

PNNL-37249
NREL/TP-5000-92549

Evaluating Tools and Technologies for Monitoring Baleen Whales During Offshore Wind Foundation Installation

January 2025

- 1 Angela R. Szesciorka
- 2 Mark Severy
- 3 Kristen Ampela
- 4 Cris Hein
- 5 Michael Richlen
- 6 Joseph Haxel
- 7 Jeff Clerc

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 Battelle Memorial Institute, 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, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062

www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: <http://www.ntis.gov>

Evaluating Tools and Technologies for Monitoring Baleen Whales During Offshore Wind Foundation Installation

January 2025

1 Angela R. Szesciorka¹

2 Mark Severy²

3 Kristen Ampela³

4 Cris Hein³

5 Michael Richlen²

6 Joseph Haxel²

7 Jeff Clerc³

Prepared for
the U.S. Department of Energy
under Contracts DE-AC05-76RL01830 and DE-AC36-08GO28308

¹ California Ocean Alliance, 9099 Soquel Drive, Suite 8, Aptos, CA 95003

² Pacific Northwest National Laboratory, 902 Battelle Blvd., Richland, WA 99354

³ National Renewable Energy Laboratory, 15013 Denver W Pkwy, Golden, CO 80401

Executive Summary

Offshore wind energy facilities are currently being planned and installed in United States (U.S.) waters in accordance with regulations and processes designed to avoid negative effects on federally-protected baleen whales and other marine mammals. The Bureau of Ocean Energy Management (BOEM) oversees planning, leasing, construction, operations, and eventual decommissioning of offshore wind energy projects in federal waters. In this role, BOEM consults and coordinates with other state and federal agencies, including with National Oceanic and Atmospheric Administration (NOAA) Fisheries, to meet requirements of the Endangered Species Act. Offshore wind developers consult with NOAA Fisheries directly to obtain authorizations under the Marine Mammal Protection Act. To minimize potential impacts of offshore wind energy activities—including site characterization surveys, vessel traffic, and pile driving during construction—on baleen whales and other marine mammals, BOEM and NOAA Fisheries require project proponents to submit detailed marine mammal monitoring and mitigation plans, under all visibility conditions, prior to project start. If pile driving activities are proposed in low- and no-visibility conditions (e.g., darkness, rain, fog), the project proponent must submit an Alternative Monitoring Plan (AMP) that describes in detail the proposed monitoring protocols and equipment to be used in these conditions. At the time of writing, no AMPs for initiating pile installation at night have been approved for an offshore wind project in the U.S. The ability to install offshore wind foundations in low visibility and in darkness is of interest to offshore wind energy developers and other stakeholders because of increased project schedule flexibility and compression of construction time windows, thereby reducing the overall duration of in-water foundation installation activities, which may reduce potential impacts to marine life. Several available monitoring technologies have the potential to detect whales in low- and no-light conditions, but no standardized approach currently exists to characterize the functional performance of these technologies.

Impacts to baleen whales are of particular concern in the context of offshore wind development because of their conservation status, the overlap of their habitat with current and planned offshore wind energy facilities, and their expected sensitivity to low-frequency sounds, such as those produced during pile driving. Therefore, this report focuses on monitoring technologies capable of detecting this species group. The scope of this report is to assess the current state of technologies and methods for monitoring baleen whales in low- and no-light conditions before and during offshore wind turbine foundation installation.¹ This work considers the performance of various sensing technologies that can be used to observe the presence of baleen whales near offshore construction activities. The scope of work includes:

1. **Technology Performance Metrics** – Identify metrics that can be used to assess technology performance.
2. **Technology Evaluation** – Synthesize the past performance of technologies with respect to the performance metrics based on available literature.
3. **Technology Characterization Framework** – Describe a potential framework to characterize technology performance with a standardized approach.
4. **Technology Research and Development Needs** – Identify future needs and opportunities for research, development, and deployment to improve technologies.

To accomplish this scope, staff from the Pacific Northwest National Laboratory (PNNL) and the National Renewable Energy Laboratory (NREL) engaged with relevant groups via scoping and

¹ The full breadth of offshore wind facility construction and operations were not the focus of this effort.

steering committee meetings, participated in public workshops organized by the Regional Wildlife Science Collaborative (RWSC), and recruited subject matter experts for interviews and peer review. Based on this feedback, the PNNL and NREL team developed technology-specific documents and standardized approaches to help guide offshore wind energy project developers and technology designers as they navigate technological challenges. This report presents the outputs from our analysis, including documentation about key performance metrics, a synthesis of the state of monitoring technology performance, a potential framework to standardize technology characterization, and recommendations for research, development, and deployment.

The scope of this effort is specific to baleen whale monitoring during foundation installation in low- and no-light conditions. However, the monitoring technologies we describe, as well as our proposed characterization framework, could be applicable to other marine mammal species and aspects of offshore wind development, with the caveat that other marine mammal species have behaviors (e.g., surfacing and calling) that differ from baleen whales and may require a modified approach.

Technology Performance Metrics

Baleen whale monitoring technologies are designed to determine the presence of these whales in a particular area. The outputs from different technologies carry different levels of specificity, including detection of a baleen whale, species classification, and localization of that detection to a particular area. For a monitoring technology to be useful in the context of offshore wind foundation installation, it must be able to accurately detect baleen whales with few missed detections, detect whales over a relevant monitoring distance, and transfer that information reliably to a human operator within a pre-defined decision-making time window.

Evaluating the performance of a sensor system requires the application of metrics that quantify the ability of the system to perform its intended task with a consistent and standardized approach. Technology performance metrics should quantify the basic functionality of a sensor system in terms of accuracy, range, latency, and reliability with the following key measurements:

- **Efficacy** – ability of a sensor to correctly and repeatably identify the presence or non-presence of a baleen whale
 - **Precision** – fraction of detections or classifications that were correct, i.e., number of true positives divided by the total of true and false positives.
 - **Recall** – fraction of whale occurrences that were detected, i.e., number of true positives divided by the total number of whale occurrences (true positives plus false negatives).
 - **Probability of Missed Mitigation** – likelihood that a baleen whale entered a mitigation zone² without being detected.
- **Range** – physical distance at which a sensor can observe an animal cue
 - **Reliable Detection Range** – distance at which a whale occurrence can reliably be detected.
- **Data Delivery** – ability of system to reliably deliver information to a human operator
 - **Detection Time Latency** – average time duration between a whale cue and information about the detection being received by a human operator.
 - **System Reliability** – fraction of time that the system is operational during the planned operational time window.

² A mitigation zone is defined as an area in which the presence of a baleen whale may trigger an operational decision, such as a pause in pile driving. Clearance zones and shutdown zones are examples of mitigation zones.

These key performance metrics depend on the system design, deployment location, species, and environmental factors during operation. Importantly, evaluation of each metric should demonstrate how the performance varies relative to environmental and biological conditions.

Technology Evaluation

Various technologies and tools for monitoring baleen whales have been developed, each with different approaches and characteristics. Eight of these technologies were identified that could be relevant for monitoring baleen whales around offshore wind foundation installations:

- **Satellite imagery** – identifies whales at or near the water surface in satellite imagery
- **Optical cameras** – identifies whale blows or bodies at or near the water surface using optical imagery sensor systems
- **Infrared imagery** – identifies thermal signature of whale blows or bodies at or near the water surface using infrared sensor systems
- **Telemetry** – tracks whale movement using animal-borne tags
- **Dimethyl sulfide (DMS) concentration** – concentration of DMS in marine surface waters that may be associated with whale occurrence because DMS concentration is correlated to areas of high productivity or zooplankton concentration where whales may be feeding
- **Environmental deoxyribonucleic acid (eDNA)** – identifies genetic material left behind by whales in seawater
- **Active acoustics** – uses reflections from emitted underwater sound sources (e.g., sonar) to detect whales in the water column
- **Passive acoustic monitoring (PAM)** – records underwater sound to identify whale calls

Each technology was reviewed in detail to summarize its past performance in the context of baleen whale monitoring. Documented results vary between studies and technologies, highlighting the difficulties inherent in characterizing the performance of these technologies and the lack of consistent and standardized evaluation methods among studies. All but one (DMS concentration) monitoring method demonstrated successful detections of baleen whales. Results clearly show that technology performance varies significantly with system design, local conditions, and automated detection and classification methods.

Different whale monitoring technologies are designed to detect different animal cues. Visual monitoring approaches involving Protected Species Observers (PSOs) are designed to detect visual animal cues, such as whale surfacing behavior and blows. PAM approaches are, by contrast, designed to detect sounds made by calling animals, regardless of surfacing behavior. There have been several attempts to evaluate the effectiveness of various whale monitoring technologies by comparing detection probability in controlled experimental situations, where the number and location of whales is known. For example, the ability of PAM, aerial surveys, and infrared imaging to concurrently detect baleen whales can range from 10% to 100%. This range can be attributed to the availability of relevant animal cues, and the detectability of these cues on different monitoring systems. This range of results also highlights an opportunity to use complementary monitoring technologies in concert to maximize the probability of whale detection using a variety of animal cues.

The capabilities of each technology were assessed relative to the application of monitoring baleen whales in near real time around offshore wind foundation installation. Using four criteria specific to this use case, PAM and infrared imaging emerged as potentially suitable primary technologies. Both of these monitoring technologies have been employed during offshore wind

foundation installation in the U.S. (Macrander et al. 2022; South Fork Wind, 2024). No single system has been shown to reliably detect all baleen whales in a particular area, but technology efficacy can be improved using multi-modal systems. Multi-modal systems combine multiple technologies to improve detection rates.

Further testing of monitoring technologies in relevant environments would help better characterize expected system performance. Across the suite of technologies, key performance metrics are not reported consistently in terms of precision and recall, and metrics about detection distance, detection time latency, and system reliability are often not reported. In addition, few studies have reported the overall performance of multi-modal systems.

Technology Testing and Research and Development Needs

Given the difficulties in assessing technology performance and the lack of consistent metrics and methods, we developed an example framework for technology characterization. The potential framework is designed to characterize the performance of a baleen whale monitoring system with a specific monitoring objective in relevant marine conditions considering location and species. The framework encompasses three phases: (1) System Design: describe the monitoring use case and system specifications, (2) Pre-Test Assessment: evaluate system specifications and develop a field characterization plan, and (3) Field Characterization and Modeling: evaluate the application of the technology through field observations and subsequent system modeling (Figure ES.1). Through this framework, a technology provider can characterize the performance of their sensor using relevant field observations, then model the performance of a system that combines multiple sensors or technologies across a wider geographic area.

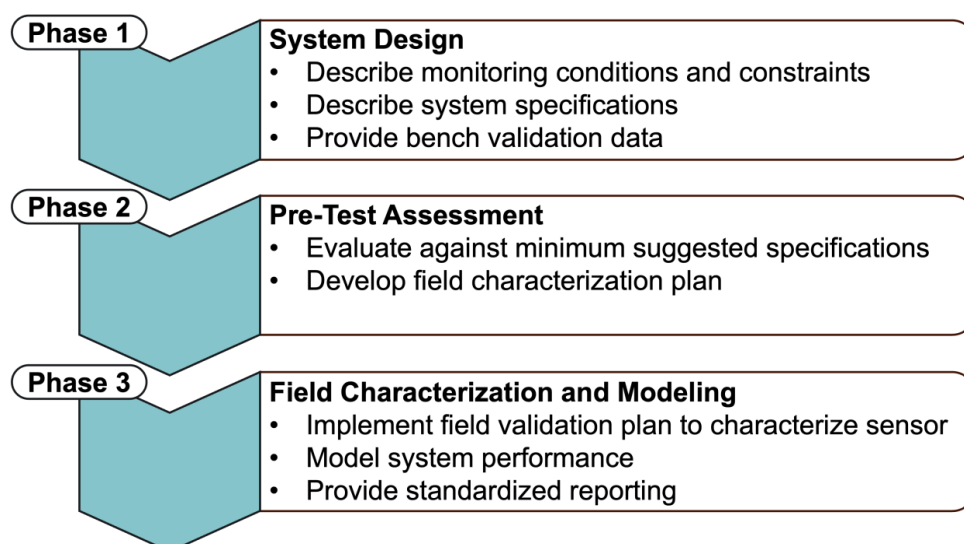


Figure ES.1 Summary of three phases of potential framework to characterize system performance

In order to advance the state of technologies and tools for monitoring baleen whales around offshore wind development activities, we recommend the following actions.

Research Recommendations

- Implement a standardized approach for technology characterization, such as the potential framework described above.
- Develop an instrumented field testing site for technology characterization.

- Quantify standardized metrics that are applicable to operational settings.
- Develop a standard system modeling tool with input data libraries.
- Increase focus on performance of multi-modal systems.

Development Recommendations

- Improve technology performance and automation through advancement in automated detection and classification systems, and increasing power and onboard data processing capabilities (including ability to store and send snippets of raw data with detections).
- Expand capabilities and use of real-time ready systems.
- Develop standardized and publicly available software and species-specific call libraries.

Deployment Recommendations

- Invest in emerging platforms and technologies for use in conjunction with the installation of offshore wind turbine foundations and future monitoring needs, including the use of multi-modal systems in operational settings.
- Reduce costs and increase efficiency of deployment and recovery of monitoring device systems.
- Test new technologies or multi-modal systems opportunistically alongside operational monitoring campaigns.
- Apply lessons learned from other industries, such as oil and gas and defense.

Although the scope of this report is specific to baleen whale monitoring technologies for use before and during fixed-foundation installation, the monitoring technologies we describe, as well as our proposed evaluation framework, could be applicable to other marine mammal species and aspects of offshore wind development, with the caveat that other marine mammal species have behaviors (e.g., surfacing and calling) that differ from baleen whales and may require a modified approach.

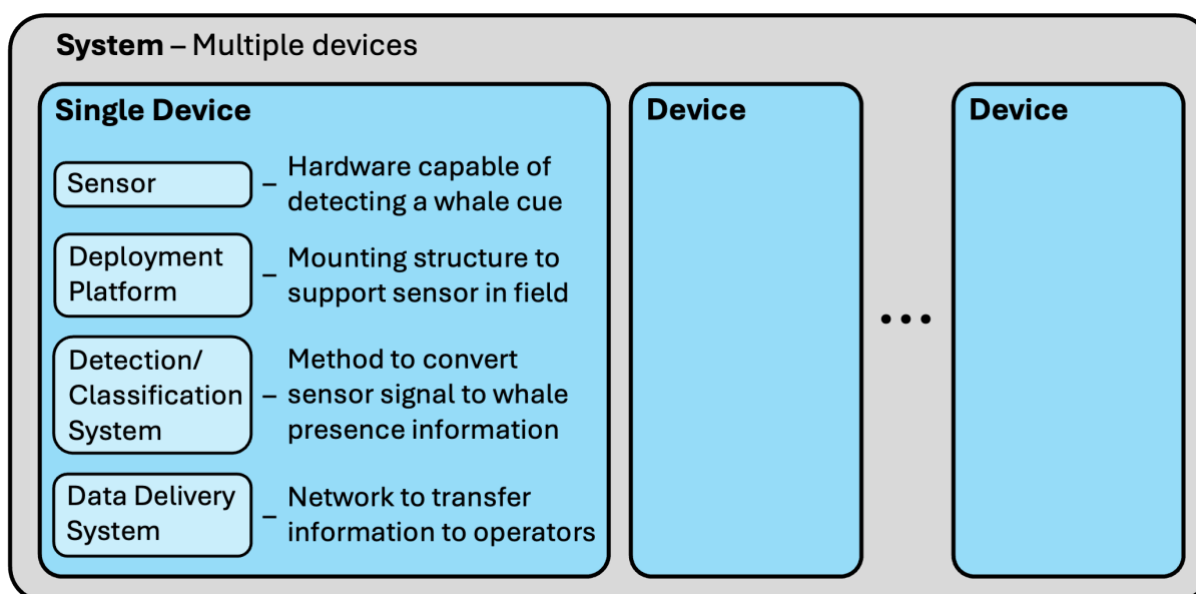
Acknowledgments

The authors thank the many people who provided their guidance and insight to define the scope of this report and provide feedback throughout the development of the draft report. In particular, we thank individuals from the National Oceanic and Atmospheric Administration, the Regional Wildlife Science Collaborative, the U.S. Department of the Interior Bureau of Ocean Energy Management, and the U.S. Department of Energy Wind Energy Technologies Office. We appreciate the one-on-one conversations with subject matter experts, including those from the wind industry, technology developers, non-governmental organizations, academia, and research institutions. Finally, we thank the workshop participants who provided useful input to help improve the report.

Funding for this report came from the U.S. Department of Energy Wind Energy Technologies Office.

Glossary of Terms and Abbreviations

Terms used to describe the various components of a whale monitoring system (i.e., instrumentation used to identify a whale's presence under a set of specific conditions) are defined below. A whale monitoring system can include a single device or multiple devices. Each device is made up of a sensor, deployment platform, detection/classification system, and data delivery system, as described in the figure below. Multiple devices can be used together to form a monitoring system. The devices in a system can be the same type of sensor or may combine different types of sensors (i.e., multi-modal system). This report categorizes technologies by different sensors or sensing techniques (e.g., passive acoustic monitoring or infrared imaging), and provides information about platforms, detection/classification systems, and data delivery systems, as appropriate to describe technology performance.



Acronyms used in the report are defined below.

Short Form	Long Form
AMP	Alternative Monitoring Plan
BOEM	Bureau of Ocean Energy Management
CO ₂	carbon dioxide
DIFAR	Directional Frequency Analysis and Ranging
DMS	dimethyl sulfide
DOE	U.S. Department of Energy
eDNA	environmental deoxyribonucleic acid
ESA	Endangered Species Act
IR	Infrared
IRC	Infrared Camera
ITA	Incidental Take Authorizations

Short Form	Long Form
ITDP	In-time Detection Probability
ITS	Incidental Take Statements
MMPA	Marine Mammal Protection Act
MMO	marine mammal observer
ML	Machine learning
MZ	Mitigation Zone
NARW	North Atlantic right whale
NOAA	National Oceanic and Atmospheric Administration
NETD	Noise equivalent temperature difference
NMFS	National Marine Fisheries Service, also NOAA Fisheries
NREL	National Renewable Energy Laboratory
OSW	offshore wind
PAM	passive acoustic monitoring
PAMO	passive acoustic monitoring system operator
PMM	Probability of missed mitigation
PSO	Protected Species Observers
RDR	Reliable Detection Range
RWSC	Regional Wildlife Science Collaborative for Offshore Wind
SL	Source levels
SME	subject matter experts
TDOA	time difference of arrival
TL	Transmission loss
TRL	Technological Readiness Levels
μPa	Micropascal
VHR	Very high-resolution

Contents

Executive Summary	ii
Acknowledgments	vii
Glossary of Terms and Abbreviations	viii
1.0 Background	1
1.1 Baleen Whale Monitoring Before and During Foundation Installation	1
1.2 Scope of this Report	3
1.3 Approach	3
2.0 Baleen Whale Monitoring System Performance Metrics	5
2.1 Efficacy	5
2.2 Range	6
2.3 Data Delivery	7
2.4 Variation in Performance Relative to Conditions	8
3.0 Technology Evaluation and Capabilities	10
3.1 Infrared Imaging	13
3.2 Passive Acoustics	16
3.3 Multi-modal Monitoring	20
3.4 Technology Readiness and Limitations	22
4.0 Potential Framework to Characterize Technology Performance	24
4.1 Phase 1: System Design	26
4.2 Phase 2: Pre-Test Assessment	27
4.3 Phase 3: Modeling and Field Characterization	29
5.0 Research and Development Recommendations	33
6.0 Conclusions and Next Steps	36
7.0 References	38
Appendix A – Technology Summary	A.1
Appendix B – Technology Performance Summary	B.1
Appendix C – Technology Profiles	C.1
Appendix D – Considerations Related to Design Specifications	D.1
Appendix E – Example Field Characterization Test Matrix Template	E.1
Appendix F – Example Field Characterization Recommendations	F.1

Figures

Figure 1. Confusion matrix	6
Figure 2. Reliable detection range. Assuming equal distribution of whales (left), predicated detection would increase at further distances from the sensor. The reliable detection range (right) is the distance where the number of detections start to decrease.	7

Figure 3. Factors influencing system performance	9
Figure 4. Summary of three phases of potential framework to characterize system performance.....	24
Figure 5. Detailed framework that can be used to characterize the performance of any chosen technology system. Colors indicate technology provider responsibility (blue) and templates that should be provided as a guide to the technology provided (orange).....	25
Figure 6. Flow diagram for Phase 1: System Design.....	26
Figure 7. Flow diagram for Phase 2: Pre-Test Assessment	28
Figure 8. Flow diagram for Phase 3: Field Characterization and Modeling.....	29
Figure 9. Example diagram of Field Characterization and System Modeling.	31

Tables

Table 1. Summary of RWSC workshops.....	4
Table 2. Evaluation of technology types based on three criteria using technology descriptions from Appendix A.	12
Table 3. Representative literature review of infrared imagery	14
Table 4. Representative literature review of PAM.....	17
Table 5. Technology capabilities summary table	A.1
Table 6. Representative literature review of satellite imagery	B.2
Table 7. Representative literature review of optical cameras	B.4
Table 8. Representative literature review of baleen whale tagging studies	B.7
Table 9. Representative literature review of DMS concentration.	B.8
Table 10. Representative literature review of eDNA.....	B.9
Table 11. Representative literature review of active acoustic monitoring	B.11
Table 12. Example checklist of design specifications for a passive acoustic monitoring system.....	D.1
Table 13. Example checklist of design specifications for an infrared camera system.	D.2
Table 14. Example field characterization test matrix for passive acoustic monitoring	E.1
Table 15. Example field characterization test matrix for infrared imaging	E.2
Table 16. Example summary of performance metrics calculated from field characterization testing.....	F.1

1.0 Background

Industrial activities in the ocean pose risks to marine mammals, including vessel strikes from maritime traffic and potential behavioral changes and/or hearing loss resulting from anthropogenic noise (Gulland et al. 2022). Impacts to baleen whales are of particular concern in the context of offshore wind development because of their conservation status, the overlap of their habitat with current and planned offshore wind energy facilities, and their expected sensitivity to low-frequency sounds, such as those produced during pile driving. For these reasons, baleen whales have been prioritized for focused monitoring and research during foundation installation and other offshore wind development activities (Southall et al. 2021). Installation of fixed-bottom offshore wind turbine foundations produces noise with the potential to impact these animals through hearing impairment, site avoidance, behavioral changes, and stress responses if an animal is near the sound source.

In the U.S., all marine mammals, including baleen whales, are protected under the Marine Mammal Protection Act (MMPA). Marine mammals that are listed as threatened or endangered are also protected under the Endangered Species Act (ESA). Protecting marine mammals around noise-producing construction activities, such as impact pile driving for offshore wind turbine foundations, requires a suite of monitoring and mitigation measures, including the establishment of project-specific clearance and shutdown zones. Visual monitoring of these zones is conducted by Protected Species Observers (PSOs), who are professionals trained to visually detect marine mammals and other threatened or endangered species, using their unaided eye and/or supporting tools such as binoculars. However, the ability of PSOs to detect these animals is limited by environmental conditions that affect visibility of the ocean surface such as ambient light, fog, precipitation, and waves.

The ability to install offshore wind foundations in low visibility and in darkness is of interest to offshore wind energy developers and other stakeholders because of increased project schedule flexibility and compression of construction time windows, thereby reducing the overall duration of in-water foundation installation activities, which may reduce potential impacts to marine life. Advancement and application of automated technologies to detect marine mammals creates an opportunity to supplement the ability of PSOs to detect marine mammals with more certainty in a variety of viewing conditions. Several available monitoring technologies have the potential to detect whales in low- and no-light conditions, but no standardized approach currently exists to characterize the functional performance of these technologies.

The scope of this effort is specific to baleen whale monitoring technologies for use in conjunction with installation of fixed-bottom offshore wind turbine foundations. Nevertheless, we acknowledge that the monitoring technologies described here, as well as our proposed evaluation framework, could be applicable to other marine mammal species and aspects of offshore wind development, with the caveat that other marine mammal species have behaviors (e.g., surfacing and calling) that differ from baleen whales and may require a modified approach.

1.1 Baleen Whale Monitoring Before and During Foundation Installation

In the U.S., all marine mammals, including baleen whales, are protected under the MMPA. Several baleen whale species, such as the blue, fin, sei, Rice's, and North Atlantic right whales (NARW), are listed as endangered and protected under the ESA. BOEM is the lead regulatory

agency responsible for offshore wind development in U.S. federal waters. BOEM consults with NOAA Fisheries to meet ESA requirements, while the offshore wind project developer seeks an authorization under the MMPA. Consultations and authorizations define clearance and shutdown zones for baleen whales prior to, and during, pile driving. The size of these zones vary depending on the known hearing frequency of the species group. Baleen whale are considered low-frequency marine mammals, so an observation of a baleen whale in the low-frequency clearance and/or shutdown zone will trigger these respective actions. Clearance zones are monitored prior to the start of pile driving, and the observation of one or more baleen whales in a clearance zone delays pile driving until the animal(s) are observed departing the zone or are not seen for a set amount of time. Shutdown zones are monitored during active pile driving, and the observation of one or more baleen whales in the shutdown zone halts pile driving after it has begun. The size of these zones is estimated during the regulatory process via site-specific sound propagation modeling and frequency-specific acoustic exposure criteria. Once construction has begun, modeled zone sizes are validated in the field via sound field verification studies. If measured zone sizes exceed the modeled ranges, clearance and shutdown zones may be adjusted to align with these in-situ measurements, additional noise mitigation requirements may be implemented, and/or additional monitoring must be conducted.

During pile driving activities, monitoring of clearance and shutdown zones is typically accomplished using several modalities. These include visual monitoring by trained PSOs and acoustic monitoring by passive acoustic monitoring (PAM) system operators (PAMOs), who listen for marine mammal vocalizations. This type of multi-modal approach is currently accepted as the best available method and is consistent with research indicating that no single monitoring approach or technology can be relied upon to detect all animals in a given area, under all conditions, and monitoring is most effective using complementary methods (Verfuss et al. 2018; Macrander et al. 2022). In the U.S., monitoring technologies proposed for use during offshore wind energy development are described in a series of activity-specific monitoring plans, which must be submitted to regulatory agencies for review and approval well in advance of initiating construction. If construction activities are proposed in low visibility (e.g., rain, fog) and/or no-visibility (i.e., darkness), the project proponent must also prepare an Alternative Monitoring Plan (AMP), which describes in detail the proposed monitoring protocols and equipment to be used to effectively monitor mitigation (i.e., clearance and shutdown) zones in these conditions. At the time of writing, no AMPs for initiating pile installation at night have been approved.

A requirement of MMPA Incidental Take Authorizations (ITA) and ESA Incidental Take Statements (ITS) is the reporting of all observations of protected species recorded by PSOs and PAMOs during construction activities. These PSO/PAM reports document all visual and acoustic observations of protected species, including those made in low visibility conditions. In these reports, there is no requirement for an evaluation of the performance of monitoring technologies (e.g., estimating the number of missed detections). However, detection results in low- and no-visibility conditions can provide general information about the relative efficacy of alternative monitoring technologies (South Fork Wind, 2024). For example, although pile driving was only permitted during daylight hours during construction of the South Fork Wind Farm, PSOs monitored for marine mammals opportunistically in periods of darkness (when pile driving was not occurring) using two models of infrared (IR) camera systems mounted on different project vessels. Distances to nighttime detections of baleen whales ranged from 0.6 to 7.5 km (South Fork Wind, 2024). Several studies have assessed the performance of low- and no-visibility monitoring technologies more directly by attempting to estimate detection probability (number of animals detected/total number of animals present). However, this probability varies substantially based on environmental conditions, animal-dependent factors, and other considerations (Verfuss et al. 2018; Smith et al. 2020; Zitterbart et al. 2020).

1.2 Scope of this Report

The scope of this report is to synthesize the state of the science regarding technologies, tools, and methods for detecting baleen whales before and during offshore wind pile driving, with a focus on techniques that can be used in low- and no-light conditions. This report aims to support the application of technologies for making operational decisions, and is not intended to provide regulatory guidance, agency position, or authority. The objectives of the report are to describe:

1. **Performance Metrics** – Identify metrics that can be consistently applied to characterize technology performance (Section 2.0).
2. **Technology Effectiveness** – Document the effectiveness of monitoring technologies based on published literature and data (Section 3.0).
3. **Characterization Framework** – Develop a technical framework for characterizing the performance of emerging and existing technologies (Section 4.0).
4. **Research and Development Needs** – Outline research and development recommendations that can further improve the ability to detect, monitor, and avoid negative interactions of baleen whales with offshore wind facility construction (Section 5.0).

The technologies and approaches discussed herein could be more broadly applicable to other marine mammals and offshore activities; and the terminology, such as ‘marine mammals’, ‘cetaceans’, or ‘construction activities’ may be used. However, the scope of this work is intended to be more narrowly focused on baleen whales and foundation installation activities (pile driving) during offshore wind energy development.

This report is not a regulatory document and is not intended as such. Agencies with regulatory authority over protection of marine mammals and construction of offshore wind energy facilities will independently assess the ability of monitoring technologies to perform their intended task.

1.3 Approach

This work, at the direction of the Department of Energy, was led by the Pacific Northwest National Laboratory (PNNL) and the National Renewable Energy Laboratory (NREL). External input and oversight were provided by a steering committee with representatives from federal agencies, including the Bureau of Ocean Energy Management (BOEM), the National Oceanographic and Atmospheric Administration (NOAA), and the Department of Energy, as well as the Regional Wildlife Science Collaborative for Offshore Wind (RWSC), and the Marine Technology Society. Additionally, the steering committee provided guidance on the scope and review of milestone deliverables. Subject matter experts, including technical professionals from the offshore wind energy industry, technology developers, academics, researchers, and conservation groups, helped focus this effort in the early stages to identify meaningful outputs and review the final report. Views of external stakeholders were obtained on an ad hoc basis and not through the establishment of a group designed to provide consensus advice to any federal agency for its own operations or activities.

Information regarding this reporting, including links to all the reference literature are available at <https://tethys.pnnl.gov/publications/evaluating-tools-technologies-monitoring-baleen-whales-offshore-wind>.

A parallel effort, coordinated by RWSC in partnership with the Consensus Building Institute, TurnForward, and the Marine Technological Society, hosted a series of public workshops to

enable broader engagement and feedback on the direction and findings of this work (Table 1). Participants in the workshop series included offshore wind stakeholders, monitoring technology developers, state and federal agencies, researchers, and conservation organizations.

Table 1. Summary of RWSC workshops

Workshop	Date	Attendees	Proceedings
Workshop 1: Proposed approach and scope	Apr 18, 2024	225	RWSC & MTS, (2024a)
Workshop 2: Present initial framework	Jun 26, 2024	131	RWSC & MTS, (2024b)
Workshop 3: Final outcome review	Nov 19, 2024	143	RWSC & MTS, (2024c)

2.0 Baleen Whale Monitoring System Performance Metrics

Monitoring technologies are designed to maximize their abilities to determine the presence of marine mammals in a particular area. The outputs from different modalities carry different levels of specificity about their observations including detection, classification, and localization abilities.

- **Detection** – the ability of a sensor to determine if a baleen whale is within the observational range of the sensor.
- **Classification** – the ability to determine what species of baleen whale is detected.
- **Localization** – the ability to determine the spatial location of the baleen whale that is detected.

For a technology to effectively monitor baleen whales around offshore wind construction activities, it must be able to make detections with sufficient spatial resolution and then transfer that information from the sensor to a human operator within a decision-making time window. Several components of the technology are needed to process this information flow. The sensor collects the data from the environment followed by one or more steps of signal/data processing to determine a detection, species classification, or localization of the observed detection(s). Lastly, the information is relayed to an operator to inform construction mitigation decisions.

Evaluating the performance of a sensor system requires the application of metrics that quantify the ability of the sensor to perform its intended task with a consistent and standardized approach. Performance metrics must quantify the basic functionality of a sensor in terms of accuracy, range, latency, reliability, and variation in performance relative to environmental conditions. The key performance metrics for baleen whale monitoring technologies are summarized in the following subsections in categories of efficacy (Section 2.1), range (Section 2.2), and data delivery (Section 2.3). The quantification of each performance metric changes based on the environmental conditions during testing. Importantly, each metric should show the variation in performance relative to environmental conditions (see Section 2.4).

2.1 Efficacy

Detection efficacy is the ability of a sensor to correctly and repeatably identify the presence and non-presence of a baleen whale. Detection capability is understood through a *confusion matrix* (Figure 1), which is a table that groups detections into four categories of true positive (TP = sensor makes a detection, and an animal is present), false positive (FP = sensor makes a detection, but an animal is not present), true negative (TN = sensor does not make a detection, and an animal is not present), and false negative (FN = sensor does not make a detection, but an animal is present). Information from a confusion matrix is used to evaluate the performance of a sensor system by calculating the *precision* and *recall* (Hildebrand et al. 2022). Further, for use in a construction setting, the performance of a system should be measured in relation to its ability to inform correct operational decisions, which in this case would be implementing a mitigation action if a whale is observed. If a whale is within a mitigation zone¹ but not detected, then a missed mitigation can have potential consequences for that individual. The potential for missed mitigation can be quantified as the *probability of missed mitigation*.

¹ A mitigation zone is defined as an area in which the presence of a baleen whale may trigger an operational decision, such as a pause in pile driving. Clearance zones and shutdown zones are examples of mitigation zones.

- **Precision** – Measures the fraction of all detections and classification that were correct. Calculated as the number of true positives over the number of total positives for both detection and classification, as applicable.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall** – Measures the fraction of whale occurrences that were detected. Calculated as the number of true positives over the number of actual whale occurrences.

$$Recall = \frac{TP}{TP + FN}$$

- **Probability of Missed Mitigation (PMM)** – Probability that a baleen whale entered the mitigation zone but was not detected (Baumgartner 2024). Calculated as the number of false negatives in the shutdown zone over the total number of positives in the mitigation zone.

$$PMM = \frac{FN_{MZ}}{TP_{MZ} + FN_{MZ}}$$

		Actual Conditions	
		Present	Not Present
Sensor Observation	Detected	True Positive (TP)	False Positive (FP)
	Not Detected	False Negative (FN)	True Negative (TN)

Figure 1. Confusion matrix

Quantifying these performance metrics requires field testing, simulating, or modeling of the sensor system to gather information about detections and non-detections over a period of time. Then, the sensor observations need to be classified as true or false based on information about the actual conditions. Depending on the method of data collection, knowing the actual conditions can be difficult. For example, if sensor observations are collected in the field from natural whale movements, supplementary monitoring system(s) are needed for comparison or some other mechanism of knowing the actual conditions. If testing is done using artificial whale cues, then information about the actual conditions aligns with when the artificial cue was activated.

2.2 Range

Detection range is the physical distance at which a sensor can observe an animal cue. As the distance between an animal and the sensor increases, the strength of noise relative to the strength of the signal increases until the signal cannot be discerned from the background noise. Assuming that there is an equal distribution of whales around a sensor (Figure 2, left), we would

expect the number of whale detections to increase at locations further from the sensor because the sensor is observing a larger portion of the environment (Figure 2, right, dashed line). However, at some distance away from the sensor, the number of whale detections starts to decrease indicating that the sensor is not reliably making all detections (Figure 2, right, solid line).

The distance where detections start to decrease is known as the *reliable detection range* (Zitterbart et al. 2020). The reliable detection range is not universally calculated for a sensor system. Rather, the reliable detection range varies based on site-specific conditions, levels of background noise or interference created by environmental conditions (e.g. visibility, rain), and the species-specific cues that are being detected.

- **Reliable Detection Range** – Distance at which detections of baleen whales starts to decrease. Calculated through collecting observational data.

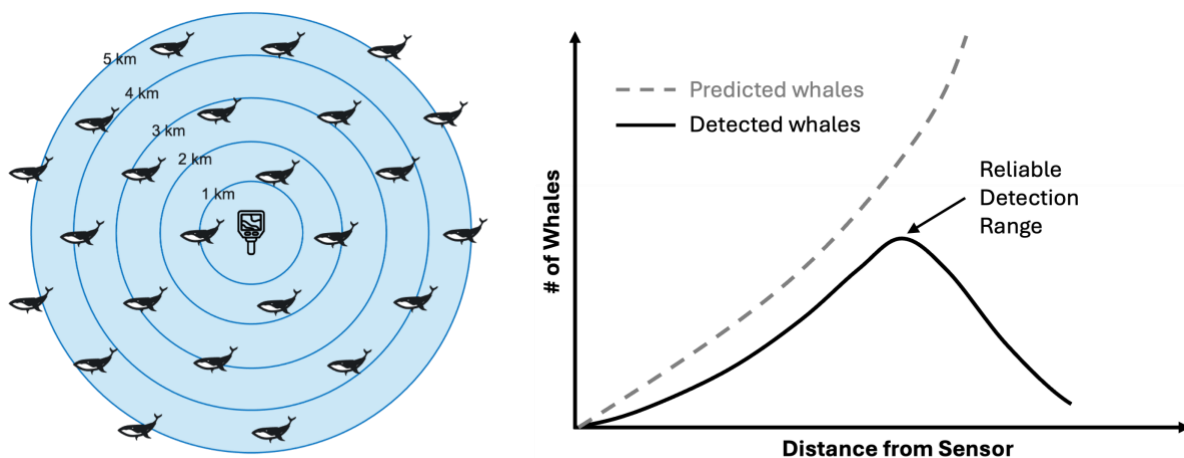


Figure 2. Reliable detection range. Assuming equal distribution of whales (left), predicted detection would increase at further distances from the sensor. The reliable detection range (right) is the distance where the number of detections start to decrease.

In addition to the magnitude of the distance, the spatial resolution of the detection range should be reported. The spatial resolution of detection range refers to the smallest measurable change in distance. For example, the measurement resolution will define whether the detection distance reported to the nearest 0.1 km, 1 km, 10 km, etc.

2.3 Data Delivery

Key metrics related to data delivery include the time required for information transfer and the amount of time that the system is online and functional. Transmitting information in near real-time is critical to the functional performance of a system in an operational setting because any observation could trigger a delay or shutdown procedure. An observation must be transmitted reliably and rapidly to PSOs or PAMOs so they can make appropriate responses to avoid impacts on baleen whales. *Detection latency* is the time duration between a baleen whale detection within a sensor's range and information about the whale's presence being transmitted to a human operator. Detection latency encompasses the time for: (1) the whale cue to travel to the sensor (e.g., considering the speed of sound traveling through water for calls), (2) the signal to be processed by the sensor, (3) an automated or manual process to perform detection,

classification, and/or localization of the signal, (4) additional human review or verification of automated detections, classifications, or localizations, and (5) transfer of this information to the lead PSO to implement mitigation measures (i.e., delay or shutdown) if appropriate. Any of these steps may include time to send information through satellite, wired, or wireless networks.

Information about whale occurrence can only be exchanged when the system is fully operational. *System availability* is the proportion of time that the system is actually operational, compared to the amount of time it is expected to be operational. An operational system means that the system is able to make observations, interpret observations, and transfer information to an outside entity. Information about system availability informs how much down time would be expected.

- **Detection Latency** – Average time duration for information about a whale occurrence to be processed, verified, and sent to an operator when the sensor system is operational.
- **System Reliability** – Proportion of time that the system is operational vs. planned time for the system to be operational.

$$\text{System Reliability} = \frac{\text{Time}_{\text{operational}}}{\text{Time}_{\text{planned}}}$$

In addition to the measuring the average detection latency, the resolution of the temporal measurement should also be reported. The temporal resolution refers to the smallest measurable change in reporting the time of a whale occurrence. For example, the temporal resolution will define whether a whale occurrence can be attributed to a 1 second window, 15 minute time window, 1 hour time window, etc.

2.4 Variation in Performance Relative to Conditions

The performance of a sensor system in a given context is distinct from the technological readiness of that system. Other efforts, including efforts by NOAA and The MITRE Corporation, are underway to define the Technological Readiness Levels of general monitoring approaches, including PAM, visual surveys, and telemetry. Actual performance of a given system is context-specific and depends on the conditions under which it is tested or deployed, including the sensor and platform configuration, environmental factors, and animal-dependent factors (Figure 3). The performance metrics described above will not have a single value that can be universally applied to all conditions, target species, and configurations. Rather, performance metrics should be reported within the context of their testing conditions and not extrapolated to other conditions. With these considerations in mind, we propose a potential framework to characterize the performance of a given sensor system in the context of relevant site conditions and target species.

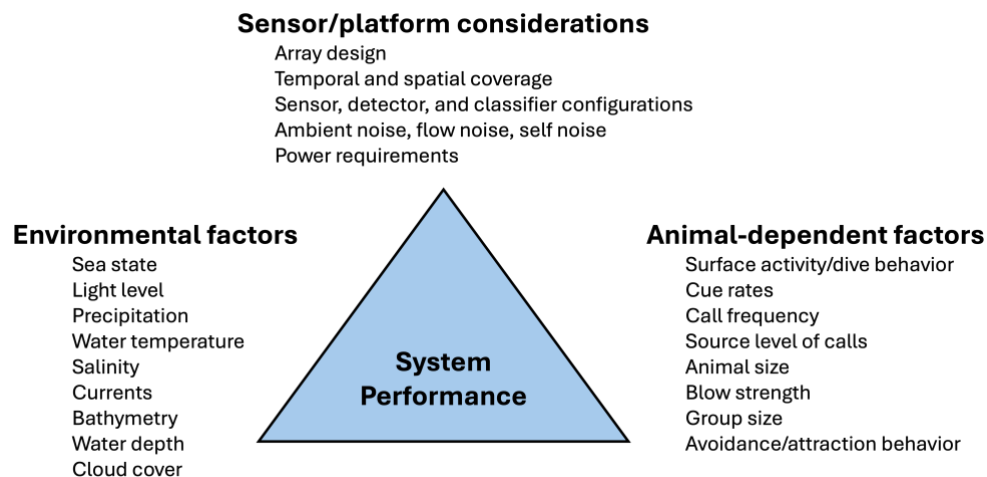


Figure 3. Factors influencing system performance

3.0 Technology Evaluation and Capabilities

Technologies have been used to monitor baleen whales for decades with varying levels of success. Different technologies and tools have been developed and demonstrated, each with different approaches and characteristics. Information about past performance has been documented through technical reports and peer-reviewed literature.

Eight technologies or tools were identified that could be relevant for monitoring baleen whales around offshore wind foundation installation. Technologies were identified through a literature review, input from subject matter experts, and from engagement during public workshops (RWSC, 2024a). The monitoring technologies identified were:

- **Satellite imagery** – identifies whales at or near the water surface
- **Optical cameras** – identifies whale blows or bodies at or near the water surface
- **Infrared imagery** – identifies thermal signature of whale blows or bodies at or near the water surface
- **Telemetry** – animal-borne tags to track whale movement
- **Dimethyl sulfide (DMS) concentration** – concentration of DMS in marine surface waters has been correlated to areas of high productivity or zooplankton concentration where whales may be feeding
- **Environmental deoxyribonucleic acid (eDNA)** – identifies genetic material left behind by animals in water samples
- **Active acoustics** – uses reflections from emitted underwater sound sources (e.g., sonar) to detect whales in the water column
- **Passive acoustic monitoring (PAM)** – measures underwater sound to identify whale calls.

The eight technology types described above are existing or emerging methods to detect the presence of baleen whales. Selecting appropriate technologies depends on the goals of the monitoring campaign, the timing and location of the monitoring campaign, and the behavioral characteristics of the species of interest. No single technology is capable of detecting all individual baleen whales within a given area during all times and weather conditions; therefore, we evaluated the capabilities of each technology type to meet the specific monitoring focus of this report, which is detecting and providing real-time or near real-time information on the presence and location of baleen whales, specifically during offshore wind foundation installation activities under low- and no-light conditions. Each technology was evaluated against four criteria that are critical to perform the intended function of baleen whale detection near offshore wind foundation installation activities (Table 2). The four criteria are:

1. Can this technology record the time and general location of a baleen whale detection in relation to certain offshore wind construction activities?
The purpose of this criterion is to evaluate whether the technology can provide relevant information to inform operational construction decisions. If a whale is detected, an operator would need to know if that animal is within a clearance zone or shutdown zone, and when the animal entered/exited that area. To meet this criterion, a technology should be able to provide spatial resolution on the order of 0.5 to ~10 km and temporal resolution of less than 1 hour. This criterion also signifies whether the technology has the ability to perform localization of an observed animal to a particular location in space and time.

2. Can baleen whale cues be detected and delivered within a decision-making time window?
The purpose of this criterion is to evaluate whether the technology is capable of supporting decision-making (implementing mitigation) in near- real time. If a whale is detected, the information needs to be processed, packaged, and communicated to a PSO and construction operator. To meet this criterion, the technology should be able to deliver information to a human within approximately 10 minutes or less of the baleen whale exhibiting the behavioral cue.
3. Can this technology detect and classify any baleen whale that enters the zone of perception and displays the necessary cue?
The purpose of this criterion is to evaluate whether the technology has the functional ability to detect any baleen whales and classify them to an adequate level of taxonomical classification for management purposes. For example, is the technology able to distinguish between a school of fish and a baleen whale or a toothed whale and a baleen whale? For the technology to be able to classify *any* baleen whale, the technology must not rely on individual-specific markings or tags (such as telemetry) that create a perception bias to only detecting specific individuals.
4. Can this technology detect baleen whale cues in low- or no-light conditions?
The purpose of this criterion is to evaluate whether the technology can detect animals in low light conditions relevant to offshore wind turbine foundation installation.

Using these four evaluation criteria and the past performance of each technology, we determined that PAM and IR imaging met the basic needs for monitoring baleen whales around offshore wind foundation installation activities in low- and no-light condition. Given their capabilities, these two technologies, when designed appropriately, can be suitable as a primary modality for detecting baleen whales around offshore wind foundation installation activities. The technologies that did not meet the criteria could be suitable to function in tandem with one or more of the primary technologies as part of a multi-technology monitoring system.

Table 2. Evaluation of technology types based on three criteria using technology descriptions from Appendix A.

Technology Category	Ability to record time and location of detection?	Detection relayed in relevant time window? ^[a]	Ability to detect and classify whale(s) in zone?	Functions in low- to no-light?	Summary
Active Acoustics	Yes	Yes	No	Yes	Limited detection range and ability to classify species. Adds noise to environment.
DMS Concentration	No	No	No	Yes	Predictive technology; no detection of individuals.
eDNA	No	No	Yes	Yes	Detection and classification possible. Cannot assign a time and place of animal occurrence.
Infrared Imaging	Yes	Yes	Yes	Yes	Detection, classification, and localization possible in all light conditions.
PAM	Yes ^[b]	Yes	Yes	Yes	Detection, classification, and localization ^[b] possible in all light conditions.
Optical Camera	Yes	Yes	Yes	No	Detection, classification, and localization possible in good light conditions.
Satellite	Yes	No	Yes	No	Detection, classification, and localization possible in good light conditions. Satellite image temporal availability not currently suitable for near real time.
Telemetry	Yes ^[c]	Yes ^[d]	No	Yes	Only informative for tagged animals.

^[a] A decision-making time window is defined as an amount of time that allows implementation of a mitigation measure or measures that prevent harmful interaction with a baleen whale after the animal exhibits a detectable cue. A reasonable time window is considered here to be approximately 10 minutes.

^[b] The ability of PAM systems to localize detections is dependent on the type, number, and placement of hydrophones.

^[c] For telemetry data transmitted via the Argos satellite network, the number of satellite fixes upon tag surfacing determines the accuracy of location estimation.

^[d] Tools such as goniometers or other RF tracking may be able to augment ability to detect animals without requiring data processing/filtering of Argos (satellite) position data.

A literature review was conducted for each technology to evaluate past performance across the key performance metrics of efficacy, range, data delivery, and variation. Literature was collected from peer-reviewed journal articles, technical reports, consultant reports, and regulatory documents, as applicable and available. Specific focus was put on literature that included detection of baleen whales in marine environments.

The function and performance of IR imaging and PAM are summarized in Sections 3.1 and 3.2, respectively, followed by a summary of multi-modal technologies in Section 3.3. A literature review about the performance of the six other monitoring technologies is provided in Appendix B. Each technology literature review includes a table that summarizes relevant studies, including the technology used, study location, species, environmental conditions, detection time and distance, and results are presented for each technology.²

In addition, we developed technology profiles for PAM and IR imaging to provide a detailed description of the design and use of the technology in Appendix C. The technology profiles provide background information to describe how the technology functions, how design choices and environmental conditions affect performance, how to approach characterizing system performance, a case study on system characterization, and recommendations for research and development.

3.1 Infrared Imaging

An IR camera is similar to a standard camera, except it captures long-wave IR light (8–14 μm) and does not require visible light. A cue can be a body part of an animal or its breath, but to be detected, it must be at or above the surface of the water (Baldacci et al. 2005) and have greater thermal energy than the water. To avoid blurry imaging, cameras need to be stabilized on a platform at a high enough elevation to maximize the field of view while still allowing for near-field detections. Performance of IR imagery for detecting baleen whales is summarized in Table 3.

Camera performance is greatly impacted by weather and ocean conditions. Fog and precipitation can mask cues and reduce detection distances (Richter et al. 2024; Verfuss et al. 2018). Sea state, white caps, and breaking waves can increase IR clutter and false detection rates (Baldacci et al. 2005; Graber et al. 2011; Verfuss et al. 2018). For platforms at sea (vessels), the sea conditions can also affect sensor stabilization (Smith et al. 2020) and higher wind speeds will cause whale blows to dissipate quicker (Richter et al. 2024). Humidity may also cause condensed water vapor to accumulate on the sensor lenses (Baldacci et al. 2005). Glare can resemble warm anomalies (Zitterbart et al. 2013; Verfuss et al. 2018), especially at incidence angles $>60^\circ$ (Graber et al. 2011). IR imaging is typically more effective at night, when there is less reflected radiation (Verfuss et al. 2018; Zitterbart et al. 2013). Finally, warmer temperatures may reduce the temperature differential and thus the contrast between animals and their surroundings (Verfuss et al. 2018), especially with uncooled sensors (Verfuss et al. 2016, 2018; Horton et al. 2019).

Camera performance also differs depending on the characteristics of the camera, including pixel size (smaller pixels have greater spatial resolution but lower thermal sensitivity), focal length (longer focal lengths have greater spatial resolution but smaller field of view), frame rate (tradeoff between storage size, processing power, and ability to detect a moving animal), and whether cooled or uncooled (uncooled have shorter detection distances).

² Each table presents a representative sample of the literature that best documents system performance and is not intended to be fully a comprehensive survey of every existing study.

Table 3. Representative literature review of infrared imagery

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
Real-time cooled thermal imaging system 8 m above waterline	Estimate the thermal infrared radiation of whales	Svalbard; minke, blue, fin, sperm, humpback whales	All observations during daytime in calm, slow ocean swells, Beaufort 0–3; SST 2.7–10.1°C	Real-time	20–150 m; blue whale blow at 1 km	Temperature difference depends on incidence angle and distance	Cuyler et al. 1992
Real-time cooled thermal imaging system 22 m above sea level	Diel differences in migrating	California; gray whales	Fair to excellent based on sea state and wind	Recorded on VHS tape for review later	4–5.4 km	Visible day and night; migration rate and offshore distances greater at night	Perryman et al. 1999
Calibrated, cooled handheld infrared binoculars	Infrared detection of marine mammals	Italy; fin, sperm, and pilot whale, dolphin spp.	Day and night tests across a range of good and bad weather/sea states	Real-time	150 m to 1 km (max for blue whale blows)	Limited by FOV; clutter from waves and rain and fog affected images	Baldacci et al. 2005
Aerial (down-looking) and cliff-mounted systems (height unknown)	Observe marine mammals	Hawaii; humpback	Not stated	Not stated	Ranges up to 12.9 km	Bit size on aerial sensor too small	Schoonmaker et al. 2008
Uncooled infrared camera 13 m above sea level with automatic detection software	Infrared detection of marine mammals	Washington; killer whales	Day and night in clear skies, calm seas, and wind speed 0–4 m/s	150 sec manual review; 9 sec automatic detection	Distances ranging from 42 to 162 m	Detections day, night, twilight; incidence angle affected detections; 42 false detections	Graber et al. 2011
Stabilized rotating (360°) thermographic scanner with automatic detection software	24/7 monitoring for mitigation during seismic surveys	Arctic and Southern Oceans; mainly humpback, fin, minke whales	Day and night with sea surface temps mostly -1.8 to +10°C, wind speeds below 7 Beaufort	Not specifically stated but present probability of timely detection based on distance and species-specific dive times	Distances of up to 5,500 m.	Detected more than 4,500 whale blows; Classifier AUC was 0.99 for training set, and 0.98 for test set; ROC curve better at night (0.98 vs 0.90)	Zitterbart et al. 2013

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
Uncooled Forward Looking Infrared camera 2–10 m above sea level	Quantify temp anomalies, blow height, effects of emissivity	Alaska; Humpback whales	Daytime only in Beaufort 2–4	Not stated	Blows imaged up to 150 m	Extreme emissivity b/c data collected at sub-horizontal angles; temps measured precise but not accurate	Horton et al. 2017
Shore-based infrared 28.1 m above sea level	Migrating gray whale call and blow rates	California; gray whales	Day and night in good weather conditions, little wind	Not stated	479 m to 5.8 km; Prob. of detection decreased beyond 2.1 km offshore. Probability of infrared blow detection was the same night and day.	Detecting distant blows from a video difficult due to decreasing image resolution as range increases; still need human observers	Guazzo et al. 2019
Stabilized rotating infrared camera at 7.8 m above water line with automated detection algorithms	Compare marine mammal monitoring methods	Nova Scotia and Newfoundland and, Canada; many marine mammal species	Day and night in mixed weather, including rain, drizzle, fog; half time Beaufort ≤ 3	Not stated; real-time	Not stated; only marine mammal observer (MMO) distances shown	PAM > MMO & IR if precipitation; prob det by MMOs and IR 20–34%; high number false positives with IR	Smith et al. 2020
Rotating cooled line scanner on tripod; range of heights (16 to 51.3 m) with automated detection algorithms	Evaluate infrared automatic detection technology	Australia; mainly humpback and minke whales	Day and night in 12°C vs 10°C and 21°C vs 25°C (air vs sea surface temperature), Beaufort <4	Real-time	Whales perceptible in >70% of images up to 3 km; Distance to detection peaks ranged 0.5 to 3 km	Height, wind, and fog impacted detections. Visibility generally higher during nighttime; variable false alerts	Zitterbart et al. 2020

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
Uncooled thermal imaging cameras mounted 4 and 15 m above MSL with automated detection algorithms	effective infrared detection systems for smaller marine mammals	British Columbia, Canada; humpback whale, killer whale, pinniped and porpoise	Day and night observations. Environmental conditions not stated (recorded 22 mos)	Not stated but detection algorithm ran in real-time	Detections increased then decreased with distance; Det. functions differ by species and height/location of cameras; max 2100 m	Still used human operator for validation	Richter et al. 2024

IR cameras only detect animals at or above the surface, thus, long deep-diving animals may have lower detection probabilities. Depending on species, the body or blow might be more visible (Graber et al. 2011; Verfuss et al. 2016, 2018). Species identification is also not possible, unless there is a strong, defining characteristic that can be captured with thermal imaging. Different species may also have different detection functions from day to night (e.g., Perryman et al. 1999; Richter et al. 2024; Zitterbart et al. 2020).

If a human operator is making observations alongside an IR camera, the distance of a detection can be determined with a range finder (Yonehara et al. 2012). Horizontal distance ranging using IR cameras can be estimated in a similar manner as binoculars that use distance below the horizon (reticles) and the bearing (degrees) relative to the survey platform's position (e.g., if vessel-mounted, this refers to the vessel's position) as well as platform and camera lens height (Michel, 2015). Distance estimations typically use spherical triangulation (Lerczak and Hobbs 1998). However, this mathematical approach does result in distance-dependent error. For example, Zitterbart et al. (2013) found that for distances less than 5 km, relative errors are within 12%.

Cameras that can be directly connected to a computer running automatic detection software (e.g., Smith et al. 2020; Zitterbart et al. 2013) allow for real-time monitoring and mitigation. Data recovery and download (e.g., from mobile platforms such as drones) may add additional time to the data delivery process (e.g., Schoonmaker et al. 2008). There are currently proprietary automatic detection and classification software programs (e.g., Horton et al. 2017, 2019) and custom software (e.g., Richter et al. 2024; Zitterbart et al. 2013, 2020) that use a machine learning based classification system to identify significant thermal anomalies relative to the surrounding water surface. Moreover, because automatic IR detections include true positives (meaning a cue was present and the automatic detector identified the cue) and false positives (meaning a cue was not present but the automatic detector identified something in the image as a cue), they must be confirmed by a human operator (Zitterbart et al. 2013).

3.2 Passive Acoustics

PAM uses hydrophones to record underwater sounds in the environment such as baleen whale calls. Acoustic data can be analyzed through manual review or automated techniques to

perform detection, classification (e.g., Mellinger 2002; Baumgartner and Mussoline, 2011; Helble et al. 2012; Gillespie et al. 2013; Allen et al. 2021; Rasmussen and Širović 2021), and localization of baleen whales (e.g., Širović et al. 2007; Helble et al. 2015; Guazzo et al. 2017; Gervaise et al. 2021).

Hydrophones can be deployed from towed arrays; stationary platforms at the surface, moored in the water column or on the seafloor; or mobile platforms that are powered, autonomous, or free-drifting (Table 4). To perform near real-time monitoring, PAM systems need to have communication systems and adequate power and processing capabilities to transfer acoustic information to an operator instead of archiving the data onboard for later analysis (Baumgartner et al. 2013, 2014; Kowarski et al. 2020; Premus et al. 2022).

Most baleen whales produce species-specific calls that are unique in duration, frequency range, and source level. The ability to detect baleen whales using PAM depends on individual animals producing calls within range of a hydrophone (availability bias) and the underwater acoustic properties, which enable the call to be distinguished above background noise (perception bias). Most baleen whales produce low-frequency calls ranging from 10 Hz to 10 kHz (Thompson et al. 1979), with some calls, including some from NARW reaching frequencies above 20 kHz (e.g., Clark 1990; Matthews and Parks 2021). Baleen whales do not produce calls in predictable time intervals; instead, it depends on their behavior and environmental conditions, such as prey availability and anthropogenic noise. For example, call rates can vary as widely as 0 to 200 calls per hour for NARW (Parks et al. 2011) and 19 to 122 calls per hour for sei whales (Baumgartner and Fratantoni 2008).

Table 4. Representative literature review of PAM

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
PAM, towed hydrophone array. PAMGuard software	Compare marine mammal monitoring methods	Nova Scotia and Newfoundland, Canada; many marine mammal species	Mixed; rain, drizzle, fog; half time Beaufort ≤3	Not stated; real-time	Not stated; only MMO distances shown	Baleen whale calls were masked by vessel noise; could not be detected	Smith et al. 2020
PAM, moored buoy	Evaluate performance of moored PAM buoy for baleen whales compared to aerial survey	Massachusetts; North Atlantic right, fin, humpback, sei	All conditions	“Near real-time”; exact latency not reported	Not assessed. Estimated 9 km for NARW & humpback; 10–15 km for sei; 10s of km for fin (all with reference to other literature)	Human analyst performance compared to aerial surveys, assessed in 15-min windows, not individual detections: Precision = 100% Recall = 33–58%	Baumgartner et al. 2019

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
PAM, array of bottom-mounted recorders	Evaluate presence of baleen whales off coastal Virginia	Virginia; North Atlantic right, fin, blue, minke	All conditions	Recordings, not real-time	0.4 km (minke whales in loud periods) to >150 km (fin whale in quiet periods)	Relative Precision: Fin: 99% NARW: 85% Minke: 87.5% Automated detector performance compared to manual review	Salisbury et al. 2018
PAM, glider	Evaluate performance of near real time monitoring of marine mammals from gliders	Gulf of St. Lawrence; North Atlantic right, fin, blue	All conditions, 45 days	15 min to 3 h 15 min	Not assessed; (estimated 30 km for NARW in Laurinolli et al. 2006)	Automated detector performance compared to human analyst: Precision = 10–71% Recall = 11–98%	Kowarski et al. 2020
PAM, moored acoustic recorders 1.5–2 m above sea floor	Test performance of fully automatic detection and classification methods	Gulf of Maine; sei, North Atlantic right	Not reported. All conditions during 4 multi-day campaigns	Recordings, not real-time	n/a	Automated detector performance compared to 1 of 3 manual reviews: Sei 33% missed detection, 58% false detections NARW: 63% missed detections, 58% false detections	Baumgartner and Mussolinie 2011
PAM, underwater recorders with surface communication buoy	Describe technology used for real-time acoustic monitoring during seismic surveys	Sakhalin Island, Russia; Gray	Not reported	Not reported	Not reported	Implemented 4 survey shutdowns for approaching gray whales. Whale locations determined by MMO, sound exposure level determined with near real-time PAM measurements	Rutenko et al. 2022
PAM, gliders	Compare acoustic and visual surveys under various simulations	Gulf of Saint Lawrence, Canada; North Atlantic right whales	Not reported	<5 h for visual versus 39 h for acoustic detection with a prob of 0.5	Prob det 0.5 at 10 km (nominal max 20 km)	Gliders performed better; acoustic glider detected under every scenario but visual surveys only reliable (>0.5) when 20 or more whales present	Ceballos et al. 2022

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
PAM, seafloor mounted recorders	Assess performance of PAM as an alternative to aerial visual surveys	Cape Cod Bay, MA; North Atlantic right whales	Aerial surveys <5 ft, vis >1.2km, no fog, rain, snow	Not reported	9 km	Acoustic detections 100% of days when detected with visual or acoustic methods but only 62% of days with visual methods	Clark et al. 2010

PAM enables detection of baleen whales in variable lighting, cloud, and environmental conditions that would limit visibility needed for visual or optical technologies. If baleen whale calls are distinct and stereotypic, PAM can detect and classify baleen whales to the species level. Near real-time detections can be made if the PAM system is equipped with appropriate processing and communication software and hardware to transmit signals to a human operator.

In specific instances, localization has been done using a single hydrophone (Tiemann et al. 2006; Bonnel et al. 2014; 2020) or DIFAR (Directional Frequency Analysis and Ranging) sonobuoy(s) (Miller 2012; Rone et al. 2012; Bonnel et al. 2020). However, localization is typically done using a hydrophone array, which is a collection of hydrophones attached to each other at known fixed distances. These arrays can be fixed (Stanistreet et al. 2013; Tremblay et al. 2019) or mobile, such as gliders (Baumgartner et al. 2020) and towed arrays with left-right ambiguity (Marques et al. 2013; Zimmer, 2013; Van Parijs et al. 2009; von Benda-Beckmann et al. 2013; Yack et al. 2013). With linear arrays, even just two hydrophones, time differences in sound arrival on each hydrophone allow orientation of the call's direction to estimate location. With nonlinear configurations (i.e., four or more hydrophones), three-dimensional localization (range and depth) can be determined (Wahlberg et al. 2001).

The typical methodology uses time difference of arrival (TDOA) of the signal to multiple time-synchronized hydrophones to estimate the location. However, localization is not possible in every environment and for every whale species. Large errors in estimated ranges of deep diving animals can also occur, which could erroneously estimate an animal outside of a mitigation zone. Additionally, acoustic localizations are only possible if the call is received on multiple hydrophones. With few hydrophones and large-aperture (widely spaced) arrays, this lowers the likelihood of receiving a call on multiple hydrophones, especially baleen whale calls. Increased noise levels during pile driving may mask calls, further reducing the capabilities of localization (Verfuss et al. 2018). Multiple overlapping calls from many/distance animals can also reduce the capabilities of localization (de Castro et al. 2024). Finally, few monitoring systems are currently capable of real-time detection with localization.

The ability of a PAM system to distinguish a whale call depends on the ambient sound levels in the ocean and the position of the hydrophone in the water column, specifically relative to the surface and bottom. Background noise levels, and thus the ability to detect a whale call, vary with dynamic environmental conditions as well as human activities (e.g., vessels, sonar, and construction). Whale calls can propagate over 150 km in quiet conditions, as little as 0.4 km in noisy conditions (Salisbury et al. 2018), or can be completely masked by nearby vessel traffic (Smith et al. 2020), wave action, or other sound sources (e.g., Rankin et al. 2020). The deployment and characteristics of a PAM system need to be designed with recognition of the local sound propagation and acoustic environment. Even in suitable conditions, the ability to

perceive a whale with PAM requires that an individual is calling when located within the detectable range, but whale calling rates change based on life stage, activity, and behavior (e.g., Parks et al. 2011; Thomas and Marques 2012; Booth et al. 2017; Rice et al. 2022).

3.3 Multi-modal Monitoring

Each monitoring modality described above and in Appendix B has unique advantages and limitations. Applying two or more monitoring approaches together in a complementary manner can improve the effectiveness and reliability of detection, classification, and localization of baleen whales during offshore wind foundation installation. For example, monitoring of clearance and shutdown zones during pile driving is typically achieved by a combination of visual observations by trained PSOs and listening in real-time for whale calls using PAM systems (Macrander et al. 2022). In addition to these two approaches, IR camera systems may also be employed in daylight hours when visibility is obscured by rain or fog. These are examples of a multi-modal monitoring system because they use multiple sensing technologies or approaches to achieve the same monitoring goal.

Monitoring baleen whales using multi-modal approaches has also been investigated directly and indirectly through scientific and engineering studies in peer-reviewed literature. Several studies that deployed multiple monitoring technologies simultaneously at a single location are summarized below:

- Clark et al. (2010) investigated the performance of PAM as an alternative to aerial visual surveys. During 58 days of concurrent surveys, whales were detected on 49 days by PAM and 30 days by aerial survey. The study did not find a correlation between number of PAM detections and number of visual detections, which may be attributed to each modality observing different animal cues.
- Baumgartner et al. (2019) compared the performance of an automated PAM detection and classification system to aerial visual surveys. The PAM detections agreed with >99.9% precision with aerial detections of right, humpback, sei, and fin whales within 15-minute time windows, indicating that when whales are seen in an aerial survey, they are also heard on the PAM system. Recall of the aerial surveys ranged from 33% (humpback) to 73% (fin), showing that the aerial survey missed some detections that were made by the PAM system within 15-minute time windows.
- Guazzo et al. (2019) used PAM, IR imaging, and PSOs to independently measure cue rates from migrating gray whales to support estimations of population-level abundance calculations (i.e., calculate blows/hour and calls/hour). This study did not compare the performance of the three methods with each other, but the authors note the benefit of using multiple monitoring modalities to observe whales.
- Smith et al. (2020) used PAM, IR imaging, and PSOs to observe cetaceans from a vessel. During times of concurrent monitoring, each modality detected some of the same whale encounters and many different whale encounters. Comparing PSOs to IR, 34% of baleen whale encounters detected by PSO were also detected by IR imaging; and 21% of baleen whale encounters detected by IR imaging were also detected by PSO. Similarly for PAM, 28% of cetaceans³ detected by PSOs were also detected by PAM,

³ Cetacean detections are reported instead of baleen whale detections from the PAM system because most low-frequency baleen whale vocalizations were masked by vessel noise.

and 42% of PAM detections were also detected by PSOs. And lastly, PAM detected 12% of the IR detections of cetaceans, while IR imaging detected 13% of PAM detections of cetaceans. Their results show that one modality alone could not detect the majority of whale encounters; instead, different modalities detect different encounters.

- Ceballos et al. (2022) developed a computer simulation to compare the detection probability of NARW from acoustic gliders, and visual surveys from aircraft and vessels. The model results indicate that a slow-moving acoustic glider can detect a single right whale in a mitigation zone while visual survey techniques require multiple transits across the zone or multiple whales to be present to have a detection probability greater than 50%. While this model compared the performance of different monitoring platforms, it did not investigate the performance of a multi-modal system.

Results from multi-modal studies show a range of possible outcomes, with shared detections anywhere between 10% and 100% for two technologies. While some studies above are using detections from one technology to confirm detections from another technology, these systems can also be used in a complementary fashion to increase the number of detections. Across all studies, it is clear that simultaneous use of multiple monitoring modalities increases the overall probability of detection because of the ability to track multiple animal cue types. Low agreement between two technologies (e.g., 12–13% agreement between PAM and IR imaging in Smith et al. (2020)), indicates that multi-modal approaches are effective to observe animals that may be displaying different cues, thereby reducing the number of false negative detections.

Multi-modal approaches are also used to detect and locate whales in other monitoring and research contexts. For example, a combination of visual sightings and PAM are sometimes used to locate target animals for tag deployment during U.S. Navy-funded field studies (Baird et al. 2024; Henderson et al. 2024). In any whale tagging study, the availability of wild, free-ranging animals suitable for tagging is often unpredictable and locating these whales is time- and labor-intensive. There are also typically short weather windows in which to deploy tags successfully. To maximize the chances of locating whales and deploying tags successfully, researchers conducting small-boat tagging surveys in and near at sea Navy ranges—some of which are instrumented with underwater hydrophone arrays—often coordinate remotely with the PAMOs operating these arrays. If a target species is detected acoustically on the array, the PAMO relays species information and the approximate location of the vocalization to the boat crew, thereby directing the tagging team to the general location of the animal(s) (Baird et al. 2024; Henderson et al. 2024). These efforts are conducted in daylight conditions because sufficient ambient light is required for accurate tag deployment.

A multi-modal approach to baleen whale detection is therefore consistent with accepted best practices in a variety of monitoring and research contexts. Available research indicates that no single monitoring approach or technology can be relied upon to detect all animals in a given area, under all conditions, and monitoring is most effective using complementary methods (Verfuss et al. 2018; Smith et al. 2020). In this report, we do not attempt to quantify the change in detection probability resulting from any particular combination of technologies because the performance of multi-modal monitoring approaches has not been adequately assessed. Nevertheless, we recommend the use of multi-modal monitoring approaches to maximize the efficacy of monitoring efforts and technology applications.

3.4 Technology Readiness and Limitations

Each monitoring technology comes with its own set of benefits and limitations. Each technology is limited by availability bias, which is the likelihood that an animal displays a detectable cue within the sensor's range. For example, PAM relies on baleen whales calling to be detected and classified. Calling rates are highly variable based on an individual's behavior (e.g., they can range from 0 to 200 calls per hour for NARW (Parks et al. 2011)). Therefore, false negative detections are possible if an individual swims through a sensor's range while silent. Each technology is also limited by perception bias, which is the likelihood that the sensor can distinguish a cue from the background environmental noise or signal. For example, IR imaging is affected by atmospheric and sea surface conditions. Fog or precipitation limits the ability of IR cameras to detect whales at the sea surface, which can lead to false negative detections. Waves, white caps, and other animals, such as birds, can create thermal anomalies that can frequently be interpreted by IR cameras as false positive detections. Overall, the availability bias and perception bias of a technology depend on many factors including system or sensor design, biological conditions, and environmental conditions. Ultimately, the result of these conditions and design choices impact technology performance across the key metrics, through the following mechanisms:

- **System or Sensor Design** – Sensors that monitor the presence of baleen whales are designed to detect the presence of an animal based on the behavior produced by an individual or group (e.g., surfacing, calling). Cues represent observable features that the individual or group displays. All sensors have limitations for animal detectability—for example hydrophones are sensitive to specific frequencies and cameras are sensitive to atmospheric conditions. The capabilities of a sensor change based on the site characteristics and environmental conditions where it is deployed.
- **Biological Conditions** – Monitoring technologies are only effective at detecting the specific cues they are designed to detect. The ability of a given system to detect a baleen whale depends on whether the animal exhibits that cue (and frequency of exhibiting such cue) within the detection range of a given sensor. For example, a whale must call to be detected by a PAM system, and likewise must be at or above the surface to be detected by an imaging system. The behavioral characteristics of breathing, breaching, swimming, diving, and calling all affect the ability of a sensor detect an animal.
- **Environmental Conditions** – The conditions, including bathymetry, seafloor sediment, water temperature, salinity, wave height, wind, precipitation, water depth, cloud cover, and visibility, can influence the detection capability of a given monitoring system. For example, wind-driven surface waves can increase ambient noise levels recorded by PAM systems that are deployed at or near the surface resulting in masking of whale calls and likewise these waves can result in false cue detections by the IR camera.

Technology performance is influenced and limited by system design, biological conditions, and environmental conditions. The effects of different conditions can be measured with intentional studies. However, consistent metrics or methods have not been used across multiple studies to allow comparison of technology performance across the different conditions described above.

As described in Section 2.0, key performance metrics include measurements of efficacy (i.e., precision, recall, PMM), range (i.e., reliable detection range), data delivery (i.e., latency, availability), and how these metrics change across a variety of conditions. Detection range and data delivery metrics are not frequently reported in studies. The lack of reporting may be attributed to the research objectives of many studies, which investigate whale behavior and

populations with biological applications instead of focusing on technology development with engineering or operational applications.

Similarly, the methods used to measure efficacy are not easily comparable between studies. This is due in part to the difficulty in accurately assessing false negative and false positive rates, which are required to calculate precision and recall. In several studies, precision and recall of one method are calculated relative to missed detections and false detections from another method. For example, assessments of precision and recall of automated detectors and classifiers have compared results to those from an analyst manually reviewing the same data to identify false or missed detections (e.g., Baumgartner and Mussoline 2011; Salisbury et al. 2018; Kowarski et al. 2020). However, this comparison approach inherently overlooks availability bias of the sensor and focuses on evaluating perception bias from different analytical methods. Other approaches have focused on the conditional probability of one technology making the same detection as another methodology (e.g., comparing IR imaging, PSO, and PAM in Smith et al. 2020; or IR imaging and PSO in Zitterbart et al. 2020), which incorporates the availability bias of different modalities.

Considering the literature on PAM and IR imaging as a whole, technology advancements have been made for automated detectors and classifiers to perform nearly as well as a human review of the data in terms of recall (i.e., low missed detection rates not considering availability bias). High numbers of false detections are possible, particularly from IR imaging detecting birds and waves, but adding a human-in-the-loop to quickly review and classify automated detections in real time can significantly improve precision (Smith et al. 2020).

When one technology is compared to another, detections agree within <10% to 60% (based on the information compiled in Section 3.0). Using field observations, there is not a perfect—or “true”—baseline for comparison because there is always a possibility of an unknown false negative detection where a baleen whale is present but not available for detection. Whether testing a technology in the field or using it in an operational setting, multiple modalities will yield better outcomes in terms of system-wide probability of detection and accuracy of classification because each modality can detect and classify different animal cues (e.g., an IR signature, an optical image, and a call).

A comprehensive characterization of system performance is not available from past studies because: (1) performance varies with system design, biological factors, and environmental factors unique to a monitoring location, (2) consistent methods have not been used across field studies, (3) metrics on detection range and data delivery are missing from many studies, and (4) performance of multi-modal systems have not been thoroughly evaluated to see real-world monitoring performance of baleen whales. To advance the state of technology for this application, a framework is needed to characterize the performance of a single technology or multi-modal system with a consistent and standardized approach.

4.0 Potential Framework to Characterize Technology Performance

A potential framework⁴ is described below to characterize technology performance with data collection and modeling through a consistent and standardized approach. The framework is designed to characterize the performance of a baleen whale monitoring system with a specific monitoring objective in relevant marine conditions considering location and species. The framework encompasses three phases (summarized in Figure 4): (1) system design to describe the monitoring use case and system specifications, (2) pre-testing assessment to evaluate system specifications and develop a field characterization plan, and (3) field characterization and modeling to evaluate the application of the technology through field observations and/or modeling.

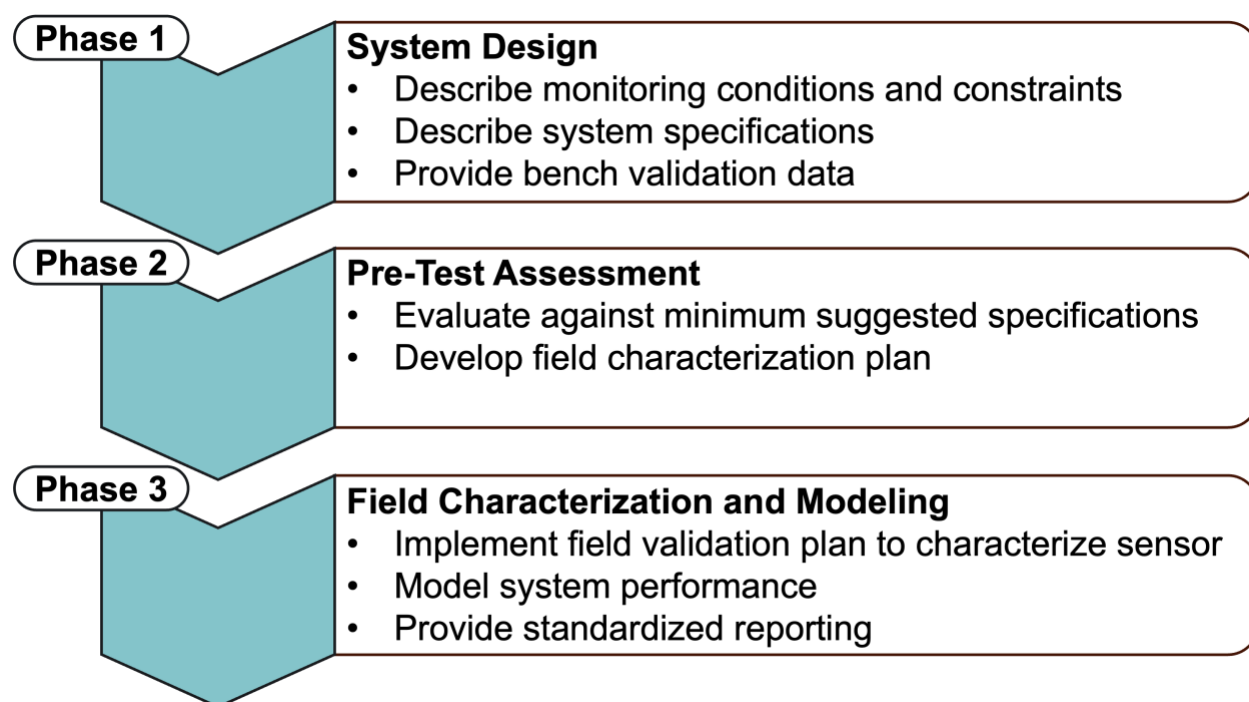


Figure 4. Summary of three phases of potential framework to characterize system performance.

The process by which a technology might be characterized using this framework is detailed in Figure 5. The framework involves a technology developer or operator who is performing the system characterization and a third party who is reviewing the system design and field characterization plan. The technology developer or operator is intended to lead the characterization framework throughout all three phases with input from a third party. Single technologies or multi-modal systems can be evaluated through the framework. The process and objective of each phase of the framework is described in the following subsections.

⁴ The potential framework should not be interpreted as any type of regulatory requirement in any context. The objective of this framework is to present potential evaluation approaches for a variety of different actors to consider, including offshore wind developers, technology providers, or regulatory agencies.

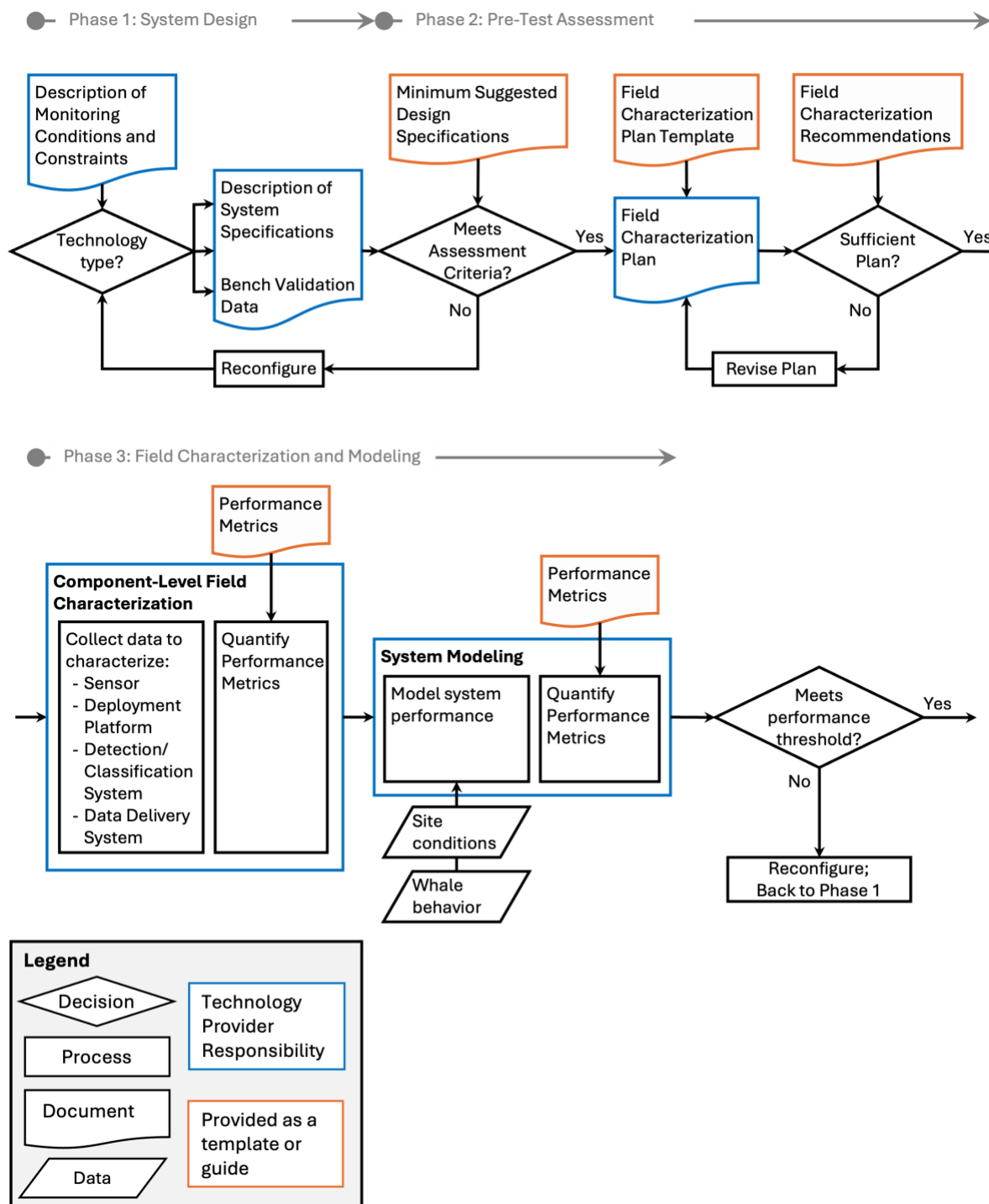


Figure 5. Detailed framework that can be used to characterize the performance of any chosen technology system. Colors indicate technology provider responsibility (blue) and templates that should be provided as a guide to the technology provided (orange).

4.1 Phase 1: System Design

The purpose of Phase 1 is for the technology provider to define the eventual use case of the monitoring technology and document the system specifications with supporting bench data. (Figure 6).

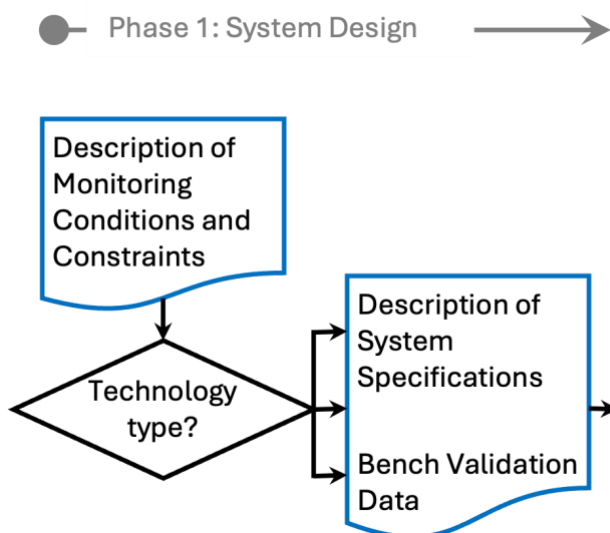


Figure 6. Flow diagram for Phase 1: System Design

The first step within Phase 1 is to clearly define the Monitoring Conditions and Constraints, which outlines the objectives and use case for the technology. The use case for the technology is its eventual application in an operational setting, that is, when the technology will be used in the field for operational monitoring purposes. The use case is not the pre-operation field characterization test, which is defined in Phase 2 and executed in Phase 3. The **Monitoring Conditions and Constraints** include the following information:

- Timing and location of proposed monitoring
- Relevant environmental conditions at the proposed location
- Relevant anthropogenic conditions at the proposed location, including construction noise, vessel use, and other factors that could affect monitoring system performance
- Biological conditions, including which species are to be monitored and their behavioral characteristics
- Definition of the monitoring and detection targets the technology provider plans to meet, including maximum notification timing duration, detection distance and range, spatial resolution (i.e., confidence intervals, as applicable to localization), and whether species classification is used.

Next, the technology provider describes the System Specifications for the system that is designed to meet the expected monitoring conditions. The **System Specifications** will include a description of:

- Sensor type and technical specifications
- Deployment platform type and specifications
- Detection and classification approach

- Data delivery specifications, network diagram, and communication latency, including information about how information is communicated from the sensor to an operator that can implement mitigation decisions.

To supplement the system specifications, **Bench Validation Data** can be provided to illustrate that the system meets the technical specifications. Providing Bench Validation alongside the Systems Specifications helps prove that a system or sensor can meet the documented technical specifications. For example, if a PAM system is specified to have a frequency range from 10 Hz to 100 kHz or an IR imaging system is specified to have a 25-degree horizontal field of view, then supplementary Bench Validation Data can show the system's capabilities for meeting those specifications from a controlled laboratory environment. As appropriate, Bench Validation Data can document the basic functional performance of the technology. Bench Validation Data can include calibration data (e.g., pistonphone calibration for hydrophones), laboratory performance, and documented detections (e.g., output from a system detecting known sounds, IR images, and accompanying classifications of known animals).

4.2 Phase 2: Pre-Test Assessment

The objective of Phase 2 is to assess the proposed system specifications and develop a plan to characterize its performance (Figure 7).

The first step in Phase 2 is to assess the System Specifications provided in Phase 1. The System Specifications document is reviewed by a third party against a set of **Minimum Suggested Design Specifications** to evaluate whether the proposed system has the technical capabilities of performing under the expected Monitoring Conditions and Constraints. This assessment is an early check to confirm that a proposed system is suitable for use for the proposed monitoring objective. This assessment is based solely on technical specifications and supporting Bench Validation Data and occurs prior to any field testing. The Minimum Suggested Design Specifications provide a guide for evaluating whether a system has suitable technical specification to perform the intended task. The Minimum Suggested Design Specifications would include a simple checklist of yes or no questions to easily evaluate the proposed system design. Examples of Minimum Suggested Design Specifications are provided in Appendix D for both PAM and IR imaging. If a third party determines that a proposed system does not meet the Minimum Suggested Design Specifications, then the technology developer can reconfigure the system and return to Phase 1.

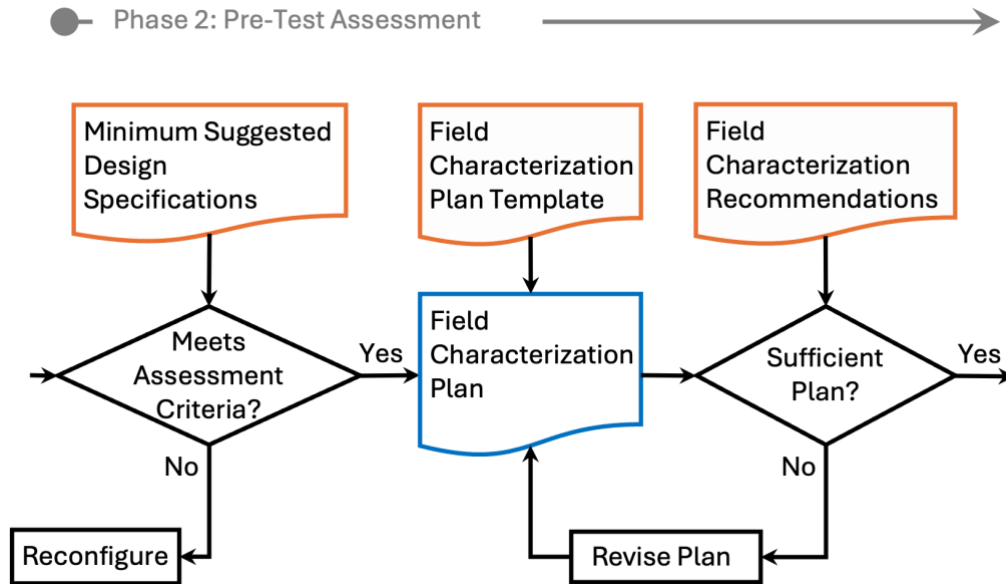


Figure 7. Flow diagram for Phase 2: Pre-Test Assessment

If the proposed System Specifications meet the Minimum Suggested Design Specifications, the technology developer can proceed to developing a Field Characterization Plan. The **Field Characterization Plan** clearly describes the testing objectives, conditions, and methods under which a system or sensor will be characterized. The purpose of the Field Characterization plan is to (1) describe the testing objectives, (2) illustrate that the methods and test conditions are relevant to the expected use case defined in Phase 1, and (3) demonstrate that testing activities will collect enough data to quantify the key performance metrics. An example Field Characterization Plan is outlined in Appendix E. The Field Characterization Plan should include:

- Description of testing objectives and what is trying to be achieved with the testing.
- Conditions for testing
 - Duration and location
 - Expected environmental conditions
 - Biological factors of target species.
- Testing methodology
 - Methods for operation of the technology and communication system
 - Mechanisms to simulate anthropogenic factors (such as increased noise from pile driving) that replicate the expected operational setting of the future use case
 - Whether tests will use real or artificial targets (e.g., playbacks of animal calls with realistic source levels or simulations of whale blows)
 - How the data will be processed
 - Calculation procedures that clearly indicate how each of the key performance metrics will be quantified based on field observations
 - Methods for supplementary data collection or analysis to evaluate false positive and false negative rates.

- Testing matrix that summarizes the conditions and duration under which the tests will be conducted. For example, if the technology provided proposes to operate in sea state 1 through 4, then the test matrix should describe the expected duration of data collection in each applicable sea state.
 - Mechanisms to simulate anthropogenic factors (such as increased noise from pile driving) that replicate the expected operational setting of the future use case.

The Field Characterization plan should be evaluated by a third party using the **Field Characterization Recommendations** as guidance for review. The purpose of the Field Characterization Recommendations is to provide a transparent outline that describes how testing plans will be evaluated. If the Field Characterization Plan is considered sufficient, the technology provider should proceed to Phase 3. If the Field Characterization Plan is not sufficient, then the technology provider should revise the plan and have it reviewed again.

4.3 Phase 3: Modeling and Field Characterization

The objective of Phase 3 is to quantify key performance metrics associated with the proposed monitoring system by gathering data in the field and then modeling performance of a multi-sensor or multi-modal system, as applicable (Figure 8).

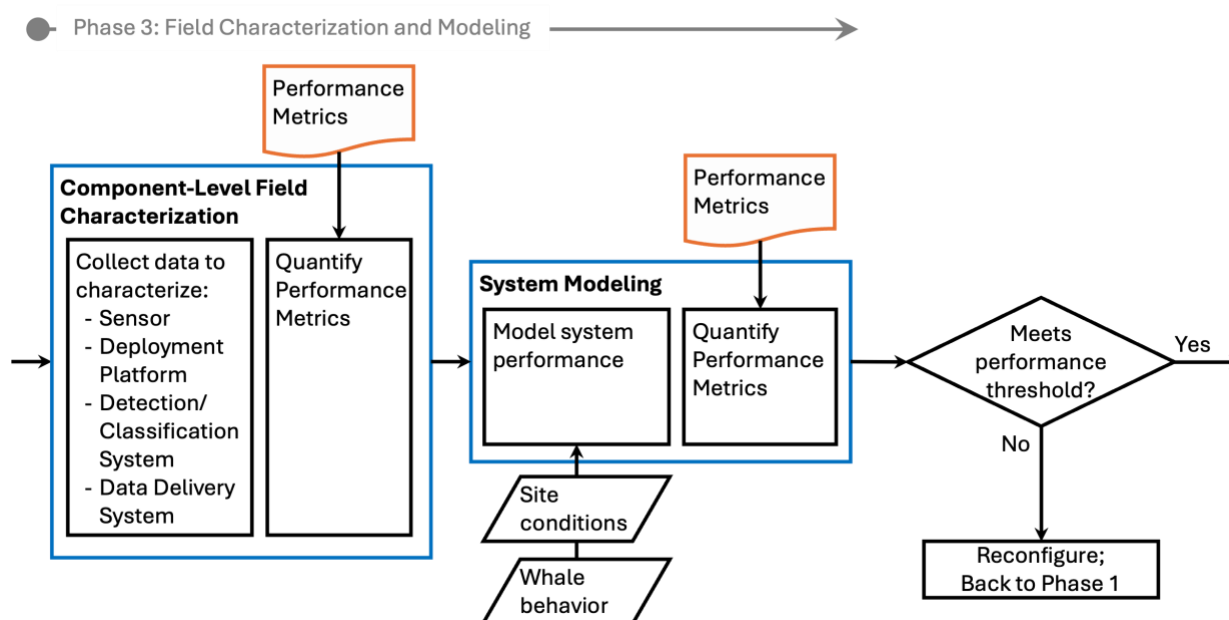


Figure 8. Flow diagram for Phase 3: Field Characterization and Modeling

During Phase 3, the **Component⁵-Level Field Characterization** is conducted to gather observations using the methods defined in the Field Characterization Plan. Data are collected in the Component-Level Field Characterization about the performance of the sensor, deployment platform, detection/classification system, and the data delivery system for a single component, or monitoring device. Observations from the Field Characterization will be used to quantify the

⁵ 'Component' refers to a single monitoring system that includes the sensor, deployment platform, detection and classification system, and data delivery system. Multiple components of the same monitoring technology or different monitoring technologies can be combined into a larger system. See Glossary for a definition.

performance metrics at a component-level, not at the full system level if multiple technologies are proposed. A summary of testing results and performance metrics can be provided in a format similar to the example shown in 0.

Component-Level Field Characterization can be carried out using naturally occurring baleen whale cues or simulated whale cues, provided that this approach was considered appropriate in the testing plan review. Characterizing naturally occurring baleen whale cues would involve operating the monitoring system in a representative environment and observing for baleen whales. Supplementary observations alongside the proposed sensor will be needed in order to evaluate false positive and false negatives and quantify precision and recall. Supplementary observations could come from PSOs, imaging systems, PAM, or other technologies capable of detecting baleen whales. If artificial baleen whale cues are used, for example if recordings of baleen whale calls are played underwater, supplementary monitoring systems may not be needed during the test because there will be accurate knowledge about when and where an artificial baleen whale cue is played. The use of artificial cues overcomes some challenges for quantifying precision and recall but bring about other considerations, including designing representative artificial cues, determining appropriate source levels for artificial baleen whale calls, approvals associated with using artificial cues, and others. In addition to precision and recall, the Component-Level Field Characterization will also observe system performance with respect to the other key performance metrics of reliable detection range, detection latency, and system reliability. Two detailed examples of real-world testing scenarios can be found in the Technology Profiles (Appendix B).

Field tests have limitations. For example, a technology cannot feasibly be tested under every possible environmental condition. The focus should be on conditions under which offshore wind foundation installation will occur. Animals may behave differently during foundation installation (e.g., due to pile driving and vessel noise), and in the case of PAM, source levels and frequencies of animal calls vary, which will affect detection probability. For IR camera systems, the ground truth does not typically involve the true number of animals, but a relative comparison to the performance of a human observer.

After data are collected through the Component-Level Field Characterization, **Performance Metrics** must be measured to document the observed performance of the component and associated subsystems in a standardized approach. To document the results of the field characterization, the technology provider should write a technical report that quantifies the performance metrics, describes the testing conditions, and provides raw and processed observational data that support the calculations.

After the Performance Metrics are calculated based on the Component-Level Field Characterization Tests, information about the sensor's performance can be used to predict the success of implementing mitigation measures in an operational setting. If the technology provider proposes using an array of multiple components or multiple technologies, characterizing the performance of the full system in the field may be untenable due to time, logistics, and permitting constraints. Instead, data from the field characterization should be used to describe the performance of each component in the proposed system as part of a simulated model (Figure 9).

In the **System Modeling** step, the performance of the system is evaluated by simulating realistic whale behavior in expected environmental/ambient conditions. System Modeling will allow a technology provider to investigate the performance of different system configurations by

varying the spacing, location, and number of components similar to models described by Baumgartner (2024), SMRU Consulting (2024), and Ceballos et al. (2023).

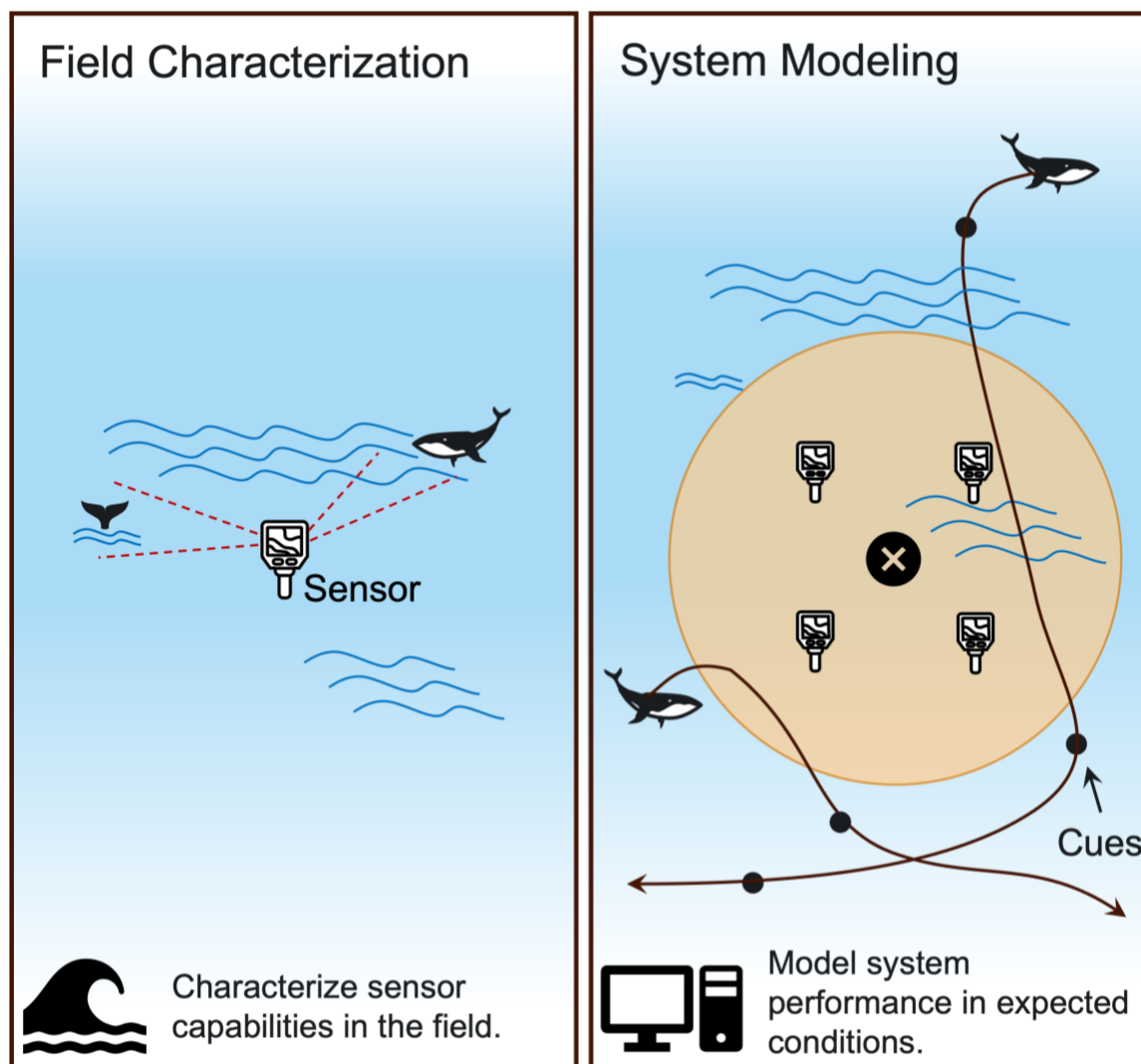


Figure 9. Example diagram of Field Characterization and System Modeling.

The purpose of System Modeling is to evaluate the ability of a multi-component or multi-technology system to use detections to inform appropriate mitigation decisions. The System Model simulates the array of sensors deployed in representative conditions alongside the movement and behavioral characteristics of the expected whale species. Inputs to the model include (1) characteristics of whale behavior including their movement patterns (i.e., swimming location) and relevant cueing patterns (e.g., call or breaching patterns) and (2) environmental conditions, such as wave spectra, ambient noise levels (including those right before and during pile driving), visibility, bathymetry, and other factors relevant to modeling the proposed sensor type. Characteristics of the sensors (e.g., response to cues at different ranges) will be taken directly from the Performance Metrics calculated as part of the Component-Level Field Characterization Plan and input into the System Model for evaluation.

The System Model will simulate baleen whale behavior around an offshore wind foundation installation activity with associated shutdown and clearance zones. The System Model will determine if the baleen whales are detected by the system, and if that detection would trigger a mitigation action. Outputs from the System Model would show how many baleen whales were detected within a mitigation zone (TP_{MZ}), how incorrect detections were made when a baleen whale was not present within a mitigation zone (FP_{MZ}), and how many baleen whales entered the mitigation zone without being detected (FN_{MZ}). These results are used to calculate the Probability of Missed Mitigation.

The placement of components in the array can be rearranged into different configurations as part of the System Modeling to find the best results. Results from the System Modeling should be documented within a technical report that describes the system configuration, the simulation results, and includes documentation about the modeling code and input data.

After calculating the performance metrics using results from Field Characterization and System Modeling, the overall performance of the system can be evaluated for the intended application. Performance metrics would be compared against a set of performance thresholds to determine if the system is suitable for operational use. The performance thresholds needed for operational use are not defined within this document because it is outside the scope of this effort to set performance targets.

5.0 Research and Development Recommendations

In order to advance the state of technologies and tools for monitoring baleen whales around offshore wind development activities, we recommend the following actions.

Technology Research Recommendations

- Implement a standardized approach for technology characterization

The performance of different technologies has not been evaluated using a consistent and standardized approach. Technology developers should adhere to a standardized approach to characterize the performance of their technology, such as the framework described in Section 4.0. Results collected during technology characterization tests should be shared publicly to allow knowledge transfer and assessment.

- Develop an instrumented field testing site for technology characterization

Conducting field testing is costly and requires facilities suited for this type of research and development, and in some cases, permits that may not be practical for individual technology developers. Establishing an instrumented, pre-permitted testing site will significantly reduce barriers for technology developers to characterize the performance of their technology and create consistency for comparison.

- Quantify standardized metrics that are applicable to operational settings

Studies that use baleen whale monitoring technologies do not consistently report metrics related to efficacy, detection distance, data delivery (latency), system reliability (down time) and variability. We recommend that technical reports, scientific studies, or other applications of baleen whale monitoring technologies quantify the performance of the technology using standardized metrics that describe the details and metrics described in Section 2.0.

- Develop a system modeling tool

Recognizing the limitations of field characterization testing, a system modeling tool should be developed that can evaluate the configuration and performance of a multi-sensor and multi-technology monitoring system in a simulated operational setting. This modeling tool should be able to simulate baleen whale behavior(s) and evaluate detections that would result in a mitigation action around a monitoring zone. The system modeling tool would be standardized for use across multiple technologies, publicly accessible, and open-source.

- Develop standardized input data libraries for a system modeling tool

A system modeling tool will use data about species-specific baleen whale behavior, environmental conditions, sound propagation and other factors. These input data should be standardized and packaged together into a publicly accessible, open-source library that can be easily used with a system modeling tool. Any information contained within the data libraries should be cited and based on peer-reviewed literature, when available.

- Increased focus on performance of multi-modal systems

Multi-modal systems provide a greater likelihood of observing a baleen whale because they use different cues for detection, classification, and localization. Report and studies of technology performance has largely focused on the application of a single modality or using two to make a relative comparison. Recognizing the benefit of multiple complementary technologies, future research should focus on evaluating different combinations, and applications, of multi-modal systems. In particular, these should quantify the performance

metrics of the multi-modal system as a whole, in addition to quantifying the performance of each component. This research should also focus on identifying and addressing challenges of multi-modal and multi-vendor systems such as time/data synchronization, data transfer, communication, and automated operation.

Technology Development Recommendations

- Improve technology performance and automation

Technology capabilities have been documented indicating their ability to detect baleen whales, but more advancement to improve performance, reliability, and automation is needed. This report does not intend to provide a comprehensive list of technology development needs; however, two examples may include:

- Automated detection and classification

Proprietary and custom software and algorithms for detection and classification of baleen whales have been developed for PAM and IR cameras; however, many tend to be developed for a specific research project (i.e., research goal, place, time, species, and/or cue). While the field is growing, machine learning techniques are still in their infancy both in terms of development, availability, and adoption by the larger community. This is an identified gap and should be a priority for research and development, standardization, and open-source platforms.

- Power and onboard data processing

Many autonomous deployment platforms (e.g., buoys or autonomous underwater vehicles), are not currently capable of complex real-time onboard processing needed for detection and classification at the time scales required for operational use. For example, PAM systems have been successfully used for real or near real-time data delivery; however, some only allow for summary information to be transmitted. Data delivery methods should allow for raw or processed data to be archived and transmitted to a human operator to confirm detections before making operational decisions. Further development to reduce latency but also increase data delivery would be recommended.

Technology Deployment Recommendations

- Use multi-modal system in operational setting

Multi-modal systems should allow for improved monitoring efficacy across the variety of offshore environmental conditions by reducing availability and perception bias. The use of multi-modal systems should be closely considered in most operational settings.

- Test new technologies or multi-modal systems opportunistically alongside operational monitoring campaigns

When monitoring is required in an operational setting, such as installing offshore wind foundations, when feasible, opportunities to evaluate performance of concurrent technologies in real-world conditions is recommended. These opportunities should be explored to evaluate the performance and use of new technologies or new multi-modal approaches. Implementation of any opportunistic testing should not interfere with the required monitoring but can provide supplementary data that can be used for evaluating

efficacy of application. Information collected during observations should be shared publicly through reports so lessons can be learned and shared.

- Apply lessons learned from other industries

Monitoring baleen whales around offshore wind development should not be limited to the offshore wind industry but can leverage other related offshore activities such as shipping, oil and gas, and defense industries. A dedicated effort should be initiated to engage with other industries to share information that may not be easily accessible in the public domain. Knowledge transfer should emphasize lessons learned about suitable deployment configurations, application of monitoring technologies in varying environmental conditions, and best practices for measuring performance in operational settings. In addition, sharing observational data from monitoring operations from other industries can inform the design, evaluation, and application of baleen whale monitoring systems to the offshore wind industry.

6.0 Conclusions and Next Steps

The public workshops and the literature review demonstrated that there is no “one size fits all” solution to real-time baleen whale monitoring in conjunction with offshore wind construction activities. The environment in which offshore wind development occurs varies substantially by location. Varying environmental, bathymetric, and oceanographic conditions influence species composition within each region, which can further vary seasonally and annually. While this document does not define required detection distance, performance thresholds, or readiness levels, the guidance presented here can be used to facilitate the process of evaluating technology performance by providing a standardized approach for selecting, designing, testing, and reporting on the performance of a technology or suite of technologies.

Although the scope of this report is specific to baleen whale monitoring technologies before and during fixed-foundation installation, the monitoring technologies we describe, as well as our proposed evaluation framework, could be applicable to other marine mammal species, with the caveat that other species have behaviors (e.g., surfacing and calling) that differ from baleen whales and may require a modified approach. PAM and IR camera systems can be used to monitor toothed whales (odontocetes), pinnipeds (seals, sea lions, and walrus), and sirenians (manatees and dugongs), in addition to baleen whales. Our findings could also be applicable to other aspects of offshore wind development. For example, instead of fixed foundations, offshore wind farm construction in deeper waters (60–1,000+ m) would involve installation of floating platforms and anchoring systems. Although the need for pile driving during construction of floating offshore wind farms would be eliminated or greatly reduced, real-time monitoring, in a variety of lighting conditions, would still be required during anchor emplacement and other construction activities potentially impactful to marine mammals. Our findings could also be applicable to marine mammal monitoring during offshore wind farm decommissioning, as the removal of in-water structures may result in underwater noise and other stressors similar to those produced during construction. Our approach provides an example framework that can be used to characterize the performance of any chosen monitoring technology, using numerical modeling and/or field characterization, as described above for PAM and IR camera systems. This includes simple guidelines for minimum suggested design specifications, field characterization recommendations, and potential performance metrics. Our approach also includes suggestions for data collection, data storage, data analysis, and data reporting standards.

We emphasize that no single monitoring technology exists that can detect every whale present in a given area. For example, PAM cannot detect a whale if it does not call, and some whale species have highly variable call rates. Calling animals may be masked by the noise of pile driving or may reduce their calling rates in response to the increased noise levels. IR cameras cannot detect animals that spend little time, or exhibit a low profile, at the surface. IR cameras are limited operationally based on environmental conditions and may not be able to identify individual whales at the species level.

Although unaided visual observers are a crucial component of monitoring and mitigation, they are unable to identify marine mammals at night or under low visibility conditions. Both PAM and IR also have their constraints, thus a multi-modal approach can improve overall detection with each technology compensating for limitations of the other. While this report does not consider the probability of detection for any combination of technologies, a multi-modal system will increase the overall detection probability.

Several monitoring tools did not meet our inclusion criteria (Table 5, Appendix A), including telemetry and satellite imagery. However, these are valuable tools that could be used to aid in the characterization of the monitoring technology systems and to understand the long-term presence and distribution of marine mammal species. For example, while satellite imagery can identify whales, the approach is opportunistic and can result in gaps in coverage and lags in data delivery. However, this technology could be developed to allow a more frequent delivery of images if satellites have a higher revisit rate over the monitoring area. These tools will be helpful in understanding the broader distribution of whales in time and space. Thus, we place an emphasis on longer-term research and development of not only the technologies we assessed in this report, but additional tools to develop a suite of monitoring systems. These tools, in combination with PAM and IR imaging, can be used to detect the time and location of any marine mammal within a predetermined zone of perception during offshore wind construction activities, specifically during foundation installation activities under low- and no-light conditions, and the information be delivered within a practical decision-making time window.

7.0 References

Abileah, R. 2002. "Marine mammal census using space satellite imagery." *US Navy Journal of Underwater Acoustics*, 52(3), 709–724.

Allen, A. N., M. Harvey, L. Harrell, A. Jansen, K. P. Merkens, C. C. Wall, J. Cattiau, and E. M. Oleson. 2021. "A Convolutional Neural Network for Automated Detection of Humpback Whale Song in a Diverse, Long-Term Passive Acoustic Dataset." *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.607321>.

Andrews, R. D., R. W. Baird, J. Calambokidis, C. E. C. Goertz, F. M. D. Gulland, M. P. Heide-Jørgensen, S. K. Hooker, M. Johnson, B. Mate, Y. Mitani, D. P. Nowacek, K. Owen, L. T. Quackenbush, S. Paverty, J. Robbins, G. S. Schorr, O. V. Shpak, F. I. Townsend Jr., M. Uhart, R. S. Wells, and A. N. Zerbini. 2019. "Best Practices Guidelines for Cetacean Tagging." *Journal of Cetacean Research and Management* 20: 27–66. <https://doi.org/10.47536/jcrm.v20i1.237>.

Andriolo, A., F. R. de Castro, T. Amorim, G. Miranda, J. Di Tullio, J. Moron, B. Ribeiro, G. Ramos, and R. R. Mendes. 2018. "Marine Mammal Bioacoustics Using Towed Array Systems in the Western South Atlantic Ocean." In *Advances in Marine Vertebrate Research in Latin America: Technological Innovation and Conservation*, edited by M. R. Rossi-Santos and C. W. Finkl. 113–147. Springer International Publishing. https://doi.org/10.1007/978-3-319-56985-7_5.

Andruszkiewicz, E. A., K. M. Yamahara, C. J. Closek, and A. B. Boehm. 2020. "Quantitative PCR Assays to Detect Whales, Rockfish, and Uncommon Murre Environmental DNA in Marine Water Samples of the Northeastern Pacific." *PLOS ONE* 15 (12): e0242689. <https://doi.org/10.1371/journal.pone.0242689>.

Andruszkiewicz, E. A., H. A. Starks, F. P. Chavez, L. M., Sassoubre, B. A. Block, and A. B. Boehm. 2017. "Biomonitoring of Marine Vertebrates in Monterey Bay using eDNA Metabarcoding." *PLOS ONE*, 12 (4): e0176343. <https://doi.org/10.1371/journal.pone.0176343>.

Araújo, V. M., A. Shukla, C. Chion, S. Gambs, and R. Michaud. 2022. "Machine-Learning Approach for Automatic Detection of Wild Beluga Whales from Hand-Held Camera Pictures." *Sensors* 22 (11): 4107. <https://doi.org/10.3390/s22114107>.

Aschettino, J. M., D. T. Engelhaupt, A. G. Engelhaupt, A. DiMatteo, T. Pusser, M. F. Richlen, and J. T. Bell. 2020. "Satellite Telemetry Reveals Spatial Overlap Between Vessel High-Traffic Areas and Humpback Whales (*Megaptera novaeangliae*) Near the Mouth of the Chesapeake Bay." *Frontiers in Marine Science* 7. <https://doi.org/10.3389/fmars.2020.00121>.

Au, W. W. L. 1996. "Acoustic Reflectivity of a Dolphin." *The Journal of the Acoustical Society of America* 99 (6): 3844–3848. <https://doi.org/10.1121/1.415002>.

Baird, R. W., A. E. Harnish, R. D. Andrews, J. K. Lerma, M. A. Mohler, J. E. Phipps, and S. D. Mahaffy. 2024. *Small Boat Surveys and Satellite Tagging of Cetaceans on the Pacific Missile Range Facility, Kaua'i, in February 2024*. Field survey report to U.S. Pacific Fleet by HDR, under Federal contract number N62470-20-D0016, Task Order No. N62742-24-F0102. [Baird et al 2024 Kauai.pdf](#).

Baldacci, A., M. J. Carron, and N. Portunato. 2005. *Infrared Detection of Marine Mammals*. NATO Undersea Research Centre SR-443. La Spezia, Italy.

<https://www.cmre.nato.int/research/publications/latest-techreports/426-infrared-detection-of-marine-mammals/file>

Bamford, C. C. G., N. Kelly, L. Dalla Rosa, D. E. Cade, P. T. Fretwell, P. N. Trathan, H. C. Cubaynes, A. F. C. Mesquita, L. Gerrish, A. S. Friedlaender, and J. A. Jackson. 2020. "A Comparison of Baleen Whale Density Estimates Derived from Overlapping Satellite Imagery and a Shipborne Survey." *Scientific Reports* 10 (1): 12985. <https://doi.org/10.1038/s41598-020-69887-y>.

Barlow, J., and R. Gisiner. 2006. "Mitigating, Monitoring, and Assessing the Effects of Anthropogenic Sound on Beaked Whales." *Journal of Cetacean Research and Management* 7 (3): 239–249. <https://doi.org/10.47536/jcrm.v7i3.734>.

Barlow, J., and B. L. Taylor. 2005. "Estimates of Sperm Whale Abundance in the Northeastern Temperate Pacific from a Combined Acoustic and Visual Survey." *Marine Mammal Science* 21 (3): 429–445. <https://doi.org/10.1111/j.1748-7692.2005.tb01242.x>.

Baumgartner, M. 2024. Evaluating the efficacy of real-time passive acoustic monitoring near offshore wind energy development activities to help mitigate risks to North Atlantic right whales. 2024 State of the Science Workshop: Taking an Exosystem Approach: Integrated Offshore Wind, Wildlife, and Fisheries (Conference Presentation). Long Island, NY. 17 July 2024. https://youtu.be/_t3Od7djeMo?si=Ax6pk9g22UW5ydQo&t=1275

Baumgartner, M. F., K. Ball, J. Partan, L.-P. Pelletier, J. Bonnell, C. Hotchkin, P. J. Corkeron, and S. M. Van Parijs. 2021. "Near Real-Time Detection of Low Frequency Baleen Whale Calls from an Autonomous Surface Vehicle: Implementation, Evaluation, and Remaining Challenges." *The Journal of the Acoustical Society of America* 149 (5): 2950–2962. <https://doi.org/10.1121/10.0004817>.

Baumgartner, M. F., J. Bonnell, S. M. Van Parijs, P. J. Corkeron, C. Hotchkin, K. Ball, L.-P. Pelletier, J. Partan, D. Peters, J. Kemp, J. Pietro, K. Newhall, A. Stokes, T. V. N. Cole, E. Quintana, and S. D. Kraus. 2019. "Persistent Near Real-Time Passive Acoustic Monitoring for Baleen Whales from a Moored Buoy: System Description and Evaluation." *Methods in Ecology and Evolution* 10 (9): 1476–1489. <https://doi.org/10.1111/2041-210X.13244>.

Baumgartner, M. F., and D. M. Fratantoni. 2008. "Diel Periodicity in both Sei Whale Vocalization Rates and the Vertical Migration of their Copepod Prey Observed from Ocean Gliders." *Limnology and Oceanography* 53 (52): 2197–2209. https://doi.org/https://doi.org/10.4319/lo.2008.53.5_part_2.2197.

Baumgartner, M. F., D. M. Fratantoni, T. P. Hurst, M. W. Brown, T. V. N. Cole, S. M. Van Parijs, and M. Johnson. 2013. "Real-Time Reporting of Baleen Whale Passive Acoustic Detections from Ocean Gliders." *The Journal of the Acoustical Society of America* 134 (3): 1814–1823. <https://doi.org/10.1121/1.4816406>.

Baumgartner, M. F., and S. E. Mussoline. 2011. "A Generalized Baleen Whale Call Detection and Classification System." *The Journal of the Acoustical Society of America* 129 (5): 2889–2902. <https://doi.org/10.1121/1.3562166>.

- Baumgartner, M. F., K. M. Stafford, P. Winsor, H. Statscewich, and D. M. Fratantoni. 2024. "Glider-Based Passive Acoustic Monitoring in the Arctic." *Marine Technology Society Journal* 48 (5): 40–51. <https://doi.org/10.4031/MTSJ.48.5.2>.
- Beier, K., and H. Gemperlein. 2004. "Simulation of Infrared Detection Range at Fog Conditions for Enhanced Vision Systems in Civil Aviation." *Aerospace Science and Technology* 8 (1): 63–71. <https://doi.org/10.1016/j.ast.2003.09.002>.
- Bergler, C., H. Schröter, R. X. Cheng, V. Barth, M. Weber, E. Nöth, H. Hofer, and A. Maier. 2019. "ORCA-SPOT: An Automated Killer Whale Sound Detection Toolkit Using Deep Learning." *Scientific Reports* 9 (1): 10997. <https://doi.org/10.1038/s41598-019-47335-w>.
- Bernasconi, M., R. Patel, L. Nøttestad, and A. S. Brierley. 2013. Fin Whale (*Balaenoptera physalus*) Target Strength Measurements. *Marine Mammal Science* 29 (3): 371–388. <https://doi.org/10.1111/mms.12032>.
- Boebel, O., and D. P. Zitterbart. 2013. "24/7 Automatic Detection of Whales Near Seismic Vessels Using Thermography." 75th European Association of Geoscientists and Engineers (EAGE) Conference and Exhibition, London, UK, June 10–13, 2013. <https://doi.org/10.3997/2214-4609.20131189>.
- Bonnel, J., A. M. Thode, S. B. Blackwell, K. Kim, and A. M. Macrander. 2014. "Range Estimation of a Bowhead Whale (*Balaena mysticetus*) Calls in the Arctic Using a Single Hydrophone." *The Journal of the Acoustical Society of America* 136 (1): 145–155. <https://doi.org/10.1121/1.4883358>.
- Bonnel, J., A. M. Thode, D. Wright, and R. Chapman. 2020. "Nonlinear Time-Warping Made Simple: A Step-by-Step Tutorial on Underwater Acoustical Modal Separation with a Single Hydrophone." *The Journal of the Acoustical Society of America* 147 (3): 1897–1926. <https://doi.org/10.1121/10.0000937>.
- Booth, C. G., C. S. Oedekoven, D. Gillespie, J. Macaulay, R. Plunkett, R. Joy, D. Harris, J. Wood, T. A. Marques, L. Marshall, U. K. Verfuss, P. Tyack, M. Johnson, and L. Thomas. 2017. *Assessing the Viability of Density Estimation for Cetaceans from Passive Acoustic Fixed Sensors Throughout the Life Cycle of an Offshore E&P Field Development*. IOGP Sound and Marine Life Joint Industry Programme SMRUC-OGP-2017-001. SMRU Consulting. Friday Harbor, WA. https://gisserver.intertek.com/JIP/DMS/ProjectReports/Cat4/Other/Booth2017_CetaceansandPAM.pdf.
- Bouchard, B., J.-Y. Barnagaud, M. Poupard, H. Glotin, P. Gauffier, S. T. Ortiz, T. J. Lisney, S. Campagna, M. Rasmussen, and A. Célrier. 2019. "Behavioral Responses of Humpback Whales to Food-Related Chemical Stimuli." *PLOS ONE* 14 (2): e0212515. <https://doi.org/10.1371/journal.pone.0212515>.
- Bazhyna, A. 2009. "Image Compression in Digital Cameras." PhD thesis. Department of Signal Processing, Tampere University of Technology, Tampere, Finland. <https://trepo.tuni.fi/bitstream/handle/10024/113958/bazhyna.pdf?sequence=1&isAllowed=y>.

- Cauchy, P. K. J. Heywood, N. D. Merchant, D. Risch, B. Y. Queste, and P. Testor. 2023. "Gliders for Passive Acoustic Monitoring of the Oceanic Environment." *Frontiers in Remote Sensing* 4. <https://doi.org/10.3389/frsen.2023.1106533>.
- Ceballos, V., C. Taggart, and H. Johnson. 2022. "Comparison of Visual and Acoustic Surveys for the Detection and Dynamic Management of North Atlantic Right Whales (*Eubalaena glacialis*) in Canada." *Conservation Science and Practice* 5. <https://doi.org/10.1111/csp2.12866>.
- Cheeseman, T., K. Southerland, J. Park, M. Olio, K. Flynn, J. Calambokidis, L. Jones, C. Garrique, A. F. Jordán, A. Howard, W. Reade, J. Neilson, C. Gabriele, and P. Clapham. 2022. "Advanced Image Recognition: A Fully Automated, High-Accuracy Photo-Identification Matching System for Humpback Whales." *Mammalian Biology* 102 (3): 915–929. <https://doi.org/10.1007/s42991-021-00180-9>.
- Cheeseman, T., K. Southerland, J. M., Acebes, K. Audley, J. Barlow, C. Birdall, A. L. Bradford, J. K. Byington, J. Calambokidis, R. Cartwright, J. Cedarleaf, A. J. G. Chavez, J. J. Curie, J. De Weerd, N. Doe, T. Doniol-Valcroze, K. Dracott, O. Filatova, ...P. Clapham. 2023. "A Collaborative and Near-Comprehensive North Pacific Humpback Whale Photo-ID Dataset." *Scientific Reports* 13 (1): 10237. <https://doi.org/10.1038/s41598-023-36928-1>.
- Clark, C. W. 1990. "Acoustic Behavior of Mysticete Whales." In *Sensory Ability of Cetaceans: Laboratory and Field Evidence*, edited by J. A. Thomas and R. A. Kastelein, 571–583. U.S.: Springer. https://doi.org/10.1007/978-1-4899-0858-2_40.
- Clark, C. W., M. W. Brown, and P. Corkeron. 2010. "Visual and Acoustic Surveys for North Atlantic Right Whales, *Eubalaena glacialis*, in Cape Cod Bay, Massachusetts, 2001–2005: Management Implications." *Marine Mammal Science* 26 (4): 837–854. <https://doi.org/https://doi.org/10.1111/j.1748-7692.2010.00376.x>.
- Closek, C. J., J. A. Santora, H. A. Starks, I. D. Schroeder, E. A. Andruszkiewicz, K. M. Sakuma, S. J. Bograd, E. L. Hazen, J. C. Field, and A. B. Boehm. 2019. "Marine Vertebrate Biodiversity and Distribution Within the Central California Current Using Environment DNA (eDNA) Metabarcoding and Ecosystem Surveys." *Frontiers in Marine Science* 6. <https://doi.org/10.3389/fmars.2019.00732>.
- Corrêa, A. A., J. H. Quoos, A. S. Barreto, K. R. Groch, and P. P. B. Eichler. 2022. "Use of Satellite Imagery to Identify Southern Right Whales (*Eubalaena australis*) on a Southwest Atlantic Ocean Breeding Ground." *Marine Mammal Science* 38 (1): 87–101. <https://doi.org/10.1111/mms.12847>.
- Cubaynes, H. C., P. T. Fretwell, C. Bamford, L. Gerrish, and J. A. Jackson. 2019. "Whales From Space: Four Mysticete Species Described Using New VHR Satellite Imagery." *Marine Mammal Science* 35 (2): 466–491. <https://doi.org/10.1111/mms.12544>.
- Cuyler, L. C., R. Wiulsrød, and N. A. ØRitsland. 1992. "Thermal Infrared Radiation from Free Living Whales." *Marine Mammal Science* 8 (2): 120–134. <https://doi.org/10.1111/j.1748-7692.1992.tb00371.x>.
- de Castro, F. R., D. V. Harris, S. J. Buchan, N. Balcazar, and B. S. Miller. 2024. "Beyond Counting Calls: Estimating Detection Probability for Antarctic Blue Whales Reveals Biological

Trends in Seasonal Calling.” *Frontiers in Marine Science* 11.
<https://doi.org/10.3389/fmars.2024.1406678>.

Dekeling, R. P. A., M. L. Tasker, A. J. Van der Graaf, M. A. Ainslie, M. H. Andersson, M. André, J. F. Borsani, K. Brensing, M. Castellote, D. Cronin, J. Dalen, T. Folegot, R. Leaper, J. Pajala, P. Redman, S. P. Robinson, P. Sigray, G. Sutton, F. Thomsen, S. Werner, D. Wittekind, and J. V. Young. 2014. *Monitoring Guidance for Underwater Noise in European Seas, Part II: Monitoring Guidance Specifications*. Joint Research Centre (JRC): Institute for Environment and Sustainability Scientific and Policy Report EUR 26555 EN, Publications Office of the European Union. Luxembourg. <https://doi.org/10.2788/27158>.

Dooling, R. J. 2019. “The Impact of Urban and Traffic Noise on Birds.” *Acoustics Today* 15 (3): 19. <https://doi.org/10.1121/AT.2019.15.3.19>.

Dooling, R. J. and M. R. Leek. 2018. “Communication Masking by Man-Made Noise.” In *Effects of Anthropogenic Noise on Animals, Volume 66*, edited by H. Slabbekoorn, R. J. Dooling, A. N. Popper, and R. R. Fay. New York: Springer. https://doi.org/10.1007/978-1-4939-8574-6_2.

Dove, A. D. M. 2015. “Foraging and Ingestive Behaviors of Whale Sharks, *Rhinocodon typus*, in Response to Chemical Stimulus Cues.” *The Biological Bulletin* 228 (1): 65–72.
<https://doi.org/10.1086/BBLv228n1p65>.

Dunn, J. L. 1969. “Airborne Measurements of the Acoustic Characteristics of a Sperm Whale.” *The Journal of the Acoustical Society of America* 46 (4B): 1052–1054.
<https://doi.org/10.1121/1.1911803>.

Durden, J. M., T. Schoening, F. Althaus, A. Friedman, R. Garcia, A. G. Glover, J. Greinert, N. J. Stout, D. O. B. Jones, A. Jordt, J. W. Kaeli, K. Koser, L. A. Kuhn, D. Lindsay, K. J. Morris, T. W. Nattkemper, J. Osterloff, H. A. Ruhl, H. Singh, M. Trann, and B. J. Bett. 2016. “Perspectives in Visual Imaging for Marine Biology and Ecology: From Acquisition to Understanding.” In *Oceanography and Marine Biology: An Annual Review, Vol. 54*, edited by R. N. Hughes, D. J. Hughes, E. P. Smith, and A. C. Dale. 1–72. Florida: CRC Press.

Formel, N., I. C. Enochs, C. Sinigalliano, S. R. Anderson, and L. R. Thompson. 2021. “Subsurface Automated Samplers for eDNA (SAsE) for Biological Monitoring and Research.” *HardwareX* 10: e00239. <https://doi.org/10.1016/j.ohx.2021.e00239>.

Fretwell, P. T., I. J. Staniland, and J. Forcada. 2014. “Whales from Space: Counting Southern Right Whales by Satellite.” *PLOS ONE* 9 (2): e88655.
<https://doi.org/10.1371/journal.pone.0088655>.

Galí, M., M. Levasseur, E. Devred, R. Simó, and M. Babin. (2018). “Sea-Surface Dimethylsulfide (DMS) Concentration from Satellite Data at Global and Regional Scales.” *Biogeosciences* 15 (11): 3497–3519. <https://doi.org/10.5194/bg-15-3497-2018>.

Geoggy, M., S. Rousseau, F. R. Knudsen, and L. Fortier. 2016. “Target Strengths and Echotraces of Whales and Seals in the Canadian Beaufort Sea.” *ICES Journal of Marine Science* 73 (2): 451–463. <https://doi.org/10.1093/icesjms/fsv182>.

Gervaise, C., Y. Simard, F. Aulanier, and N. Roy. 2021. “Optimizing Passive Acoustic Systems for Marine Mammal Detection and Localization: Application to Real-Time Monitoring North

Atlantic Right Whales in Gulf of St. Lawrence.” *Applied Acoustics* 178 (107949).
<https://doi.org/10.1016/j.apacoust.2021.107949>.

Gillespie, D., M. Caillat, J. Gordon, and P. White. 2013. “Automatic Detection and Classification of Odontocete Whistles.” *The Journal of the Acoustical Society of America* 134 (3): 2427–2437.
<https://doi.org/10.1121/1.4816555>.

Gillespie, D., G. Hastie, J. Montabaranom, E. Longden, K. Rapson, A. Holoborodko, and C. Sparling. 2023. “Automated Detection and Tracking of Marine Mammals in the Vicinity of Tidal Turbines Using Multibeam Sonar.” *Journal of Marine Science and Engineering* 11 (11): 2095.
<https://doi.org/10.3390/jmse11112095>.

Gillespie, D., D. K. Mellinger, J. Gordon, D. McLaren, P. Redmond, R. McHugh, P. Trinder, X. Deng, and A. Thode. 2009. “PAMGUARD: Semiautomated, Open-Source Software for Real-Time Acoustic Detection and Localization of Cetaceans.” *The Journal of the Acoustical Society of America* 125 (4_Supplement): 2547. <https://doi.org/10.1121/1.4808713>.

Gillespie, D., M. Oswald, G. Hastie, and C. Sparling. 2022. “Marine Mammal HiCUP: A High Current Underwater Platform for the Long-Term Monitoring of Fine-Scale Marine Mammal Behavior Around Tidal Turbines.” *Frontiers in Marine Science* 9.
<https://doi.org/10.3389/fmars.2022.850446>.

Gillespie, D., L. Palmer, J. Macaulay, C. Sparling, and G. Hastie. 2020. “Passive Acoustic Methods for Tracking the 3D Movements of Small Cetaceans Around Marine Structures.” *PLOS ONE* 15 (5): e0229058. <https://doi.org/10.1371/journal.pone.0229058>.

Gold, Z., J. Sprague, D. J. Kushner, M. E. Zerecero, and P. H. Barber. 2021. “eDNA Metabarcodings as a Biomonitoring Tool for Marine Protected Areas.” *PLOS ONE* 16 (2): e0238557. <https://doi.org/10.1371/journal.pone.0238557>.

Goldbogen, J. A., B. L. Southall, S. L. DeRuiter, J. Calambokidis, A. S. Friedlaender, E. L. Hazen, E. A. Falcone, G. S. Schorr, A. Douglas, D. J. Moretti, C. Kyburg, M. F. McKenna, and P. L. Tyack. 2013. “Blue Whales Respond to Simulated Mid-Frequency Military Sonar.” *Proceedings of the Royal Society B* 280 (1765). <https://doi.org/doi:10.1098/rspb.2013.0657>.

Graber, J., J. Thomson, B. Polagye, and A. Jessup. 2011. “Land-Based Infrared Imagery for Marine Mammal Detection.” In *Proceedings of SPIE, Remote Sensing and Modeling of Ecosystems for Sustainability VIII*. SPIE. <https://doi.org/10.1117/12.892787>.

Griffiths, E. T., and J. Barlow. 2015. *Equipment Performance Report for the Drifting Acoustic Spar Buoy Recorder (DASBR)*. NOAA Southwest Fisheries Science Center. La Jolla, CA.
<http://doi.org/10.7289/V5/TM-SWFSC-543>.

Guazzo, R. A., T. A. Helble, G. L. D’Spain, D. W. Weller, S. M. Wiggins, and J. A. Hildebrand. 2017. “Migratory Behavior of Eastern North Pacific Gray Whales Tracked Using a Hydrophone Array.” *PLOS ONE* 12 (10): e0185585. <https://doi.org/10.1371/journal.pone.0185585>.

Guazzo, R. A., D. W. Weller, H. M. Europe, J. W. Durban, G. L. D’spain, and J. A. Hildebrand. 2019. “Migrating Eastern North Pacific Gray Whale Call and Blow Rates Estimated from Acoustic Recordings, Infrared Camera Video, and Visual Sightings.” *Scientific Reports* 9 (1): 12617. <https://doi.org/10.1038/s41598-019-49115-y>.

Guirado, E., S. Tabik, M. L. Rivas, D. Alcaraz-Segura, and F. Herrera. 2019. "Whale Counting in Satellite and Aerial Images with Deep Learning." *Scientific Reports* 9 (1): 14259. <https://doi.org/10.1038/s41598-019-50795-9>.

Gulland, F. M. D., J. D. Baker, M. Howe, E. LaBrecque, L. Leach, S. E. Moore, R. R. Reeves, and P. O. Thomas. 2022. "A Review of Climate Change Effects on Marine Mammals in United States Waters: Past Predictions, Observed Impacts, Current Research, and Conservation Imperatives." *Climate Change Ecology* 3. <https://doi.org/10.1016/j.ecochg.2022.100054>.

Hart, K. M., and K. D. Hyrenbach. 2009. "Satellite Telemetry of Marine Megavertebrates: The Coming of Age of an Experimental Science." *Endangered Species Research* 10: 9–20. <https://doi.org/10.3354/esr00238>.

Hastie, G. D., M. Bivins, A. Coram, J. Gordon, P. Jepp, J. MacAulay, C. Sparling, and D. Gillespie. 2019a. "Three-Dimensional Movements of Harbour Seals in a Tidally Energetic Channel: Application of Novel Sonar Tracking System." *Aquatic Conservation: Marine and Freshwater Ecosystems* 29 (4): 564–575. <https://doi.org/https://doi.org/10.1002/aqc.3017>.

Hastie, G. D., G.-M. Wu, S. Moss, P. Jepp, J. MacAulay, A. Lee, C. E. Sparling, C. Evers, and D. Gillespie. 2019b. "Automated Detection and Tracking of Marine Mammals: A Novel Sonar Tool for Monitoring Effects of Marine Industry." *Aquatic Conservation: Marine and Freshwater Ecosystems* 29 (S1): 119–130. <https://doi.org/https://doi.org/10.1002/aqc.3103>.

Helble, T. A., G. R. Ierley, G. L. D'Spain, M. A. Roch, and J. A. Hildebrand. 2012. "A Generalized Power-Law Detection Algorithm for Humpback Whale Vocalizations." *The Journal of the Acoustical Society of America* 131 (4): 2682–2699. <https://doi.org/10.1121/1.3685790>.

Helble, T. A., G. R. Ierley, G. L. D'Spain, and S. W. Martin. 2015. "Automated Acoustic Localization and Call Association for Vocalizing Humpback Whales on the Navy's Pacific Missile Range Facility." *Journal of the Acoustical Society of America* 137 (1): 11–21. <https://doi.org/10.1121/1.4904505>.

Henderson, E. E., M. A. Kratofil, R. W. Baird, C. R. Martin, A. E. Harnish, S. W. Martin, and B. L. Southall. 2024. *Assessing Exposure and Responses of Satellite-Tagged Blainville's Beaked Whales on the Pacific Missile Range Facility, Hawai'i*. U.S. Pacific Fleet Commander, NIWC Pacific. Pearl Harbor, HI. [Microsoft Word - Henderson et al 2024 Assessing Exposure Response Tagged Blainvilles on PMRF.docx](#).

Hildebrand, J. A., K. E. Frasier, T. A. Helble, and M. A. Roch. 2022. "Performance Metrics for Marine Mammal Signal Detection and Classification." *The Journal of the Acoustical Society of American* 151 (1): 414–427. <https://doi.org/10.1121/10.0009270>.

Hendricks, A., C. M., Mackie, E. Luy, C. Sonnichsen, J. Smith, I Grundke, M. Tavaoli, A. Furlong, R. G. Beiko, J. LaRoche, and V. Sieben. 2023. "Compact and Automated eDNA Sampler for In Situ Monitoring of Marine Environments." *Scientific Reports* 13 (1): 5210. <https://doi.org/10.1038/s41598-023-32310-3>.

Hodul, M., A. Knudby, B. McKenna, A. James, C. Mayo, M. Brown, D. Durette-Morin, and S. Bird. 2022. "Individual North Atlantic Right Whales Identified from Space." *Marine Mammal Science* 39 (1): 220–231. <https://doi.org/10.1111/mms.12971>.

- Horton, T. W., N. Hauser, S. Cassel, K. F. Klaus, T. Fettermann, and N. Key. 2019. "Doctor Drone: Non-Invasive Measurement of Humpback Whale Vital Signs Using Unoccupied Aerial System Infrared Thermography." *Frontiers in Marine Science* 6. <https://doi.org/10.3389/fmars.2019.00466>.
- Horton, T. W., A. Oline, N. Hauser, T. M. Khan, A. Laute, A. Stoller, K. Tison, and P. Zawar-Reza. 2017. "Thermal Imaging and Biometrical Thermography of Humpback Whales." *Frontiers in Marine Science* 4. <https://doi.org/10.3389/fmars.2017.00424>.
- Hou, F., Y. Zhang, Y. Zhou, M. Zhang, B. Lv, and J. Wu. 2022. "Review on Infrared Imaging Technology." *Sustainability* 14 (18): 11161. <https://doi.org/10.3390/su141811161>.
- Huang, H. C., J. Joseph, M. J. Huang, and T. Margolina. 2016. "Automated Detection and Identification of Blue and Fin Whale Foraging Calls by Combining Pattern Recognition and Machine Learning Techniques." IEEE OCEANS 2016 MTS/IEEE, Monterey, CA, September 19–23, 2016. <https://doi.org/10.1109/OCEANS.2016.7761269>.
- Irvine, L. M. Winsor, T. Follett, B. Mate, and D. Palacios. 2020. "An At-Sea Assessment of Argos Location Accuracy for Three Species of Large Whales, and the Effect of Deep-Diving Behavior on Location Error." *Animal Biotelemetry* 8. <https://doi.org/10.1186/s40317-020-00207-X>.
- Kaarna, A. 2007. "Compression of Spectral Images." In *Vision Systems: Segmentation and Pattern Recognition*, edited by G. Obinata and A. Dutta, 269–298. Austria: I-Tech Education and Publishing. <https://doi.org/10.5772/4964>.
- Keenan, G., C. Sparling, H. Williams, and F. Fortune. 2011. *SeaGen Environmental Monitoring Programme: Final Report*, Royal Haskoning, Marine Current Turbines (MCT). <https://tethys.pnnl.gov/publications/seagen-environmental-monitoring-programme-final-report>.
- Khan, C. B., K. T. Goetz, H. C. Cubaynes, C. Robinson, E. Murnane, T. Aldrich, M. Sackett, P. J. Clarke, M. A. LaRue, T. White, K. Leonard, A. Ortiz, and J. M. L. Ferres. 2023. "A Biologist's Guide to the Galaxy: Leveraging Artificial Intelligence and Very High-Resolution Satellite Imagery to Monitor Marine Mammals from Space." *Journal of Marine Science and Engineering* 11 (3): 595. <https://doi.org/10.3390/jmse11030595>.
- Klinck, H. and D. K. Mellinger. 2011. "The Energy Ratio Mapping Algorithm. A Tool to Improve the Energy-Based Detection of Odontocete Echolocation Clicks." *The Journal of the Acoustical Society of America* 129 (4): 1807–1812. <https://doi.org/10.1121/1.3531924>.
- Knudsen, F. R., O. B. Gammelsæter, P. H. Kvadsheim, and L. Nøttestad. 2008. "Evaluation of Fisheries Sonar for Whale Detection in Relation to Seismic Survey Operations." *Bioacoustics* 17(1–3): 326–328. <https://doi.org/10.1080/09524622.2008.9753864>.
- Kowarski, K. A., B. J. Gaudet, A. J. Cole, E. E. Maxner, S. P. Turner, S. B. Martin, H. D. Johnson, and J. E. Moloney. 2020. "Near Real-Time Marine Mammal Monitoring From Gliders: Practical Challenges, System Development, and Management Implications." *The Journal of the Acoustical Society of America* 148 (3): 1215. <https://doi.org/10.1121/10.0001811>.
- Küsel, E. T., D. K. Mellinger, L. Thomas, T. A. Marques, D. Moretti, and J. Ward. 2011. "Cetacean Population Density Estimation from Single Fixed Sensors Using Passive Acoustics."

The Journal of the Acoustical Society of America 129 (6): 3610–3622.
<https://doi.org/10.1121/1.3583504>.

Levenson, C. 1974. “Source Level and Bistatic Target Strength of the Sperm Whale (*Physeter catodon*) Measured from an Oceanographic Aircraft.” *The Journal of the Acoustical Society of America* 55 (5): 1100–1103. <https://doi.org/10.1121/1.1914660>.

Lerczak, J. A., and R. C. Hobbs. 1998. “Calculating Sighting Distances from Angular Readings During Shipboard, Aerial, and Shore-Based Marine Mammal Surveys.” *Marine Mammal Science* 14 (3): 590–599. <https://doi.org/10.1111/j.1748-7692.1998.tb00745.x>.

Love, R. H. 1973. “Target Strengths of Humpback Whales *Megaptera novaeangliae*.” *The Journal of the Acoustical Society of America* 54 (5): 1312–1315.
<https://doi.org/10.1121/1.1914428>.

Macrander, A. M., L. Brzuzy, K. Raghukumar, D. Preziosi, and C. Jones. 2022. “Convergence of Emerging Technologies: Development of a Risk-Based Paradigm for Marine Mammal Monitoring for Offshore Wind Energy Operations.” *Integrated Environmental Assessment and Management* 18: 939–949. <https://doi.org/10.1002/ieam.4532>.

Madhusudhana, S. G. Pavan, L. A. Miller, W. L. Gannon, A. Hawkins, C. Erbe, J. A. Hamel, and J. A. Thomas. 2022. “Choosing Equipment for Animal Bioacoustic research.” In *Exploring Animal Behavior Through Sound: Volume 1: Methods*, edited by C. Erbe and J. A. Thomas, 37–85. Basel: Springer International Publishing. https://doi.org/10.1007/978-3-030-97540-1_2.

Markowitz, T. M., A. Harlin, and B. Würsig. 2003. “Digital Photography Improves Efficiency of Individual Dolphin Identification.” *Marine Mammal Science* 10 (1): 217–223.
<https://doi.org/https://doi.org/10.1111/j.1748-7692.2003.tb01103.x>.

Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, J. A. Ward, D. J. Moretti, D. Harris, and P. L. Tyack. 2013. “Estimating Animal Population Density Using Passive Acoustics.” *Biological Reviews* 88 (2): 287–309. <https://doi.org/https://doi.org/10.1111/brv.12001>.

Marques, T. A., L. Thomas, J. Ward, N. DiMarzio, and P. L. Tyack. 2009. “Estimating Cetacean Population Density Using Fixed Passive Acoustic Sensors: An Example with Blainville’s Beaked Whales.” *The Journal of the Acoustical Society of America* 125 (4): 1982–1994.
<https://doi.org/10.1121/1.3089590>.

Matthews, L. P., and S. E. Parks. 2021. “An Overview of North Atlantic Right Whale Acoustic Behavior, Hearing Capabilities, and Responses to Sound.” *Marine Pollution Bulletin* 173.
<https://doi.org/10.1016/j.marpolbul.2021.113043>.

Mazzoil, M., S. McCulloch, R. Defran, and E. M. Titcomb. 2004. “Use of Digital Photography and Analysis of Dorsal Fins for Photo-Identification of Bottlenose Dolphins.” *Aquatic Mammals* 30: 209–219. <https://doi.org/10.1578/AM.30.2.2004.209>.

Mellinger, D. K. 2002. *Ishmael 1.0 User’s Guide*. Pacific Marine Environmental Laboratory, NOAA Technical Memorandum OAR PMEL-120.
<https://repository.library.noaa.gov/view/noaa/11056>.

Mellinger, D. K., and C. W. Clark. 2000. "Recognizing Transient Low-Frequency Whale Sounds by Spectrogram Correlation." *The Journal of the Acoustical Society of America* 107: 3518–3529. <https://doi.org/10.1121/1.429434>.

Mellinger, D. K., S. W. Marin, R. P. Morrissey, L. Thomas, and J. J. Yosco. 2011. "A Method for Detecting Whistles, Moans, and Other Frequency Contour Sounds." *The Journal of the Acoustical Society of America* 129 (6): 4055–4061. <https://doi.org/10.1121/1.3531926>.

Michel, H. 2015. "Analysis of the Behavioural Response of Fin and Humpback Whales to an Icebreaker Using a Thermal Imaging Based Whale Detection System." Master's thesis, Climate Sciences, Carl von Ossietzky University, Oldenburg. <https://epic.awi.de/id/eprint/52685/>.

Miller, J. H., and D. C. Potter. 2001. "Active High Frequency Phased Array Sonar for Whale Shipstrike Avoidance: Target Strength Measurements." In *An Ocean Odyssey: Oceans 2001 MTS/IEEE Conference Proceedings*, Honolulu, HI. <https://doi.org/10.1109/OCEANS.2001.968324>.

Miller, J. H., D. C. Potter, T. Weber, and J. Felix. 1999. "The Target Strength of the Northern Right Whale (*Eubalaena glacialis*)." *The Journal of the Acoustical Society of America* 105 (2_Supplement): 992. <https://doi.org/10.1121/1.424796>.

Miller, M. S. 2012. "Real-Time Tracking of Blue Whales Using DIFAR Sonobuoys." In *Proceedings of Acoustics 2012*. Acoustics 2012, Australian Acoustical Society, Fremantle. November 21–23, 2012.

Mordy, C. W., E. E. Cokelet, A. De Robertis, R. Jenkins, C. E. Kuhn, N. Lawrence-Slavas, C. L. Berchok, J. L. Crance, J. T. Sterling, J. N. Cross, P. J. Stabeno, C. Meinig, H. M. Tabisola, W. Burgess, and I. Wangen. 2017. "Advances in Ecosystem Research: Saildrone Surveys of Oceanography, Fish, and Marine Mammals in the Bering Sea." *Oceanography* 30 (2): 113–115. <https://doi.org/10.5670/oceanog.2017.230>.

Ntalampiras, S. 2017. "Hybrid Framework for Categorising Sounds of Mysticete Whales." *IET Signal Processing* 11 (4): 349–355. <https://doi.org/10.1049/iet-spr.2015.0065>.

Okane, D., E. P. Koveke, K. Tashima, K. Saeki, S. Maezono, T. Nagahata, N. Hayashi, K. Owen, D. P. Zitterbart, S. -I. Ohira, and K. Toda. 2019. "High Sensitivity Monitoring Device for Onboard Measurement of Dimethyl Sulfide and Dimethylsulfoniopropionate in Seawater and an Oceanic Atmosphere." *Analytical Chemistry* 91 (16): 10484–10491. <https://doi.org/10.1021/acs.analchem.9b01360>.

Oliver, J. S., and R. G. Kvitek. 1984. "Side-Scan Sonar Records and Diver Observations of the Gray Whale (*Eschrichtius robustus*) Feeding Grounds." *The Biological Bulletin* 167 (1): 264–269. <https://doi.org/10.2307/1541353>.

Owen, K., K. Saeki, J. D. Warren, A. Bocconcelli, D. N. Wiley, S. -I. Ohira, A. Bombosch, K. Toda, and D. P. Zitterbart. 2021. "Natural Dimethyl Sulfide Gradients Would Lead Marine Predators to Higher Prey Biomass." *Communications Biology* 4 (1): 149. <https://doi.org/10.1038/s42003-021-01668-3>.

Palacios, D. M., L. M. Irvine, B. A. Lagerquist, J. A. Fahlbusch, J. Calambokidis, S. M. Tomkiewicz, and B. R. Mate. 2022. "A Satellite-Linked Tag for the Long-Term Monitoring of

Diving Behavior in Large Whales.” *Animal Biotelemetry* 10 (1): 26.
<https://doi.org/10.1186/s40317-022-00297-9>.

Palacios, D. M., B. R. Mate, C. S. Baker, C. E. Hayslip, T. M. Follett, D. Steel, B. A. Lagerquist, L. M. Irvine, and M. H. Winsor. 2019. “Tracking North Pacific Humpback Whales To Unravel Their Basin-Wide Movements.” Final Technical Report. Prepared for Pacific Life Foundation. Marine Mammal Institute, Oregon State University. Newport, Oregon, USA. 30 June 2019.
<https://doi.org/10.5399/osu/1117> .

Parks, S. E., A. Searby, C. Aurélie, M. P. Johnson, D. P. Nowacek, and P. L. Tyack. 2011. “Sound Production Behavior of Individual North Atlantic Right Whales: Implications for Passive Acoustic Monitoring.” *Endangered Species Research* 15 (1): 63.
<https://doi.org/10.3354/esr00368>.

Perryman, W. L., M. A. Donahue, J. L. Laake, and T. E. Marin. 1999. “Diel Variation in Migration Rates of Eastern Pacific Gray Whales Measured with Thermal Imaging Sensors.” *Marine Mammal Science* 15 (2): 426–445. <https://doi.org/10.1111/j.1748-7692.1999.tb00811.x>.

Podobna, Y., J. Schoonmaker, C. Boucher, and D. Oakley. 2009. “Optical Detection of Marine Mammals.” In *Proceedings Volume 7317, Ocean Sensing and Monitoring*. SPIE Defense, Security, and Monitoring. Orlando, FL, April 13–17, 2009. <https://doi.org/10.1117/12.818359>.

Premus, V. E., P. A. Abbot, V. Kmelnitsky, C. J. Gedney, and T. A. Abbot. 2022. “A Wave Glider-Based, Towed Hydrophone Array System for Autonomous, Real-Time, Passive Acoustic Marine Mammal Monitoring.” *The Journal of the Acoustical Society of America* 152 (3): 1814–1828. <https://doi.org/10.1121/10.0014169>.

Pyć, C. D., M. Geoffroy, and F. R. Knudsen. 2016a. “A Summary Comparison of Active Acoustic Detections and Visual Observations of Marine Mammals in the Canadian Beaufort Sea.” In *The Effects of Noise on Aquatic Life II*, edited by A. Popper, and A. Hawkins, 879–884. New York: Springer. https://doi.org/10.1007/978-1-4939-2981-8_108.

Pyć, C. D., Geoffroy, M., Knudsen, F. R. 2016b. “An evaluation of active acoustic methods for detection of marine mammals in the Canadian Beaufort Sea.” *Marine Mammal Science* 32: 202–219. <https://doi.org/10.1111/mms.12250>.

Rankin, S., C. Oedekoven, and F. Archer. 2020. “Mark Recapture Distance Sampling: Using Acoustics to Estimate the Fraction of Dolphins Missed by Observers During Shipboard Line-Transect Surveys.” *Environmental and Ecological Statistics* 27 (2): 233–251.
<https://doi.org/10.1007/s10651-020-00443-7>.

Rasmussen, J. H., and A. Širović. 2021. “Automatic Detection and Classification of Baleen Whale Social Calls Using Convolutional Neural Networks.” *The Journal of the Acoustical Society of America* 149 (5): 3635. <https://doi.org/10.1121/10.0005047>.

Read, A. J. 2009. “T –Telemetry.” In *Encyclopedia of Marine Mammals (Second Edition)*, edited by W. F. Perrin, B. Würsig, and J. G. M. Thewissen, 1153–1156. Massachusetts: Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-373553-9.00265-0>.

Renò, V., G. Dimauro, G. Labate, E. Stella, C. Fanizza, G. Cipriano, R. Carlucci, and R. Maglietta. 2019. “A SIFT-Based Software System for the Photo-Identification of the Risso’s

Dolphin.” *Ecological Informatics* 50: 95–101.

<https://doi.org/https://doi.org/10.1016/j.ecoinf.2019.01.006>.

Rice, A., A. Širović, J. A. Hildebrand, M. Wood, A. Carbaugh-Rutland, and S. Baumann-Pickering. 2022. “Update on Frequency Decline of Northeast Pacific Blue Whale (*Balaenoptera musculus*) Calls.” *PLOS ONE* 17 (4): e0266469. <https://doi.org/10.1371/journal.pone.0266469>.

Richardson, J. W., C. R. Greene, C. I. Malme, and D. H. Thompson, eds. 1995. *Marine Mammals and Noise*. San Diego: Academic Press.

<https://doi.org/10.1371/journal.pone.0266469>.

Richter, S., H. Yurk, A. Winterl, E. Chmelnitsky, N. Serra, P. D. O’Hara, and D. Zitterbart. 2024. “Coastal Marine Mammal Conservation Using Thermal Imaging-Based Detection Systems (pre-print).” *bioRxiv*. <https://doi.org/10.1101/2023.08.25.554754>.

Rone, B. K., C. L. Berchok, J. L. Crance, and P. J. Clapham. 2012. “Using Air-Deployed Passive Sonobuoys to Detect and Locate Critically Endangered North Pacific Right Whales.” *Marine Mammal Science* 28 (4): E528–E538. <https://doi.org/https://doi.org/10.1111/j.1748-7692.2012.00573.x>.

Rutenko, A. N., M. M. Zykov, V. A. Gritsenko, M. Y. Fershalov, M. R. Jenkerson, D. S. Manulchev, R. Racca, and V. E. Nechayuk. 2022. “Acoustic Monitoring and Analyses of Air Gun, Pile Driving, Vessel, and Ambient Sounds During the 2015 Seismic Surveys on the Sakhalin Shelf.” *Environmental Monitoring and Assessment* 194 (Supplement 1) (744). <https://doi.org/10.1007/s10661-022-10021-y>.

RWSC & MTS. 2024a. *Collaborative Whale Detection Technology Evaluation Virtual Workshop Series: Session 1* (p. 24). Regional Wildlife Science Collaborative for Offshore Wind. <https://mtsociety.memberclicks.net/assets/docs/Collaborative%20Technology%20Workshop%20Session%201%20Proceeding.pdf>

RWSC & MTS. 2024b. *Collaborative Whale Detection Technology Evaluation Virtual Workshop Series: Session 2* (p. 16). Regional Wildlife Science Collaborative for Offshore Wind. https://tethys.pnnl.gov/sites/default/files/publications/RWSC_MTS_2024.pdf

RWSC & MTS. 2024c. *Collaborative Whale Detection Technology Evaluation Virtual Workshop Series: Session 3* (p. 11). <https://rwscorg.sharepoint.com/:b:/r/sites/TechnologySubcommittee/Shared Documents/Public Technology Subcommittee Folder/Collaborative Technology Workshop Series 2024/Session 3 - Summary of Findings and Next Steps for Evaluating Realtime Marine Mammal Monitoring Tools and Technologies- November 19 2024/Final Technology Workshop Session 3 Proceeding v1.pdf?csf=1&web=1&e=dO0b9E>

Salisbury, D., B. Estabrook, H. Klinck, and A. Rice. 2018. *Understanding Marine Mammal Presence in the Virginia Offshore Wind Energy Area*. Bureau of Ocean Energy Management (BOEM), Cornell Laboratory of Ornithology, BOEM 2019-007.

Savoca, M. S., and G. A. Nevitt. 2014. “Evidence that dimethyl sulfide facilitates a tritrophic mutualism between marine primary producers and top predators.” *PNAS* 111 (11): 4157–4161. <https://doi.org/doi:10.1073/pnas.1317120111>.

- Schoonmaker, J., J. Dirbas, Y. Podobna, T. Wells, C. Boucher, and D. Oakley. 2008. "Multispectral Observations of Marine Mammals." *Electro-Optical and Infrared Systems: Technology and Applications V*, 7113: 276–284. <https://doi.org/10.1117/12.800024>.
- Seymore, A. C., J. Dale, M. Hammill, P. N. Halpin, and D. W. Johnston. 2017. "Automated Detection and Enumeration of Marine Wildlife Using Unmanned Aircraft Systems (UAS) and Thermal Imagery." *Scientific Reports* 7 (1): 45127. <https://doi.org/10.1038/srep45127>.
- Shiu, Y., K. J. Palmer, M. A. Roch, E. Fleishman, X. Liu, E. -M. Nosal, T. Helble, D. Cholewiak, D. Gillespie, and H. Klinck. 2020. "Deep Neural Networks for Automated Detection of Marine Mammal Species." *Scientific Reports* 10 (1): 607. <https://doi.org/10.1038/s41598-020-57549-y>.
- Shiu, Y., Palmer, K. J., Roch, M. A., Fleishman, E., Liu, X., Nosal, E.-M., Helble, T., Cholewiak, Silber, G. K., S. Bettridge, and D. Cottingham. 2009. *Report of a Workshop to Identify and Assess Technologies to Reduce Ship Strikes of Large Whales*. U. S. Department of Commerce, NOAA Technical Memo, NMFS-OPR-42. Providence, RI.
- Sirović, A., J. A. Hildebrand, and S. M. Wiggins. 2007. "Blue and Fin Whale Call Source Levels and Propagation Range in the Southern Ocean." *The Journal of the Acoustical Society of America* 122 (2): 1208–1215. <https://doi.org/10.1121/1.2749452>.
- Smith, H. R., D. P. Zitterbart, T. F. Norris, M. Flau, E. L. Ferguson, C. G. Jones, O. Boebel, and V. D. Moulton. 2020. "A Field Comparison of Marine Mammal Detections via Visual, Acoustic, and Infrared (IR) Imaging Methods Offshore Atlantic Canada." *Marine Pollution Bulletin* 154: 111026. <https://doi.org/10.1016/j.marpolbul.2020.111026>.
- SMRU Consulting. 2024. *Detection Probability Simulation* [Computer software]. <https://smru-consulting.shinyapps.io/BearingFlowerSimulation/>.
- South Fork Wind. 2024. *South Fork Wind Marine Mammal and Sea Turtle Monitoring During Windfarm Construction*. <https://www.fisheries.noaa.gov/s3/2024-06/SouthForkWind-ConstructionPSOPAMReport-PlusAppendices-OPR1.pdf>.
- Southall, B., L. Morse, K. A. Williams, and E. Jenkins. 2021. *Marine Mammals Workgroup Report for the State of the Science Workshop on Wildlife and Offshore Wind Energy 2020: Cumulative Impacts*. Report to the New York State Energy Research and Development Authority (NYSERDA). Albany, NY. <https://www.nyetwg.com/2020-workgroups>.
- Southall, B. L., A. N. Allen, J. Calambokidis, C. Casey, S. L. DeRuiter, S. Fregosi, A. S. Friedlander, J. A. Goldbogen, C. M. Harris, E. L. Hazen, V. Popov, and A. K. Stimpert. 2023. "Behavioural Responses of Fin Whales to Military Mid-Frequency Active Sonar." *Royal Society Open Science* 10 (12): 231775. <https://doi.org/doi:10.1098/rsos.231775>.
- Stafford, K. M., C. G. Fox, and D. S. Clark. 1998. "Long-Range Acoustic Detection and Localization of Blue Whale Calls in the Northeast Pacific Ocean." *The Journal of the Acoustical Society of America* 104: 3616–3625. <https://doi.org/10.1121/1.423944>.
- Stanistree, J. E., D. Risch, and S. M. Van Parijs. 2013. "Passive Acoustic Tracking of Singing Humpback Whales (*Megaptera novaeangliae*) on a Northwest Atlantic Feeding Ground." *PLOS ONE* 8 (4): e61263. <https://doi.org/10.1371/journal.pone.0061263>.

- Suarez-Bregua, P., M. Alvarez-González, K. M. Parsons, J. Rotlant, G. J. Pierce, and C. Saavedra. 2022. "Environmental DNA (eDNA) for Monitoring Marine Mammals: Challenges and Opportunities [Mini Review]." *Frontiers in Marine Science* 9. <https://doi.org/10.3389/fmars.2022.987774>.
- Thode, A., J. Skinner, P. Scott, J. Roswell, J. Straley, and K. Folkert. 2010. "Tracking Sperm Whales with a Towed Acoustic Vector Sensor." *The Journal of the Acoustical Society of America* 128 (5): 2681–2694. <https://doi.org/10.1121/1.3495945>.
- Thomas, L., and T. A. Marques. 2012. "Passive Acoustic Monitoring for Estimating Animal Density." *Acoustics Today* 8 (3): 35–44. <https://doi.org/10.1121/1.4753915>.
- Thompson, T. J., H. E. Winn, and P. J. Perkins. 1979. "Mysticete Sounds." In *Behavior of Marine Animals: Current Perspectives in Research*, edited by H. E. Winn and B. L. Olla. U.S.: Springer. https://doi.org/10.1007/978-1-4684-2985-5_12.
- Tiemann, C. O., A. M. Thode, J. Straley, V. O'Connell, and K. Folkert. 2006. "Three-Dimensional Localization of Sperm Whales Using a Single Hydrophone." *The Journal of the Acoustical Society of America* 120 (4): 2355–2365. <https://doi.org/10.1121/1.2335577>.
- Tremblay, C. J., S. M. Van Parijs, and D. Cholewiak. 2019. "50 to 30-Hz Triplet and Singlet Down Sweep Vocalizations Produced by Sei Whales (*Balaenoptera borealis*) in the Western North Atlantic Ocean." *Journal of the Acoustical Society of America* 145 (6): 3351–3358. <https://doi.org/10.1121/1.5110713>.
- Würsig, B., and M. Würsig. 1977. "The Photographic Determination of Group Size, Composition, and Stability of Coastal Porpoises (*Tursiops truncatus*)." *Science* 198 (4318): 755–756. <https://doi.org/doi:10.1126/science.198.4318.755>.
- Valsecchi, E., A. Arcangeli, R. Lombardi, E. Boyse, I. M. Carr, P. Galli, and S. J. Goodman. 2021. "Ferries and Environmental DNA: Underway Sampling from Commercial Vessels Provides New Opportunities for Systematic Genetic Surveys of Marine Biodiversity." *Frontiers in Marine Science* 8. <https://doi.org/10.3389/fmars.2021.704786>.
- Van der Schaar, M., M. André, E. Delory, D. Gillespie, and J-F. Rolin. 2017. *Passive Acoustic Monitoring from Fixed Platform Observatories*. IFREMER for FixO3, Deliverable 12.6. Brest, France. <http://dx.doi.org/10.25607/OBP-967>.
- Van Parijs, S. M., C. W. Clark, R. S. Sousa-Lima, S. E. Parks, S. Rankin, D. Risch. D. and I. C. V. Opzeeland. 2009. "Management and Research Applications of Real-Time and Archival Passive Acoustic Sensors Over Varying Temporal and Spatial Scales." *Marine Ecology Progress Series* 395: 21–36. <https://doi.org/10.3354/meps08123>.
- Van Parijs, S. M., K. Baker, J. Carduner, J. Daly, G. E. Davis, C. Esch, S. Guan, A. Scholik-Schlomer, N. B. Sisson, and E. Staatterman. 2021. "NOAA and BOEM Minimum Recommendations for Use of Passive Acoustic Listening Systems in Offshore Wind Energy Development Monitoring and Mitigation Programs." *Frontiers in Marine Science* 8. <https://doi.org/10.3389/fmars.2021.760840>.
- Verfuss, U. K., D. Gillespie, J. Gordon, T. A. Marques, B. Miller, R. Plunkett, J. A. Theriault, D. J. Tollit, D. P. Zitterbart, P. Hubert, and L. Thomas. 2018. "Comparing Methods Suitable for

Monitoring Marine Mammals in Low Visibility Conditions During Seismic Surveys.” *Marine Pollution Bulletin* 126: 1–18. <https://doi.org/10.1016/j.marpolbul.2017.10.034>.

Verfuss, U. K., D. Gillespie, J. Gordon, T. Marques, B. Miller, R. Plunkett, J. Theriault, D. Tollit, D. P. Zitterbart, P. Hubert, and L. Thomas. 2016. *Low Visibility Real-Time Monitoring Techniques Review*. SMRU Consulting, SMRUM-OGP2015-002. Friday Harbor, WA.

von Benda-Beckmann, A. M., S. P. Beerens, and S. P. van IJsselmuide. 2013. “Effect of Towed Array Stability on Instantaneous Localization of Marine Mammals.” *The Journal of the Acoustical Society of America*, 134 (3): 2409–2417. <https://doi.org/10.1121/1.4816553>.

Wade, P., M. P. Heide-Jørgensen, K. Shelden, J. Barlow, J. Carretta, J. Durban, R. LeDuc, L. Munger, S. Rankin, A. Sauter, and C. Stinchcomb. 2006. “Acoustic Detection and Satellite-Tracking Leads to Discovery of Rare Concentration of Endangered North Pacific Right Whales.” *Biology Letters* 2 (3): 417–419. <https://doi.org/doi:10.1098/rsbl.2006.0460>.

Wahlberg, M. B. Møhl, M. P. Teglberg. 2001. “Estimating Source Position Accuracy of a Large-Aperture Hydrophone Array for Bioacoustics.” *Journal of the Acoustical Society of America* 109: 397–406. <https://doi.org/10.1121/1.1329619>.

Yack, T. M. 2013. “The Development of Automated Detection Techniques for Passive Acoustic Monitoring as a Tool for Studying Beaked Whale Distribution and Habitat Preferences in the California Current Ecosystem.” PhD dissertation, University of California, Davis. <https://www.proquest.com/docview/1426390336>

Yonehara, Y., L. Kagami, H. Yamada, H. Kato, M. Terada, and S. Okada. 2012. “Feasibility on Infrared Detection of Cetaceans for Avoiding Collision with Hydrofoil.” *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* 6 (1): 149–154.

Zhong, M., M. Torterotot, T. A. Branch, K. M. Stafford, J. -Y. Royer, R. Dodhia, and J. L. Ferres. 2021. “Detecting, Classifying, and Counting Blue Whale Calls with Siamese Neural Networks.” *The Journal of the Acoustical Society of America* 149 (5): 3086–3094. <https://doi.org/10.1121/10.0004828>.

Zimmer, W. M. X. 2013. “Range Estimation of Cetaceans with Compact Volumetric Arrays.” *The Journal of the Acoustical Society of America* 134 (3): 2610–2618. <https://doi.org/10.1121/1.4817892>.

Zitterbart, D. P., L. Kindermann, E. Burkhardt, and O. Boebel. 2013. “Automatic Round-the-Clock Detection of Whales for Mitigation from Underwater Noise Impacts.” *PLOS ONE* 8 (8): e71217. <https://doi.org/10.1371/journal.pone.0071217>.

Zitterbart, D. P., H. R. Smith, M. Flau, S. Richter, E. Burkhardt, J. Beland, L. Bennett, A. Cammareri, A. Davis, M. Holst, C. Lanfredi, H. Michael, M. Noad, K. Owen, A. Pacini, and O. Boebel. 2020. “Scaling the Laws of Thermal Imaging-Based Whale Detection.” *Journal of Atmospheric and Oceanic Technology* 37 (5): 807–824. <https://doi.org/10.1175/JTECH-D-19-0054.1>.

Appendix A – Technology Summary

The capabilities of baleen whale monitoring technologies are summarized in Table 5 based on the information described in Section 3.0 from publicly available information and documents. Information from this table was used to support the technology evaluation described in Section 3.0, Table 2.

Table 5. Technology capabilities summary table

Technology	Ability to record time and location of an animal?	Can cue be detected and delivered within decision-making time window?	Practical ability to detect <i>any</i> individual whale that enters the zone of perception in ideal conditions and displays/produces the cue?	Mechanism of Observation	Detection Cue	Automation Capability	Real-time Integration / Reporting Latency	Deployment Platform Type(s)	Success Affected by	Detection range
Satellite Imagery	Yes	No (opportunistic platform)	Yes	Satellite imagery	Body above or near surface	- Automated analysis of images is possible - ML in various stages of development/validation - Human review/confirmation is needed	Hours to months	Satellite	- Visibility (darkness, precipitation, fog, sea state, ice cover) - Size, species-specific features, protrusion (part of body), and behavior of animal impacts detection - Speed of satellite (faster than visual surveys) - Limited capacity to differentiate among species	No theoretical limit
Optical Camera	Yes	Yes	Yes	Optical camera	Body or blow above or near surface	- Automated analysis of images is possible - ML in various stages of development/validation - Human review/confirmation is needed	Minutes to months	- Vessel Fixed (land-based or offshore stationary vantage point) - Mobile (ASV, drone, aircraft)	- Visibility (darkness, precipitation, fog, sea state) - Aerial (plane or drone) cameras point primarily below (do not scan horizon)	0 to ~12 km (horizon), limited by mounting height and visibility up to 360 degrees Range (distance and angle) is limited by focal length and magnification

Technology	Ability to record time and location of an animal?	Can cue be detected and delivered within decision-making time window?	Practical ability to detect any individual whale that enters the zone of perception in ideal conditions and displays/produces the cue?	Mechanism of Observation	Detection Cue	Automation Capability	Real-time Integration / Reporting Latency	Deployment Platform Type(s)	Success Affected by	Detection range
Infrared Imaging	Yes	Yes	Yes	Infrared camera	Body or blow above or near surface	- Automated cameras are possible - Automated detection is possible - ML in various stages of development/validation - Human review/confirmation is needed	Minutes	- Vessel (handheld, or mounted) - Fixed (land-based or offshore stationary vantage point) - Mobile (ASV/drone, aircraft)	- Height and angle on platform - Camera stabilization - Visibility (darkness, precipitation, fog, sea state) - Animal must have detectable difference in temp - Limited capacity to differentiate among species	Detections up to 5.5 km possible. Reliable detection up to 4 km (in good visibility; ≤ 4 Beaufort, no fog, no rain) with a good cooling system; 1-2 km with high-mounted uncooled system; 350 m with handheld system (vessel-based/eye-level) 360 degrees, depending on obstructions
Telemetry	Yes	Yes	No (all whales would need to be tagged)	Animal-borne tag	Animal above surface	- Automation detection is possible (but limited by satellite passes/coverage) - Argos satellite positions require processing and filtering (requires human review)	Seconds to minutes	Two platforms: Transmitter on animal Receiver on vessel, fixed, mobile (goniometer, VHF), or satellite (Argos)	- Cannot tag every animal or animals <1 yr - Longevity and durability of tags - Permitting (risk of injury/infection) and public opinion - Availability of tagging experts	VHF can track out to 4 to 10 km Argos-based lat/lon position uncertainty depends on number of satellites present (position error up to 1.5 km)
Dimethyl sulfide (DMS) Concentration	No	No	No (no consistent/predictable response to DMS)	DMS concentration	Concentration threshold (assumes relationship)	- Automated sample and analysis on vessel is possible - Automated analysis of satellite image is possible	Days+	Vessel (for sampling) Satellite (for imaging)	- Response of whales to DMS concentrations	Uncertain - not a direct measurement of immediate individual presence. Satellite DMS measurement has only been demonstrated to 1/12 degree resolution

Technology	Ability to record time and location of an animal?	Can cue be detected and delivered within decision-making time window?	Practical ability to detect any individual whale that enters the zone of perception in ideal conditions and displays/produces the cue?	Mechanism of Observation	Detection Cue	Automation Capability	Real-time Integration / Reporting Latency	Deployment Platform Type(s)	Success Affected by	Detection range
eDNA	No	No	Yes (within day-week window)	DNA analysis	Recently shed genetic material	Full automation not possible: -Automated eDNA samplers are possible -Manual processing in lab is needed -Manual review of data is needed	Days+	Vessel (for sampling)	- Sampling location and time	Uncertain - eDNA moves within water before sample collection. Detected eDNA could have been transported (depends on local hydrodynamics)
Active Acoustic Monitoring	Yes	Yes	No (cannot reliably classify an echo as whale)	Imaging sonar	Strength of echo above a threshold	- Automation detection of a specific target strength is possible - ML in various stages of development/validation - Human review/confirmation is needed	Hours	- Vessel (hull or towed) - Fixed (moored, bottom-mounted) - Mobile (AUVs, USVs)	- Difficult to ID to species or broader taxonomic group - Risk of sound exposure and behavioral response - Target strength variable: depends on animal distance, depth, orientation, and behavior and sound speed profile, surface/bottom reverberations, angle of incidence	Highly dependent on instrument, platform, environment, species: 30-55 m for smaller animals <250-350 generally for larger whales (but a few reports of 1000-2000 m)
Passive Acoustic Monitoring	Yes	Yes	Yes	Hydrophone	Call	- Automated detection and classification is possible (must consider onboard power constraints) - ML in various stages of development/validation - Human review/confirmation is needed	Minutes	- Vessel (towed) - Fixed (moored, bottom-mounted) - Mobile (AUV, ASV, drifter)	- Ambient noise levels, including whether or not pile driving is actively occurring - Propagation effects (environment, bathymetry, bottom type, depth of animal & hydrophone, transmission loss, signal distortion) - Call rate - Performance of automatic classifier - Experience of PAM system operator	<1 to 150 km depending on environment, equipment, species, and level of ambient noise (before vs. during pile driving) 360 degrees

Technology	Ability to record time and location of an animal?	Can cue be detected and delivered within decision-making time window?	Practical ability to detect <i>any</i> individual whale that enters the zone of perception in ideal conditions and displays/produces the cue?	Mechanism of Observation	Detection Cue	Automation Capability	Real-time Integration / Reporting Latency	Deployment Platform Type(s)	Success Affected by	Detection range
Marine Mammal Observer	Yes	Yes	Yes	Human eye	Body or blow above or near surface	Automation not possible	Seconds	- Vessel - Fixed - Mobile (aerial)	- Vessel obstructions - Visibility (darkness, precipitation, sea state) - Observer experience and fatigue	0 to ~12 km (horizon), depending on height of viewing bridge and visibility 360 degrees, depending on vessel obstructions

Appendix B – Technology Performance Summary

This Appendix provides a summary of past performance of whale monitoring technologies that did not meet the screening criteria described in Section 3.0. A literature review was conducted for each technology to evaluate past performance across the key performance metrics of efficacy, range, data delivery, and variation. Literature was collected from peer-reviewed journal articles, technical reports, consultant reports, and regulatory documents, as applicable and available. Specific focus was put on literature that included detection of baleen whales in marine environments.

The following subsections describe how each technology functions and documents past performance. Tables summarizing relevant studies, including the technology used, study location, species, environmental conditions, detection time and distance, and results are presented for the following technologies:

- Appendix B.1 – Satellite imagery
- Appendix B.2 – Optical cameras
- Appendix B.3 – Telemetry
- Appendix B.4 – Dimethyl sulfide (DMS) concentration
- Appendix B.5 – Environmental deoxyribonucleic acid (eDNA)
- Appendix B.6 – Active acoustics

B.1 Satellite Imagery

Individual whales can be observed in very high-resolution (VHR) satellite imagery (Hodul et al. 2022; Khan et al. 2023). Whales have been detected from satellite imagery with resolution less than 1 m (Abileah 2002); however, satellites have much greater detection capabilities with resolutions of 50 cm or less. There are currently nine commercial VHR satellite systems with resolution less than 50 cm in the Earth's orbit with at least three planned or in the process of full deployment (Khan et al. 2023). VHR satellites provide imagery of the Earth's surface in a panchromatic spectrum, where visible light is combined into a single intensity of solar radiation value, and multispectral, where the intensity of solar radiation is measured across discrete spectral bands.

Satellite imagery can detect whales at or near the surface. Detection ability is limited by an availability bias, where a whale needs to be at a detectable depth when the satellite sweeps that area of the ocean, and by perception bias, where the sea states, clouds, and environmental conditions provide a suitable view of a calm ocean (Hodul et al. 2022). The abilities for satellite imagery to be used to identify baleen whales is summarized in Table 6.

Abileah (2002) first proved that humpback whales could be detected in satellite imagery, and recent advancements in VHR satellites and image processing have allowed for detection of multiple species (Cubaynes et al. 2019), making population density estimates (Bamford et al. 2020; Corrêa et al. 2022), and the application of automated detection algorithms (Fretwell et al. 2014; Guirado et al. 2019) possible. Past studies demonstrated a correlation between whale detections made by manual review of VHR satellite imagery and observations from a vessel or aerial survey (Bamford et al. 2020; Corrêa et al. 2022). Two studies have applied automated detection algorithms to VHR satellite imagery and found precision and recall to between 76%

and 99% when comparing the automated system to manual image review (Fretwell et al. 2014; Guirado et al. 2019). Detection of different species has been demonstrated by Cubaynes et al. (2019) with VHR satellite imagery at 50 m resolution by looking at imagery from four different locations each with a different predominant whale species. A good understanding of species composition of a study area is needed to aide species classification, which is limited with current approaches (Bamford et al. 2020).

Table 6. Representative literature review of satellite imagery

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
HR Satellite, IKONOS-2	Evaluate feasibility of satellite imaging for marine mammal detection	Maui, Hawaii Humpback	Relatively cloud free	Not reported	Extent of satellite image; 100s of km ²	Simulations show that 14 m whale can be detected 20 m below surface. Predicted threshold of 1 false positive per 100 km ²	Abileah 2002
VHR Satellite	Compare baleen whale density estimate from satellite imagery and ship survey	Antarctica Humpback	"ideal" to "difficult" sea states	Not reported	Extent of satellite image; 100s of km ²	Density estimates from satellite were same order of magnitude as ship surveys (0.13 to 0.33 whales/km ² , respectively)	Bamford et al. 2020
MR, HR, VHR Satellite	Compare number of whale sightings between satellite images and aerial survey	Brazil Southern Right	Beaufort scale 1–3, visibility good to optimal, limited cloud cover	Not reported	Extent of satellite image; 100s of km ²	Found correlation between aerial sightings and satellite detection. Satellite images detected 43 whale like objects. Aerial survey detected 80 individual whales in 32 groups	Corrêa et al. 2022

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
VHR Satellite, WorldView 3	Demonstrate manual detection of whales from VHR satellite and characterize body measurements and radiance values	Mexico, Hawaii, Argentina, Ligurian Sea Fin, Humpback, Southern Right, Gray	Not reported	Not reported	4 images, range in extent from 80 km ² to 4,230 km ²	4 sites. Number of whales range from 34–62 per site. Proportion of detections that were definitive range from 36–76%. Not verified with alternative method	Cubaynes et al. 2019
VHR Satellite, WorldView-2	Develop and test automated whale detection using VHR satellite imagery	Argentina Southern Right	Cloud free and calm sea state	Not reported	Image covering 113 km ²	55 probable whales, 23 possible whales, 13 other objects by coastal band. Manual v automated image analysis: Precision = 76% Recall = 85%	Fretwell et al. 2014
VHR Satellite (WorldView-3, 15 cm product)	Evaluate feasibility of NARW detection with satellite	Cape Cod Bay North Atlantic Right	Cloud free, 10 cm wave height, minimal sun glint		Imagery covering 200 km ²	31 whale-like objects, 25 of which are definite whales	Hodul et al. 2022
VHR Satellite, Google Earth (combination of satellites)?	Develop and validate a machine learning approach to automated whale detection of VHR satellite imagery	Global, with validation at ten hotspots All whales	Not reported	Not reported	Global imagery	Detection success for automated method versus manual image analysis Precision = 86% Recall = 80% Counting success: Precision = 99% Recall = 89%	Guirado et al. 2019

Current state-of-the-art, limitations, and challenges of VHR satellite whale detection are well described by the Geospatial Artificial Intelligence for Animals (GAIA) initiative (see Khan et al. 2023). Overall, VHR satellites are a promising data source for detecting baleen whales for surveys of distribution and population estimates. Satellite offers a lower cost mechanism to monitor remote or difficult to access regions of the globe compared to vessel or aerial survey techniques. However, the expense of commercial satellite imagery can still be cost prohibitive and have data sharing limitations (Khan et al. 2023). Satellite detection of whales has largely been explored for survey and conservation observations without a particular focus on real-time monitoring due to limitations on satellite passes of an area of the Earth. As of 2023, VHR satellites revisited areas of the globe between every 12 hours to 3 days; next generation VHR satellite systems currently planned or in progress such as WorldView Legion and Pelican may revisit survey areas between 15–30 times daily, which opens new possibilities for higher frequency monitoring if the conditions are suitable (Khan et al. 2023). Automated analysis techniques could allow for faster processing and detection times compared to the multi-hour manual review time required for image analysis, but faster analysis alone will not overcome barriers to achieving near real-time monitoring without satellite systems with higher revisit rates.

B.2 Optical Cameras

An optical camera captures wavelengths of light from 400–700 nm, the same region of the electromagnetic spectrum detectable by the human eye. Because they are designed to create images that replicate human vision, they require visible light. Similar to the cues that a PSO would use, a cue can be a body part of an animal or its breath at or above the surface of the water.

Optical cameras that record in the visible light spectrum can detect baleen whale occurrence, classification, and in some cases, localization. Because the camera sees as a human might, camera performance is greatly affected by environmental conditions that would impact visual based marine mammal monitoring, including sea state, glare, fog, and cloud cover. This limits their ideal application to daytime under clear conditions, especially if using unmanned aerial vehicles, which cannot operate in poor weather.

Performance of optical cameras for detecting baleen whales is summarized in Table 7. Optical cameras for monitoring marine mammals have been deployed from airplanes (Podobna et al. 2009; Smultea et al. 2022), vessels (Baldacci et al. 2005), and stationary locations such as a bridge (Podobna et al. 2009) or a building (Richter et al. 2024). As with any camera, stabilization is required if not mounted on a stationary platform. If scanning the horizon, platform height needs to balance field of view with near-field detectability. Camera performance also differs depending on the characteristics of the camera (e.g., pixel size, focal length, frame rate; Durden et al. 2016).

Table 7. Representative literature review of optical cameras

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
Optical camera on tripod on vessel	Camera used to supplement IR imaging	Italy Fin, sperm, pilot whale; dolphin spp.	Range of good and bad weather/sea states	Real-time	150 m to 1 km (max for blue whale blows)	Visible light camera detections not published explicitly	Baldacci et al. (2005)

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
4-camera system, each with different focal length, from airplane and bridge	Test optical camera system	Canada, Hawaii, California Humpback, fin, minke, dolphin	Clear days; conditions not reported	Manual identification, latency not described	Not reported; within field of view	Successfully identified baleen whales. Precision or recall not investigated	Podobna et al. (2009)
Video camera from airplane; specifications not provided	Opportunistic documentation of blue whale behavior	Washington Blue	Not reported	n/a	Not reported	Blue whale mother-calf pair were observed	Smultea et al. (2022)
2-camera system; paired with IR	Evaluate performance of IR system, use optical camera for verification	Salish Sea, Canada Orca, Humpback, Minke, pinniped, porpoise	Not stated (recorded 22 mos)	Not stated but det. algorithm ran in real-time	Detections increased then decreased with distance; Det. functions differ by species and height/location of cameras; max 2100 m	Optical detections agreed with IR; precision or recall not reported	Richter et al. (2024)

The performance and detection capabilities of optical cameras to perform near real-time detection and classification have not been well characterized. Optical cameras are used to document survey results (e.g., resulting in opportunistic data recordings from Smultea et al. (2022)) or as a supplementary monitoring system to verify detections from IR cameras (e.g., Baldacci et al. (2005); Richter et al. (2024)). The function of an optical camera system matches the cues that are detected by PSOs and has similar requirements for visibility. Because optical photographs have traditionally been used to identify individuals (e.g., Würsig and Würsig 1977; Markowitz et al. 2003, Mazzoil et al. 2004) most automated detection algorithms have been developed for this purpose (e.g., Kahn et al. 2022; Cheeseman et al. 2022, 2023) rather than real-time species identification. There has been limited development of automated technologies for identifying whales in the wild from optical photographs (e.g., Renò et al. 2019; Araújo et al. 2022). Automated detection could aid PSOs by reducing the number of missed detections caused by PSO fatigue; however, real time automated detection and classification systems need further development and testing.

B.3 Telemetry

Telemetry refers to the use of animal-borne tags that gather information, primarily about a whale's location. Additional tag data may also include swim speed; orientation; direction of travel; acceleration; dive, foraging, and acoustic behavior; received sound levels, and/or environmental data of interest. Tags may either be archival (logging data until physically recovered) or transmit data while the tag is still attached to the animal (Read 2009). The most common data transmission system for baleen whale tags is via satellite, which allows the

tagged animal to be tracked remotely over long distances (Hart and Hyrenbach 2009; Read 2009). Depending on battery life, storage capacity, tag sensors, sampling strategy, attachment method, and the study question being asked, tags may record data over periods of days, weeks, or months. Some of the most common tag types deployed on baleen whales are shown in Table 8. At the time of writing, very little information was available about baleen whale tagging during offshore wind foundation installation.

Telemetry data are collected almost continuously and can therefore provide insight into a whale's movements over the length of the tag's operation, at all hours of the day and in difficult environmental conditions (Hart and Hyrenbach 2009). Tag data collection is independent of human observers and can therefore elucidate sub-surface behaviors otherwise unable to be seen by PSOs (Andrews et al. 2019). Because of their substantial spatial coverage, satellite-tracked tags are important tools for understanding the movement ecology and habitat use of baleen whales, many species of which are highly migratory (Palacios et al. 2019). These data can help answer questions about whales' displacement from, and avoidance of, habitat exposed to environmental stressors from offshore wind development. Certain tags are also capable of recording fine-scale behaviors related to baleen whale foraging, and whether their foraging is interrupted by anthropogenic disturbance (Goldbogen et al. 2013; Palacios et al. 2022). Whales' responses to anthropogenic stressors have been shown to be dependent upon behavioral state, including foraging status (Southall et al. 2023; Goldbogen et al. 2013).

In some circumstances, environmental conditions can affect the quality of, and ability to collect, telemetry data. Tags can only be deployed successfully in favorable environmental conditions, but once attached to the animal, they collect and transmit data regardless of sea state, time of day, rain, fog, precipitation, water depth, and cloud cover. Archival acoustic tags can record the calls (and lack thereof) of the whales to which they are attached, as well as those of conspecifics in close proximity. These data elucidate baleen whale acoustic ecology, call rates, and call source levels (Matthews and Parks 2021). The sounds recorded by acoustic tags, primarily those intended to be analyzed for received levels and acoustic behavior, are subject to environmental conditions including water depth, water temperatures, salinity, and thermocline dynamics which impact the resulting sound speed profiles (Parks et al. 2011). Tissue samples obtained during tag deployment/recovery can provide important ancillary information about genetic stock structure, sex, reproductive status, health, stress hormones, and diet.

Baleen whale tagging requires extensive tag deployment expertise, permitting, and logistical planning. There are typically short weather windows in which to deploy tags successfully. Locating animals suitable for tagging is labor-intensive, and the availability of target animals for tagging is unpredictable. As a result, sample sizes of tagged animals tend to be relatively small when compared with other monitoring approaches. Implantable tags that are designed to anchor below the blubber layer, while potentially providing longer-term datasets, are by nature invasive and can cause injury if not properly deployed (Andrews et al. 2019). Conversely, non-invasive tags have very short deployment durations, limiting the temporal coverage to monitor baleen whale behavior in conjunction with construction activities.

Archival tag data can only be retrieved after the tags are physically recovered and are therefore unavailable to the researcher for the duration of the tag deployment. Tag data transmitted via the Argos satellite network, which is commonly used for marine mammal telemetry, are limited by the presence of satellites overhead when the whale surfaces. Data cannot be transmitted when the animal is below the surface. The number of satellite fixes upon tag surfacing determines the accuracy of location estimation. Depending on these factors, estimates of

latitude and longitude can be off by up to 10 km (Irvine et al. 2020). The placement of the tag on the animal's body (i.e., dorsal surface) can also affect the efficiency of data transmission.

Despite having numerous advantages, and potential applicability to other aspects of offshore wind monitoring, telemetry did not meet our inclusion criteria specific to real-time baleen whale monitoring during foundation installation. Reasons include:

- substantial detection latency for both archival and satellite-tracked tags, making real-time monitoring in conjunction with construction activities infeasible with this technology
- impracticability of tagging all whales that may enter a clearance and/or shutdown zone, making telemetry an unsuitable technology for real-time construction monitoring
- variable operational duration of tags, which may not align with the duration of the stressor of interest (i.e., pile driving).

Table 8. Representative literature review of baleen whale tagging studies

Technology Type and Tag Attachment Method	Study Objective	Location and Species	Tag Duration	Env. Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
Archival DTAG; suction cup	Baleen whale sound production behavior	U.S. East Coast, NARW	Hours–days	Shallow water (~200 m); in summer months, a near-surface thermocline leads to a downward-refracting sound speed profile	Duration of deployment (once tag is recovered)	N/A	Only once tag is recovered	Parks et al. 2011
Satellite-tracked LIMPET tag; anchored in blubber layer	Baleen whale habitat use	U.S. East Coast, humpback whales	Days–weeks	Not stated	Dependent on animal surfacing frequency, satellite coverage, and data processing time	Extent of satellite coverage; 100s of km ²	Only when animal surfaces and a satellite receiver is overhead	Aschettino et al. 2020
Satellite-tracked implantable tag; anchored in muscle layer	Baleen whale habitat use	U.S. West Coast, humpback and blue whales	Weeks–months	Not stated	Dependent on animal surfacing frequency, satellite coverage, and data processing time	Extent of satellite coverage; 100s of km ²	Only when animal surfaces and a satellite receiver is overhead	Palacios et al. 2022

B.4 Dimethyl Sulfide (DMS) Concentration

The concentration of DMS in marine surface waters is correlated to areas of high productivity and areas of high zooplankton concentration (Savoca et al. 2014; Owen et al. 2021). DMS concentration in surface waters can be estimated through satellite observations of chlorophyll *a* and photosynthetically available radiation (Gali et al. 2018) or through automated water sampling from a vessel (Okane et al. 2019). Some baleen whales feed on zooplankton, which are correlated with areas of higher DMS concentration, and some studies have suggested that some marine mammals are attracted to areas with high DMS concentration in search of prey (Dove 2015; Savoca et al. 2014). However, physiological studies have not proven that baleen whales have the ability to sense DMS concentrations, and behavioral studies have not proven that baleen whales actively seek areas of higher DMS concentration. Bouchard et al. (2019) found that humpback whales change their acoustic activity near higher DMS concentrations, but they did not observe increased exploratory behavior (see Table 9).

Table 9. Representative literature review of DMS concentration.

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Range	Detection Capability	Citation
DMS Concentration ^[a]	Observe whale response to increased concentration of krill extract and DMS	Madagascar, Iceland, Antarctica Humpback	No precip Visibility > 5 km Beaufort wind < 4 Swell < 1.5 m			Humpbacks modified acoustic activity in response to an increase in DMS concentration, but did not exhibit an exploratory behavior.	Bouchard et al. 2019

^[a] Only one study is presented in this table because it was the only published literature identified that attempts to physically observe baleen whale response to DMS.

While monitoring DMS concentration near construction activities may predict areas where baleen whale occurrence may be related to foraging opportunities, more research is needed in this area to understand any potential correlation and response. DMS concentration cannot be used as a real-time detection of individual baleen whales, only an indicator of potential increased encounter rates due to animals utilizing a possible prey source.

B.5 Environmental Deoxyribonucleic Acid (eDNA)

Environmental DNA (eDNA) is genetic material left behind by organisms that are present in the environment. eDNA sampling is a technique that involves analyzing environmental samples to identify trace amounts of eDNA that indicate the presence of a species. In marine environments, water samples are gathered then extracted, amplified, and sequenced to determine if any DNA from a particular species was present in the sample. Water samples are typically collected manually, but automated samplers have been recently demonstrated (e.g., Formel et al. 2021; Hendricks et al. 2023) to preserve DNA in the field for later laboratory analysis.

The use of eDNA for monitoring marine mammals can be done for single species detection using specific assays or for multiple species detection and classification using metabarcoding (Table 10). Studies have shown that single species eDNA monitoring can be useful to detect elusive or endangered species (see Suarez-Bregua et al. 2022). With a different approach to

DNA analysis, multiple species can be detected from a single sample simultaneously using metabarcoding, which can be useful for biodiversity monitoring for marine vertebrates.

Environmental DNA sampling provides information for detection and classification of baleen whales because traces of eDNA in a sample indicate that a specific species shed that genetic material into the environment. Environmental DNA sampling cannot provide precise localization of an individual animal in space or time because eDNA travels through the dynamic marine environment and degrades over time at different rates depending on local conditions. Information about the persistence, dilution, dispersion, and degradation of eDNA is not well understood (Suarez-Bregua et al. 2022), thus detection of eDNA in a sample does not provide information about where or when an animal was present.

Table 10. Representative literature review of eDNA

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
eDNA metabarcoding	Demonstrate ability of eDNA metabarcoding	Monterey Bay National Marine Sanctuary Marine vertebrates	3-day survey. Conditions not reported			eDNA identified humpback whale	Andruszkiewicz et al. 2017
eDNA metabarcoding	Characterize biodiversity and distribution within National Marine Sanctuaries in California Current	Central California Marine vertebrates	Samples collected on multi-day surveys during 2 years in all conditions. During year 1, supplementary data from marine mammal observers in year 1 in daylight during vessel transit between stations			Year 1: eDNA detected humpback; MMO detected humpback and blue/fin whale. Year 2: eDNA detected humpback, gray, and blue or fin whale; No MMO in year 2	Closek et al. 2019
eDNA PCR assay	Design assays to detect humpback whale with lab samples, then demonstrate with field samples	Monterey Bay, California Humpback whale	Not reported			Lab testing on known samples: Precision = 44% Recall = 100% False positives for minke and gray whale	Andruszkiewicz et al. 2020
eDNA metabarcoding	Compare eDNA and underwater visual surveys to assess fish communities	Santa Cruz Island, California Marine vertebrates	Environmental conditions not reported			eDNA detected blue and humpback whale	Gold et al. 2021

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Summary of Results	Citation
eDNA metabarcoding	Demonstrate feasibility of eDNA sampling from commercial vessels	Ligurian/Tyrrhenian Seas, Italy Marine vertebrates	Water samples taken at fixed points and after visual sighting of cetacean. MMO conducted when wind strength < 3 Beaufort			Fin whale DNA found at 4 of 16 sampling sites. Fin whale DNA found in samples after 1 of 2 MMO sightings	Valsecchi et al. 2021

Results from eDNA sampling for marine mammals indicate that specific tests can be used to detect a single species with high recall (Andruszkiewicz et al. 2020) or detect multiple baleen whale species in a single sample (Closek et al. 2019; Gold et al. 2021). However, non-detection of eDNA in a sample does not mean that baleen whales are absent from an area and there is a risk of a false negative as a true negative detection. During field trials, eDNA samples failed to detect baleen whales (i.e., false negatives) that were present to visual observers at the same location and time of sampling (Closek et al. 2019; Valsecchi et al. 2021). eDNA methods are looking for the specific cue of shed genetic material in a water sample, and false negative detections are possible due to dilution, degradation of DNA, or movement of eDNA away from the sampling location. False positives are also possible by misidentification of a DNA sequence or by detection of eDNA from a deceased animal. Lastly, eDNA is not capable of delivering real-time information about whale presence with enough temporal or spatial specificity.

B.6 Active Acoustics

Active acoustic monitoring includes using imaging sonars (e.g., fisheries sonars, imaging sonars/acoustic cameras, or echosounders) to detect a marine mammal through the strength of the signal reflected from the encountered target (i.e., the animal's body). A transducer emits a sound pulse and then detects the echoes. The time delay of the echoes can be used to estimate the range and bearings to form two- or three-dimensional images (i.e., echograms). The use of active acoustic for whale detection is summarized with example literature in Table 11.

Many early studies of active acoustics focused on estimating a target strength for individual species (e.g., Dunn 1969; Levenson 1974; Love 1973; Oliver and Kvitek 1984; Au 1996; Miller et al. 1999; Miller and Potter 2001), especially as they vary depending on the tilt angle of the instrument, incident angle of the target, and changes in lung compression and swim speed (see Bernasconi et al. 2013; Geoffroy et al. 2016). This is an important first step in target acceptance/rejection for different species (Bernasconi et al. 2013). Few studies have reported details on detection time, detection range, and detection capabilities, especially under the context of offshore wind energy construction and operation. However, active acoustic monitoring has been used in the United Kingdom to monitor for seals around tidal turbines (e.g., Keenan et al. 2011; Hastie et al. 2019a; Gillespie et al. 2022), including automated detection and tracking (Hastie et al. 2019b; Gillespie et al. 2023). A tradeoff between sonar frequency and ping duration, the maximum detection of seals and small cetaceans was around 50–60 m (Gillespie et al. 2020, 2022).

Table 11. Representative literature review of active acoustic monitoring

Technology Type and Description	Study Objective	Location and Species	Environmental Conditions	Detection Latency	Detection Distance	Detection Capability	Citation
Simrad SP90 (20–30 kHz) and SH80 (110–122 kHz)	Detect whales in relation to seismic surveys	Norway; killer whales	Various conditions (calm, rain, and storms); wind speeds 3–7 m/s; SST 6–7°C	Not stated	Max detection range 1,500 m; reliable detection to 400; max depths to 400 m; reliable depth 200 m	Performance (depth and distance) differed depending on instrument used; TS changed depending on orientation, size, and depth	Knudsen et al. 2008
downward-looking Simrad EK60 38 and 120 kHz split-beam scientific echosounder	Target strengths of Arctic whales and seals	Canadian Beaufort Sea; Bowhead whales	Good weather (for MMO to ID species)	Not stated	Max detection range 2,000 m (in an acoustic duct); typical TS at <300 m	Similar TS variations as other studies; observers critical to ID to species	Geoffroy et al. 2016

Because marine mammals are large reflective targets (i.e., due to their size and air-filled lungs), they have high target strengths, making them ideal for detection using active acoustics. Additionally, sonars do not rely on the animal producing a sound or surfacing for visual detection; therefore, not limited by light level, visibility, and weather conditions. However, the physical properties of the water column, as well as presence of other organisms, can result in high false detection rates (Barlow and Gisiner 2006; Silber et al. 2009; Knudsen et al. 2008). Identifying a target to the species level is also difficult and can be confounded by animals traveling in close proximity, particularly large groups of smaller cetaceans (e.g., dolphins). While maximum detection ranges have been reported out to more than 3,000 m for some large whales (reviewed by Geoffrey et al. 2016), reliable ranges have only been within 500 m (Knudsen et al. 2008), limiting the distance that can be monitored. In some shallow water environments, reverberation clutter can obscure echoes reflecting off whales (e.g., Pyć et al. 2016a, 2016b). No fully automated detection system currently exists (Pyć et al. 2016a; Silber et al. 2009), and a human operator is still required to monitor and confirm any identified targets. Additionally, it can be challenging to receive regulatory approval to introduce more sound into the environment, not only to prevent risk of auditory impairment, but also to avoid using frequencies that may overlap with the vocal ranges of some marine mammal species (Richardson et al. 1995) possibly resulting in a behavioral response or interference with communication (Pyć et al. 2016b; Silber et al. 2009).

Appendix C – Technology Profiles

The technology profiles in this Appendix provide an in-depth examination of the system design and deployment platforms; signal detection and classification algorithms; data delivery platforms; and the impacts of sensor capabilities and settings on sensor performance. These profiles outline approaches for characterizing the performance of the technology system, detector and/or classifier, and data delivery, which may include modeling or field assessment under ideal conditions, and research and development recommendations for furthering the performance of a given technology system. Technologies profiles are included for passive acoustic monitoring (PAM) (Appendix **Error! Reference source not found.**) and infrared (IR) imaging (Appendix C.2).

C.1 Passive Acoustic Monitoring Technology System

C.1.1 Technology System Description

Passive Acoustic Monitoring (PAM) involves the use of one or multiple hydrophones to record underwater sounds in the environment, usually within a specific frequency range. PAM is useful in the marine environment where animals are often easier to hear than see at a multitude of monitoring ranges. It can operate independent of time of day or year, weather conditions, or light level, making PAM an important method for detecting, monitoring, and where necessary localizing, calling marine mammals around offshore wind construction activities.

C.1.1.1 Signal Description

For the purposes of detecting animals that enter a pre-defined area around offshore wind construction activities, a positive detection is defined as an individual marine mammal call detected via PAM sensor. Unique acoustic signals, or “calls” made by marine mammals, typically allow them to be identified to the species level. Most baleen whales produce low-frequency sounds ranging from 10 Hz to 10 kHz (Thompson et al. 1979), with some signals, including those from North Atlantic right whale, reaching frequencies of >20 kHz (e.g., Clark 1990; Matthews and Parks 2021).

C.1.1.2 Sensor Description

When selecting a PAM system for monitoring offshore wind foundation installation, it should have the appropriate dynamic range and sensitivity to detect and classify calls, without saturating when maximum levels are received at close range. Commercial digital recorders typically operate with 16- or 24-bit resolution. Because only 20 bits are effectively used in a 24-bit system and extra storage space required (Madhusudhana et al. 2022), a 16-bit resolution system is often sufficient. Hydrophone sensitivity will vary by manufacturer and model. Generally, for 16-bit systems, combined receiving voltage sensitivity and amplifier gains should be between -165 dB to -145 dB re 1 V/μPa. If higher bit systems are required, combined sensitivities should be between -185 dB to -145 dB re 1 V/μPa. The hydrophone response should be omnidirectional in the horizontal plane to ±5 dB over the specified frequency range. There is not a similar requirement for vertical directionality.

C.1.1.3 Platform Descriptions

This section describes the platforms for sensor deployment, signal detection and classification, and data delivery.

Sensor Deployment Platforms

Below are the typical components, spatial and temporal capabilities, and limitations of the vessel-based (e.g., towed arrays and dipping hydrophones), fixed (e.g., moored), and mobile (e.g., drifting and autonomous) platforms on which a PAM system can be deployed.

Vessel-Based Platforms

Towed arrays, which consist of several hydrophone sensors towed in a variety of configurations behind a vessel, are a widely used tool to monitor marine mammals (Andriolo et al. 2018; Van Parijs et al. 2009). With the appropriate configuration, towed arrays can be used to localize calling marine mammals by measuring relative time difference of arrivals of signals (Van Parijs et al. 2009; Yack 2013) using software developed for public use (e.g., Gillespie et al. 2009; Mellinger 2002).

While a linear towed hydrophone array (i.e., multiple, equally spaced hydrophones) can determine the bearing angle of a detected animal, it is unable to differentiate whether the animal is on the left or the right side of the array (e.g., Barlow and Taylor 2005). Changing course can resolve left-right ambiguity but may not be feasible depending on the platform. Low-frequency signals (e.g., baleen whale) have a less precise bearing than high-frequency signals (e.g., toothed whale). Noise contamination can happen when towed array(s) are deployed less than a kilometer (Thode et al. 2010) from the towing vessel and noise generated by the vessel can mask some marine mammal sounds, especially low-frequency sounds. Vector sensors (e.g., Wade et al. 2006; Miller 2012; Thode et al. 2010; Rone et al. 2012), three-dimensional arrays (e.g., Zimmer, 2013), and nonlinear warping methods (Bonnell et al. 2014; 2020) may allow for range and depth estimation with just one hydrophone, but only for some species and some locations. The typical methodology using time difference of arrival (TDOA) of the signal requires multiple time-synchronized hydrophones to estimate location. Noise reduction designs may be used for surface and sub-surface components to limit noise interference with the hydrophone.

Mainly used for opportunistic data collection, dipping hydrophones are connected to a cable that is then plugged into an onboard digital recorder. The hydrophone is lowered over the side of a boat a few meters below the sea surface and records the signals on the recorder. Dipping hydrophones are typically intended to record marine mammal sounds with concurrent visual observation. However, they cannot be deployed while underway. Additionally, performance limitations such as calibration, variability in sensitivity and directionality, and low signal-to-noise ratios mean additional considerations must be made to be able to use the data for sound analysis purposes.

Fixed Platforms

Fixed platforms include inline vertical moorings and instrument frames and platforms attached on or near the seafloor. Typical components of inline moorings include anchors or weights at the seafloor, an acoustic release, a hydrophone recorder, and flotation that are attached to a line. Bottom-mounted packages work similarly, but the hydrophone sensor is generally located closer to the seabed and connected directly to the anchor or ballast platform or frame.

Stationary platforms that are moored on the seafloor are generally unaffected by metocean forcing (e.g., wind and waves). However, any moorings with a surface or near-surface expression can be impacted by wave action and therefore cannot operate in high sea states. However, this can be addressed with special mooring components designed to specifically reduce system component noise (Baumgartner et al. 2019). And because the spatial position of the platform does not change over time, this provides a consistent spatial point of reference for detection and classification of calling animals with a detection radius that only varies with changes in propagation conditions and ambient noise levels.

The number of hydrophone stations and spatial configuration of an effective PAM array depends on the study area, species, and monitoring questions of interest. Deployment locations may be influenced by the depth-rating of the instrument and oceanographic conditions, such as depth, temperature, substrate type, ambient noise level, and source level, calling frequency and calling depth of the species of interest.

Deployment location and configuration also depend on whether (and which) other species may be present, which may impact detection of the species of interest if the different species overlap in frequency range and have similar calls that might confound automatic detection.

Mobile Platforms

Mobile platforms, including free-drifting or autonomous systems, allow for extended spatial coverage using a single hydrophone unit or multiple hydrophones on a vehicle or towed array. While mobile platforms can extend the spatial coverage, they simultaneously allow gaps in temporal coverage because the platform is moving. Multiple mobile platforms can significantly reduce the number of sensors required to monitor a larger area. Mobile platforms include systems with surface expressions (e.g., free-drifting buoys and wave- or wind-propelled gliders) and fully submerged platforms (e.g., battery- and buoyancy-powered ocean gliders and profiling floats). Both free-drifting and autonomous systems' spatial positions change over time. Free-drifting platforms are largely driven by local currents and wind (Griffiths and Barlow 2015) and autonomous systems move through the water column through buoyancy engines or propulsion. Therefore, neither are designed to collect long-term data at a single location.

Drifting system trajectories cannot be controlled, which may require multiple deployments during monitoring. Alternatively, the movement of autonomous platforms can be coordinated by buoyancy engines and thrusters, but trackline coverage can also be influenced by metocean forcing.

Some platforms can produce self-noise (e.g., propulsion noise) or contribute to flow noise, resulting in masking and limiting monitoring periods. Wind-powered platforms generate significant flow noise while moving through the water (Baumgartner et al. 2021; Mordy et al. 2017), which could compromise the platform's utility for PAM applications. However, most platforms are slow-moving and generate little flow noise (Cauchy et al. 2023). Noise reducing designs may be used for surface or sub-surface components to reduce noise interference with the hydrophone.

Signal Detection, Classification, and Localization Platforms

There are numerous open-source and proprietary automatic detection and classification bioacoustics software programs. Automated classification involves detecting a potential signal of

interest, extracting relevant spectro-temporal features, and classifying the sounds to species, where possible. Detection and classification approaches typically involve using known species calls to train algorithms to detect and classify signals of interest in new data.

Detection and Classification

Common automated detectors and classifiers include amplitude or energy summation (e.g., Helble et al. 2012; Klinck and Mellinger 2011) to identify peaks above background noise in frequency bands of interest; spectrogram correlation (e.g., Mellinger and Clark 2000) and matched filtering (e.g., Stafford et al. 1998), which calculates cross-correlation to construct a detection function using a kernel (defined in the frequency and time domain, respectively); and contour and pitch tracking (e.g., Baumgartner and Mussoline 2011; Gillespie et al. 2013; Mellinger et al. 2011), which generate contour lines that trace tonal sounds. These methods use various thresholds or statistical analyses to assign a call to a known call type.

Model based classifiers include unsupervised (e.g., clustering algorithms, dimension reduction techniques, Gaussian mixture models, dynamic time-warping, Hidden Markov Models) and supervised (e.g., random forest, support vector machines, convolutional neural network machine learning models). Deep learning, a subset of machine learning, has been used increasingly in PAM studies on cetaceans, particularly with convolutional neural networks, which are artificial neural networks that can extract and classify objects from images (e.g., Allen et al. 2021; Bergler et al. 2019; Huang et al. 2016; Ntalampiras 2017; Rasmussen and Širović 2021; Shiu et al. 2020; Zhong et al. 2021).

However, these algorithms are species and call specific and have been tested and trained with specific call libraries. Detection algorithms may not be applicable to new regions, even using the same species and call types, and will need to be developed and tested prior to use under new conditions. Each use may also require different levels of additional verification. Additionally, because animals produce so many different types of sounds, developing algorithms to detect, recognize, and classify multiple acoustic signals can be challenging.

Localization

Vector sensors (e.g., Wade et al. 2006; Miller 2012; Thode et al. 2010; Rone et al. 2012), three-dimensional arrays (e.g., Zimmer 2013), and nonlinear warping methods (Bonnel et al. 2014; 2020) may allow for range and depth estimation with just one hydrophone, but only for some species and some locations. The typical methodology using TDOA of the signal requires multiple time-synchronized hydrophones to estimate location. Noise reduction designs may be used for surface and sub-surface components to limit noise interference with the hydrophone.

Data Delivery Platforms

Because vessel-based towed arrays are connected to equipment onboard the ship, they can collect and deliver real-time data streams. Moored systems are capable of real-time data delivery via cable, iridium satellite link, or VHF antenna. However, data transmission via satellite does not support high data transfer rates, which may (but not always) result in truncated data transfer (e.g., pitch tracks instead of raw acoustic data). For mobile and fixed platforms, the ability to transfer data back to shore have made many systems effective for real-time monitoring (Baumgartner et al. 2013, 2014; Kowarski et al. 2020; Premus et al. 2022).

Other important considerations, especially for real-time detection and classification, are the power consumption and onboard processor capabilities required for real-time processing. While archival PAM requires adequate battery and storage space, real-time processing needs may limit what can be done with currently available fixed and mobile deployment platforms (e.g., both detection and classification in real-time and in some cases on remote systems with the ability to store raw or short spectrograms/pitch tracks to be sent to shore for review and confirmation by a human operator).

C.1.2 Technology System Performance

Below we discuss the considerations (i.e., sensor and deployment, environmental, and biological) that may affect the performance and detection capabilities of a chosen PAM system.

C.1.2.1 Sensor Settings and Deployment Considerations

The performance of a PAM system will differ depending on the characteristics of the hydrophone(s), including frequency sensitivity, which may affect probability of detection. Variation in frequency response (within and across different hydrophones) can affect the detectability of certain species if the frequency of their calls falls outside the optimal sensitivity range. While many hydrophones come with manufacturer calibration information, and should be stable over time, hydrophones need to be calibrated/assessed prior to each deployment. This is especially important if multiple hydrophones are used, as variations among hydrophones may affect the total area in which marine mammals can be detected. The position and configuration of the hydrophone(s), including how they are moored (e.g., the type of attachment cables used) and their position in the water column relative to the surface and bottom, will also affect detection, as mooring components rubbing against each other, surface noise from wind and waves, and surface roughness and bottom substrate can raise ambient levels (e.g., Dekeling et al. 2014).

C.1.2.2 Environmental Considerations

The performance of a PAM system will differ depending on the environment at the deployment location. The environmental conditions at the deployment location(s) can influence how sound propagates through the water, some of which vary spatially and temporally. Major characteristics to be considered include bathymetry, type and thickness of seafloor sediment or hard substrate, properties affecting the speed of sound (e.g., temperature and salinity), depth of the PAM system, and depth of the calling marine mammal.

Additionally, ambient noise levels change spatially and temporally due to contributions from natural sources (e.g., wind, waves, rain), biological contributions (e.g., mammals, fish, invertebrates), and human inputs (e.g., construction vessels, dynamic positioning thrusters, and bubble curtains). For example, while PAM can be used in any sea state, detection and classification decreases dramatically at Beaufort sea states greater than 5 (e.g., Rankin et al. 2020). Thus, a signal to noise level of +3 dB is generally necessary for discrimination of a signal above the basic detection level of any sound above ambient levels at a particular receiver (Dooling and Leek 2018).

C.1.2.3 Biological Considerations of Signal Producer

The ability of a PAM system to detect the presence of a baleen whale depends on whether the animal is calling. The probability of an animal calling depends on many factors, including the

species, the regionally specific population, individual differences, group dynamics, behavior, time of day, time of year, age, sex, reproductive status, availability of prey, and changes in response to anthropogenic noise (e.g. Booth et al. 2017; L. Thomas and Marques 2012). For some animals calling (e.g., NARW) rates can be highly variable (e.g., Parks et al. 2011). Some animals change their sound repertoires over time, further complicating their classifications (e.g., Rice et al. 2022). Thus, call rate estimation, which will impact detection, will require a deeper understanding of the species of interest as well as an additional component of research involving tools beyond the scope of PAM systems, such as paired tagging and visual observations.

C.1.3 Validating Technology System Performance

Prior to monitoring, the PAM system must be calibrated and tested to ensure the sensors are recording data correctly and to estimate the detection range. Calibration allows for the frequency response and system sensitivity of the PAM system to be determined or validated against manufacturer reported values, which may change over time. In addition to model simulations of performance, in-situ testing should be conducted in the region where the PAM system will be deployed and under different noise levels and environmental conditions using frequencies and ranges of interest. This will also allow the PAM system's hardware to be optimized for low self-noise, including the mooring system, and setting the appropriate gain (i.e., essentially an amplifier that increases the signal) levels for the particular environment. In addition to data collection protocols, calibration and testing should be standardized across systems (including assumptions), study sites, and use among operators.

C.1.3.1 Detection Range Estimates

Propagation loss model simulations (e.g. parabolic equation, normal mode) will allow users to estimate the detection range of a signal of interest. This approach uses data about the source level, depth, and frequency characteristics of the target signal; environmental data on the bathymetry, seafloor characteristics, and varying sound speed depth profile along the range from source to receiver; variation in ambient sound levels in the area of interest, and information about the efficiency of the detection and classification system to estimate the probability of detection (Küsel et al. 2011). In-situ field tests can allow for real-world estimates of detection ranges under various conditions (e.g., ambient noise levels) (Marques et al. 2009). The goal in evaluating the system's performance for monitoring around offshore wind foundation installation is to determine how well users can detect a calling baleen whale using this technology, assuming that cue is present. This information can be used to optimize performance and minimize missed detections in real-world applications with the range of background and ambient noise present.

C.1.3.2 Evaluation of Detection and Classification

In addition to evaluating the performance of the PAM system to observe a signal of interest in a particular environment, future efforts will need to include an evaluation of the performance of the automated detection and classification algorithms, if they are being included in the monitoring plan. This is not necessary if a human operator is reviewing spectrograms in real-time assuming there is a rotation plan for operators to reduce fatigue and missed detections. Estimates of the true positives, false positives, true negatives, and false negatives (Figure 1) are used to calculate performance metrics, including precision and recall (which should be maximized) as well as rates of false detection, false positive, false omission, and missed detection (which should be minimized) (Baumgartner et al. 2019; Van Parijs et al. 2021). Evaluations of detection

and classification performance must include consideration of signal to noise ratios at the receiver.

C.1.4 Technology System Validation Case Study

Below is a generalized example of how noise propagation modeling can be used to estimate detection ranges using a PAM system followed by field validation using a low-frequency acoustic projector for selected locations, frequency ranges, environmental conditions, and noise levels to fully characterize a chosen PAM system.

C.1.4.1 Modeling

Transmission loss (TL) is modeled using appropriate models for the region (e.g., parabolic equation, normal mode) along a chosen number of degree bearings (e.g., eight 45° bearings) from a source location out to a maximum range (e.g., 20 km). Inputs to the model include time dependent sound speed profiles, seafloor characteristics (e.g. hard substrate or sediment type and thickness), high-resolution bathymetry, and additional geo-acoustic parameters (e.g., sheer and compressional attenuation and speed). Some of this information can be found in published peer-reviewed literature or on NOAA data repositories. For example, sound speed profiles can be calculated using temperature and salinity data for specific dates and locations modeled (every 1/12 degree) using the Hybrid Coordinate Ocean Model. Two dimensional fields of transmission loss along the range from source to receiver are then extracted from the model with the PAM system at a selected and stationary depth at frequencies of interest (using source levels (SL) for whale calls measured or from published literature), assuming a common/average depth of a calling animal. The results provide range dependent estimates of acoustic energy loss along each bearing surrounding the PAM system. The transmission loss curves can be applied to different ambient noise level conditions (AL; e.g., 90th, 50th, 10th percentiles) to assess how that will impact the ability of a system to detect and classify the signal. Using the passive sonar equation, the maximum transmission loss for a positive detection and classification (with a minimum SNR of +3 dB) would be: $\text{Max TL} = \text{SL} - \text{AL} - 3 \text{ dB}$ (Dooling and Leek 2018; Dooling 2019). The first crossing of the maximum TL number along the TL curve provides the estimated detection range.

C.1.4.2 Field Validation

A field test should be conducted to test the performance of the PAM system in real-world settings. This entails deploying the hydrophone(s) at the depth it will be used during monitoring, suspending a vessel-based calibrated transducer at realistic depth(s) where the animals would be calling, and projecting a known number of recorded sounds with known source levels of interest through a transducer at prescribed distances (to at least, if not beyond, the required detection distance during monitoring).

This approach ensures that the “true” number of simulated calls present during the validation of the PAM system is as accurate as possible. The field test should also be conducted in a region with conditions as close to the target species’ habitat where offshore wind construction activities will take place and under different oceanographic conditions. If active pile driving is not occurring in the test area during the technology evaluation, data modeling from playbacks performed in ambient conditions could be used to assess the effective detection range during higher ambient conditions that includes pile driving.

Source levels should be as representative as possible, which can be determined from the literature. The signal should also be played back at a range of representative source levels, to assess the effect of varying intensity on detection probability. This requires a transducer and vessel-based computer set-up with enough power to project at the appropriate source levels.

Questions that should be answered during this field evaluation include and for each relevant source level:

- How many times was a recorded whale call played in the test area?
- How many whale calls did your sensor record in the test area and at what pre-determined distances to generate a detection function?
- If using an automatic detector/classifier algorithm, how well did it work?
- At what distance were calls reliably detected (i.e., what is the reliable detection range)?
- If human operators are reviewing spectrograms in real-time, is there an adequate rotation plan to reduce fatigue and missed detections?
- How did your detector work across different oceanographic conditions (sea state/bathymetry etc.)?
- How well did the human confirmation component work?
- How fast was a detection confirmed and communicated to proper channels?

Assessing how many baleen whale calls the PAM system captured at different distances will provide the estimate of true positives, which will indicate how well the PAM system performed. The classifier error rate will indicate how well the automatic detector worked after being validated by human screening. The field evaluation can be used to model the probability of the detection function and the reliable detection function range for each species of interest, location, deployment platform of interest, and under different oceanographic conditions. Finally, this includes assessing the human performance in assessing each automatically identified cue to confirm whether the detections were valid, how quickly they were able to communicate the detection, and how much time between a detection and mitigation decision is made and executed (e.g., the in-time detection probability).

C.1.5 Research and Development Needs

C.1.5.1 Sensor Research and Development

While the cost of acoustic sensors has declined significantly in the last decade, state-of-the-art sensors are still costly; thus, establishing a PAM system for monitoring around offshore wind foundation installation may come with large initial expenses. Additionally, sensor performance can change over time and under different environmental conditions, including the biofouling of instruments that are deployed long-term. Regular maintenance of instruments and calibrations are recommended, and field tests should accompany numerical simulations to continuously assess performance especially under different conditions and in different regions.

C.1.5.2 Platforms Research and Development

Sensor Deployment

Sensor deployment can become costly, especially if a PAM system needs to be built piecewise or where multiple systems are needed. Another potential limitation of sensor

deployment is in the recovery and maintenance of a system or systems, especially if the systems need a lot of turn around (i.e., battery replacement, manual data downloads, etc.). Developers might consider investing in systems and methods that allow for easier deployment and recovery.

Signal Detection and Classification

Whale call libraries are region specific and often small, limiting their usefulness as training data for state-of-the-art deep learning methods that require large training datasets. This lack of resources represents a major current gap in the field. Additionally, power requirements for onboard processing are still a challenge, which limits the endurance of current systems. Complex, real-time processing needs for onboard detection and classification require power consumption and onboard processor capabilities that not all fixed and mobile deployment platforms are currently capable of. Low-power algorithms or processing hardware needs are a priority.

Data Delivery

While most sensor platforms present the opportunity for real or near real-time data delivery, there are some limitations. For example, some methods only allow for summary information to be transmitted, which may result in missed or false detections. Data delivery methods should allow for raw or processed data (e.g., short spectrogram or pitch tracks) to be archived and transmitted to shore so a human operator may confirm detections before making mitigation decisions. Some deployment platforms are capable of detection and classification but not in real time. In other cases, real-time detection and classification may not be able to be implemented on a remote system (e.g., drone or AUV).

C.1.6 Conclusions

PAM systems provide a robust way to monitor calling marine mammals under many conditions, including at night and during poor visibility. However, there are some limitations, which have been discussed above. Some of these limitations can be addressed with additional research and development. PAM systems should be considered within a suite of monitoring tools to account for non-calling individuals, irregular cue rates, and high ambient noise areas. The performance of PAM systems should be regularly assessed in the field, especially under different environmental conditions. Assessing real-world performance also includes an understanding of how animals may change their calling behavior in the presence of vessel traffic and pile driving around offshore wind activities. It is critical to evaluate the efficacy of proposed PAM system/configuration(s) and decide for each species, call type, and situational conditions, what level of certainty is acceptable with respect to initiating and/or ceasing pile driving activities.

C.2 Infrared Imaging Technology Systems

C.2.1 Technology System Description

Infrared Camera (IRC) systems using single or multiple components involve the use of thermal cameras to capture images using infrared radiation. The process is similar to standard camera imaging, except IRC systems detect a different range of wavelengths (8–14 μm) and do not require visible light to function. Because it can operate day and night, is not masked by ambient light, and can detect animals at greater distances than human observers, this makes IRC an important method for detecting, monitoring, and localizing surfacing marine mammals around offshore wind construction activities.

C.2.1.1 Signal Description

A cue can be a body part of an animal or its breath, but to be detected, it must be at or above the surface of the water (Baldacci et al. 2005) and have greater thermal energy than the water. For the purposes of detecting animals that enter a pre-defined area around offshore wind construction activities, a positive detection occurs when an image of the cue is captured by the IRC and exceeds a pre-defined threshold of the difference between the thermal energy or heat emitted by the body part of a surfacing animal (e.g., dorsal fin, tail, breach) or its respiration (i.e., blow) and the temperature of the surrounding water.

C.2.1.2 Sensor Description

When selecting a system for monitoring around offshore wind construction activities, the IRC should have the appropriate concurrent ocean coverage, be mounted at the appropriate elevation, and be stabilized on the platform to avoid blurry imaging. Additionally, the IRC system should optimize for a larger focal length, which has a smaller field of view but greater spatial resolution, allowing for cues to be more accurately imaged and detected. Although scanning can achieve a 360° field of view, coverage of the full field of view is not simultaneous and can therefore result in missed detections. Multiple cameras with larger focal lengths and narrower fields of view (e.g., 6–15°) will achieve the desired ocean coverage (e.g., >50%) thereby increasing the probability of detection (Verfuss et al. 2018).

C.2.1.3 Platform Descriptions

This section describes the platforms for sensor deployment, signal detection and classification, and data delivery.

Sensor Deployment Platforms

Below are the typical components, spatial and temporal capabilities, and limitations of handheld, fixed (e.g., vessel-based, shore-based), and mobile (e.g., gliders, floats) platforms on which IRC systems can be deployed.

Handheld Platforms

Handheld platforms involve positioning an IR camera at eye-level by a human operator onboard a vessel. The higher the human operator is on a vessel, the further the distance at which they can scan. The operator scans the horizon by panning the camera left and right, similar to the process of scanning with the unaided eye or binoculars. During panning however, the operator will miss detections outside their field of view. Handheld

devices do not work inside the vessel's wheelhouse, or in areas with heat source interference. Additionally, surveying for thermal images in real-time can lead to survey fatigue, resulting in missed detections (Boebel and Zitterbart 2013). As a result of the constraints associated with handheld IR cameras, this platform is not a recommended platform for detecting marine mammals around offshore wind construction activities (Boebel and Zitterbart 2013).

Fixed Platforms

Fixed platforms include shore-based platforms (e.g., attached to a lighthouse or on a cliff) and vessel-based platforms (e.g. mounted on a vessel). Fixed platforms provide more stability and, in some cases (i.e., shore-based), do not require the use of a gimbal for stability (e.g., Zitterbart et al. 2020). However, vessel-mounted cameras, especially if the vessel is moving requires active stabilization. IRC systems also can be attached to fixed buoys but will require stabilization and may not be high enough to be used for detecting baleen whales at the surface.

Mobile Platforms

IRC systems can be deployed on gliders and profiler floats. However, these platforms will not be high enough to be used for detecting marine mammals at the surface during offshore wind construction activities. However, these can be used to gather baseline information of species composition. Mobile aerial drones present an opportunity for surveying at an adequate height provided stabilization (i.e., gimbal) is possible (e.g., Horton et al. 2019).

Signal Detection and Classification Platforms

Some studies have relied on manual image analysis (e.g., Guazzo et al. 2019; Perryman et al. 1999). IRC systems can also be reviewed in real-time using a human operator; however, missed detections have been reported due to PSO fatigue (Boebel and Zitterbart 2013). There are proprietary automatic detection and classification software programs (e.g., Horton et al. 2017, 2019) and custom software (e.g., Richter et al. 2024; Zitterbart et al. 2013, 2020). These algorithms use a machine learning based classification system to identify significant thermal anomalies relative to the surrounding water surface. However, automatic IR detections include true and false positives and must be confirmed by a human operator prior to triggering any decision-making during offshore wind construction activities (Zitterbart et al. 2013). Due to compression, captured IR images also tend to have blurry edges (Hou et al. 2022; Kaarna et al. 2007; Bazhyna 2009), which may reduce detection capability around frame edges.

Data Delivery Platforms

For fixed platforms, IRC systems can be directly connected to a computer that is running automatic detection software (e.g., Smith et al. 2020; Zitterbart et al. 2013) for real-time monitoring and mitigation. For mobile platforms, the data need to be downloaded on a field computer for processing, which may add additional time to the data delivery process (e.g., Schoonmaker et al. 2008).

C.2.2 Technology System Performance

Below we discuss the conditions that may affect the performance and detection capabilities of a chosen IRC system.

C.2.2.1 Sensor Settings and Deployment Considerations

The performance of an IRC system will differ depending on the characteristics of the thermal imaging device, including field of view, sensor size, focal length, sample rate, whether the system is cooled, and deployment considerations, including height and stabilization, all which may affect probability of detection.

The field of view (i.e., the size of the space that can be imaged) is determined by the size of the thermal sensor and the focal length of the lens. The relationship can be calculated as: $\text{fov} = 2 \cdot \tan^{-1}(h/2 \cdot f)$, where fov is field of view, h is sensor size, and f is the focal length. The thermal sensor size is determined by the number and size of the pixels of the thermal sensor, whereas focal length is determined by the distance from the lens's optical center to the camera's image plane when the lens is focused at infinity. Sensor size, focal length, and field of view are reported by the manufacturer. However, these settings and capabilities can be selected to optimize for marine mammal monitoring and mitigation.

Thermal sensor resolution (i.e., number of pixels) has increased since earlier studies that were less than 320x240 pixels (e.g., Cuyler et al. 1992; Graber et al. 2011) to more than 640x480 pixels (Guazzo et al. 2019; Horton et al. 2017; Seymour et al. 2017). Sensor size continues to increase, with some sensors reaching 2048x1536 pixels. However, the size of each pixel (i.e., pixel pitch) directly impacts spatial resolution. Smaller pixels have a higher spatial resolution, but receive less IR radiation, and are less able to detect small differences in temperature (i.e., reduced thermal sensitivity). Thermal resolution is also affected by noise equivalent temperature difference (NETD), the smallest temperature difference that the sensor can perceive (lower NETD has higher sensitivity). The size of individual pixels in thermal sensors have rarely been reported; however, some have reported the degree relative to the field of view that each pixel captured, ranging from 0.01° to 0.05° in the horizontal and 0.3° in the vertical (Richter et al. 2024; Smith et al. 2020; Zitterbart et al. 2013, 2020).

Previous studies reported lens focal lengths ranging from 18 mm to 200 mm (Graber et al. 2011; Guazzo et al. 2019; Horton et al. 2017, 2019; Richter et al. 2024). However, many studies did not report lens focal length. A shorter focal length offers a wider field of view, however spatial resolution decreases proportionally. Increasing focal length will allow for a greater spatial resolution during monitoring, albeit within a smaller angular field of view. These studies also reported narrow (6–15°; Cuyler et al. 1992; Guazzo et al. 2019; Richter et al. 2024; Zitterbart et al. 2013, 2020) to mid-range (25°; Graber et al. 2011; Horton et al. 2017; Richter et al. 2024) horizontal fields of view. Vertical fields of view were not often included; however reported values ranged from 5° to 19° (Guazzo et al. 2019; Horton et al. 2017; Smith et al. 2020; Zitterbart et al. 2013, 2020).

Cooled systems reduce thermal noise below the level of thermal radiation of an object of interest whereas uncooled systems use a material with temperature dependent resistance (Verfuss et al. 2016, 2018). Older cooled IRC systems used liquid nitrogen (Cuyler et al. 1992) but newer systems use semiconductor detectors (e.g., Sterling coolers) to cool the IRC systems to 84°K (e.g., Smith et al. 2020; Zitterbart et al. 2013, 2020). Cooled IRC systems have a better signal-to-noise ratio but come at a higher cost, whereas uncooled systems are lighter and cheaper. Thermal resolution differs by a factor of 3–5 between cooled and uncooled systems (Verfuss et al. 2018) with uncooled systems having reduced detection capabilities as a result. However, IRC systems using both cooled and uncooled systems have been used to successfully detect signals from surfacing whales (e.g., Cuyler et al. 1992; Horton et al. 2017; Richter et al. 2024; Zitterbart et al. 2013).

Regarding the deployment platform, the height of the camera and the amount of stabilization on the platform will have a significant impact on detection probability. Height gives an IRC system a greater vertical field of view. Studies reporting the height of IRC systems have ranged from 2 to 52 m above sea level (Cuyler et al. 1992; Graber et al. 2011; Guazzo et al. 2019; Horton et al. 2017; Richter et al. 2024; Smith et al. 2020; Zitterbart et al. 2013, 2020) with an average height of 15.6 m. For mounted systems, the suggested elevation is between 50 and 100 m and at certain heights the system cannot resolve near-field detections (Hou et al. 2022). For example, a system mounted at 800 m was not able to detect marine mammals closer than 5000 m. Stabilization on land-based platforms can be achieved with mounting devices such as tripods. However, vessel-mounted and mobile platforms need active stabilization with a gimbal. Gimbals measure and compensate for a vessel's roll and pitch using a gravitational sensor or gyroscopes (Verfuss et al. 2016). For vessel-mounted platforms, rolling seas may exceed the limit of stabilization by the gimbal (e.g., Smith et al. 2020). This will result in periods where automatic detection performs poorly (e.g. increased false positives and decreased true positives).

Sample rate (or frame rate) should also be considered, as the number of images captured during a monitoring period will involve a tradeoff between storage size, processing power, and the ability to detect a moving animal. Previous studies have reported 5–30 Hz (or frames per second) (Cuyler et al. 1992; Graber et al. 2011; Guazzo et al. 2019; Horton et al. 2017, 2019; Richter et al. 2024; Smith et al. 2020; Zitterbart et al. 2013, 2020).

C.2.2.2 Environmental Considerations

While sensor settings can greatly impact detection, weather and oceanographic conditions (e.g., temperature, visibility, sea state, humidity, glare, light level, aerosols, snow) can also impact detection probability and automatic detection performance (Beier and Gemperlein 2004; Verfuss et al. 2016, 2018). Graber et al. (2011) defined ideal conditions as clear skies, calm seas, and wind speed (0–4 m/s). Ideal operating conditions have been summarized as a Beaufort sea state (BSS) ≤ 4 , winds < 17 – 21 knots, waves < 6 – 8 feet, whitecaps common with little to no spray, and little to no fog or rain.

- **Temperature:** While warm temperature could impact thermal sensitivity by reducing the contrast between whales and their surroundings (Verfuss et al. 2018), Horton et al. (2017) showed that blows, dorsal fins, flukes and rostrums in tropical and sub-polar regions displayed similar magnitude brightness temperature anomalies with blows appearing 3°C warmer than surrounding waters. Horton et al. (2019) noted, however, that uncooled sensors, are sensitive to temperature and environmental conditions and that built-in auto-corrections can cause large shifts in the recorded temperature coming from the cue. Similarly, Zitterbart et al. (2013, 2020) showed that whale blows were perceptible in more than 70% of thermal images up to 3 km distance with sea surface temperatures ranging from 10 – 25°C , and atmospheric temperatures from 12 – 21°C . Graber et al. (2011) also noted that at above 60° incidence angle between the camera and sea surface, sky temperatures begin to dominate over sea surface temperatures resulting in an apparent decrease in sea surface temperature with increasing incidence angles, especially from 50 – 70° .
- **Visibility:** Low visibility caused by fog or rain has the potential to reduce the distance at which whales can be detected. Heavy rain compromises IRC performance by masking cues and fog has a strong impact on detection probability (Richter et al. 2024; Verfuss et al. 2018). Zitterbart et al. (2020) reported detection distances dropped to less than 500 m with visibility less than 5 km, whereas when visibility was greater than 7 km, detection distances reached 2 km.

- **Sea State:** Rough seas with white caps and breaking waves can hinder the detection of marine mammals by increasing IR clutter (Baldacci et al. 2005; Graber et al. 2011; Verfuss et al. 2018) leading to false detections. Baldacci et al. (2005) reported decrease in effectiveness when sea states were above 2 or 3, greatly declining at Beaufort sea state greater than 5. Others suggested that Beaufort sea state <3 are optimal for maximizing detection distances (Smith et al. 2020; Zitterbart et al. 2020), although detections of blows have been made up to Beaufort sea state 7 (Zitterbart et al. 2013). Additionally, strong wind will dissipate whale blows quicker thus reducing detectability (Richter et al. 2024). Finally, waves can also exceed the limit of mechanical stabilization by a gimbal, which creates frames that cannot be processed through automatic detectors (Smith et al. 2020).
- **Humidity:** High relative humidity may attenuate signals in thermal images. Michel (2015) reported no effect of humidity up to 91% and Zitterbart et al. (2020) reported no effect on perceptibility under relative humidity ranging from 60-90%. However, Baldacci et al. (2005) reported that some IR systems fail when covered by condensed water vapor.
- **Glare:** Glare can impact detection, especially in the higher frequency IR bands (Verfuss et al. 2018) because it can resemble warm anomalies, rendering the detection of blows less likely (Zitterbart et al. 2013). Glare can also impact detection at incidence angles greater than 60° (Graber et al. 2011).
- **Light Level:** While IRC systems can be used at any time of day, they are typically more effective at night, when there is less reflected radiation from the sun and sky than during the day (Verfuss et al. 2018; Zitterbart et al. 2013). This difference is reflected in a decreased probability of detection from night to day (Zitterbart et al. 2020).

C.2.2.3 Biological Considerations of Surfacing Animal

A cue may include a body part or blow. However, for an animal to be detected by an IRC system, it must be at or above the surface with high enough signal to noise (SNR) ratio of body/blow temperature compared to the water. Water is not transparent to thermal radiation and IRCs cannot see even a few microns below the surface (Baldacci et al. 2005). Because IRCs will only detect animals at or above the surface, long deep-diving animals (e.g., sperm and beaked whales) may have very few opportunities for detection even when they are within a monitoring zone.

Depending on the species, the body or blow of an animal might be more visible. For example, killer whale dorsal fins protrude above the water (see: Graber et al. 2011) and may be more detectable than a small animal or animal with no dorsal fin (e.g., gray whale). For some species, a whale's blow might be several meters high (e.g., blue whale) and remain visible for several seconds depending on weather conditions (Verfuss et al. 2016, 2018). However, detecting smaller, low-surfacing odontocetes and sea turtles is likely not possible. Species identification is also possible, unless there is a strong defining characteristic (i.e., killer whale dorsal) that can be captured with IRC.

C.2.3 Validating Technology System Performance

C.2.3.1 Detection Range Estimates

The probability of detection during monitoring should be close to 1 across the entire monitoring zone if all animals are to be detected (Verfuss et al. 2018). However, because the probability of

detection is influenced by sensor settings and platform, environmental conditions, and the biology of the animal of interest, it will be important to identify how each influences detection probability.

While a survey area should increase linearly with distance, studies have shown that IRC detection (i.e., signal-to-noise ratio) decreases with distance (Beier and Gemperlein 2004; Richter et al. 2024). IRC systems with uncooled sensors often result in shorter detection distances than those with cooled sensors.

Thus, there needs to be a distinction between the maximum distance at which an animal can be detected by the IRC system and the reliable detection range (RDR), which is the distance with the peak number of detections. This peak can be considered the range at which all whale cues (i.e., a body or blow at the surface) are detected; that is, the probability of detection is assumed to be equal to 1 (e.g., Guazzo et al. 2019; Baile and Zitterbart 2022; Richter et al. 2024). After the peak number of detections, the IRC system will begin to miss detections.

The probability of detection also depends on the species and environmental conditions, especially the conditions in which construction activities might take place. For example, different species have different blow characteristics and surface behaviors that influences their ability to be reliably detected (e.g., Perryman et al. 1999; Richter et al. 2024; Zitterbart et al. 2020). Behavior may also change based on time of day, season, location, life stage, or other factors that can change detection probability.

C.2.3.2 Evaluation of Detection and Classification

Cue detection should be evaluated in terms of how well the IRC system performs and how well the automated detector performs. The IRC system performance should be tested against a real-world scenario where the true number of whales is known. In the absence of a “true” number the IRC system can be evaluated against the performance of a protected species observer (see Zitterbart et al. 2020); however, this will only be an assessment of how much IRC systems perform relative to human observers. Previous studies have found the conditional probability that detections made by marine mammal observers were also made by the IR system, and vice versa, ranged from 20% to 50% (Smith et al. 2020; Zitterbart et al. 2020).

The automated detection algorithm should be evaluated against a scenario where the true number of whales detected by the IRC system is known. In this case, a human analyst would go through the images captured by the IRC system and manually annotate all the detections. Metrics comparing the performance of the automated detection algorithm would include true positive (the IRC system detected the cue), true negative (there was no cue to detect), false positive (the IRC system detected a cue that was not real), and false negative (the IRC system missed the cue).

Because the sea is in constant motion, automatic detectors will result in many false detections, especially in rougher sea states (Graber et al. 2011). Thresholds can be applied to reduce false detections; however, for monitoring and mitigation, the probability of detection during monitoring should be as close to 1 as possible (zero missed detections). Therefore, the threshold for detection should be low enough to reduce missed detections, which means that the use of a human will be necessary to view captured images that an algorithm flagged as possible detections.

C.2.4 Technology System Validation Case Study

A field test of performance validation should include a real-world scenario where whales are present and can be captured not only by the IRC system but by other detection modalities so that the system can be validated. This requires knowledge of how many whales are in the test area. This approach would include land and vessel-based human observers conducting visual surveys, aerial surveys, passive acoustic monitoring, and any other additionally available methods to ensure that the “true” number of whales present during the validation of the IRC system is as accurate as possible. This field test can be conducted with the species of interest or with a proxy species that has similar behaviors to the target species but is found more reliably in greater numbers. The field test should also be conducted in a region with conditions as close to the target species’ habitat where offshore wind construction activities will take place and under varying oceanographic conditions.

Questions that should be answered during this field evaluation include:

- How many cues did the IRC system observe in the test area?
- How well did the automatic detector algorithm work?
- At what distance were whales reliably detected?
- How well did the human confirmation component work?
- How fast was a detection confirmed and communicated to the proper channels?

Assessing how many cues the IRC system captured will provide the estimate of true positives, which will indicate how well the IRC system performed. The classifier error rate will indicate how well the automatic detector worked after being validated by human screening. The field evaluation can be used to model the probability of detection function and the reliable detection function range for each species of interest, location, deployment platform of interest, and under different oceanographic conditions. Finally, this includes assessing the human performance in assessing each automatically identified cue to confirm whether the detections were valid, how quickly they were able to communicate the detection, and how much time between a detection and mitigation decision is made and executed (e.g., the in-time detection probability).

C.2.5 Research and Development Needs

C.2.5.1 Sensor Research and Development

One of the biggest hurdles for IRC systems are the costs associated with high-resolution sensors, especially cooled systems. Because multiple sensors with larger focal lengths and narrower fields of view are desirable for maximizing desired ocean coverage and increasing the probability of detection, these costs add up.

C.2.5.2 Platforms Research and Development

Sensor Deployment

Thus, the biggest hurdle for sensor deployment is the cost of the equipment required. For example, drone-deployable IR systems are much more expensive than systems that can be mounted on a vessel (Horton et al. 2019). Cooled IRC systems cost more than uncooled systems. And stabilization equipment can also be expensive, especially when using drones capable of carrying gimbaled IRC systems during long-distance flights

(e.g., Horton et al. 2019). This is an area where monitoring technology developers might consider investing to reduce costs but not at the cost of performance.

Signal Detection and Classification

Few public research studies have used automatic detection algorithms to identify whales, either using proprietary (e.g., Horton et al. 2017, 2019) or custom (e.g., Richter et al. 2024; Zitterbart et al. 2013, 2020) software. This is an area where standardization in analyses could be developed as well as open-source code sharing. Additionally, automated detectors currently available are not able to identify whales to the species level. Future algorithms could invest in incorporating Bayesian approaches that provide probabilities of species identification based on time of year, location, and potentially different cues.

Current technology validations assume that all animals at the surface will be detected. There is also a need to invest in methods to determine what proportion of animals were missed because they were not at the surface. This requires additional multi-modal testing but will be important for getting as close to a probability of detection of 1 for any animal within the monitoring zone whether or not they display a specific cue.

Data Delivery

Very few research studies have used IRC systems with real- or near-time data delivery systems (e.g., Smith et al. 2020; Zitterbart et al. 2013). In addition to automated detection algorithms, this presents a large opportunity for developers for making more real-time ready IRC systems.

C.2.6 Conclusions

IRC systems provide a robust way to monitor surfacing marine mammals under many conditions, regardless of available light and background noise levels. However, this technology has limitations as discussed above. Some of these limitations can be addressed with additional research and development. Because IRC can only detect whales at or above the surface, a combination of different detection modalities will improve overall detection probability for real-time monitoring (e.g., Richter et al. 2024; Verfuss et al. 2018). The performance of IRC systems should be assessed not only in the validation phase but during operations, especially if environmental conditions change and as equipment ages. Assessing real-world performance also includes an understanding of how animals may change their surfacing behavior in the presence of vessel traffic and pile driving around offshore wind activities. It is critical to evaluate the efficacy of proposed IRC configuration(s) and decide for each species what level of certainty (e.g., missed detections) is acceptable with respect to initiating and/or ceasing pile driving activities.

Appendix D – Considerations Related to Design Specifications

An example checklist of design considerations for passive acoustic monitoring (Table 12) and infrared imaging (Table 13) are provided in this Appendix. The purpose of this Appendix is to provide an example of design specifications of a monitoring system that can be summarized in a simple to use checklist. This checklist could be used by a third party as in Phase 2 of the potential framework (Section 4.2) to assess whether a monitoring system has the technical capabilities to meet the proposed monitoring objectives. The checklist does not need to have only ‘Yes’ responses for the monitoring system to be technically capable. Rather, the responses to the checklist allow a third party to quickly evaluate the capabilities of a monitoring system in an easy to use and easy to understand format.

Table 12. Example checklist of design specifications for a passive acoustic monitoring system.

Design Specification; Passive Acoustic Monitoring	Yes	No	Reference Literature
Sensor			
Hydrophone is omnidirectional including is part of a vector sensor			Robinson et al. 2014; van der Schaar et al. 2017; Verfuss et al. 2018; Browning et al. 2020; Van Parijs et al. 2021; Madhusudhana et al. 2022; Pavan et al. 2022
Sample rate capable of twice the maximum frequency of signal(s) of interest			
Sampling is continuous (not duty-cycled)			
Sensitivity is high enough to detect cue above background noise			
Gain is appropriate for environment (not so high that signal is diluted/clipped?)			
Bit depth functionally high without requiring additional storage			
Hydrophone has been or will be calibrated			
Platform			
Hydrophone at sufficient depth to avoid surface noise			van der Schaar et al. 2017; Baumgartner et al. 2019; Van Parijs et al. 2021; Cauchy et al. 2023
Detection range can reach required monitoring distance			
Noise reduction design for surface and sub-surface component(s) or moving platforms			
Platform easy to access for replacing batteries, servicing, downloading data			
Signal Detection and Classification			
Call library available for signal(s) of interest			Booth et al. 2017; van der Schaar et al. 2017; Browning et al. 2020; Oswald et al. 2022; Madhusudhana et al. 2022; Pavan et al. 2022; Silber et al. 2023
Automatic detector available for signal(s) of interest			
Description of methods and training data used for development of automated detector/classifier			
Adequate power for data acquisition and analysis			
Adequate processor for automated detection and classification analysis			
Ability to classify to broad taxonomy			
Ability to classify to species			

Design Specification; Passive Acoustic Monitoring	Yes	No	Reference Literature
Data Delivery			
Capable of real-time or near real time delivery			van der Schaar et al. 2017; Van Parijs et al. 2021;
Adequate storage space for saving raw and processed data			Gannon et al. 2022
Human needed to confirm detection			

Table 13. Example checklist of design specifications for an infrared camera system.

Design Specification; Infrared Imaging	Yes	No	Reference Literature
Sensor			
Camera operates within IR wavelengths (8–14 μm)			Cuyler et al. 1992; Baldacci et al. 2005; Graber et al. 2011; Zitterbart et al. 2013, 2020; Horton et al. 2017, 2019; Verfuss et al. 2018; Guazzo et al. 2019; Smith et al. 2020; Richter et al. 2024
Can provide continuous spatial coverage (non-panning)			
Sample rate high of 0.2 frames per second to capture cue			
Focal length reaches maximum distance of monitoring zone			
Field of view can capture required radial coverage (e.g., 0–360°)			
IR sensor has been or will be calibrated			
Image compression will not be used			
Platform			
Height necessary to reach required monitoring distance			Cuyler et al. 1992; Zitterbart et al. 2013, 2020; Graber et al. 2011; Horton et al. 2017; Verfuss et al. 2016; Smith et al. 2020; Richter et al. 2024
Can be stabilized if on non-stationary platform (i.e., gimbal)			
Platform easy to access for downloading data			
Signal Detection and Classification			
Automatic detector available for signal(s) of interest			Zitterbart et al. 2013, 2020; Horton et al. 2017, 2019; Richter et al. 2024
Description of methods and training data used for development of automated detector/classifier			
Adequate power for data acquisition and analysis			
Adequate processor for automated detection and classification analysis			
Ability to classify to broad taxonomy			
Ability to classify to species			
Data Delivery			
Capable of real-time or near real time delivery			Zitterbart et al. 2013; Horton et al. 2017; Verfuss et al. 2018; Smith et al. 2020
Lossless compression			
Adequate storage space for saving raw and processed data			
Human needed to confirm detection			

Appendix E – Example Field Characterization Test Matrix Template

This Appendix provides an example table that could be used to describe the testing conditions for a field characterization test proposed under the potential framework (Section 4.2). Example tables are included for passive acoustic monitoring (Table 14) and infrared imaging (Table 15). The purpose of these tables is to provide an example of how to summarize the plans for a field characterization test under the potential framework described in Section 4.2.

Table 14. Example field characterization test matrix for passive acoustic monitoring

Test Condition; Passive Acoustic Monitoring	Response	Additional Notes
Location		
Geographic location		
Coordinates		
Distance from future construction		
Timing		
Estimated month(s) of tests		
Dates of expected construction		
Number of tests		
Conditions (expected range of tests)		
Temperature-Salinity profiles available?		
Bathymetry		
Type of sea floor sediment		
Thickness of sea floor sediment		
Ambient noise profiles available?		
Target		
Species of interest		
Cue of interest		
Test cue (natural or artificial)		
Expected cue rate from literature		
Biological considerations that might affect cue		
Depth range of test cues		
Distance range of test cues		
Sensor		
Number of hydrophones		
Sample rate		
Bit rate		
Sensitivity		
Calibration curve available		
Modeled detection radius of each hydrophone		
Combined detection radius of all hydrophones (if using multiple)		
Platform		
Number of platforms		
Type(s) of platform		
Depth of platform		
Duration of battery		
Signal Detection and Classification		
Spectrogram assessment plan or automatic detection software citation		
Classification software citation (if different)		

Test Condition; Passive Acoustic Monitoring	Response	Additional Notes
Off the shelf or bespoke?		
Validated for cue?		
Validated for species of interest?		
Will a human be used to confirm detection and classification?		
Expected detection function from literature		
Power needed?		
Onboard processor capabilities required?		
Data Delivery and Storage		
Real-time or near real time available?		
Ability to store raw data?		
Ability to store detections and classifications?		
Expected time to data delivery		
Expected time to confirm sighting		
Expected time relay confirmed sighting to POC		
Data storage (archival) plan – full continuous recordings (NCEI) and detections (PACM)		
Metadata format		

Table 15. Example field characterization test matrix for infrared imaging

Test Condition; Infrared Imaging	Response	Additional Notes
Location		
Geographic location		
Coordinates		
Distance from future construction		
Timing		
Estimated month(s) of tests		
Dates of expected construction		
Number of tests		
Conditions (expected range of tests)		
Temperature		
Visibility		
Sea state		
Humidity		
Glare		
Day/Night/Both		
Target		
Species of interest		
Cue of interest		
Test cue (natural or artificial)		
Expected cue rate from literature		
Biological considerations that might affect cue		
Distance range of tests		
Sensor		
Number of cameras and types if different models		
Sample rate (frames per second)		
Focal length of camera		
Field of view of camera		
Total ocean coverage (%)		
Cooled or uncooled		
Platform		
Number of platforms		
Type(s) of platform		

Test Condition; Infrared Imaging	Response	Additional Notes
Height(s) of platform		
Type of stabilizer used and specification, if applicable		
Signal Detection and Classification		
Automatic detection software citation		
Off the shelf or bespoke?		
Validated for cue?		
Validated for species of interest?		
Will a human be used to confirm?		
Expected detection function from literature		
Power needed?		
Onboard processor capabilities required?		
Data Delivery and Storage		
Real-time or near real time available?		
Ability to store raw data?		
Expected time to data delivery		
Expected time to confirm sighting		
Expected time to relay confirmed sighting to POC		
Data storage (archival) plan		
Metadata format		

Appendix F – Example Field Characterization Recommendations

Information in this Appendix provides an example of a summary of performance metrics from a field characterization test. A technology-agnostic summary table is provided in Table 16 as an illustrative example of the types of information that could be included to support the quantification of the key performance metrics.

Table 16. Example summary of performance metrics calculated from field characterization testing

Results Summary	Response	Additional Notes
Field Characterization		
Number of true or estimated cues from field validation		
Method(s) employed if cues are simulated, including source levels/characteristics of simulated cues and citations if based on literature		
Tested distance from cues to sensor		
Sensor performance (requires human validation)		
Number of signals correctly marked as a detection (True pos)		
Number of signals correctly marked as not a detection (True neg)		
Number of signals incorrectly marked as true (False neg)		
Number of signals incorrectly marked as false (False pos)		
Precision: Ratio of true positives to true positives and false positives		
Recall: Ratio of true positives to true positives and false negative		
Detector performance		
Number of signals automatic detector correctly marked as a detection (True pos)		
Number of signals automatic detector correctly marked as not a detection (True neg)		
Number of signals automatic detector incorrectly marked as true (False pos)		
Number of signals automatic detector incorrectly marked as false (False neg)		
Precision: Ratio of true positives to true positives and false positives		
Recall: Ratio of true positives to true positives and false negatives		
Classifier performance		
Number classifier correctly identified species/cue/signals (True pos)		
Number classifier incorrectly marked species/cue/signals (False pos)		
Detection Range		
Reliable detection range (RDR; peak detection)		
Maximum detection range		
Data Delivery		
Average time required for automatic detector to detect cue		
Average time required for delivery of detection to shore/human analyst		
Average time required for human verification		
Detection Latency: Total time between detection, confirmation, and communication of detection to POC		
Planned operational time		
Actual operational time		
System Reliability: Ratio of planned over actual operational time		

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov