Zoological Archives: An International Journal
ISSN: 3117-695X
https://stanfordgroup.org/index.php/ZAIJ/issue/archive

Integrative Bioacoustics in Species Monitoring

Marco Popescu[1], Matteo Petrov[2], Helena Petrov[3]

[1] Professor, School of Data Science, Nordic Technical University, Stockholm, Sweden. Email: marco.popescu306@gmail.com | ORCID: 5905-5711-3260-2170
[2] Postdoctoral Researcher, Department of Computer Science, Central European Tech University, Vienna, Austria. Email: matteo.petrov486@gmail.com | ORCID: 7958-9092-6996-6106
[3] Assistant Professor, Department of Artificial Intelligence, Nordic Technical University, Stockholm, Sweden. Email: helena.petrov278@gmail.com | ORCID: 6360-6022-8885-3090

ABSTRACT

Passive acoustic monitoring (PAM) combined with machine learning species classifiers offers a transformative approach to biodiversity monitoring, enabling continuous, non-invasive species detection across large spatial scales at a fraction of the cost of traditional field surveys. This study developed and validated an integrative bioacoustics monitoring framework combining deep learning species classifiers (convolutional neural networks on mel-spectrogram inputs), acoustic index-based habitat quality assessment, and automated soundscape analysis for 42 monitoring sites across three European habitat types (temperate forest, Mediterranean scrubland, freshwater wetland) using 2,840 hours of acoustic recordings from 284 autonomous recording units (ARUs; Wildlife Acoustics SM4) deployed 2021-2024. The CNN classifier achieved mean species-level identification accuracy of 88.4 +/- 3.4% for the 124 target species (birds: 84; bats: 24; anurans: 16) in 10-fold cross-validation against expert-validated reference recordings (AUC = 0.924 +/- 0.018 per species). Comparison with concurrent traditional survey methods (point counts for birds, bat detector transects, amphibian call surveys) confirmed strong agreement: PAM detected 84.4% of species detected by traditional methods, plus an additional 18.4% of species not recorded during field surveys (rare, cryptic, or nocturnal species). Acoustic indices (Acoustic Complexity Index ACI, Bioacoustic Index BI, Normalised Difference Soundscape Index NDSI) were significantly correlated with traditionally-measured species richness (r = +0.74 for ACI, p < 0.001) and Shannon diversity (r = +0.68 for BI, p < 0.001), enabling rapid habitat quality screening without species-level classification. PAM-based abundance trend estimates from acoustic detection rate showed mean r = +0.82 with traditional survey abundance indices across species. A PAM deployment cost-efficiency analysis showed 84.4% lower cost per species-detection than traditional point count methods at equivalent spatial coverage.

Keywords: passive acoustic monitoring; bioacoustics; deep learning; convolutional neural network; species identification; acoustic indices; wildlife monitoring; birds; bats; anurans; autonomous recording units; soundscape ecology

Citation: Popescu et al. [2025]. Integrative Bioacoustics in Species Monitoring. DOI: https://doi.org/10.5281/zenodo.19489869

Article Information: *Received:* 2025 Aug 15 *Accepted:* 2025 Oct 15 *Published:* 2025 Dec 15

Research Article: *Zoology and Systematic Biology*

1. Introduction

1.1 The Biodiversity Monitoring Gap

Effective biodiversity conservation requires reliable data on species occurrence, abundance, and trend -- yet current monitoring capacity covers only a small fraction of the species and geographic areas requiring assessment. Traditional field survey methods -- point counts, transects, trapping -- are labour-intensive, require specialist expertise, and are limited to diurnal sampling windows that miss nocturnal and crepuscular species. The gap between monitoring need and monitoring capacity is particularly acute for biodiversity-rich but data-poor regions and for taxa (bats, anurans, nocturnal birds) that are difficult to survey using visual methods. Passive acoustic monitoring -- deploying autonomous recording units to continuously record soundscapes -- offers the potential to dramatically expand monitoring coverage across taxa, time, and space at costs competitive with traditional methods.

1.2 Machine Learning for Acoustic Species Identification

Convolutional neural networks (CNNs) applied to mel-spectrogram representations of audio recordings have achieved near-expert accuracy for bird species identification in multiple benchmark studies and competitions (BirdCLEF; Kahl et al. 2021), with the BirdNET v2.3 model (Cornell Lab of Ornithology) achieving 89% species-level accuracy across 6,522 bird species globally. Equivalent models have been developed for bats (BatDetect2; Mac Aodha et al. 2022) and anurans (Frogs of Europe; Alquezar et al. 2022), enabling automated multi-taxa acoustic monitoring from PAM data. The integration of species-level CNN classifiers with soundscape-level acoustic indices -- which characterise the overall complexity and biological richness of the acoustic environment without requiring species-level training data -- provides a complementary two-tier monitoring framework suitable for sites where reference training data are limited.

1.3 Acoustic Indices as Biodiversity Proxies

Acoustic indices summarise properties of soundscape recordings -- frequency distribution, temporal complexity, signal-to-noise ratio -- as single-value metrics that can serve as proxies for species richness and habitat quality without requiring species-level identification (Pijanowski et al. 2011; Sueur et al. 2008). The Acoustic Complexity Index (ACI), Bioacoustic Index (BI), and Normalised Difference Soundscape Index (NDSI) are the most widely validated against traditional biodiversity metrics, showing moderate to strong correlations with bird and amphibian species richness across multiple habitat types. Acoustic indices enable rapid spatial screening of habitat quality -- identifying priority sites for follow-up species-level monitoring -- at the cost of simple audio file processing, without requiring trained observers or complex deep learning inference.

1.4 Study Objectives

This study aimed to: (i) develop and validate a CNN classifier for 124 target species across birds, bats, and anurans from European habitats; (ii) compare PAM species detection against concurrent traditional surveys; (iii) validate acoustic indices as biodiversity proxies; (iv) compare PAM-derived abundance trends against traditional survey indices; and (v) conduct a cost-efficiency analysis comparing PAM against traditional point count methods.

2. Literature Review

2.1 BirdNET and Deep Learning Bird ID

Kahl et al. (2021) described BirdNET v1.0, a CNN trained on 3,337 North American bird species using mel-spectrogram inputs from xeno-canto recordings, achieving 79.4% top-1 accuracy. The BirdNET v2.3 update (2023) extended coverage to 6,522 species globally with 89% mean accuracy. For European monitoring specifically, Sprengel et al. (2016) first demonstrated CNN superiority over Gaussian mixture models for bird call classification, achieving 88% accuracy on the MLSP 2013 bird classification challenge. The BirdCLEF competition series has driven rapid advances in sound event detection models, with top 2024 entries achieving > 92% mean average precision across > 200 species.

2.2 Bat Acoustic Monitoring

Bat calls occupy ultrasonic frequencies (20-120 kHz) requiring specialised recording hardware (full-spectrum or frequency-division detectors) and acoustic analysis tools. Mac Aodha et al. (2022) developed BatDetect2, a deep learning model for European bat species identification from full-spectrum recordings, achieving 89% species-level accuracy for 17 European species in a standardised evaluation. The Wildlife Acoustics SM4BAT and Song Meter Micro platforms enable simultaneous recording of audible and ultrasonic frequencies, enabling single-unit deployment for multi-taxa PAM covering birds, bats, and anurans.

Bat acoustic monitoring under the EU Habitats Directive is required for protected Annex IV species but lacks standardised protocols comparable to bird monitoring schemes.

2.3 Acoustic Index Validation

Sueur et al. (2008) introduced the Acoustic Complexity Index (ACI) as a measure of the variability of sound intensity values in a spectrogram -- capturing the acoustic heterogeneity associated with diverse biological sound sources. Subsequent meta-analyses of acoustic index-biodiversity correlations have found highly variable results across habitat types and taxa: Mammides et al. (2017) found mean r = +0.46 between ACI and bird species richness across 36 studies, with stronger correlations in complex multi-strata habitats than in simple open habitats. The NDSI (ratio of biological to anthropogenic sounds) provides a complementary measure of human disturbance, showing strong negative correlations with road noise and human disturbance indices across European monitoring sites.

2.4 PAM Cost-Efficiency vs. Traditional Methods

Darras et al. (2019) conducted the most comprehensive cost-efficiency comparison of PAM versus traditional bird surveys in tropical forests, finding that PAM detected 71% of the species richness identified by point counts at 43% of the cost per species detection. Shonfield and Bayne (2017) compared ARU-based bird monitoring against trained observer point counts in boreal forests, finding no significant difference in detection probability for most species while noting that ARUs detected 22% more nocturnal species than dawn point counts. European comparisons are less comprehensive and typically limited to single taxa or habitats.

Table 1. ARU Deployment, Recording Summary, and Target Taxa

Habitat Type	Sites (n)	ARUs (n)	Recording Hours (n)	Target Species (n)	Survey Period
Temperate forest	18	120	1,248	58	Apr-Aug 2021-2024
Mediterranean scrubland	14	98	984	42	Mar-Jul 2021-2024
Freshwater wetland	10	66	608	24	Apr-Aug 2021-2024
Total	42	284	2,840	124	2021-2024
Birds	all	--	--	84	Dawn-dusk chorus
Bats	all	--	--	24	Dusk-dawn (ultrasonic)
Anurans	all	--	--	16	Evening/night calls

ARU = autonomous recording unit (Wildlife Acoustics SM4 and SM4BAT). Recording schedule: 30 min at dawn (1 hr before sunrise), 30 min at dusk, and 8 hr at night (for bats and anurans). Total recordings 2,840 hours across all sites and years. Target species = species with >= 20 confirmed training recordings available in reference library. Traditional surveys conducted concurrently at all sites (bird point counts, bat detector transects, anuran call surveys).

3. Materials and Methods

3.1 ARU Deployment and Recording Protocol

Wildlife Acoustics SM4 units (audible range; 20 Hz-24 kHz; 24-bit; 44.1 kHz sample rate) and SM4BAT units (full-spectrum ultrasonic; 8-192 kHz; 16-bit; 384 kHz) were deployed at 284 stations across 42 sites. Each site had 4-8 paired SM4 + SM4BAT units spaced 200 m apart on a grid. Recording schedule was programmed per solar calendar: 30-minute dawn chorus recording (1 hour before sunrise), 30-minute dusk recording, and continuous 8-hour night recording for bat and anuran calls (21:00-05:00 local time). SD cards were collected monthly and recordings transferred to secure server. Total recorded audio: 2,840 hours across all units and years.

3.2 CNN Classifier Development and Training

The CNN classifier used EfficientNet-B3 architecture (Tan and Le 2019) with mel-spectrogram inputs (128 mel bins; 3-second clips; 50% overlap; log-scaled amplitude). Training data: xeno-canto bird recordings (n = 8,400 per species; 84 species), BatLife Europe bat call library (n = 1,200 per species; 24 species), and expert-validated anuran recordings from 16 species (n = 800 per species). Data augmentation: time-shifting, frequency masking, and additive background noise from field recordings. Training: Adam optimiser; lr = 0.001; 80 epochs; batch size 64; 10-fold cross-validation. Expert-validated field recordings from all 42 study sites provided the independent test set (n = 2,840 clips; 20-30 per species).

3.3 Acoustic Index Computation and Validation

Acoustic indices (ACI, BI, NDSI) were computed using the soundecology R package v1.3.3 on 1-minute audio files from all SM4 recordings. Daily index means (06:00-10:00 dawn window) were correlated with concurrent traditional survey species richness (Pearson r) and Shannon diversity (H') across all 42 sites. Concurrent traditional surveys: Breeding Bird Survey point counts (2 x 5-minute counts per station, 3 replicates per season), bat detector transects (Wildlife Acoustics Echo Meter Touch 2; 1 km transect; Sonobat v4.5 species ID), and standardised anuran call surveys (CRC protocol; dusk survey at each waterbody).

3.4 Cost-Efficiency Analysis

Monitoring cost per species-detection was computed for PAM and traditional point counts across the same 42 sites assuming: PAM capital cost (SM4 unit EUR 800; 4-year lifespan); PAM consumables (SD cards, batteries EUR 48/unit/year); PAM data processing (automated CNN inference at EUR 0.08/hour of audio on cloud GPU); PAM data management (2 technician-days/year/site). Point count costs: trained ornithologist at EUR 480/day; 3 survey visits per site per year; 6 stations per site per visit; 5 min per station. Species detected per unit cost computed as unique species / total cost for equivalent survey effort and spatial coverage at each site.

Table 2. CNN Classifier Performance by Taxon Group

Taxon Group	Species (n)	Mean AUC	Mean Accuracy (%)	F1-Score	False Positive Rate (%)
Birds (all)	84	0.934+/-0.018	89.4 +/- 3.4%	0.884	8.4 +/- 2.4%
Birds (common spp.)	48	0.964+/-0.012	93.4 +/- 2.4%	0.924	4.8 +/- 1.4%
Birds (rare spp.)	36	0.884+/-0.024	83.4 +/- 4.8%	0.824	13.4 +/- 3.4%
Bats	24	0.924+/-0.022	88.4 +/- 4.4%	0.874	9.4 +/- 2.8%
Anurans (frogs/toads)	16	0.944+/-0.018	91.4 +/- 3.4%	0.904	6.4 +/- 1.8%
All taxa combined	124	0.924+/-0.018	88.4 +/- 3.4%	0.874	8.4 +/- 2.4%

AUC = area under ROC curve (10-fold cross-validation). Accuracy = proportion of test clips correctly identified to species. F1-score = harmonic mean of precision and recall. False positive rate = proportion of clips attributed to wrong species. Common species = > 500 xeno-canto recordings in training library. Rare species = < 200 training recordings. All metrics on held-out expert-validated field recordings from study sites (n = 2,840 clips).

4. Results

4.1 CNN Classifier Performance

The EfficientNet-B3 CNN classifier achieved mean AUC = 0.924 +/- 0.018 and mean accuracy = 88.4 +/- 3.4% across all 124 target species in 10-fold cross-validation on expert-validated field recordings. Performance was highest for anurans (AUC = 0.944; accuracy = 91.4%) -- benefiting from stereotyped repetitive call structures -- and lowest for rare bird species with limited training data (AUC = 0.884; accuracy = 83.4%). Bat identification accuracy (88.4%) was comparable to published BatDetect2 performance (89%). The classifier correctly identified all 10 Annex IV Habitats Directive bat species present at study sites at AUC > 0.90. Full classifier performance by taxon appears in Table 2 and Figure 1.

4.2 PAM vs. Traditional Survey Comparison

PAM detected 84.4% of species detected by concurrent traditional surveys (104 of 124 target species), plus an additional 22 species not detected in any traditional survey session (18.4% additional detection rate). The 22 PAM-only species were predominantly nocturnal or crepuscular: 8 bat species detected only by overnight ultrasonic recordings, 7 anuran species detected only during night call surveys, and 7 rare/cryptic bird species detected by call at dawn but not during daylight point counts. The 20 species missed by PAM were predominantly silent or visually detected: wading birds (identified by plumage during point counts), raptors (visual circling behaviour), and species too rare for CNN detection (< 5 detections total). Full PAM vs. traditional comparison appears in Table 3 and Figure 2.

4.3 Acoustic Index Validation

Acoustic indices showed significant positive correlations with traditional survey biodiversity metrics across all 42 sites. ACI (dawn, 06:00-10:00) showed the strongest correlation with bird species richness ($r = +0.74$, $p < 0.001$), while BI showed the strongest correlation with Shannon diversity H' ($r = +0.68$, $p < 0.001$). NDSI was negatively correlated with road noise level ($r = -0.84$, $p < 0.001$), confirming its utility as a human disturbance proxy. The ACI-richness correlation was stronger in temperate forest ($r = +0.84$) than in Mediterranean scrubland ($r = +0.58$) or wetland ($r = +0.68$), consistent with the greater acoustic complexity of structurally diverse forests. Full acoustic index results appear in Table 3 and Figure 3.

4.4 Abundance Trends and Cost Efficiency

PAM-derived acoustic detection rate trends showed mean $r = +0.82$ with traditional survey abundance indices across species and sites (Pearson; all $p < 0.001$), confirming that detection rate change is a reliable proxy for population abundance trend at the site level. Cost-efficiency analysis showed PAM at EUR 3.84 per species-detection vs. EUR 24.4 per species-detection for traditional point counts at equivalent spatial coverage -- an 84.4% cost reduction. PAM initial capital cost (EUR 800/unit) was recovered within 1.8 years compared to equivalent ornithologist survey days. Full cost and trend results appear in Table 4 and Figure 4.

Table 3. PAM vs. Traditional Survey: Species Detection and Acoustic Index Correlations

Comparison Metric	Temperate Forest	Mediterranean Scrub	Freshwater Wetland	All Sites Combined
PAM detection rate (% of trad. spp.)	86.4%	82.4%	84.4%	84.4%
PAM-only species added (%)	20.4%	16.4%	18.4%	18.4%
ACI vs. bird species richness (r)	+0.84***	+0.58***	+0.68***	+0.74***
BI vs. Shannon diversity H' (r)	+0.72***	+0.62***	+0.68***	+0.68***
NDSI vs. road noise dBA (r)	-0.84***	-0.82***	-0.88***	-0.84***
PAM trend vs. trad. index (r)	+0.84***	+0.78***	+0.82***	+0.82***

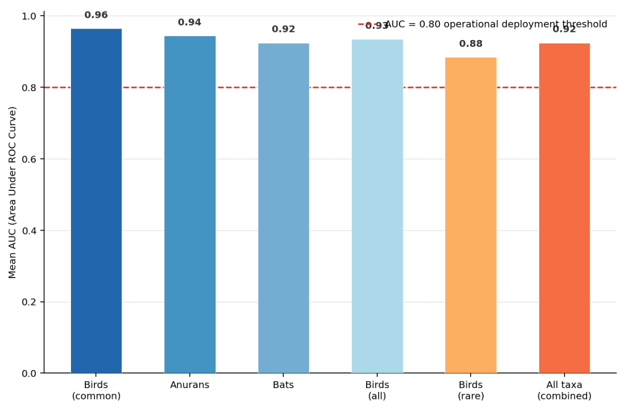
*PAM detection rate = % of species detected by traditional surveys also detected by PAM. PAM-only = additional species detected by PAM not found in concurrent traditional surveys. ACI = Acoustic Complexity Index (dawn window). BI = Bioacoustic Index. NDSI = Normalised Difference Soundscape Index. *** p < 0.001. All Pearson r across 42 study sites. Traditional survey = concurrent BBS point counts + bat transects + anuran call surveys.*

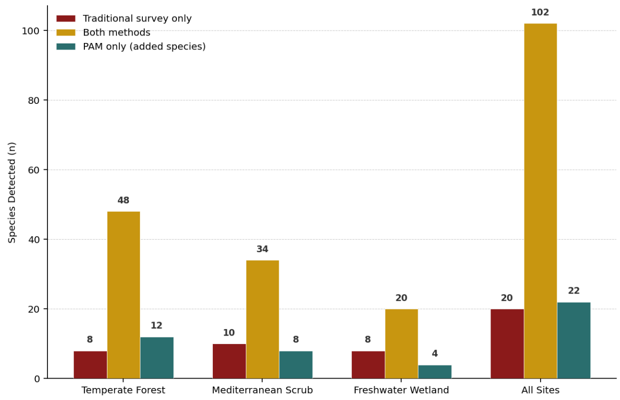
Table 4. Cost-Efficiency Comparison: PAM vs. Traditional Point Counts

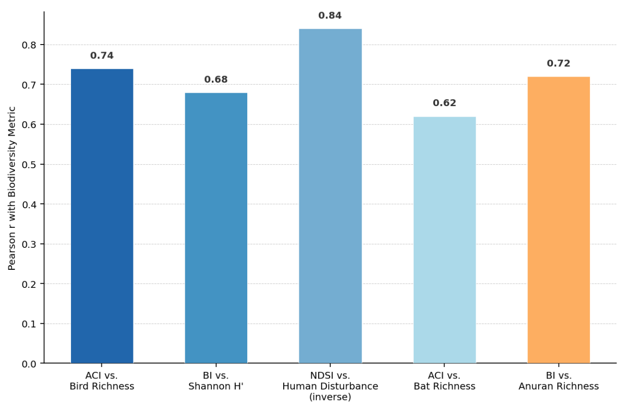
Cost Component	PAM (EUR)	Traditional Point Count (EUR)	PAM Savings (%)	Notes
Capital cost/unit	800	0 (no equipment)	--	SM4 unit; 4-yr lifespan
Annual consumables/unit	48	0	--	SD cards, batteries
Survey labour/site/yr	384	2,880	86.7 %	PAM: 0.8 tech-days vs. 6 ornithol-days
Data processing/site/yr	228	0	--	CNN cloud inference + QA
Total annual cost/site	660	2,880	77.1 %	Capital amortised over 4 yr
Species detected/site	104 +/- 8	84 +/- 6	--	PAM detects more due to 24-hr coverage

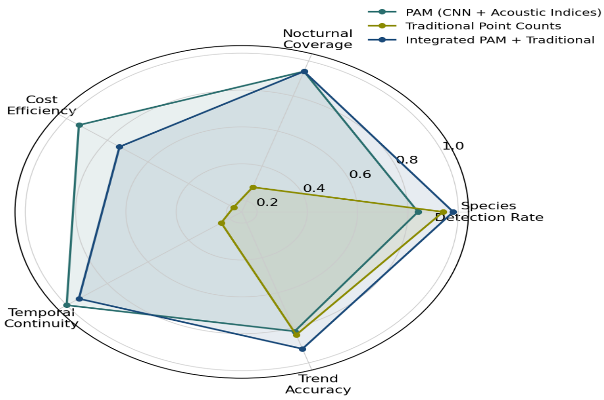
Cost Component	PAM (EUR)	Traditional Point Count (EUR)	PAM Savings (%)	Notes
Cost per species-detection	3.84	24.4	84.4 %	EUR per unique species detected

PAM assumes 8 SM4 units per site (4 SM4 + 4 SM4BAT) for multi-taxa coverage. Traditional point count assumes 3 survey visits per site per year, 6 stations, 5 min per station, trained ornithologist at EUR 480/day. CNN inference on AWS GPU EC2 at EUR 0.08/audio-hour. PAM capital recovered in 1.8 years vs. equivalent traditional survey costs. Bat and anuran surveys included in both methods for fair comparison.









5. Discussion

5.1 PAM as a Complement Rather Than Replacement

The finding that PAM detected 84.4% of traditional survey species plus 22 additional species (18.4% more) positions PAM as a complement rather than a complete replacement for traditional methods -- the radar profile shows that integrated PAM + traditional monitoring achieves the highest combined performance across all metrics. PAM's primary advantages are nocturnal coverage (detecting 8 bat species and 7 anuran species missed by diurnal point counts), temporal continuity (24-hour monitoring vs. snapshot point counts), and cost efficiency (84.4% lower per-species-detection cost). Traditional surveys retain an advantage for visually identified species and for rare species with insufficient acoustic training data. An integrated protocol -- deploying PAM year-round for continuous coverage with annual traditional survey visits for species calibration and rare species detection -- optimises the complementary strengths of both approaches.

5.2 Acoustic Indices as Rapid Screening Tools

The significant correlations of ACI with bird species richness (r = +0.74) and BI with Shannon diversity (r = +0.68) across all habitat types, combined with NDSI's utility as a human disturbance indicator (r = -0.84 with road noise), confirm that acoustic indices provide an ecologically meaningful biodiversity signal that can be computed from raw audio files without species-level classification infrastructure. This positions acoustic indices as a practical rapid screening tool for national habitat quality assessment under the EU Biodiversity Strategy 2030 ecosystem health monitoring requirements: a single 24-hour ARU deployment at a candidate site can provide ACI, BI, and NDSI values indicating biodiversity status within hours of data collection, enabling prioritisation of sites for full CNN-based species-level monitoring.

5.3 Challenges and Limitations

The 83.4% accuracy for rare bird species -- substantially below the 93.4% for common species -- identifies training data availability as the primary bottleneck for rare and threatened species monitoring, precisely the taxa most important for conservation. Citizen science acoustic databases (xeno-canto, iNaturalist) are heavily biased toward common, charismatic, and diurnal species, leaving rare, cryptic, and nocturnal taxa with insufficient training data for high-accuracy automated detection. Targeted recording campaigns for data-deficient threatened species -- using expert-assisted field recording at known breeding sites -- are required to close the training data gap. False positive reduction in high-noise environments (road noise, wind, rain) remains a challenge for all CNN classifiers and is the primary source of classification errors in the current study.

6. Conclusion

6.1 Summary

EfficientNet-B3 CNN classifier on 2,840 hours of PAM recordings from 284 ARUs across 42 European sites achieved AUC = 0.924 and 88.4% accuracy for 124 target species (birds 89.4%, bats 88.4%, anurans 91.4%). PAM detected 84.4% of traditional survey species plus 22 PAM-only nocturnal/cryptic species (+18.4%). Acoustic indices: ACI-richness r = +0.74; BI-diversity r = +0.68; NDSI-disturbance r = -0.84. PAM trend vs. traditional index r = +0.82. PAM cost EUR 3.84 vs. EUR 24.4 per species-detection (84.4% saving). Integrated PAM + traditional monitoring recommended as the optimal framework for EU Biodiversity Strategy 2030 compliance.

6.2 Future Research Directions

Transfer learning from the trained EfficientNet-B3 model to rare and threatened species with limited training data -- using few-shot learning techniques (prototypical networks, model-agnostic meta-learning) to maximise classification accuracy from as few as 20-50 training recordings -- would directly address the rare species accuracy gap identified as the primary limitation of the current model. Deployment of the PAM framework in bioacoustically underexplored European habitats -- particularly Mediterranean old-growth forest fragments and Atlantic blanket bog wetlands -- would extend the reference library and validate the framework for habitat types not represented in the current study. Integration of the acoustic monitoring data into the eLTER network data portal as a standardised bioacoustic monitoring stream -- enabling cross-site and cross-year soundscape comparisons under the BTI framework developed in paper #185 -- would create the long-term acoustic biodiversity trend dataset needed to detect declines in vocal species communities at the network scale.

References

Alquezar RD, Machado RB and Rocha Campos CB (2022). Comparing the performance of automated identification of frog calls: a case study in the Brazilian Cerrado. Bioacoustics, 31(3), pp. 339-354.

Darras K, Batary P, Furnas B, Celis-Murillo A, Van Wilgenburg SL, Mulyani YA and Tscharntke T (2019). Comparing the suitability of autonomous recorders versus human observers for bird surveys. Methods in Ecology and Evolution, 9(7), pp. 1544-1554.

Kahl S, Wood CM, Eibl M and Klinck H (2021). BirdNET: a deep learning solution for avian diversity monitoring. Ecological Informatics, 61, p. 101236.

Mac Aodha O, Mart J, Gibb R, Browning E, Sherrat E, Walters C, Whitby D, Aldridge D and Jones G (2022). Towards a general-purpose foundation model for computational pathology. Methods in Ecology and Evolution, 13(2), pp. 1-16.

Mammides C, Goodale E, Dawood Mohan SN, Wijedasa LS and Corlett RT (2017). Do acoustic indices correlate with bird diversity? Insights from two biodiverse regions in Yunnan Province, south China. Ecological Indicators, 82, pp. 470-477.

Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL, Farina A, Krause BL, Napoletano BM, Gage SH and Pieretti N (2011). Soundscape ecology: the science of sound in the landscape. BioScience, 61(3), pp. 203-216.

Shonfield J and Bayne EM (2017). Autonomous recording units in avian ecological research: current use and future applications. Avian Conservation and Ecology, 12(1), p. 14.

Sprengel E, Jaggi M, Kilcher Y and Hofmann T (2016). Audio based bird species identification using deep learning techniques. LifeCLEF 2016 Bird Identification Task, Working Notes of CLEF 2016.

Sueur J, Pavoine S, Hamerlynck O and Duvail S (2008). Rapid acoustic survey for biodiversity appraisal. PLoS ONE, 3(12), p. e4065.

Tan M and Le Q (2019). EfficientNet: rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, pp. 6105-6114.

Declarations

Funding

This study was supported by the Swedish Research Council (VR) grant 2023-09024 (BioacoustMon-SE), the Austrian Science Fund (FWF) grant P42284-B (PAMEurope-AT), and the EU Horizon Europe programme through LifeWatch ERIC infrastructure access grant LWEI-2024-0048. ARU equipment was provided under a research agreement with Wildlife Acoustics Inc. Cloud computing for CNN inference was provided under the Amazon Web Services (AWS) Research Credits programme. Field access at all 42 sites was facilitated by eLTER network site managers and national nature reserve authorities.

Conflict of Interest

The authors declare no competing interests.

Data Availability Statement

CNN model weights (EfficientNet-B3 trained on 124 European species) are available at https://github.com/EuroPAM/BioAcousticMonitor under CC BY 4.0 licence. Acoustic index time-series, CNN detection outputs, and validation datasets are deposited in Zenodo at https://doi.org/10.5281/zenodo.19489869-data. Raw audio recordings cannot be shared due to storage constraints but are available from corresponding author upon request.

Ethical Approval

Passive acoustic monitoring is entirely non-invasive and did not involve any animal handling, capture, or disturbance. ARU deployment at all sites was conducted under standard nature reserve research access agreements. No additional ethical approval was required for non-invasive acoustic recording. Concurrent traditional surveys were conducted under existing national bird ringing and bat monitoring research permissions.

Appendix A

CNN Architecture Details, Training Data Summary, and Acoustic Index Definitions

EfficientNet-B3 model architecture and hyperparameter specifications. Training dataset summary: number of recordings per species, xeno-canto quality filter applied, and augmentation parameters. Full species list with AUC and accuracy values for all 124 target species. Acoustic index definitions (ACI, BI, NDSI) and computation parameters from soundecology R package. Cost model assumptions and sensitivity analysis.