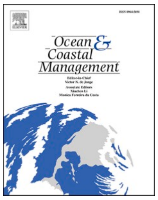
Ocean and Coastal Management 276 (2026) 108147

Contents lists available at ScienceDirect

Ocean and Coastal Management

journal homepage: www.elsevier.com/locate/ocecoaman



A call to standardize metrics for monitoring baleen whales near marine construction activities

Angela R. Szesciorka [a,*], Mark Severy [a], Kristen Ampela [b], Cris Hein [b], Michael Richlen [a], Joseph Haxel [a], Jeff Clerc [b]

[a] Pacific Northwest National Laboratory, 902 Battelle Blvd., Richland, WA, 99354, USA
[b] National Laboratory of the Rockies, 15013 Denver W Pkwy, Golden, CO, 80401, USA

ABSTRACT

Effective monitoring is necessary to protect marine mammal species during the construction of offshore infrastructure. The tools for detecting or monitoring marine mammals span traditional (e.g., visual observers, optical cameras), to newer (e.g., passive acoustic monitoring, infrared cameras, tags), and emerging (e.g., satellite imagery, environmental DNA, dimethyl sulfide concentration) technologies. Some are better suited for use during offshore development; however, peer-reviewed literature does not typically evaluate and report on the performance of these various technologies. We define a minimum set of metrics related to efficacy (i.e., confusion matrix, precision and recall, probability of missed mitigation), detection range (i.e., maximum and reliable detection range, spatial resolution), and data delivery (i.e., detection latency, system reliability, temporal resolution) that we recommend are needed to assess the utility of monitoring technologies for this purpose. Following a literature review of relevant studies, we highlight which publications reported these metrics and used multiple technologies to compare relative performance. We also emphasize the benefits of multi-modal approaches and recommend performance assessments through modeling or large-scale collaborative field testing. These metrics will standardize data collection, reporting, and analysis; promote consistent and comparable results; and foster collaboration among developers, regulatory agencies, and scientists. This may lead to the co-development of technology that achieves multiple goals, has greater application, and can answer research questions while collecting data to fulfill permitting requirements. These metrics may also inform decisions on what systems regulatory agencies might consider using and reduce monitoring costs, which is critical to support the marine sector's rapid growth alongside marine mammal conservation.

1. Background

As the ocean-based economy rapidly expands, energy needs are growing (Jouffray et al., 2020). The development of offshore infrastructure is necessary to support this growth in the marine sector (Novaglio et al., 2022). However, infrastructure development can result in the overlap of human activities with wildlife, including rare or protected species during construction and operation (Brill et al., 2025). The overlap, which may result in vessel strikes or exposure to noise during construction (National Marine Fisheries Service, 2024), involves marine mammal species that are legally protected in many countries worldwide. For example, marine mammals are protected by the Habitats Directive and Marine Strategy Framework Directive in the European Union and in the United States (U.S.) through the Marine Mammal Protection Act and Endangered Species Act. To minimize or avoid potential impacts from offshore development, which can halt or limit development activities, monitoring and mitigation plans are required by developers prior to start of the project. These plans typically include the use of monitoring (using visual observers or technology-aided observations) of project-specific clearance and shutdown zones to detect threatened or endangered marine mammal species. Guidelines and regulators vary by country. In the U.S., project approvals and authorizations are administered by the Bureau of Ocean Energy Management and National Oceanic and Atmospheric Administration's (NOAA's) National Marine Fisheries Service.

Worldwide, numerous technologies are available for marine mammal monitoring (Würsig et al., 2009), each with advantages and disadvantages; some are more conducive for use during offshore infrastructure development. While not an exhaustive list, some of the traditional methods include visual observation, the most common for meeting regulatory compliance needs (Baker et al., 2013; Joint Nature Conservation Committee, 2017), and optical cameras, which identify marine mammals in the visible light spectrum (Podobna et al., 2009; Aniceto et al., 2018). Newer technologies refined over the last few decades include passive acoustic monitoring (PAM), which can detect and identify vocalizing marine mammals (Sousa-Lima, 2013; Kowarski et al., 2020) and animal-borne tags for tracking individual animals (Watanabe and Papastamatiou, 2023). Emerging technologies include infrared (IR)

---

* Corresponding author.
E-mail address: angela.szesciorka@pnnl.gov (A.R. Szesciorka).

https://doi.org/10.1016/j.ocecoaman.2026.108147
Received 8 October 2025; Received in revised form 12 January 2026; Accepted 19 February 2026

cameras, which identify animals at or above the water's surface (Smith et al., 2020; Richter et al., 2024; Zitterbart et al., 2020); satellite imagery, which can visually identify multiple species (Fretwell et al., 2014; Cubaynes et al., 2019; Guirado et al., 2019); and chemicals like dimethyl sulfide (DMS) and environmental DNA (eDNA), which can attract or identify species, respectively (Baker et al., 2018; Suarez-Bregua et al., 2022).

The platforms for deploying the technologies (e.g., moorings, fixed infrastructure, autonomous aerial and underwater vehicles, piloted aircraft and vessels), data analyses (e.g., automated detection and/or classification), and delivery methods (e.g., near real-time) have continued to improve (Nowacek et al., 2016). However, some technologies may be better suited for operational deployment during offshore development (Macrander et al., 2022), where the technology in use must accurately detect marine mammals with few missed detections over a relevant distance and transfer that information reliably to a human operator within a pre-defined decision-making time window.

With that specific use case in mind, Szesciorka et al. (2025) reviewed the international literature (spanning at least 17 countries and the poles) on technologies for monitoring whales during offshore wind foundation installation and found that most studies do not typically or consistently report performance metrics needed to evaluate the utility of monitoring technologies. Specifically, the studies rarely evaluated or reported on the performance of monitoring tools related to efficacy (i.e., precision and recall, probability of missed mitigation), detection range (i.e., maximum and reliable detection range, spatial resolution), and data delivery (i.e., detection latency, system reliability, temporal resolution); metrics that we argue are critical for evaluating whether the technology is suitable for use in a mitigation context (Szesciorka et al., 2025).

After a critical review of the 38 peer-reviewed journal articles, technical reports, consultant reports, and regulatory documents focused on the detection of baleen whales in marine environments (methods described in Szesciorka et al. (2025)), we conclude that a better understanding of the performance of these technologies using these metrics would be valuable for diverse users. Standardizing data collection, reporting, and analysis, with other standardized procedures (Van Parijs et al., 2021; Hildebrand et al., 2022) can promote consistency and comparable results, allowing research on the use of monitoring technology to have a greater reach in application. In turn, this may remove barriers to progress related to marine sector growth, including expediting or streamlining licensing/permitting and reducing project costs. Here, we provide recommendations for a set of metrics needed to compare the performance of monitoring technologies specifically for monitoring during offshore development, regardless of jurisdiction or species, and highlight which publications from the Szesciorka et al. (2025) literature review reported on these metrics. We also emphasize the benefits of multi-modal approaches and recommend performance assessments through modeling and field testing.

2. Performance reporting

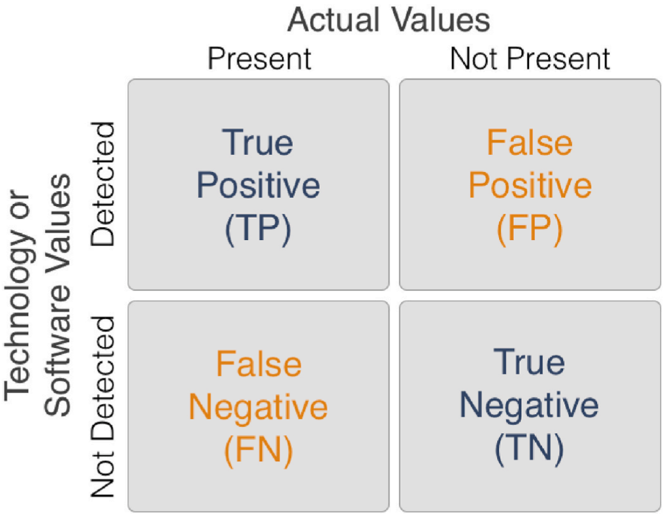
For successful marine mammal monitoring in a mitigation context, the technology must be able to observe an animal within detection range of a sensor, classify it to the species level (which may involve a human operator), and determine its position (e.g., localize within a mitigation zone or provide a relative range/bearing). The probability of detection should be close to 1 across the entire monitoring area if all animals potentially at risk are to be detected. However, because of availability bias, animals may be missed if they are present but not producing the appropriate cue (e.g., they are not visible at the surface nor producing sound). They may also be missed because of perception bias, in which the system cannot detect the cue because of factors like ambient noise, poor weather, or observer fatigue (Marsh and Sinclair, 1989). To determine whether a technology is suitable for use in a mitigation context, the tools need to be assessed for efficacy, detection range, and data delivery based on the particular context of the mitigation plan. Below we describe these metrics with the assumption that the cue (e.g., vocalization or body surfacing) is present and detectable; however, availability and perception biases should be considered in technology performance assessments. While seven of the studies we assessed discussed availability or perceptibility bias, only three modeled one or the other (Zitterbart et al., 2013, 2020; Ceballos et al., 2023).

2.1. Efficacy

The ability to correctly and repeatably identify animals can be evaluated for a technology, automatic detection, and/or classification software. Some measurements of performance that can be considered in an offshore development setting involve the construction of a confusion matrix (Fig. 1), which summarizes **true positive** (TP; detection is made and an animal is present), **false positive** (FP; detection is made but no animal is present), **true negative** (TN; no detection is made and an animal is not present), and **false negative** (FN; no detection is made but an animal is present). These are used to calculate **precision**, which is the fraction of correct detections or classifications, and **recall**, which is the fraction of occurrences that were detected (Van Parijs et al., 2021; Hildebrand et al., 2022; Fawcett, 2006).

Half of the studies we assessed presented or discussed components of the confusion matrix: 39% TP, 34% TN, 44% FP, and 37% FN. Only 10% calculated precision or recall. Most reported these metrics in relation to automated detection/classification algorithms for PAM (Kowarski et al., 2020; Smith et al., 2020; Baumgartner and Mussoline, 2011; Baumgartner et al., 2019; Clark et al., 2010; Rutenko et al., 2022; Salisbury et al., 2018), IR cameras (Smith et al., 2020; Richter et al., 2024; Zitterbart et al., 2013, 2020; Guazzo et al., 2019), satellite imagery (Fretwell et al., 2014; Guirado et al., 2019; Abileah, 2002; Corrêa et al., 2022), or conditionally in comparison with the performance of one or multiple monitoring technologies (Kowarski et al., 2020; Smith et al., 2020; Richter et al., 2024; Zitterbart et al., 2013, 2020; Fretwell et al., 2014; Ceballos et al., 2023; Clark et al., 2010; Corrêa et al., 2022; Baldacci et al., 2005; Graber et al., 2011; Bamford et al., 2020). Studies using eDNA used positive and negative controls but did not provide the results (Andruszkiewicz et al., 2017, 2020; Closek et al., 2019; Valsecchi et al., 2022; Gold et al., 2021), and studies using optical cameras, active acoustics, DMS, and tags did not include metrics. Only two of the studies provided definitions for the metrics or the formulas used to calculate them (Kowarski et al., 2020; Baumgartner et al., 2019).

FPs resulting in unnecessary mitigation and FNs resulting in missed mitigation are considered failures that have potentially negative consequences for offshore development or the animal, respectively. Only



Technology or Software Values	Actual Values: Present	Actual Values: Not Present
Detected	True Positive (TP)	False Positive (FP)
Not Detected	False Negative (FN)	True Negative (TN)

Fig. 1. Confusion matrix constructed to assess performance.

one study discussed the concepts behind and calculated the **probability of missed mitigation** (Baumgartner, 2025a), or the likelihood that a whale entered a mitigation zone without being detected. Thus, probability of missed mitigation is important to assess when evaluating a technology or detection/classification software, especially if there is no other effective method of monitoring.

2.2. Detection range

Maximum detection range, or the farthest distance that a sensor can detect an animal, is important for evaluating the performance of a technology sensor. While the detection probability functions differ depending on the technology (e.g., PAM's detection probability decreases with distance while IR's initially increases before decreasing), at some distance the signal's (detected observation) detectability decreases, and it may become more indistinguishable due to reduced resolution or increased background levels. Thus, another important metric is **peak or reliable detection range**, the distance at which a whale occurrence can reliably be detected (Zitterbart et al., 2013; Verfuss et al., 2016, 2018), and to a lesser extent spatial resolution, the smallest measurable change in distance (e.g., nearest 0.1 km, 1 km, 10 km). A sensor's ability to detect, classify, and/or localize can change based on the local geophysical characteristics (e.g., bathymetry, sediment structue), weather or oceanographic conditions, sensor configuration and settings, the species (e.g., frequency range and source levels), and other factors, thus reports should incorporate this information into detection range estimates, including uncertainty around the animal's location if localization is possible.

We found that 45% of the studies reported a maximum detection range but only 32% reported the variability of detection range as it related to an environment or temporal variable, site location or platform type, relative to other monitoring methods, and as it related to different species, group size, and behavior. Studies where detection distance estimates are not feasible or useful, including those using DMS, eDNA, satellite, and tags did not include these metrics. Studies that reported maximum detection range were those that used PAM (Ceballos et al., 2023; Baumgartner et al., 2019; Clark et al., 2010; Salisbury et al., 2018), IR cameras (Richter et al., 2024; Zitterbart et al., 2013, 2020; Guazzo et al., 2019; Baldacci et al., 2005; Graber et al., 2011; Cuyler et al., 1992; Horton et al., 2017; Perryman et al., 1999; Schoonmaker et al., 2008), optical cameras (Richter et al., 2024; Baldacci et al., 2005; Smultea Environmental Sciences, 2022), and active acoustics (Geoffroy et al., 2016; Knudsen et al., 2008).

Studies that reported variability in detection included those that used PAM (Smith et al., 2020; Ceballos et al., 2023; Baumgartner et al., 2019; Salisbury et al., 2018) and IR (Smith et al., 2020; Richter et al., 2024; Zitterbart et al., 2020; Guazzo et al., 2019; Baldacci et al., 2005; Graber et al., 2011; Cuyler et al., 1992; Horton et al., 2017; Perryman et al., 1999). Of these studies, 77% reported detection range and only six papers presented the probability of detection function. Some mentioned detection distance and in some cases reported mean values but did not include the detection range (Smith et al., 2020; Salisbury et al., 2018). Only 13% (n = 5) reported peak or reliable detection range, including studies using PAM (Clark et al., 2010), IR cameras (Richter et al., 2024; Zitterbart et al., 2020; Guazzo et al., 2019), optical cameras (Richter et al., 2024), and active acoustics (Geoffroy et al., 2016). None reported spatial resolution.

2.3. Data delivery

The ability of a system to reliably deliver information to a human operator in near real time to validate a detection and take the appropriate response is critical, as any observation could trigger a delay or shutdown. **Detection latency** includes the time it takes for a detection to travel to the sensor, the signal to be processed by the sensor, an automated or manual process to determine the detection including classification and/or localization, human review or verification of automated detections, and transfer of this information to the appropriate person to implement mitigation measures. Additional time may be needed to transmit information through various communication networks (e.g., satellite, digital, or radio).

While 45% (n = 17) of studies we assessed using PAM, IR camera, optical camera, and active acoustics were made in real or near-real time, only 8% (n = 3) of all studies discussed detection time latency. All were PAM studies, one of which modeled the time required to detect whales based on transit speeds of the survey platform (e.g., vessel, glider, aircraft) (Ceballos et al., 2023). The other two involved delivery of automated detections from a moored buoy that were reviewed by a human operator in 15 min increments (Baumgartner et al., 2019) and from a glider ranging 15 min to 3 h, 15 min, depending on the glider dive cycle and reporting schedule and weather (Kowarski et al., 2020). For some technologies, detection latency depends on data access. For example, requests for satellite imagery data within a specific time frame can vary by 21 days (Khan et al., 2023), and although satellite-linked tag studies can schedule data transmission (e.g., every 6 or 12 h) (Palacios et al., 2022), for archival tags, detection latency includes deployment duration, recovery, and analysis. Finally, while onboard measurements are possible (Okane et al., 2019), most studies using eDNA or DMS require transportation and analysis in a lab.

Other metrics such as **system reliability**, or the fraction of time a system is operational during the planned operational time, how much downtime to expect, and **temporal resolution**, or the smallest measurable change in reporting time (e.g., nearest 1 s, 15 min, 1 h), should also be reported. Only 8% (n = 3) reported periods when the technology was not functional due to weather conditions or when the gimbal was not stabilized (Smith et al., 2020; Baldacci et al., 2005). One reported 14 shutdowns resulting in 8 h of downtime (Smultea Environmental Sciences, 2022); however, none reported system latency under normal use and none reported temporal resolution.

3. Multi-modal systems

Across the studies we evaluated, no single technology reliably detected all animals in a given area and under all conditions (Smith et al., 2020). However, studies using multiple technologies have allowed their performance to be compared against one another (see summary table in Szesciorka et al. (2025)). For example, in some studies IR cameras had a similar number of detections (Zitterbart et al., 2013; Baldacci et al., 2005) and increased the number of hours of possible detection (Graber et al., 2011) compared to visual observations. In one study, IR cameras outperformed experienced observers at short ranges (2–3 km) (Zitterbart et al., 2020). Detections from satellite imagery matched the variation in number of detections from aerial surveys (Corrêa et al., 2022), with automated analysis finding most features identified manually (Fretwell et al., 2014). Density estimates made from satellite imagery data, once corrected for weather and surface availability, were found to be on the same order of magnitude as line-transect estimates (Bamford et al., 2020). Finally, in some studies PAM detections were similar to visual detections (Baumgartner et al., 2019; Clark et al., 2010); recall between manually and automatically identified calls varied from 14% to 76% depending on species and time frame (Kowarski et al., 2020).

While comparisons may be context and species specific, in these studies, monitoring appeared most effective using complementary methods. Although not all studies compared the performance of multiple technologies against each other (e.g., Guazzo et al., 2019), the data exist to do so. For example, line transect, aerial surveys, and PAM are frequently used to locate target animals for tag deployment during U.S. Navy-funded field studies (e.g., Baird et al., 2024; Henderson et al., 2025). The simultaneous use of multiple monitoring technologies increases the overall probability, effectiveness, and reliability of detection, classification, and localization because of the ability to track multiple animal cue types (Verfuss et al., 2018). Thus, multi-modal approach is considered the best available method for observing animals that may be displaying different cues, thereby reducing the probability of missed mitigation.

There is currently no standardized approach for assessing the performance of multi-modal monitoring approaches and in some cases no direct way to estimate the change in detection probability resulting from any particular combination of technologies. Thus, we recommend an increased focus on performance of multi-modal systems with future research focusing on evaluating different combinations, and applications, of multi-modal systems, including the system as a whole and each component. This should also focus on addressing challenges of multi-modal and multi-vendor systems such as time/data synchronization, data transfer, communication, and automated operation.

4. Large-scale characterization

While performance metrics are valuable for optimizing the use of monitoring technology, it is important to consider if assessing performance is necessary in every study. It may also not always be feasible given time and funding constraints, and it puts additional burdens on the user. However, one way to make it easier to understand which technologies will work the best in a given scenario is to include an assessment of detection technology performance directly into environmental effects monitoring plans. That would encourage the industry to consider approaches for assessing performance earlier in the planning process, which could ultimately expedite permitting and reporting process.

Another approach for reducing this burden is to evaluate the performance of a monitoring technology through a coordinated effort where technology developers, regulatory agencies, and researchers work together. Not only will this standardize and streamline efforts (i.e. funding, labor, equipment), but by understanding the different technologies, it also allows each to be leveraged in the most efficient way, including identifying combinations that optimize performance and reduce project costs.

A "gold standard" approach could involve a large-scale real-world scenario where stationary and vessel-based observers conduct visual and aerial surveys, PAM, and any other available methods to ensure that the "true" number of whales present in a test area during the validation is as accurate as possible. This would allow for the assessment of multiple modalities from a species of interest or with a proxy species that has similar behaviors to the target species while allowing multimodal cross-validation for hard-to-identify species. A field test could also be conducted in a region with conditions as close to the target species' habitat where future offshore infrastructure activities will occur and under different oceanographic conditions.

Because field testing can be costly and the permit process lengthy, establishing a collaborative instrumented, pre-permitted testing site will significantly reduce barriers for technology developers to characterize the performance of their technology and create consistency for comparison. A successful coordinated approach was recently spearheaded by the NOAA's Southwest Fisheries Science Center "Glider Rodeo."[1] The project brought together researchers at Southwest Fisheries Science Center, NOAA's Fisheries Pacific Islands Fisheries Science Center, Oregon State University, and Oregon Coast Aquarium, who together conducted ocean trials simultaneously alongside research vessels to understand how autonomous vehicles could augment data, ultimately to develop marine mammal stock assessments from the best available data.

Performance metrics can also be achieved with modeling, which can simulate different numbers and configurations of sensors in representative conditions, including animal behavior; environmental conditions, and other relevant factors (Gervaise et al., 2021; SMRU Consulting, 2024). A comparison of acoustic (gliders) and visual (aircraft) surveys found that a slow-moving glider detected a single North Atlantic right whale while aerial surveys required more than 20 transits and only reliably detected whales ($p > 0.5$) when 20 were present (Ceballos et al., 2023). Another study evaluated the efficacy of real-time PAM to mitigate risks to North Atlantic right whales near wind energy industrial activities and found that as few as three PAM stations 5–10 km to a construction site and using detections only (no localization) could trigger mitigation (Baumgartner, 2025b).

5. Conclusion

There is growing interest in assessing the available marine mammal monitoring technologies to reduce human-wildlife conflicts in the marine environment. Szesciorka et al. (2025) focused on the technologies that would work best for baleen whales in low- and no-light conditions before and during offshore wind turbine foundation installation. Other organizations, like the U.S.-based nonprofit MITRE, are focused on technology development and engineering approaches related to detection of the endangered North Atlantic right whale to reduce vessel strikes and entanglement in fishing gear. In the United Kingdom, the Institute of Marine Engineering, Science, and Technology, recently released a report (Dorrian et al., 2025) aimed at advancing the understanding and application of marine mammal monitoring technologies. All of these efforts call for standardization and multi-modal monitoring approaches.

We recommend that future studies evaluate and report on the performance of monitoring tools using metrics related to efficacy (i.e., precision and recall, probability of missed mitigation), detection range (i.e., maximum and reliable detection range, spatial resolution), and data delivery (i.e., detection latency, system reliability, temporal resolution). The increased use of multi-modal approaches will improve efficacy of monitoring but still require performance assessments of different combinations (as a whole and each component) to develop standardized combinations. Finally, assessments of the performance of monitoring tools should be built into monitoring plans or conducted as large-scale characterizations in real-world or modeled scenarios, where the true number of animals present can be known.

While evaluating and reporting on the efficacy, range, and data delivery of a monitoring technology require additional steps, there are benefits beyond a single project worth considering. Evaluating performance, whether a field characterization or modeling, could standardize the way data are collected, stored, and analyzed; provide additional opportunities for applied use; and incentivize technology developers to fund and work directly with researchers to co-develop technology that achieves multiple goals. In addition to benefiting researchers and industry partners, assessing the utility of different technologies for use in a mitigation context can inform decisions on what systems regulatory agencies might consider, potentially expediting or streamlining licensing/permitting and improving cost efficiency for energy developers. Collaboration among technology developers, regulatory agencies, and researchers ensures all parties are confident in the technology systems and results.

CRediT authorship contribution statement

Angela R. Szesciorka: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mark Severy:** Writing – review & editing, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kristen Ampela:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Cris Hein:** Writing – review & editing, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Michael Richlen:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation,

---

[1] https://www.fisheries.noaa.gov/news/ocean-gliders-listen-whales-oregon-test-new-ways-count-them.

Conceptualization. **Joseph Haxel:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jeff Clerc:** Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgements

We thank the NOAA, Regional Wildlife Science Collaborative, U.S. Department of the Interior Bureau of Ocean Energy Management, and U. S. Department of Energy (DOE) who provided their guidance and insight to define the scope of the report that inspired this paper. This work was authored by staff at the Pacific Northwest National Laboratory (contract no. DE-AC05-76RL01830) and the National Laboratory of the Rockies (contract no. DE-AC36-08GO28308), operated by the DOE via Battelle and the Alliance for Energy Innovation, respectively. Funding was provided by U.S. DOE Office of Critical Minerals and Energy Innovation. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Data availability

No data was used for the research described in the article.

References

Abileah, R., 2002. Marine mammal census using space satellite imagery. U.S. Navy J. Underwater Acoustics 52, 1–16.

Andruszkiewicz, E.A., Starks, H.A., Chavez, F.P., Sassoubre, L.M., Block, B.A., Boehm, A. B., 2017. Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. PLoS One 12, e0176343. https://doi.org/10.1371/journal.pone.0176343.

Andruszkiewicz, E.A., Yamahara, K.M., Closek, C.J., Boehm, A.B., 2020. Quantitative PCR assays to detect whales, rockfish, and common murre environmental DNA in marine water samples of the Northeastern Pacific. PLoS One 15, e0242689. https://doi.org/10.1371/journal.pone.0242689.

Aniceto, A.S., Biuw, M., Lindstrøm, U., Solbø, S.A., Broms, F., Carroll, J., 2018. Monitoring marine mammals using unmanned aerial vehicles: quantifying detection certainty. Ecosphere 9, e02122. https://doi.org/10.1002/ecs2.2122.

Baird, R.W., Harnish, A.E., Andrews, R.D., Lerma, J.K., Mohler, M.A., Phipps, J.E., Mahaffy, S.D., 2024. Small Boat Surveys and Satellite Tagging of Cetaceans on the Pacific Missile Range Facility, Kauaʻi. WA. February 2024, U.S. Pacific Fleet by HDR, Olympia. https://cascadiaresearch.org/files/Baird_et_al_2024_Kauai.pdf.

Baker, K., Epperson, D., Gitschlag, G., Goldstein, H., Lewandowski, J., Skrupky, K., Smith, B., Turk, T., 2013. National Standards for a Protected Species Observer and Data Management Program: a Model Using Geological and Geophysical Surveys. National Marine Fisheries Service, Silver Spring, MD. https://repository.library.noaa.gov/view/noaa/15851.

Baker, C.S., Steel, D., Nieukirk, S., Klinck, H., 2018. Environmental DNA (eDNA) from the wake of the whales: droplet digital PCR for detection and species identification. Front. Mar. Sci. 5. https://doi.org/10.3389/fmars.2018.00133.

Baldacci, A., Carron, M., Portunato, N., 2005. Infrared Detection of Marine Mammals. NATO Undersea Research Centre, La Spezia, Italy.

Bamford, C.C.G., Kelly, N., Dalla Rosa, L., Cade, D.E., Fretwell, P.T., Trathan, P.N., Cubaynes, H.C., Mesquita, A.F.C., Gerrish, L., Friedlaender, A.S., Jackson, J.A., 2020. A comparison of baleen whale density estimates derived from overlapping satellite imagery and a shipborne survey. Sci. Rep. 10, 12985. https://doi.org/10.1038/s41598-020-69887-y.

Baumgartner, M., 2025a. Efficacy of real-time passive acoustic monitoring near wind energy industrial activities for mitigating risks to right whales. Endanger. Species Res. 57, 413–430. https://doi.org/10.3354/esr01432.

Baumgartner, M., 2025b. Efficacy of real-time passive acoustic monitoring near wind energy industrial activities for mitigating risks to right whales. Endanger. Species Res. 57, 413–430.

Baumgartner, M.F., Mussoline, S.E., 2011. A generalized baleen whale call detection and classification system. J. Acoust. Soc. Am. 129, 2889–2902. https://doi.org/10.1121/1.3562166.

Baumgartner, M.F., Bonnell, J., Van Parijs, S.M., Corkeron, P.J., Hotchkin, C., Ball, K., Pelletier, L., Partan, J., Peters, D., Kemp, J., Pietro, J., Newhall, K., Stokes, A., Cole, T.V.N., Quintana, E., Kraus, S.D., 2019. Persistent near real-time passive acoustic monitoring for baleen whales from a moored buoy: system description and evaluation. Methods Ecol. Evol. 10, 1476–1489. https://doi.org/10.1111/2041-210X.13244.

Brill, D.N., Cleary, J., Roberts, J.J., O'Brien, B.R., Halpin, P.N., 2025. Expected occurrence of wildlife in US Atlantic offshore wind areas. Front. Mar. Sci. 12. https://doi.org/10.3389/fmars.2025.1602182.

Ceballos, V., Taggart, C., Johnson, H., 2023. Comparison of visual and acoustic surveys for the detection and dynamic management of North Atlantic right whales (*Eubalaena glacialis*) in Canada. Conservat Sci and Prac 5, e12866. https://doi.org/10.1111/csp2.12866.

Clark, C.W., Brown, M.W., Corkeron, P., 2010. Visual and acoustic surveys for North Atlantic right whales, *Eubalaena glacialis*, in Cape Cod Bay, Massachusetts, 2001–2005: management implications. Mar. Mamm. Sci. 26, 837–854. https://doi.org/10.1111/j.1748-7692.2010.00376.x.

Closek, C.J., Santora, J.A., Starks, H.A., Schroeder, I.D., Andruszkiewicz, E.A., Sakuma, K.M., Bograd, S.J., Hazen, E.L., Field, J.C., Boehm, A.B., 2019. Marine vertebrate biodiversity and distribution within the central California Current using environmental DNA (eDNA) metabarcoding and ecosystem surveys. Front. Mar. Sci. 6. https://doi.org/10.3389/fmars.2019.00732.

Corrêa, A.A., Quoos, J.H., Barreto, A.S., Groch, K.R., Eichler, P.P.B., 2022. Use of satellite imagery to identify southern right whales (*Eubalaena australis*) on a Southwest Atlantic Ocean breeding ground. Mar. Mamm. Sci. 38, 87–101. https://doi.org/10.1111/mms.12847.

Cubaynes, H.C., Fretwell, P.T., Bamford, C., Gerrish, L., Jackson, J.A., 2019. Whales from space: four mysticete species described using new VHR satellite imagery. Mar. Mamm. Sci. 35, 466–491. https://doi.org/10.1111/mms.12544.

Cuyler, L.C., Wiulsrød, R., Øritsland, N.A., 1992. Thermal infrared radiation from free living whales. Mar. Mamm. Sci. 8, 120–134. https://doi.org/10.1111/j.1748-7692.1992.tb00371.x.

Dorrian, N., Ferguson, E.L., Scala, L., Wright, A.J., Noseworthy, A.S., Kitchiner, A., 2025. Marine Mammal Monitoring: Methods, Technologies, and Opportunities for Innovation, Institute of Marine Engineering, Science & Technology & Department for Environment, Food & Rural Affairs – Marine Natural Capital and Ecosystem Assessment Programme. Edinburgh, Scotland, UK.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Fretwell, P.T., Staniland, I.J., Forcada, J., 2014. Whales from space: counting southern right whales by satellite. PLoS One 9, e88655. https://doi.org/10.1371/journal.pone.0088655.

Geoffroy, M., Rousseau, S., Knudsen, F.R., Fortier, L., 2016. Target strengths and echotraces of whales and seals in the Canadian Beaufort Sea. ICES J. Mar. Sci. 73, 451–463. https://doi.org/10.1093/icesjms/fsv182.

Gervaise, C., Simard, Y., Aulanier, F., Roy, N., 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. Appl. Acoust. 178, 107949. https://doi.org/10.1016/j.apacoust.2021.107949.

Gold, Z., Sprague, J., Kushner, D.J., Zerecero Marin, E., Barber, P.H., 2021. eDNA metabarcoding as a biomonitoring tool for marine protected areas. PLoS One 16, e0238557. https://doi.org/10.1371/journal.pone.0238557.

Graber, J., Thomson, J., Polagye, B., Jessup, A., 2011. Land-based infrared imagery for marine mammal detection. San Diego, CA. https://doi.org/10.1117/12.892787, 81560B-81560B–11.

Guazzo, R.A., Weller, D.W., Europe, H.M., Durban, J.W., D'Spain, G.L., Hildebrand, J.A., 2019. Migrating eastern North Pacific gray whale call and blow rates estimated from acoustic recordings, infrared camera video, and visual sightings. Sci. Rep. 9, 12617. https://doi.org/10.1038/s41598-019-49115-y.

Guirado, E., Tabik, S., Rivas, M.L., Alcaraz-Segura, D., Herrera, F., 2019. Whale counting in satellite and aerial images with deep learning. Sci. Rep. 9, 14259. https://doi.org/10.1038/s41598-019-50795-9.

Henderson, E.E., Kratofil, M.A., Baird, R.W., Martin, C.R., Harnish, A.E., Alongi, G.C., Martin, S.W., Southall, B.L., 2025. Exposure and response of satellite-tagged Blainville's beaked whales to mid-frequency active sonar off Kauaʻi, Hawaiʻi. Movement Ecology 13, 29. https://doi.org/10.1186/s40462-025-00550-9.

Hildebrand, J.A., Frasier, K.E., Helble, T.A., Roch, M.A., 2022. Performance metrics for marine mammal signal detection and classification. J. Acoust. Soc. Am. 151, 414–427. https://doi.org/10.1121/10.0009270.

Horton, T.W., Oline, A., Hauser, N., Khan, T.M., Laute, A., Stoller, A., Tison, K., Zawar-Reza, P., 2017. Thermal imaging and biometrical thermography of humpback whales. Front. Mar. Sci. 4. https://doi.org/10.3389/fmars.2017.00424.

Joint Nature Conservation Committee, 2017. JNCC Guidelines for Minimising the Risk of Injury to Marine Mammals from Geophysical Surveys. Aberdeen, UK. https://data.jncc.gov.uk/data/e2a46de5-43d4-43f0-b296-c62134397ce4/jncc-guidelines-seismicsurvey-aug2017-web.pdf.

Jouffray, J.-B., Blasiak, R., Norström, A.V., Österblom, H., Nyström, M., 2020. The blue acceleration: the trajectory of human expansion into the ocean. One Earth 2, 43–54. https://doi.org/10.1016/j.oneear.2019.12.016.

Khan, C.B., Goetz, K.T., Cubaynes, H.C., Robinson, C., Murnane, E., Aldrich, T., Sackett, M., Clarke, P.J., LaRue, M.A., White, T., Leonard, K., Ortiz, A., Lavista Ferres, J.M., 2023. A biologist's guide to the galaxy: leveraging artificial intelligence and very high-resolution satellite imagery to monitor marine mammals from space. JMSE 11, 595. https://doi.org/10.3390/jmse11030595.

Knudsen, F.R., Gammelsæter, O.B., Kvadsheim, P.H., Nøttestad, L., 2008. Evaluation of fisheries sonar for whale detection in relation to seismic survey operations. Bioacoustics 17, 326–328. https://doi.org/10.1080/09524622.2008.9753864.

Kowarski, K.A., Gaudet, B.J., Cole, A.J., Maxner, E.E., Turner, S.P., Martin, S.B., Johnson, H.D., Moloney, J.E., 2020. Near real-time marine mammal monitoring from gliders: practical challenges, system development, and management implications. J. Acoust. Soc. Am. 148, 1215–1230. https://doi.org/10.1121/10.0001811.

Macrander, A.M., Brzuzy, L., Raghukumar, K., Preziosi, D., Jones, C., 2022. Convergence of emerging technologies: development of a risk-based paradigm for marine mammal monitoring for offshore wind energy operations. Integrated Environ. Assess. Manag. 18, 939–949. https://doi.org/10.1002/ieam.4532.

Marsh, H., Sinclair, D.F., 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. J. Wildl. Manag. 53, 1017–1024. https://doi.org/10.2307/3809604.

National Marine Fisheries Service, 2024. Update To: Technical Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammal Hearing (Version 3.0): Underwater and in Air Criteria for Onset of Auditory Injury and Temporary Threshold Shifts. National Oceanic and Atmospheric Administration, Silver Spring, MD.

Novaglio, C., Bax, N., Boschetti, F., Emad, G.R., Frusher, S., Fullbrook, L., Hemer, M., Jennings, S., van Putten, I., Robinson, L.M., Spain, E., Vince, J., Voyer, M., Wood, G., Fulton, E.A., 2022. Deep aspirations: towards a sustainable offshore Blue economy. Rev. Fish Biol. Fish. 32, 209–230. https://doi.org/10.1007/s11160-020-09628-6.

Nowacek, D.P., Christiansen, F., Bejder, L., Goldbogen, J.A., Friedlaender, A.S., 2016. Studying cetacean behaviour: new technological approaches and conservation applications. Anim. Behav. 120, 235–244. https://doi.org/10.1016/j.anbehav.2016.07.019.

Okane, D., Koveke, E.P., Tashima, K., Saeki, K., Maezono, S., Nagahata, T., Hayashi, N., Owen, K., Zitterbart, D.P., Ohira, S.-I., Toda, K., 2019. High sensitivity monitoring device for onboard measurement of Dimethyl Sulfide and dimethylsulfoniopropionate in seawater and an Oceanic atmosphere. Anal. Chem. 91, 10484–10491. https://doi.org/10.1021/acs.analchem.9b01360.

Palacios, D.M., Irvine, L.M., Lagerquist, B.A., Fahlbusch, J.A., Calambokidis, J., Tomkiewicz, S.M., Mate, B.R., 2022. A satellite-linked tag for the long-term monitoring of diving behavior in large whales. Animal Biotelemetry 10, 26. https://doi.org/10.1186/s40317-022-00297-9.

Perryman, W.L., Donahue, M.A., Laake, J.L., Martin, T.E., 1999. Diel variation in migration rates of Eastern Pacific gray whales measured with thermal imaging sensors. Mar. Mamm. Sci. 15, 426–445. https://doi.org/10.1111/j.1748-7692.1999.tb00811.x.

Podobna, Y., Schoonmaker, J., Boucher, C., Oakley, D., 2009. Optical detection of marine mammals. In: Ocean Sensing and Monitoring, pp. 126–136. https://doi.org/10.1117/12.818359. SPIE.

Richter, S., Yurk, H., Winterl, A., Chmelnitsky, E., Serra, N., O'Hara, P.D., Zitterbart, D., 2024. Coastal marine mammal conservation using thermal imaging-based detection systems. bioRxiv. https://doi.org/10.1101/2023.08.25.554754.

Rutenko, A.N., Zykov, M.M., Gritsenko, V.A., Fershalov, M.Yu, Jenkerson, M.R., Racca, R., Nechayuk, V.E., 2022. Real-time acoustic monitoring with telemetry to mitigate potential effects of seismic survey sounds on marine mammals: a case study offshore Sakhalin Island. Environ. Monit. Assess. 194, 745. https://doi.org/10.1007/s10661-022-10019-6.

Salisbury, D.P., Estabrook, B.J., Klinck, H., Rice, A.N., 2018. Understanding Marine Mammal Presence in the Virginia Offshore Wind Energy Area. Bureau of Ocean Energy Management, Sterling, VA.

Schoonmaker, J., Dirbas, J., Podobna, Y., Wells, T., Boucher, C., Oakley, D., 2008. Multispectral observations of marine mammals. In: Huckridge, D.A., Ebert, R.R. (Eds.), Electro-Optical and Infrared Systems: Technology and Applications V, SPIE. Cardiff, Wales, UK, p. 9. https://doi.org/10.1117/12.800024.

Smith, H.R., Zitterbart, D.P., Norris, T.F., Flau, M., Ferguson, E.L., Jones, C.G., Boebel, O., Moulton, V.D., 2020. A field comparison of marine mammal detections via visual, acoustic, and infrared (IR) imaging methods offshore Atlantic Canada. Mar. Pollut. Bull. 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/.

Smultea Environmental Sciences, 2022. LLC (Smultea Sciences), Protected Species Observer IHA Report for the Ørsted South Fork Wind 01, Revolution 01, and Sunrise Wind 01 High Resolution Geophysical Survey 2022. Ørsted Wind Power North America, LLC, Boston, MA.

Sousa-Lima, R., 2013. A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. Aquat. Mamm. 39, 23–53. https://doi.org/10.1578/AM.39.1.2013.23.

Suarez-Bregua, P., Álvarez-González, M., Parsons, K.M., Rotllant, J., Pierce, G.J., Saavedra, C., 2022. Environmental DNA (eDNA) for monitoring marine mammals: challenges and opportunities. Front. Mar. Sci. 9. https://doi.org/10.3389/fmars.2022.987774.

Szesciorka, A.R., Severy, M., Ampela, K., Hein, C., Richlen, M., Haxel, J., Clerc, J., 2025. Evaluating Tools and Technologies for Monitoring Baleen Whales During Offshore Wind Foundation Installation, Pacific Northwest National Laboratory. National Renewable Energy Laboratory.

Valsecchi, E., Coppola, E., Pires, R., Parmegiani, A., Casiraghi, M., Galli, P., Bruno, A., 2022. A species-specific qPCR assay provides novel insight into range expansion of the Mediterranean monk seal (*Monachus monachus*) by means of eDNA analysis. Biodivers. Conserv. 31, 1175–1196. https://doi.org/10.1007/s10531-022-02382-0.

Van Parijs, S.M., Baker, K., Carduner, J., Daly, J., Davis, G.E., Esch, C., Guan, S., Scholik-Schlomer, A., Sisson, N.B., Staaterman, E., 2021. NOAA and BOEM minimum recommendations for use of passive acoustic listening systems in offshore wind energy development monitoring and mitigation programs. Front. Mar. Sci. 8, 760840. https://doi.org/10.3389/fmars.2021.760840.

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

Verfuss, U.K., Gillespie, D., Gordon, J., Marques, T.A., Miller, B., Plunkett, R., Theriault, J.A., Tollit, D.J., Zitterbart, D.P., Hubert, P., Thomas, L., 2018. Comparing methods suitable for monitoring marine mammals in low visibility conditions during seismic surveys. Mar. Pollut. Bull. 126, 1–18. https://doi.org/10.1016/j.marpolbul.2017.10.034.

Watanabe, Y.Y., Papastamatiou, Y.P., 2023. Biologging and biotelemetry: tools for understanding the lives and environments of marine animals. Ann. Rev. Anim. Biosci. 11, 247–267. https://doi.org/10.1146/annurev-animal-050322-073657.

Würsig, B., Perrin, W.F., Thewissen, J.G.M., 2009. History of marine mammal research. In: Perrin, W.F., Würsig, B., Thewissen, J.G.M. (Eds.), Encyclopedia of Marine Mammals, second ed. Academic Press, London, UK, pp. 565–569. https://doi.org/10.1016/B978-0-12-373553-9.00131-0.

Zitterbart, D.P., Kindermann, L., Burkhardt, E., Boebel, O., 2013. Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. PLoS One 8, e71217. https://doi.org/10.1371/journal.pone.0071217.

Zitterbart, D.P., Smith, H.R., Flau, M., Richter, S., Burkhardt, E., Beland, J., Bennett, L., Cammareri, A., Davis, A., Holst, M., Lanfredi, C., Michel, H., Noad, M., Owen, K., Pacini, A., Boebel, O., 2020. Scaling the laws of thermal imaging–based whale detection. J. Atmos. Ocean. Technol. 37, 807–824. https://doi.org/10.1175/JTECH-D-19-0054.1.