment: There is a need to assess the effectiveness of PSOs with other monitoring technologies, emphasizing a multi-modal approach.
- Enhanced Data Management Platforms: PSOs use digital logging, but improvements are needed for real-time data sharing with PAM operators. Implementing QAQC processes and standardized sharing protocols, similar to the Whale Alert system, is needed for other species.
- Standardization: Standardization across data collection, sharing, and analysis processes is important to ensure consistency and facilitate collaboration among PSOs.
- Accessibility of PSO Data: PSO data can address other research questions and the data should be available through an accessible portal.

Other Comments

The group acknowledged it is more effective to use visual PSOs in combination with technology, including PAM operators. Evaluating PSOs alone is challenging, highlighting the need for complementary methods. The discussion raised questions about how to validate or calibrate PSOs in comparison to validating technology.

Other Technologies Breakout Group

Facilitator: Emily Shumchenia, Ph.D. RWSC

Subcategories within this Technology Category

Breakout room participants identified the following subcategories for other technologies: satellite imagery, environmental DNA (eDNA), Dimethyl Sulfide (DMS) Concentration, telemetry, and active acoustics.

What are the attributes of this technology that make it potentially useful for the purpose of decision-making around piledriving?

- Validation of Focal Technologies: Tools such as tagging offer key data on cue rates, aiding the analysis of passive acoustic data.
- Importance of Complementary Tools: Recognizing that passive acoustics alone may not suffice, a combination of tools like eDNA and DMS can provide more comprehensive insights.
- Role of Telemetry and Satellite Tagging: While valuable for data collection with respect to whale behavior, presence, and movement, these methods may not directly address real-time detection/mitigation challenges.

The Breakout room participants engaged in a more freeform discussion on the below topics:

- Comparing Technology Performance to Decision Time: Stated the need of aligning technology capabilities with construction team decision times for shutdowns and restarts.
- eDNA Potential and Challenges: Discussed the potential of eDNA technology, citing the need for further validation and integration with other methods to enhance effectiveness. The short window (15-30 minutes) for positive detection and the technology's ability to quickly provide definitive answers using platforms like gliders were emphasized. Below is literature to reference for eDNA:
    - Alter SE, King CD, Chou E, Chin SC, Rekdahl M and Rosenbaum HC (2022) Using Environmental DNA to Detect Whales and Dolphins in the New York Bight. Front. Conserv. Sci. 3:820377. doi: 10.3389/fcosc.2022.820377
- Active Acoustics: Raised questions about the practicality and effectiveness of active acoustics for marine mammal detection and mitigation, including considerations around addition of sound to the environment. References provided included an overview of ADCPs (https://www.nortekgroup.com/knowledge-center/wiki/guide-to-understanding-adcps) and an example of echosounder technology for plankton detection (https://ieeexplore.ieee.org/document/9706094).
- Seasonal Considerations: Considered the seasonal applicability of different technologies and the need for a holistic approach using a combination of sensors.
- Tagging: Tagging projects monitor marine mammal movements and behaviors. Challenges persist in effectively tagging animals, particularly North Atlantic right whales. Integrating tagging technology with active acoustics provides insights into prey interactions. It was noted that data sharing (crowd sourcing and sharing among private companies) has

potential to improve marine mammal detection efforts, like the benefits seen with bathymetry data.

- Satellite Tagging: Participants supported the use of satellite tagging to create a detailed monitoring tool akin to a "Marauder's Map" for tracking marine mammal movements.
- Detection Diversity: Participants highlighted the distinction between technologies for general animal presence detection and those for localized detection near specific structures. They emphasized the value of both detection types in enhancing understanding of marine mammal behavior and distribution.
- Data Delivery Timing: Discussions focused on the timing considerations between real-time and autonomous data delivery methods. Questions were raised about the timeframe encompassed by "real time" and the inevitable time lag between data collection and actionable insights. Participants acknowledged the increased cost associated with scaling up an instrument for quick data delivery.
- Advocacy for Multi-Layered Approach: Participants supported a "Swiss cheese" model involving multiple layers of technology to protect whales during construction activities. They stressed the importance of integrating validated methods alongside new approaches for comprehensive whale protection measures.

Final Plenary

Following the breakout room discussions, participants were brought together to discuss feedback on the Technology Matrix and Technical Report Containing Technology Profiles. This freeform discussion focused on three main themes:

Important Metrics for Technology Evaluation

Participants identified the timeliness of data availability as one of the most important metrics needed to evaluate the readiness and effectiveness of technology deployments before and during pile driving. Decision-makers need to know when data from these systems will be available to make mitigation decisions.

Understanding the deployment of technologies was identified as another important metric for evaluating technology. Participants noted the need to understand how long deployment will take, the types of platforms on which and where technologies can be deployed, and what is being deployed.

Detection accuracy was also flagged as an important metric in evaluating technology deployments. Participants are interested in understanding how different environmental conditions such as sea state and extreme weather affect detection accuracy.

Knowledge Gaps

Participants emphasized the need to understand how data will be used by decision-makers to allow researchers to know how best to collect data. This reinforced the need to measure the timeliness of data availability to make mitigation decisions. Participants also noted the need for standardization of methods and data when implementing these technologies. Participants suggested the development of a working group to establish this standardization.

Missing Technologies in Analysis

Some participants expected to hear discussion and evaluation of radar and night vision devices. Participants suggested that these technology types be included in PNNL and NREL's analysis.

Feedback on workshop series and technology report development process

Finally, participants asked workshop organizers to provide materials in advance of the next workshop session with sufficient time to review (i.e., at least one week). There were also requests to share the list of subject matter experts and/or peer reviewers that PNNL and NREL would be engaging to review their technical report. Participants also suggested that including the subject matter experts more seamlessly into workshop discussions may offer some advantages. Some felt that breakout group discussions would have advanced further or been more sophisticated if they were informed by experts in each technology category.

Summary of Post-Workshop Survey Results

A post-workshop survey was shared with participants to capture further discussion and feedback. Results from the survey can be found in Appendix F. These results were exported one week after the workshop date on April 26, 2024.

Survey respondents echoed key themes from the breakout groups but also added new ideas for future workshops to consider:

- Technologies related to data management and synthesis/analysis (e.g., machine learning, artificial intelligence) and challenges related to big data should be discussed.
- Include more time for breakout group / small group discussion.
- There is a need for a very basic introduction to each technology category for non-experts. If this exists, future workshops should provide the resource ahead of time.

Appendices

- Appendix A: Workshop Attendees
- Appendix B: Technology Matrix
- Appendix C: Technical Report Containing Technology Profiles
- Appendix D: Presenter Slides
- Appendix E: Breakout Worksheets
- Appendix F: Post Workshop Survey Results