Automated Detection of Fish Sounds: State of the Art, Limitations, and Research Road Map

Valentin Bordoux and Manuel Vieira

Contents

Abstract

Passive acoustic monitoring (PAM) offers a noninvasive approach for underwater monitoring at large spatial and temporal scales, including in hardly accessible areas,

OPEN ACCESS with major support from JASCO APPLIED SCIENCES

V. Bordoux (✉)
Marine Animal Ecology Group, Wageningen University & Research, Wageningen, The Netherlands

M. Vieira
MARE – Marine and Environmental Sciences Centre/ARNET – Aquatic Research Network, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal
e-mail: mavieira@ciencias.ulisboa.pt

A. N. Popper et al. (eds.), *The Effects of Noise on Aquatic Life IV*,
https://doi.org/10.1007/978-3-031-94229-7_26-1

yet it remains underutilized in fish ecology. Many fish produce various sounds associated with ecologically relevant behaviors providing valuable indicators. PAM for fish monitoring is constrained by two major challenges: the limited knowledge and availability of reference sounds in libraries for species identification and the absence of efficient, standardized data analysis methods, resulting in a dependence on labor-intensive manual analyses. Recent advances in machine learning, particularly deep learning, offer promising solutions for automated detection and classification of fish sounds, improving data analysis efficiency, scalability, and cost-effectiveness. While initial applications demonstrate clear benefits, widespread implementation remains limited. This chapter outlines a road map for large-scale adoption of PAM in fish ecology, emphasizing open datasets, benchmark collections, accessible analytical tools, and expanded reference sound libraries. Addressing PAM's interconnected main challenges will create a positive feedback loop, where improved methods accelerate data collection and, in turn, enhance methodological development. Ultimately, PAM can become a scalable, flexible, and powerful tool for fish monitoring, conservation, and management in a changing ocean.

Keywords

Passive acoustic monitoring · Deep learning · Fish bioacoustics · Underwater bioacoustics · Machine learning

Introduction

As human activities increasingly threaten marine ecosystems, effective monitoring methods are essential to assess and conserve biodiversity (Díaz et al. 2019). Traditional approaches to studying fish populations, such as fishery-based methods or visual censuses, are often limited by their invasiveness, high costs, and dependence on favorable environmental conditions. Passive acoustic monitoring (PAM) is becoming a widely used and powerful tool in the study of terrestrial and marine animals, providing insights into their presence, distribution, and behavior over large temporal and spatial scales. Its noninvasive nature, ability to operate continuously, and potential for large-scale data collection make it an attractive approach for biodiversity monitoring more broadly (Stowell and Sueur 2020).

Fish produce a remarkable diversity of sounds, with hundreds of species already known to vocalize and many more likely undiscovered (Looby et al. 2023). These sounds are tied to critical ecological behaviors, including courtship, mating, spawning, aggression, and territorial defense, making them valuable indicators of species presence and activity. Recent studies have demonstrated that PAM can be used to track spawning events or assess community composition in reef ecosystems, highlighting the clear opportunities for the use of PAM in fish ecology (Desiderà et al. 2019; Stratoudakis et al. 2024). Despite this potential, the application of PAM to fish ecology remains relatively underdeveloped, largely due to two main challenges that will be the focus here.

The first major challenge is methodological: analyzing the vast volumes of acoustic data generated by autonomous recorders deployed across diverse marine environments. Although techniques, such as long-term spectrogram visualization and acoustic indices, enable relatively rapid data processing, the ecological insights they provide are often coarse and limited. Finer-scale understanding can be achieved through the detection and analysis of individual sounds, yet this still depends heavily on manual annotation, a process that is time-consuming and unsustainable for large-scale studies. While automated approaches have advanced considerably in bird and marine mammal acoustics, comparable tools for fish remain underdeveloped. Consequently, many fish PAM studies continue to rely on subsampling and manual annotation, highlighting a critical methodological gap. Therefore, developing robust, automated data analysis methods is essential to unlock the full potential of PAM for fish monitoring at broad spatial and temporal scales.

A second major challenge in applying PAM for fish is the limited availability of reference recordings for most species, accentuated by insufficient knowledge of acoustic communities and underwater soundscapes. This lack of baseline information constrains both species-level identification and community-level interpretation. To date, the most comprehensive reference sound database for fish, FishSounds, includes recordings for only 269 fish species, compared to over 22,000 species known or expected to produce sound (Cox et al. 2023; Looby et al. 2023). This knowledge gap encompasses confirmed species-specific sounds and unidentified acoustic signals. Databases of unknown sounds can be as valuable as those of known sources, as they can later be linked to species as knowledge advances (Parsons et al. 2022). Reference sound repositories should also facilitate community-level analyses, species distribution mapping, and ecological inference, even prior to source identification. Establishing systematic, open repositories that integrate both identified and unidentified fish sounds is essential to advance automated detection, biodiversity assessment, and ecological applications. Continued reporting of fish acoustic repertoires is critical in this effort, supported by recent initiatives developing catalogues of unidentified sounds (Bolgan et al. 2022; Ríos et al. 2025), including conceptual frameworks for their use (Vieira et al. 2024).

Several recent reviews provide useful context for the application of automated data analysis methods to PAM for fish. Barroso et al. (2023) proposed a comprehensive synthesis of sound-based detection, classification, and identification approaches for fish. However, their review did not propose a road map for implementation, and the future directions outlined differ from those suggested here, which build on more recent developments. Kershenbaum et al. (2025) recommended best practices for designing PAM studies from both ecological and technical perspectives, providing broadly applicable guidelines that should also inform fish-focused studies.

Building on these contributions, this chapter proposes a complementary perspective focused on the strategic steps required for the widespread adoption of PAM in fish ecology. The first section reviews current methodologies, emphasizing recent developments and trends in fish and broader bioacoustics, and identifies the key remaining challenges. The second section proposes a road map outlining the actions needed to address these challenges and facilitate large-scale implementation. In parallel, this work stresses how advances in automated analysis can help overcome the interrelated limitation of reference sound availability, which continues to constrain the interpretation of fish PAM data.

Automated Methods: State of the Art and Challenges

From Traditional Methods to Deep Learning

Automated detection and classification of fish sounds have been explored for more than a decade, with methods evolving alongside advances in computer science. As highlighted by Barroso et al. (2023), earlier work required feature engineering, starting with Hidden Markov Models, and mel-frequency cepstral coefficients, and later with traditional machine learning techniques, such as random forest, support vector machine, and K-nearest neighbors. While choosing these features carefully can lead to good results, they remain species and context-specific.

As in the broader field of bioacoustics, the introduction of deep learning (DL), particularly convolutional neural networks (CNNs), has gradually shifted the field away from explicit feature engineering (Stowell 2022). Comparative studies generally show that DL outperforms traditional machine learning methods by learning features directly from data (e.g., Mouy et al. 2024). Across bioacoustic disciplines, DL has become the preferred approach because it offers more generalizable performance on complex, real-world recordings, eliminates the need for manual feature design, scales effectively with larger datasets, and benefits from a rapidly expanding development ecosystem. Nonetheless, most DL pipelines still depend on some degree of manual preprocessing, such as the choice of spectrogram parameters (FFT size, frequency range, and scaling), which should be clearly reported to ensure reproducibility.

Expanding the Deep Learning Toolbox

Studies have also explored deep learning approaches that directly use raw waveforms as input, including applications to fish sounds (Ibrahim et al. 2018). Although considered promising (Stowell 2022), these methods have received comparatively less attention. Their main advantage is the elimination of the manual feature extraction required for spectrogram generation. However, they cannot fully leverage the extensive developments from computer vision, such as some data augmentation techniques, per-channel energy normalization (PCEN), or image-based pretrained models. More recently, transformer-based architectures also achieved end-to-end learning from raw audio and demonstrated strong performance, including for marine mammals (Hagiwara 2022; Robinson et al. 2024). Nonetheless, these models are computationally demanding with very large numbers of parameters, require extensive training datasets, and are less flexible for fine-tuning or modification. These factors currently limit their accessibility for many bioacoustic researchers.

Reviews show that CNNs are now the most widely used approach for acoustic detection in fish and other taxa (Stowell 2022; Barroso et al. 2023). More advanced architectures, including convolutional recurrent networks that capture temporal dependencies and object-detection models adapted from computer vision, have also been explored (Parcerisas et al. 2024; Huang et al. 2025). Although these models often achieve strong performance in specific studies, their broader application in fish PAM remains limited. Supervised models trained in one context rarely generalize well to new environments or species, and their development typically demands programming expertise, significant computational resources, and large annotated datasets that are costly and time-consuming to produce. Therefore, some attention has also been turned toward unsupervised learning, which can explore data without prior labels or species-specific training data.

Unsupervised and Emerging Learning Paradigms

Unsupervised learning approaches aim to uncover structure in unlabeled acoustic data, typically by clustering recordings in low-dimensional feature spaces. HDBSCAN is a promising algorithm, though its application to soundscape analysis is not yet standardized. These methods have shown particular potential in complex systems, such as coral reef soundscapes, where the high diversity and abundance of call types make manual annotation especially challenging (Minier et al. 2025). Nevertheless, their general applicability remains uncertain, as clustering may capture biologically meaningless patterns of acoustic energy. To partially address this limitation, manual revision of clusters has been proposed to improve interpretability (Parcerisas et al. 2024). Despite these advances, the challenges of evaluating and interpreting unsupervised outputs mean that supervised learning remains the preferred approach for most researchers, often combined with strategies to reduce the burden of manual annotations.

Several paradigms developed in broader machine learning, including transfer learning, active learning, and self-supervised learning, are now being adapted for bioacoustics and have potential for fish PAM. Transfer learning leverages pretrained models from other domains, such as image recognition and general bioacoustics, fine-tuning them for new tasks. This reduces annotation needs and has improved performance in multiple studies (Munger et al. 2022; Ghani et al. 2023). Active learning iteratively selects the most informative samples for annotation. In fish PAM, detectors have been trained in under 2 hours from a single example (Bordoux et al. 2025) and combined with clustering for efficient soundscape exploration (Parcerisas et al. 2024). Self-supervised learning defines auxiliary tasks, such as masked prediction and contrastive learning, to exploit large unlabeled datasets. Self-supervised learning should better support the development of foundation models, large-scale general-purpose models that can later be fine-tuned or used for downstream tasks, and be applied in bioacoustics (Schwinger et al. 2025).

Toward Foundation Models and FAIR Data

These paradigms address specific methodological challenges, yet widespread adoption of PAM for fish also depends on the availability of accessible “off-the-shelf” tools. General-purpose models that perform reliably across species and environments offer a promising solution, as demonstrated for birds, often accompanied by user-friendly graphical interfaces. However, the creation of large, diverse annotated datasets for fish, required for training general-purpose models, is unlikely in the near term due to the diversity of sounds, environmental variability, and the logistical difficulty of recording many habitats. Progress may instead rely on self-supervised learning and foundation models trained on extensive collections of unannotated recordings. Although recent studies show that current foundation models in bioacoustics perform poorly without fine-tuning (Chen and Yang 2025), this approach typically requires less data than training from scratch. Currently, fish sounds remain largely absent from training or evaluation of foundation models, highlighting the urgent need for FAIR (Findable, Accessible, Interoperable, and Reusable), well-curated datasets.

Future Directions for Scalable Fish PAM

In summary, the limited large-scale adoption of automated analysis in fish PAM reflects three underlying barriers: accessibility, with few easy-to-use, GUI-based tools; adaptability, since most existing models remain species- or location-specific; and scarcity of openly available data, particularly the shortage of annotated datasets for training and benchmarking. Addressing these barriers will require coordinated efforts in data collection and sharing, methodological innovation, and the creation of accessible tools that lower the entry barrier for researchers in fish bioacoustics. Given these limitations, how can the field move forward toward scalable, automated fish PAM?

Road Map for Adoption

For PAM to become a widely adopted monitoring method for fish, automated data analysis methods must be both accessible, usable by ecologists with minimal technical expertise, and flexible, capable of adapting to different species, environments, and datasets with minimal retraining. Achieving this requires a coordinated effort to connect data, models, and user-friendly tools. The following road map outlines key steps toward large-scale implementation, aiming to unlock the full potential of PAM for fish monitoring and ecological research (Fig. 1).

Building a Shared Data Foundation

Automation in PAM relies on the availability of comprehensive datasets. Reference libraries of identified calls, such as FishSounds (Cox et al. 2023), are steadily expanding, yet centralized efforts to share unidentified sounds remain limited despite their ecological and methodological value. The need for a global, open-access library for underwater sound sharing, encompassing both identified and unidentified

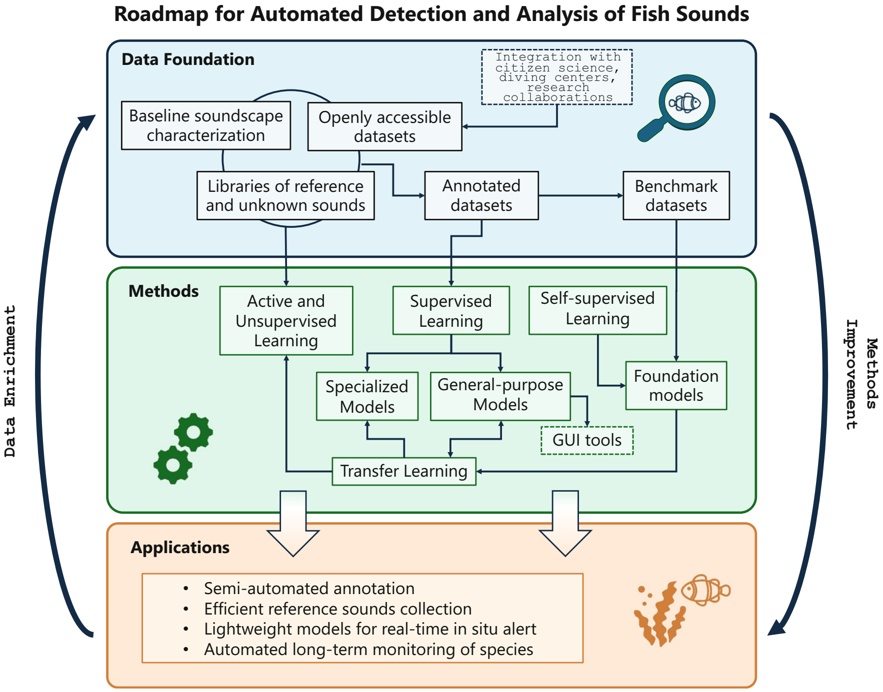

Fig. 1 Road map for large-scale adoption of PAM for fish. A strong data foundation enables the development of improved analysis methods, which in turn expand opportunities for diverse applications. The relationships between specific methods and applications are not yet well defined and remain flexible, therefore represented by broader connections. Effective application of PAM and data analysis methods will strengthen the data foundation through possible feedback processes

reference sounds as well as datasets for model training, has already been emphasized (Parsons et al. 2022). Initiatives, such as the Marine Sound Library (https://marinesoundlib.org/) developed by the Flanders Marine Institute, aim to fill this gap, and if widely adopted and well-maintained, could reduce redundancy, accelerate discovery, and strengthen collaboration.

Equally important is the open sharing of annotated datasets and additional information to avoid fragmentation and duplicated effort across research groups. Shared passive acoustic datasets should adhere to FAIR principles and include comprehensive metadata describing both the recording chain, such as hydrophone model, sensitivity, and frequency response, preamplifier settings, recorder bit depth, sampling rate, and overall dynamic range, and the deployment context, including depth, GPS location, and recording dates and times. In addition, information on the soundscape, habitat characteristics, and acoustic community should be provided within the publication associated with the data to ensure comparability and reuse.

Beyond data management, building a baseline understanding of underwater soundscapes is essential for meaningful ecological interpretation. Identifying the types of sounds present, characterizing acoustic communities, understanding their variation across environments, and linking them to specific species and behaviors provide the foundation for extracting robust ecological insights from PAM data. Ultimately, establishing open, standardized, and interoperable acoustic datasets is crucial for transforming PAM from a collection of local studies into a globally integrated tool for marine biodiversity assessment.

Expanding Data Sources

Broadening data sources represents another important step toward large-scale PAM for fish. Citizen science, for example, can provide valuable recordings, particularly in coastal areas where affordable underwater cameras with built-in microphones are increasingly accessible. Although, such recordings may be noisy or inconsistent, their utility can be enhanced through clear collection protocols and automated processing. Quality control and ethical considerations, such as minimizing the inadvertent capture of human speech, should accompany these participatory efforts.

Collaborations with diving centers, recreationally diverse, and marine biologists already using underwater audio-video systems also offer promising opportunities for data collection. When combined with automated analysis pipelines, these initiatives could substantially expand PAM coverage, supporting both species-level studies and long-term ecological monitoring.

Developing Benchmarks

Once datasets with annotations are openly shared, the next step is to develop benchmark collections. Benchmarks with standardized evaluation metrics allow systematic comparison of methods, reveal how models generalize across sites and species, and provide a foundation for creating general-purpose models. Without standardized evaluation methods, progress risks remaining fragmented, and studies become difficult to compare and contrast. Benchmarks are also essential for integrating fish sounds into the foundation models currently emerging in bioacoustics. Importantly, benchmarks do not need to be large; even a small, cross-site collection, such as ReefSet (Williams et al. 2025), can support model development and be reused.

Enhancing Accessibility and Usability

A road map for advancing underwater bioacoustics should support the parallel development of different methodological strands. Specialized models (also called domain-specific models), trained on regional or species-specific datasets, will remain essential for targeted ecological studies. General-purpose models, supported by community benchmarks and provided in accessible formats, can lower the barrier for reuse. Transfer learning links these approaches, allowing general models to be adapted to specific tasks, as demonstrated successfully in terrestrial bioacoustics with tools such as BirdNET (Kahl et al. 2021).

In underwater environments where sounds remain poorly characterized, machine learning offers powerful tools for discovering new sounds and for large-scale analysis. The choice of method depends on data availability, available expertise, and research goals: Unsupervised approaches support exploration, active learning enhances detection efficiency, and sufficient annotated data allow direct training of specialized models. As models are scaled to broader applications, transfer learning and self-supervised learning offer promising avenues for enhancing performance and generalizability. Moreover, methods are not confined to specific tasks and can be effectively combined, and their potential for diverse applications remains to be explored.

Progress toward foundation models promises robust, flexible tools for multiple applications. These models' development will require FAIR-compliant sharing of datasets, including both identified and unidentified sounds, along with standardized benchmark collections. Although supervised learning still outperforms self-supervised methods when large annotated datasets exist (hundreds of thousands of samples) (Merriënboer et al. 2025), such datasets may be unattainable for fish bioacoustics. Furthermore, while previous work has suggested that the transferability of models in bioacoustics is facilitated by shared sound production mechanisms among terrestrial vertebrates and marine mammals (Ghani et al. 2023), fish rely on fundamentally different mechanisms (Fine and Parmentier 2015). This distinction may make fish a particularly challenging group for model transfer, yet it also underscores the importance of incorporating fish vocalizations into the development of general bioacoustic foundation models. The potential of foundation models for fish remains to be demonstrated.

Accessibility is central to adoption. Automated tools should support rather than replace experts, reducing manual workload while preserving ecological relevance. Graphical User Interfaces (GUIs) provide intuitive access for general users, whereas platforms such as Jupyter notebooks, Python packages, or code repositories cater to advanced users. Recent work also explored combining a GUI with natural language interaction (Robinson et al. 2024). Not every method requires a GUI, but for widely applicable and robust tools, a GUI maximizes impact. By combining usability for nonspecialists with advanced capabilities for experts, automated approaches can move beyond fragmented case studies into widely adopted resources that support ecological understanding, conservation, and restoration.

Integrating Ecological Context and Applications

Building comprehensive reference libraries for fish sounds is accelerated by innovative methods (Dantzker et al. 2025) but remains constrained by the manual effort required. Advances in deep learning methods, such as active learning, unsupervised learning, transfer learning, and their combination, are beginning to accelerate semi-automated data analysis pipelines for faster discovery and classification of new sound types. Continued development of these approaches will be critical for scaling up species-level acoustic inventories. In species or environments where confirmation is impeded by turbidity, combining passive acoustic records with other techniques, such as fishery data or eDNA sampling, can provide a promising avenue to constrain the list of potential sound producers (Vieira et al. 2024; Watson 2025). Several species and communities may require more time-consuming and less scalable evidence of sound production; thus, the systematic collection and curation of unidentified sounds remain essential, as each addition broadens the known acoustic repertoire and increases the long-term ecological value of passive acoustic monitoring.

Studying specific species to discover new reference sounds and understand their ecological functions remains a central goal of PAM for fish. Specialized model development enables the characterization of acoustic activity patterns and assessment of species presence. While many methods can serve multiple purposes, advances in automated approaches will increasingly improve their practical applications. For example, active learning and unsupervised learning are particularly suited for semiautomated annotation workflows, which can play a major role in expanding reference sound libraries and strengthening the data foundation.

Accessible and flexible models further facilitate the use of PAM, reducing costs and enabling broader ecological insights. Promising applications include lightweight models for real-time alerts to support efficient management. Current examples already demonstrate event detection (Grabowski et al. 2020), monitoring migrations (Alcázar-Treviño et al. 2025), and detecting invasive species (Amorim et al. 2023), all of which could be implemented in near real time. In the future, fully automated long-term monitoring of specific species is feasible. These multiple applications will continue to enrich the reference data foundation, improving the interpretation of PAM recordings and creating a positive feedback loop that drives further methodological advances and ecological insights (Fig. 1).

Key Actions Moving Forward

Finally, these action points synthesize the road map toward large-scale adoption of automated detection and classification of fish sounds for realizing the full potential of PAM for fish:

- Expand annotated datasets through manual and semiautomated approaches, as well as collaborations with citizen scientists and multi- or transdisciplinary initiatives, and ensure open sharing via suitable infrastructure.
- Streamline reference sound collection with automated, accelerated data acquisition and expand soundscape baseline knowledge to support ecological interpretation.
- Aggregate datasets into benchmarks to standardize evaluation and enable inclusion of fish sounds in model development initiatives.

- Develop models and tools at specialized, regional, and general-purpose scales, including accessible GUIs for broader reusability.
- Implement automated PAM data analysis in long-term, large-scale monitoring programs to inform conservation and management decisions.

Conclusion

PAM offers a powerful, noninvasive approach to studying fish and their environment, yet its potential remains largely untapped. Progress toward large-scale adoption now depends on two interconnected advances: developing accessible automated analysis tools and enhancing knowledge about soundscape, including but not limited to cataloguing reference sounds. Together, these steps can transform PAM from fragmented applications into a scalable, reliable, and cost-effective framework for long-term monitoring of fish biodiversity. By aligning data infrastructure, methodological advances, and accessible tools, the research community can establish PAM as a central approach for detecting ecological change, informing conservation, and supporting sustainable management of marine ecosystems.

Competing Interest Declaration The author(s) has no competing interests to declare that are relevant to the content of this manuscript.

References

Alcázar-Treviño J, Korneliussen RJ, Escánez A, Aguilar de Soto N (2025) Evening choruses in deep waters are associated with mesopelagic diel vertical migrations. Mar Environ Res 211: 107440. https://doi.org/10.1016/j.marenvres.2025.107440

Amorim MCP, Wanjala JA, Vieira M, Bolgan M, Connaughton MA, Pereira BP, Fonseca PJ, Ribeiro F (2023) Detection of invasive fish species with passive acoustics: discriminating between native and non-indigenous sciaenids. Mar Environ Res 188:106017. https://doi.org/10.1016/j.marenvres.2023.106017

Barroso VR, Xavier FC, Ferreira CEL (2023) Applications of machine learning to identify and characterize the sounds produced by fish. ICES J Mar Sci 80:1854–1867. https://doi.org/10.1093/icesjms/fsad126

Bolgan M, Di Iorio L, Dailianis T, Catalan IA, Lejeune P, Picciulin M, Parmentier E (2022) Fish acoustic community structure in Neptune seagrass meadows across the Mediterranean basin. Aquat Conserv Mar Freshw Ecosyst 32:329–347. https://doi.org/10.1002/aqc.3764

Bordoux V, Parcerisas C, Debusschere E, Jälmby M, Murk AJ, van der Ven RM (2025) Rapid fish sound detection using human-in-the-loop active learning. Preprint at SSRN. https://doi.org/10.2139/ssrn.5403106

Chen C, Yang Z (2025) No free lunch from audio pretraining in bioacoustics: a benchmark study of embeddings. Preprint at ArXiv. https://doi.org/10.48550/arXiv.2508.10230

Cox KD, Looby A, Vela S, Riera A, Bravo S, Davies HL, Rountree R, Spriel B, Reynolds LK, Martin CW, Matwin S, Juanes F (2023) FishSounds version 1.1: data archive, user experience, and online resources. In: Popper AN, Sisneros J, Hawkins AD, Thomsen F (eds) The effects of noise on aquatic life. Springer International Publishing, Cham, pp 1–12. https://doi.org/10.1007/978-3-031-10417-6_35-1

Dantzker MS, Duggan MT, Berlik E, Delikaris-Manias S, Bountourakis V, Pulkki V, Rice AN (2025) Deciphering complex coral reef soundscapes with spatial audio and 360° video. Methods Ecol Evol. https://doi.org/10.1111/2041-210X.70149

Desiderà E, Guidetti P, Panzalis P, Navone A, Valentini-Poirrier C-A, Boissery P, Gervaise C, Iorio LD (2019) Acoustic fish communities: sound diversity of rocky habitats reflects fish species diversity. Mar Ecol Prog Ser 608:183–197. https://doi.org/10.3354/meps12812

Díaz SM, Settele J, Brondízio E, Ngo H, Guèze M, Agard J, Arneth A, Balvanera P, Brauman K, Butchart S, Chan KMA, Garibaldi LA, Ichii K, Liu J, Subramanian S, Midgley G, Miloslavich P, Molnár Z, Obura D, Pfaff A, Polasky S, Purvis A, Razzaque J, Reyers B, Roy Chowdhury R, Shin Y-J, Visseren-Hamakers I, Willis K, Zayas C (2019) The global assessment report on biodiversity and ecosystem services: summary for policy makers. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn

Fine ML, Parmentier E (2015) Mechanisms of fish sound production. In: Ladich F (ed) Sound communication in fishes. Springer, Vienna, pp 77–126. https://doi.org/10.1007/978-3-7091-1846-7

Ghani B, Denton T, Kahl S, Klinck H (2023) Global birdsong embeddings enable superior transfer learning for bioacoustic classification. Sci Rep 13:22876. https://doi.org/10.1038/s41598-023-49989-z

Grabowski T, Young SP, Cott PA (2020) Looking for love under the ice: using passive acoustics to detect burbot (Lota lota: Gadidae) spawning activity. Freshw Biol 65:37–44. https://doi.org/10.1111/fwb.13314

Hagiwara M (2022) AVES: animal vocalization encoder based on self-supervision. Preprint at ArXiv. https://doi.org/10.48550/arXiv.2210.14493

Huang Z, Ochs D, Amorim MCP, Fonseca PJ, Goel M, Nunes NJ, Vieira M, Lopes M (2025) Deep learning–based frameworks for the detection and classification of soniferous fish. J Acoust Soc Am 158:1060–1071. https://doi.org/10.1121/10.0038800

Ibrahim AK, Zhuang H, Chérubin LM et al (2018) Automatic classification of grouper species by their sounds using deep neural networks. J Acoust Soc Am 144:EL196–EL202. https://doi.org/10.1121/1.5054911

Kahl S, Wood CM, Eibl M, Klinck H (2021) Bird NET: a deep learning solution for avian diversity monitoring. Ecol Inform 61:101236. https://doi.org/10.1016/j.ecoinf.2021.101236

Kershenbaum A, Akçay Ç, Babu-Saheer L, Barnhill A, Best P, Cauzinille J, Clink D, Dassow A, Dufourq E, Growcott J, Markham A, Marti-Domken B, Marxer R, Muir J, Reynolds S, Root-Gutteridge H, Sadhukhan S, Schindler L, Smith BR, Stowell D, Wascher CAF, Dunn JC (2025) Automatic detection for bioacoustic research: a practical guide from and for biologists and computer scientists. Biol Rev 100:620–646. https://doi.org/10.1111/brv.13155

Looby A, Erbe C, Bravo S, Cox K, Davies HL, Di Iorio L, Jézéquel Y, Juanes F, Martin CW, Mooney TA, Radford C, Reynolds LK, Rice AN, Riera A, Rountree R, Spriel B, Stanley J, Vela S, Parsons MJG (2023) Global inventory of species categorized by known underwater sonifery. Sci Data 10:892. https://doi.org/10.1038/s41597-023-02745-4

Minier L, Rouch J, Sabbagh B, Bertucci F, Parmentier E, Lecchini D, Sèbe F, Mathevon N, Emonet R (2025) Visualization and quantification of coral reef soundscapes using CoralSoundExplorer software. PLoS Comput Biol 21:e1012050. https://doi.org/10.1371/journal.pcbi.1012050

Mouy X, Archer SK, Dosso S, Dudas S, English P, Foord C, Halliday W, Juanes F, Lancaster D, Van Parijs S, Haggarty D (2024) Automatic detection of unidentified fish sounds: a comparison of traditional machine learning with deep learning. Front Remote Sens 5. https://doi.org/10.3389/frsen.2024.1439995

Munger JE, Herrera DP, Haver SM, Waterhouse L, McKenna MF, Dziak RP, Gedamke J, Heppell SA, Haxel JH (2022) Machine learning analysis reveals relationship between pomacentrid calls and environmental cues. Mar Ecol Prog Ser 681:197–210. https://doi.org/10.3354/meps13912

Parcerisas C, Schall E, te Velde K, Botteldooren D, Devos P, Debusschere E (2024) Machine learning for efficient segregation and labeling of potential biological sounds in long-term underwater recordings. Front Remote Sens 5. https://doi.org/10.3389/frsen.2024.1390687

Parsons MJG, Lin T-H, Mooney TA, Erbe C, Juanes F, Lammers M, Li S, Linke S, Looby A, Nedelec SL, Van Opzeeland I, Radford C, Rice AN, Sayigh L, Stanley J, Urban E, Di Iorio L (2022) Sounding the call for a global library of underwater biological sounds. Front Ecol Evol 10. https://doi.org/10.3389/fevo.2022.810156

Ríos N, Pereira J, Muñoz-Duque S, Silva G, Pais MP, Fonseca PJ, Vieira M, Amorim MCP (2025) Acoustic fish community in a biogeographic transition zone of the Northeast Atlantic. ICES J Mar Sci 82:fsaf027. https://doi.org/10.1093/icesjms/fsaf027

Robinson D, Miron M, Hagiwara M, Pietquin O (2024) NatureLM-audio: an audio-language foundation model for bioacoustics. Preprint at ArXiv. https://doi.org/10.48550/arXiv.2411.07186

Schwinger R, Zadeh PV, Rauch L, Kurz M, Hauschild T, Lapp S, Tomforde S (2025) Foundation models for bioacoustics – a comparative review. Preprint at ArXiv. https://doi.org/10.48550/arXiv.2508.01277

Stowell D (2022) Computational bioacoustics with deep learning: a review and roadmap. PeerJ 10: e13152. https://doi.org/10.7717/peerj.13152

Stowell D, Sueur J (2020) Ecoacoustics: acoustic sensing for biodiversity monitoring at scale. Remote Sens Ecol Conserv 6:217–219. https://doi.org/10.1002/rse2.174

Stratoudakis Y, Vieira M, Marques JP, Amorim MCP, Fonseca PJ, Quintella BR (2024) Long-term passive acoustic monitoring to support adaptive management in a sciaenid fishery (Tagus Estuary, Portugal). Rev Fish Biol Fish 34:491–510. https://doi.org/10.1007/s11160-023-09825-z

van Merriënboer B, Dumoulin V, Hamer J, Harrell L, Burns A, Denton T (2025) Perch 2.0: the bittern lesson for bioacoustics. Preprint at ArXiv. https://doi.org/10.48550/arXiv.2508.04665

Vieira M, Ríos N, Muñoz-Duque S, Pereira J, Carriço R, Fernandez M, Monteiro JG, Pais MP, Quintella BR, Silva G, Silva RP, Fonseca PJ, Amorim MCP (2024) Cross-referencing unidentified fish sound data sets to unravel sound sources: a case study from the Temperate Northern Atlantic. Front Remote Sens 5. https://doi.org/10.3389/frsen.2024.1377206

Watson M (2025) Reef riffs: iological soundscapes & fish communities of reefs in the Wadden Sea. University of Groningen, Groningen. https://doi.org/10.33612/diss.1419294894

Williams B, van Merriënboer B, Dumoulin V, Hamer J, Fleishman AB, McKown M, Munger J, Rice AN, Lillis A, White C, Hobbs C, Razak T, Curnick D, Jones KE, Denton T (2025) Using tropical reef, bird and unrelated sounds for superior transfer learning in marine bioacoustics. Philos Trans R Soc B Biol Sci 380:20240280. https://doi.org/10.1098/rstb.2024.0280