Audio embeddings identify novel sounds and track soundscape change within a tropical restoration landscape

Jared de Silva[1*] and Cristina Banks-Leite[1]

[1] Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, UK

*Author for correspondence: j.de-silva25@imperial.ac.uk ; jareddesilva1@gmail.com

Keywords: agile modelling; audio embeddings; passive acoustic monitoring; insect sonotypes; restoration

Abstract

Passive acoustic monitoring (PAM) is increasingly being used over conventional biodiversity surveys because it is more cost-effective and implementable at a landscape level. However, the cheaper sensors that removed the barriers to PAM have now shifted the challenge from data collection to interpretation. Acoustic indices address this by condensing each recording into single summary statistics, but their relationship to biodiversity varies widely across studies. The alternative, acoustic classifiers, name what is calling. However, they are mainly trained on bird call, so exclude the insects and frogs that dominate tropical nocturnal soundscapes and serve as powerful bioindicators. At most tropical restoration sites, the non-avian vocalising community has not been catalogued, and therefore its contribution to changing soundscapes cannot be quantified meaningfully. Here, we built a pipeline using Perch v2 audio embeddings to identify novel sounds and to measure the dissimilarity of the non-avian community between reforested pasture and forest. We tested it on >3,880 hours of audio recorded at reforestation sites in the Amazon over three consecutive years. Using an agile modelling approach, we built acoustic classifiers for fifteen frog and insect sounds (sonotypes) and derived a dissimilarity in sonotype composition between pasture and forest. We also used the embeddings as an acoustic index to quantify dissimilarity between the whole night-time soundscapes of pasture and forest. Both dissimilarities showed reforested pasture soundscapes diverging from their baseline and becoming more similar to forest after two years. Two forest-associated sonotypes became more prevalent over time, while one pasture-associated sonotype declined, highlighting the sonotypes driving species turnover. Only five sonotypes could be matched to a candidate frog species from public databases, likely due to limited archive coverage. Our pipeline requires no reference library or fine-tuning, and therefore any restoration projects with recorded audio can measure change in their uncatalogued vocalising communities.

Introduction

Complex systems such as tropical rainforests are difficult to monitor given the vast number of species and their cryptic diversity (Guayasamin *et al*., 2024). Conventional field surveys used to assess them are too costly and time-consuming to meet the pace at which monitoring data is required (King & Halpern, 2025). To address this problem, passive acoustic monitoring (PAM) is increasingly being used as it is sufficiently cheap and widely available (Hill *et al*., 2018; Gibb *et al*., 2019). The limiting step is no longer data collection but extracting meaningful ecological signal from terabytes of audio (Napier *et al*., 2024). This extraction has mostly relied on acoustic indices (Somervuo *et al*., 2025) which correlate only moderately with biological diversity and perform inconsistently across studies (Alcocer *et al*., 2022; Hutschenreiter *et al*., 2026), with their predictive power weakening toward the equator (Pan *et al*., 2024). Large pre-built acoustic classifiers exist, though almost exclusively for birds, the only group with annotated libraries large enough for pretraining (Williams *et al*., 2025a). Consequently, such classifiers miss the insects that dominate tropical soundscapes (Riede & Balakrishnan, 2025) and the frogs which can serve as useful bioindicators (Wickramasingha *et al*., 2026). New bioacoustics approaches are urgently needed to monitor these overlooked communities.

Multi-taxa bioacoustics models present alternatives. Whereas traditional acoustic indices reduce a recording to a single number capturing one aspect of the sound, these models describe it with hundreds of numbers. Each number may represent different features the model learned from raw audio (McCarthy *et al*., 2025). These numbers form a vector called an embedding, which places every recording in a shared space where similar sounds lie closer together (Stowell, 2022). Position in that space differentiates soundscapes by community and habitat (Sethi *et al*., 2020) and enables unlabelled recordings to be matched against labelled ones (Peralta *et al*., 2025). Google's Perch v2 (referred to hereafter as "Perch") is the leading bioacoustics model under linear probing, wherein a simple classifier is trained on embeddings (van Merriënboer *et al*., 2025; Pickering *et al*., 2025; Schwinger *et al*., 2026). Because the model learns the general acoustic structure, embeddings can transfer to taxa and tasks far outside its training data (Ghani *et al*., 2023; Burns *et al*., 2025).

Embedding features can function as an acoustic index encapsulating animal calls and environmental noise. On three biogeographically independent reef datasets, pretrained features outperform standard indices in classifying fish community, coral cover, and depth zone, and nearly match custom models at a fraction of the computation (Williams *et al*., 2025b). However, learned features could not predict absolute species richness across datasets, but tracked community turnover within each set (Sethi *et al*., 2023). An embedding index therefore registers that a soundscape has changed without describing the change itself. Where vocalising animals are unknown, interpreting the soundscape requires a way to build classifiers for species without reference libraries.

Agile modelling is a recent approach for rapid classifier building that works directly with precomputed Perch embeddings and requires only one example of the target sound to start (Dumoulin *et al*., 2025). Since every window (five seconds long) in the archive is already embedded, that example serves as a query to retrieve, in seconds, the most acoustically similar windows from millions of candidates and can surface sound classes not known to be present beforehand (Allen-Ankins *et al*., 2025). Iterative human confirmation of these retrievals, with active learning choosing the most important clips to review next (Kath *et al*., 2024; Rauch *et al*., 2024), allows training a classifier from very few labels. Where no reference recording exists for that first example, the target becomes a sonotype, an acoustic morphotype distinguished by ear and on a spectrogram. Sonotypes track real ecological patterns, for example, insect sonotype richness following diel cycles in Brazilian tropical savanna (Ferreira *et al*., 2018) and seasonal cycles in Amazonian forest (Howells *et al*., 2025). They also work at the community level, where fish sonotype detections alone differentiated healthy, restored, and degraded reefs (Dumoulin *et al*., 2025). Thus, agile modelling offers a way forward for detecting and analysing novel calls which are often prevalent in the complex soundscapes of tropical rainforests.

Amazonian deforestation is concentrated in eastern Pará, Brazil, driven largely by pasture expansion (Cromberg *et al*., 2026). Restoration in Pará has included planting native seedlings, with success typically judged from tree composition and community structure (da Cruz *et al*., 2020). Whether fauna actually return is harder to establish because recovery rates differ between taxa and unfold over decades (Lennox *et al*., 2018). Insects dominate Pará's night-time audio (Howells *et al*., 2025), and as small ectotherms whose calling tracks microhabitat conditions, they register habitat change more precisely than vertebrates (Riede &

Balakrishnan, 2025). Frog assemblages in Amazonia are likewise structured by deforestation (Torralvo *et al*., 2022). Acoustic reference libraries for both these groups remain patchy, as under 10% of cricket and grasshopper species have documented calls (Riede, 2018), and roughly a third of Brazilian anurans lack described calls, with nearly half lacking archived recordings (Guerra *et al*., 2018). As a result, tropical insect and frog acoustics have largely been avoided. When the soundscapes of tropical restored habitat have been analysed, it has generally been through snapshots across time of avian recovery (Ramesh *et al*., 2023; Kümmet *et al*., 2025). Thus, a method of measuring acoustic change that can capture both insects and frogs may provide new insights into the effects of restoration on tropical soundscapes.

This study creates a pipeline for acoustic restoration studies that lack local reference libraries, enabling analysis of novel vocalisations and understanding changes in the soundscape beyond birds. Recordings span three years at a reforestation project in Brazil with the same pasture and forest sites sampled each year. Our first aim (a) is to identify non-avian sonotypes and use agile modelling to build classifiers that track their response to reforestation. Our second aim (b) is to investigate if the reforested pastures' soundscape and sonotype composition are becoming more similar to forest. We predict that reforestation shifts pastures' sonotype composition toward a more forest-like assemblage, with their whole soundscapes following this trajectory.

Methods

2.1 Study site and data collection

The study site lies in the Amazonian 'Arc of Deforestation' in north-eastern Pará, Brazil (Cromberg *et al*., 2026). Its climate is tropical monsoon (Köppen Am), with a mean annual temperature of 27 °C and, near the site (Mãe do Rio), a mean annual precipitation of ~2,248 mm (Climate-Data.org, n.d.). The site, a former cattle ranch, has been under active native-species restoration by the carbon-credit company Mombak since 2023. It covers ~2,933 ha, of which about 6.1% is forest and 88.5% pasture (Figure 1).

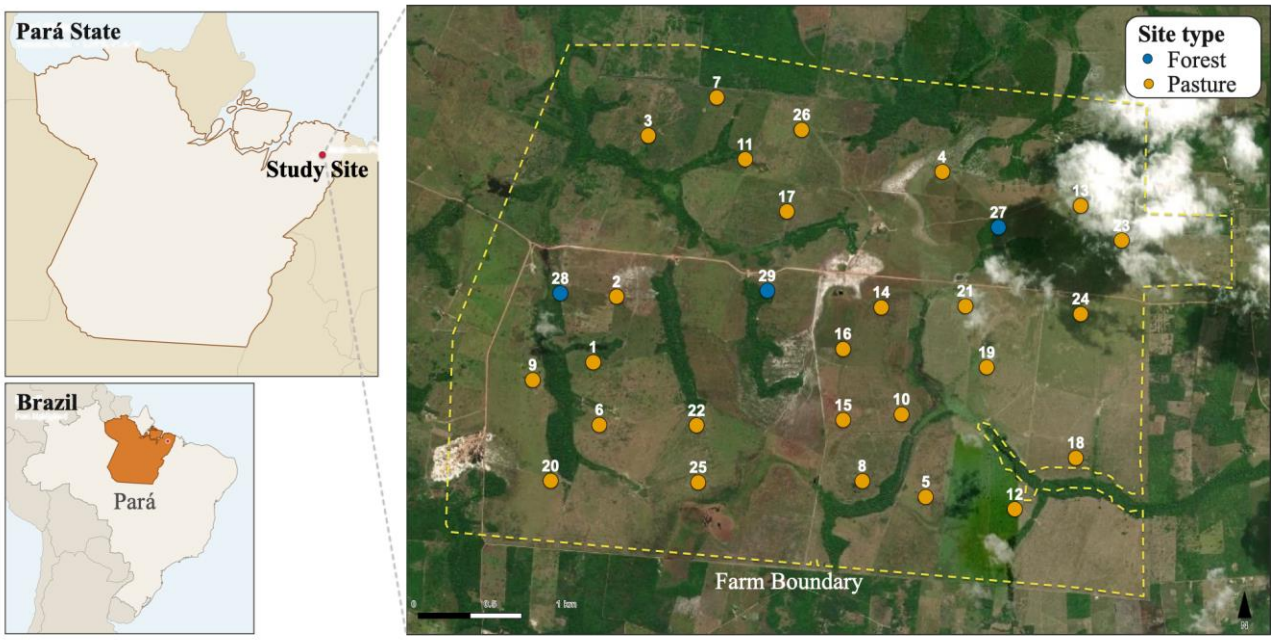


Figure 1. The 29 passive acoustic recorder sites in the study landscape, north-eastern Pará, Brazil (2°05'S, 47°28'W). Three forest reference sites and 26 pasture sites were sampled in 2023, 2024 and 2025. All pasture sites entered restoration in 2023. The property boundary (dashed yellow line) was approximated from the cadastral polygon in the project's registry entry (Isometric, 2026). Base imagery is from the ArcGIS Living Atlas Wayback service (Esri, 2026).

Acoustic data were recorded at 29 sites over three consecutive years (2023, 2024, and 2025), with 3 sites in forest patches and 26 in pasture, and a minimum distance of 500 m between sites. Each device recorded for 1 minute, every 5 minutes, 24 hours a day, over 7 days at every site, creating 48 kHz .wav files. All 29 sites recorded in 2023 used AudioMoth 1.0.0 devices (Hill *et al*., 2018; Open Acoustic Devices, n.d.) running firmware 1.7.1. Additionally, Song Meter Mini 2 Bat devices (Wildlife Acoustics, n.d.) ran alongside AudioMoths at 15 of the 29 sites in 2024 and at all 29 sites in 2025. After accounting for device failures and corrupted data, recordings were retained from 27 sites in 2023, 29 sites in 2024 and 29 sites in 2025, yielding 2,793,688 five-second acoustic windows (3,880.1 hours of audio) for embedding and analysis.

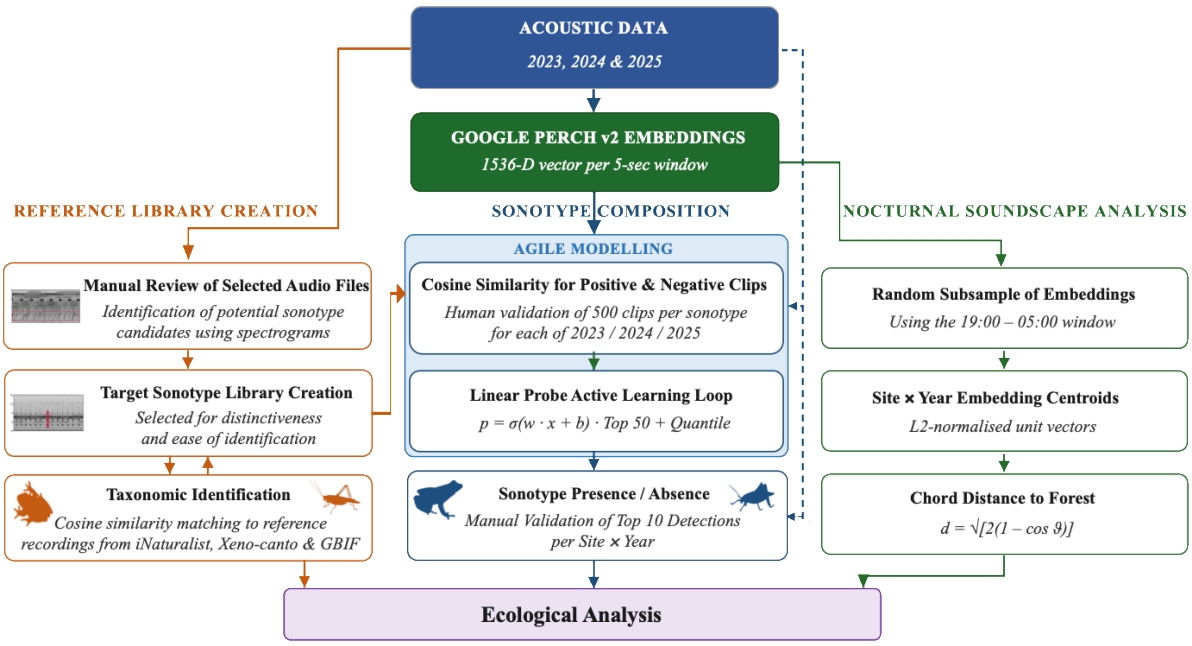
2.2 Software

We ran all analyses in R 4.5.3 (R Core Team, 2026), which called the Kaggle version of Perch v2 (van Merriënboer *et al*., 2025) and ran the linear classifier in Python through "reticulate" (Ushey *et al*., 2024). Sonotype annotation used Raven Pro 1.6 (K. Lisa Yang Center for Conservation Bioacoustics, 2019). Review workbooks were assembled from "tuneR" clips (Ligges *et al*., 2023) and "seewave" spectrograms

(Sueur *et al*., 2008). Classifier metrics were computed with "pROC" (Robin *et al*., 2011) and "PRROC" (Grau, Grosse & Keilwagen, 2015). Indicator analysis was run in "indicspecies" (De Cáceres & Legendre, 2009) with restricted permutations from "permute" (Simpson, 2022), and all mixed models were fitted with "glmmTMB" (Brooks *et al*., 2017) following Harrison *et al*. (2018).

2.3 Processing pipeline

Three workstreams run from raw audio to ecological analysis (Figure 2). Every five-second window becomes a Perch embedding, and every later step uses these as a foundation (§2.4). The first workstream surfaced the candidate sonotypes. Sonotypes were annotated by hand from spectrograms and were then compared against three public reference databases for preliminary species identifications (§2.5). The second workstream used one exemplar clip as a search query via cosine similarity to retrieve candidate clips for manual verification. Then those labels were grown over three cycles of reviewed detections and used to train one linear classifier per sonotype. The classifiers were evaluated on a validation split and dataset-corrected retrieval metrics. Each classifier was run over every site and year, and its top ten detections were manually reviewed to score presence or absence (§2.6). The third workstream summarised each site-year as one mean night-time embedding, measured by its chord distance to a fixed forest reference. The presence-absence matrix and those distances gave compositional and whole-soundscape measures that were both tested against time since restoration (§2.7).



135 **Figure 2. Processing pipeline from passive acoustic recordings to ecological analysis, built on Perch**
136 **embeddings and run as three parallel workstreams.** The workstreams are sonotype identification,
137 sonotype composition and nocturnal soundscape analysis. In sonotype identification (orange, headed
138 'Reference library creation' in the figure), sonotypes were manually found and annotated from spectrograms
139 of mainly nocturnal audio and later matched against public recordings of Neotropical frogs and insects to
140 generate candidate identifications. Sonotype composition (blue) used one exemplar clip per sonotype as a
141 search query, retrieving the most acoustically similar windows for manual validation. Those confirmed
142 windows trained a per-sonotype linear classifier, which then proposed its own next detections for review
143 across three cycles of subsequent review. Nocturnal soundscape analysis (green) averaged night-time
144 embeddings into one vector per site and year and measured its chord distance from a fixed forest reference,
145 functioning as an acoustic index. Ecological analysis interpreted the findings. Solid arrows carry
146 embeddings, dashed arrows indicate raw audio returned for human validation.

147 **2.4 Audio feature extraction with Perch**

148 Before feature extraction, every 60-second audio clip was resampled from 48 kHz to 32 kHz, converted to
149 mono float32, and, where necessary, trimmed or zero-padded into 12 non-overlapping 5-sec clips, to match
150 Perch input length. Embeddings of all clips were then computed on Imperial College's High-Performance
151 Computing (HPC) infrastructure using Perch as a frozen feature extractor applied without fine-tuning. We

mean-pooled the Perch output to create a single embedding of 1,536 dimensions (D), following van Merriënboer *et al*. (2025). Embedding creation is illustrated in Appendix A, Figure A.1.

2.5 Identifying sonotypes

In this study, a sonotype is considered a distinct acoustic signal that a non-expert can reliably recognise by ear and sight, without reference to the species producing it. Distinct candidate sonotypes were manually annotated in Raven Pro, without restriction on time of day, using bounding boxes around signals within five-second extracts from the 2023 baseline recordings, with an emphasis on night-time recordings to minimise selection of avian sonotypes. Noise and rain clips were manually identified from spectrograms of randomly selected five-second clips from all sites/years/devices for use as model negatives. For each sonotype, the exemplar clip was the five-second window that most closely overlapped its target annotation box, contained the fewest competing calls, and made the target call as dominant as possible within the window.

The retrieval and active-learning workflow followed agile modelling, with full protocols as per Dumoulin *et al*. (2025). Candidate detections were found by example, using each sonotype's exemplar clip as a search query. Cosine similarity between the query and every other window was computed in the 1,536-D embedding space as the dot product between L2-normalised vectors. A high cosine similarity meant only that Perch represented the two windows similarly, not that they contained the same sound (Steck *et al*., 2024). The top 500 windows per year across all sites were retained, giving up to 1,500 candidates per sonotype for manual review. Each retrieved window was manually verified before use. The proportion of retrieved windows confirmed as true positives was recorded per sonotype as a measure of retrieval precision. Retention was stratified by year because a query tended to retrieve windows most similar to its own year and site, so a global top 1,500 would be dominated by the query's year. This process gave each classifier positives from all three years, limiting domain shift, since a classifier trained mainly on one year may transfer poorly to the others. This year stratification of the search vector was the main deviation from published agile modelling methodology.

Review and manual validation produced 15 candidate sonotypes, which were each evaluated against three reference databases containing high-quality labelled audio recordings, namely iNaturalist (iNaturalist, 2026), Xeno-canto (Xeno-canto Foundation, 2026), and the Global Biodiversity Information Facility (GBIF, 2026).

The databases were filtered for frog (Anura) and insect (Insecta) species holding sound recordings in each of 20 South and Central American countries. Fifty recordings per species were drawn from a randomly ordered species list, capped at 20,000 recordings per database per country. Every recording (mostly mp3 & m4a) was decoded to 32 kHz mono, divided into five-second windows and embedded with Perch. Each sonotype clip was scored against the target sonotype reference library by cosine similarity, retaining the 50 most similar recordings per sonotype after pooling across countries. This produced 848 unique species, 724 Anura and 124 Insecta, against which the 15 target sonotypes were compared. The sonotypes were manually reviewed and circulated to taxonomists for review.

2.6 Classifier development, evaluation and presence-absence scoring

Fifteen per-sonotype linear classifiers were trained (Alain & Bengio, 2016), each a single dense unit with sigmoid activation and no hidden layers, fitting one weight per embedding dimension plus an intercept, which is 1,537 parameters per classifier. Each classifier was fitted with the Adam optimiser (learning rate 0.001) (Kingma & Ba, 2015) under binary cross-entropy loss for up to 200 epochs at batch size 32, using an 80:20 stratified train/validation split seeded with 42. To offset class imbalance, inverse-frequency class weights were applied (Johnson & Khoshgoftaar, 2019), early stopping was triggered on validation AUC – Area Under the Curve (patience 15, restoring best weights) (Prechelt, 1998), and the learning rate was halved after seven epochs without improvement on that metric (ReduceLROnPlateau, factor 0.5, patience 7).

Each classifier was trained against a shared pool of 3,000 manually curated negatives drawn from forest-noise and rain clips, giving positive-to-negative ratios of 1:2 to 1:39. Three active-learning cycles then grew each labelled set to roughly 600 labels, reviewing 200 detections per sonotype per cycle before cumulative retraining. The composition of the pool, the sampling rule applied at each cycle and the checks against label leakage are set out in Appendix B, Table B.3.

For evaluation, ROC-AUC, PR-AUC, and F1-max (with its threshold) were computed from the 80:20 split for each classifier. For dataset-corrected figures, plain and stratified (P(b)-weighted) retrieval-AUC were computed, with a species call-density $P(+) = \Sigma_b P(+|b)P(b)$, estimated under Beta priors with a 95% confidence interval from 10,000 bootstrap resamples of the per-bin Beta distributions (following Navine *et al*., 2024). A final review pass supplied the sample: 200 detections per sonotype from all scored windows,

the top 50 after non-maximum suppression plus 150 spread across eight log-score bins. The plain retrieval-AUC used all 200 detections; the stratified variant and the call density used only the 150 binned detections, each bin weighted by its share of all scored windows. For rare sonotypes such as sonotype M, no positives fell within the score-stratified sample, so the stratified (P(b)-weighted) retrieval-AUC could not be estimated, and the plain retrieval-AUC is reported instead. Permuted-label and random-embedding classifiers were fitted as negative controls, establishing the score a classifier achieves with no genuine acoustic signal.

Each sonotype's final trained classifier was run against all site-year-device cells except those contributing to its training positives; those cells were scored "present" by default. For every other cell, the top ten detections were retained after five-second non-maximum suppression. Scoring covered the full 24-hour recording cycle, unlike the night-restricted soundscape analysis in §2.7. Every detection was then manually reviewed, and a sonotype was scored as "present" if at least one was confirmed, and as "absent" if all ten were rejected. Where both devices ran at a site, their records collapsed into one site under a present-wins rule. Therefore, a presence-absence matrix across all sonotypes for each site-year was produced.

2.7 Quantifying sonotype composition and soundscape change

This presence-absence matrix was used to test how individual sonotypes differed between habitats and changed over time. Habitat-associated sonotypes were identified by indicator value analysis (IndVal) with restricted permutations, combining a two-group forest-versus-pasture test with the group-size-corrected φ point-biserial coefficient (signed positive toward forest, negative toward pasture), which weights the two habitats equally and so removes the bias arising from the unbalanced number of forest and pasture sites. All tests were false-discovery-rate corrected (De Cáceres & Legendre, 2009; De Cáceres *et al*., 2010). Per-sonotype temporal trends were estimated with binomial Generalized Linear Mixed Models (GLMM) of presence against centred year with a site-level random intercept, fitted separately for each sonotype on the pasture strata and false-discovery-rate corrected.

The soundscape-level analysis used site-by-year acoustic summaries computed solely from AudioMoth recordings. AudioMoth was the only device deployed in all three years, keeping differences in recorder frequency response and self-noise out of the embedding centroids. Analysis was restricted to the night since acoustic activity in Pará forest is highest between 20:00 and 05:00 (Howells *et al*., 2025). However, the

analysed period was extended by an hour earlier to capture activity onset. From each 1-minute file between 19:00 and 05:00 across all recording days at every site, one five-second window was randomly drawn and L2-normalised. To equalise sampling effort across strata, windows were rarefied to N = 417 per site-year through a night-stratified draw, with a floor of 208 windows. This rarefied N was the minimum number of windows across all sites consistent with the random draws. Each site-year centroid was the mean of its retained unit vectors, re-normalised to unit length. A fixed forest reference was the mean of the forest-site centroids pooled across all three forest sites across all three years (2023–2025), likewise re-normalised, against which every site-year was compared. Each centroid was treated as a learned summary of a site-year's soundscape.

Two distance measures to the fixed forest reference were derived from the summaries. The first was an embedding chord distance, the Euclidean distance between a site-year centroid and the fixed forest reference, both unit-length, separated by an angle θ, so that $d = \sqrt{(2(1 - \cos \theta))}$, equivalently $2 \sin(\theta/2)$. The second was a Sørensen dissimilarity (Legendre & De Cáceres, 2013), computed from each sonotype's binary presence-absence data.

Two mixed models were fitted, each with 2023 as the reference year and a site-level random intercept. The embedding chord distance to the fixed forest reference was modelled against year with a Gamma GLMM (log link), and the sonotype Sørensen distance to the same reference with a beta GLMM (logit link) after a Smithson-Verkuilen transformation to bound the response within the open unit interval (Smithson & Verkuilen, 2006). Concordance was assessed with Spearman's rank correlation between the embedding and sonotype distances-to-forest across the pasture strata, tested against a site-blocked permutation null (999 permutations).

Results

3.1 Candidate species identification

Of the 15 sonotypes (Appendix B, Table B.1 – labelled "A" to "O") compared against recordings of 724 anurans and 124 insect species from online reference databases, five returned species-level preliminary identifications, all frogs (Figure 3). The highest cosine match was not usually the candidate retained.

Leptodactylus fuscus matched its sonotype across 34 of the 50 most similar recordings, whereas *Adenomera andreae*, *Adenomera hylaedactyla*, *Dendropsophus microcephalus* and *Dendropsophus nanus* each accounted for only one to six of the 50 matches (Appendix B, Table B.2). The spectrograms of the ten sonotypes without a preliminary database match were identified as vocalising insects. No taxonomic information could be obtained for these ten sonotypes, since the regional acoustic repository for insect species identification by a non-expert did not exist. In addition, no species-level identification could be confirmed by taxonomists with the recordings alone.

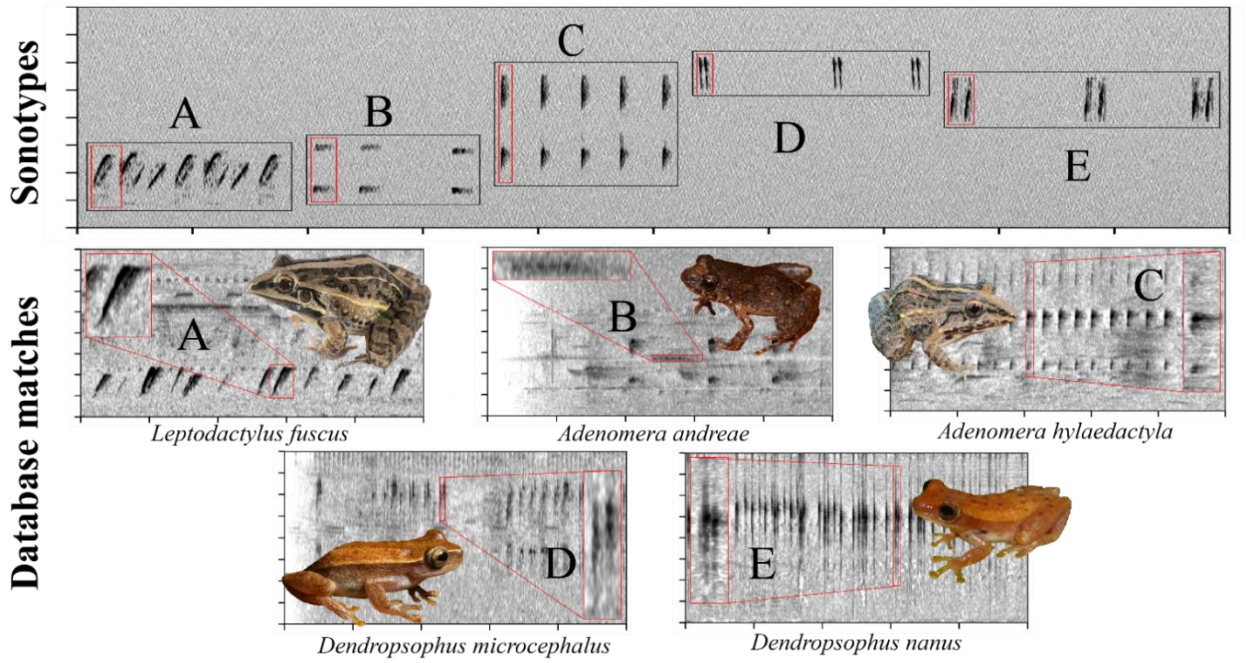


Figure 3. Sonotypes matched to candidate species by cosine similarity, for the five of 15 that returned a frog species-level candidate. The main panel is a ten-second field recording from the study site shown as a spectrogram from 0 to 8 kHz, with the five sonotype calls spliced in, and highlighted with a box with the sonotype name "A" to "E" next to it. Each spectrogram (five-second, 0 to 8 kHz) below the main panel pairs the matched referenced spectrogram with the sonotype's spectrogram, highlighting and enlarging the call, alongside an illustration of the candidate species. Candidates were ranked by cosine similarity in the Perch embedding space against 848 anuran and insect species held by iNaturalist, Xeno-canto and GBIF, keeping the 50 nearest recordings per sonotype. Support was uneven. *L. fuscus* matched 34 of its 50 retrieved

recordings to the validated candidate species, whilst the other four sonotypes matched one to six. Per-sonotype match counts and cosine similarities are given in Appendix B, Table B.2.

3.2 Classifier performance

Manual review confirmed 77 to 1,500 positives per sonotype from up to 1,500 retrieved candidates (median 562) and a median of 51 of the 200 detections reviewed per active-learning cycle (Appendix B, Table B.4). Validation ROC-AUC ranged from 0.99 to 1.00, PR-AUC from 0.96 to 1.00, with the train-validation gap near zero throughout (Figure 4; Appendix B, Table B.5). Permuted-label and random-embedding controls returned 0.49 to 0.59, at or near chance (0.50). The dataset-corrected plain retrieval-AUC was also high, 0.91 to 1.00. The stratified retrieval-AUC was unstable, ranging from 0.03 to 1.00 across the one to three usable score buckets for the 13 classifiers where it could be estimated. The max-F1 threshold ranged from 0.59 to 0.96 and call-density P(+) from 0.048 to 0.156.

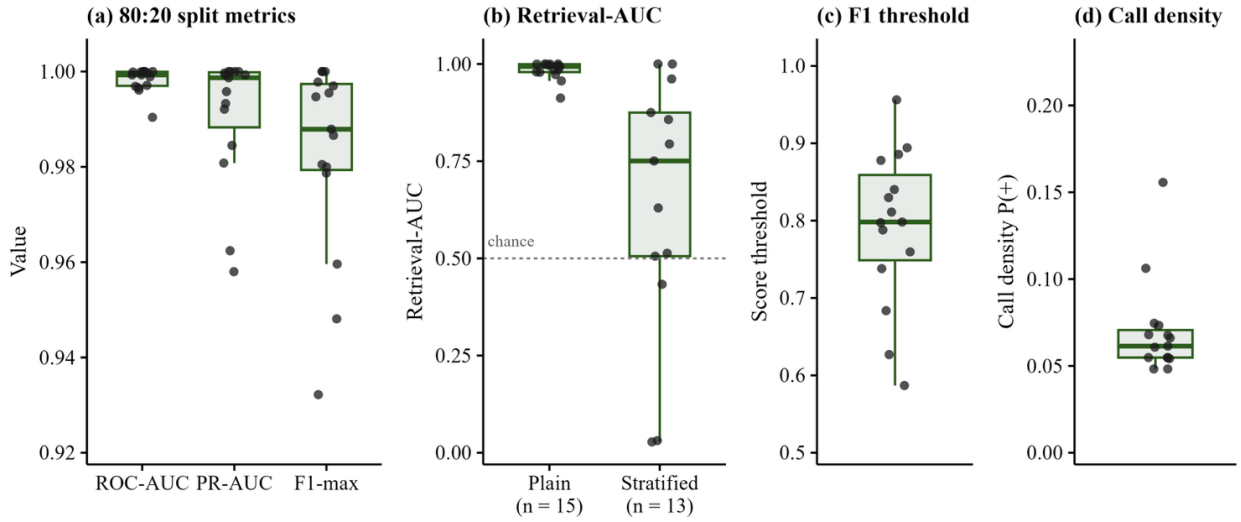


Figure 4. Validation and in-context retrieval performance of the 15 sonotype classifiers, after three active-learning cycles. Per-classifier values are in Appendix B, Table B.5. Points are individual classifiers. Boxes span the interquartile range with the median marked, and whiskers reach 1.5 times that range. Each panel has its own y-axis. (a) ROC-AUC, PR-AUC and F1-max are reported from the 80:20 stratified validation split on a restricted axis (0.92 to 1.00) as all three sit near the ceiling. (b) Dataset-corrected retrieval-AUC, plain and stratified (P(b)-weighted) are reported with the dashed line marking chance (0.5). The stratified estimate is more variable, falling below 0.04 for F and E, and undefined for two rare sonotypes (M and K) where no score bucket held both positive and negative, so n = 13. (c) The score threshold

maximising F1 on the validation split is reported. (d) Call-density P(+), is the estimated proportion of all scored windows containing the sonotype, indexing how common it is rather than how well it is classified, ranging from 0.048 (M and K) to 0.156 (A). Permuted-label and random-embedding controls (not shown) returned AUCs of 0.49 to 0.59, at chance.

3.3 Sonotype-level responses

Sonotype richness averaged 4.9 of the 15 sonotypes per forest site (SD 1.2, n = 3) and 8.9 per pasture site (SD 1.5, n = 26). In pasture, mean sonotype richness was 9.0 in 2023, 8.3 in 2024 and 9.4 in 2025. Habitat association (φ) ranged from +0.83 (B, forest) to −0.89 (O, pasture). Across the 3 forest and 26 pasture sites, φ flagged ten of the 15 sonotypes as habitat-associated after FDR correction, two being forest and eight being pasture (Figure 5). The uncorrected two-group IndVal flagged seven of those same ten, dropping sonotypes J, K and L (IndVal Padj = 0.176, 0.142 and 0.064, respectively), whose IndVal associations fall just short of significance despite passing the φ threshold.

Trends in the proportion of sites present were fitted per sonotype on 74 pasture site-years across 26 sites (binomial GLMM, log-odds per year). One of the 15, sonotype O, was present in 69 of those 74 site-years and sits near separation, so its slope is not interpretable. Three of the other 14 were significant after FDR correction: sonotype B increased at +1.38 (95% CI +0.42 to +2.34, q = 0.037) and sonotype F at +0.93 (95% CI +0.24 to +1.61, q = 0.040), while sonotype J declined at −1.15 (95% CI −1.87 to −0.43, q = 0.026); the remaining eleven sonotypes showed no detectable trend (Figure 5; Appendix B, Table B.6).

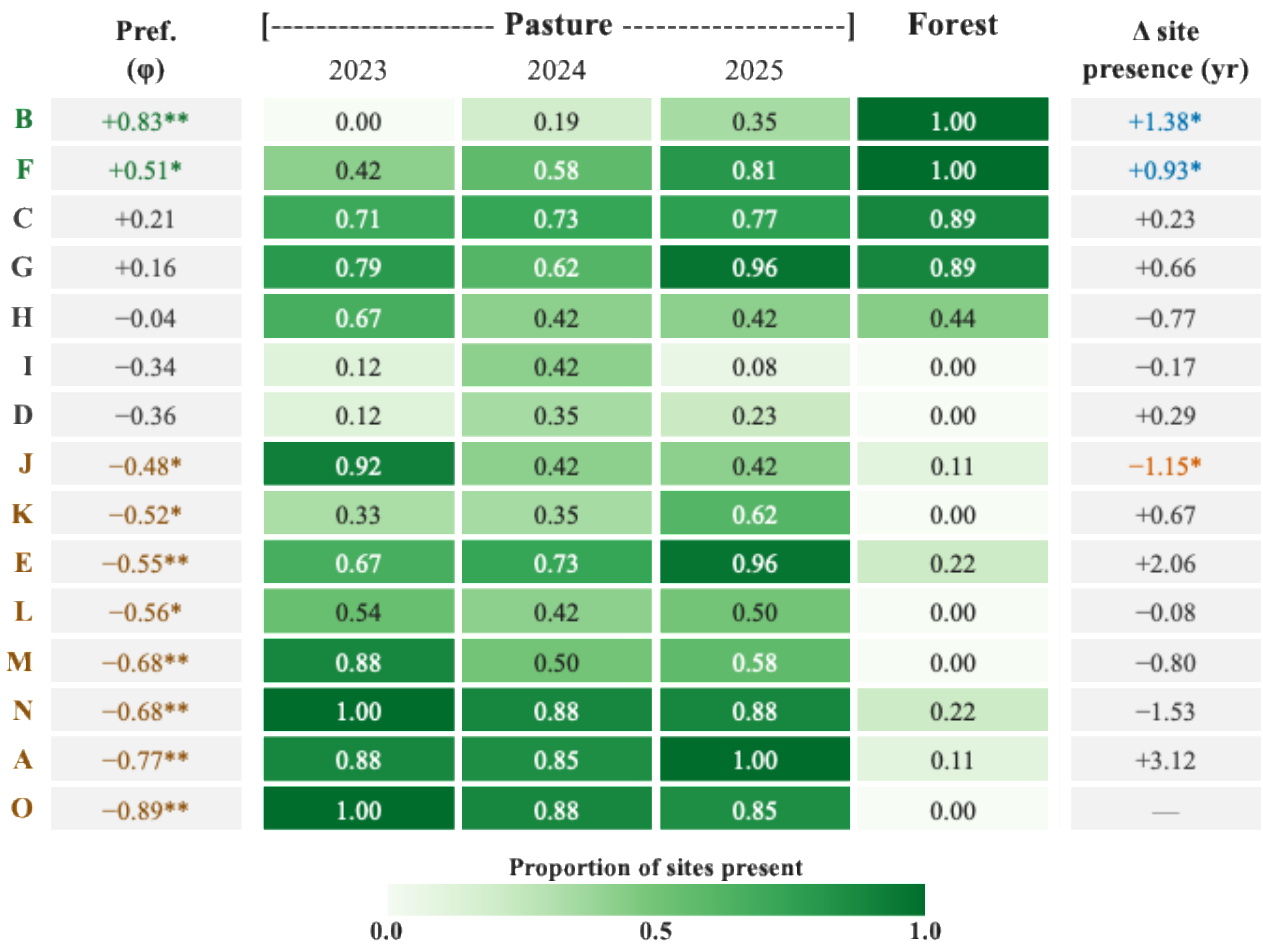


Figure 5. Proportion of sites at which each of the 15 sonotypes was detected, by habitat and year. Sonotypes are ordered by habitat association φ, with the most forest-associated at top and the most pasture associated at the bottom. The four shaded columns give the proportion of sites where each sonotype was detected, split by pasture year and pooled across forest sites. Darker shading indicates a higher proportion. φ is the group-corrected point-biserial coefficient (−1 pasture … +1 forest) and Δ site presence (yr) is the slope in log-odds of presence per year, from a per-sonotype binomial GLMM with a site-level random intercept, fitted on 74 pasture site-years across 26 sites. Stars mark FDR-corrected significance at * $q < 0.05$, ** $q < 0.01$ and *** $q < 0.001$. The sonotype O trend is unstable and not interpretable. Sonotypes were assigned alphabetically and carry no meaning.

3.4 Sonotype composition and soundscape change across time

Chord distance from each pasture site-year to the fixed forest reference, pooled from nine forest site-years across three sites, did not differ detectably from 2023 in 2024 (Gamma GLMM ratio = 0.944, 95% CI 0.866 to 1.028, $p = 0.19$) but diverged significantly by 2025 (0.885, 95% CI 0.811 to 0.965, $p = 0.006$), with the

predicted mean falling from 1.066 (95% CI 0.998 to 1.138) to 0.943 (0.883 to 1.007) (Figure 6). Sørensen distance to the same reference showed no detectable change in 2024 (beta GLMM $\beta = -0.139$, 95% CI −0.365 to 0.087, $p = 0.23$) but also diverged significantly by 2025 ($\beta = -0.419$, 95% CI −0.638 to −0.199, $p = 0.0002$), with the predicted mean falling from 0.637 (95% CI 0.597 to 0.675) to 0.536 (0.496 to 0.575). Across the 72 pasture site-years where both distances were defined, embedding and composition distances were moderately positively correlated (Spearman $\rho = 0.350$; site-blocked permutation $p = 0.033$). Full model output is given in Appendix B, Table B.7.

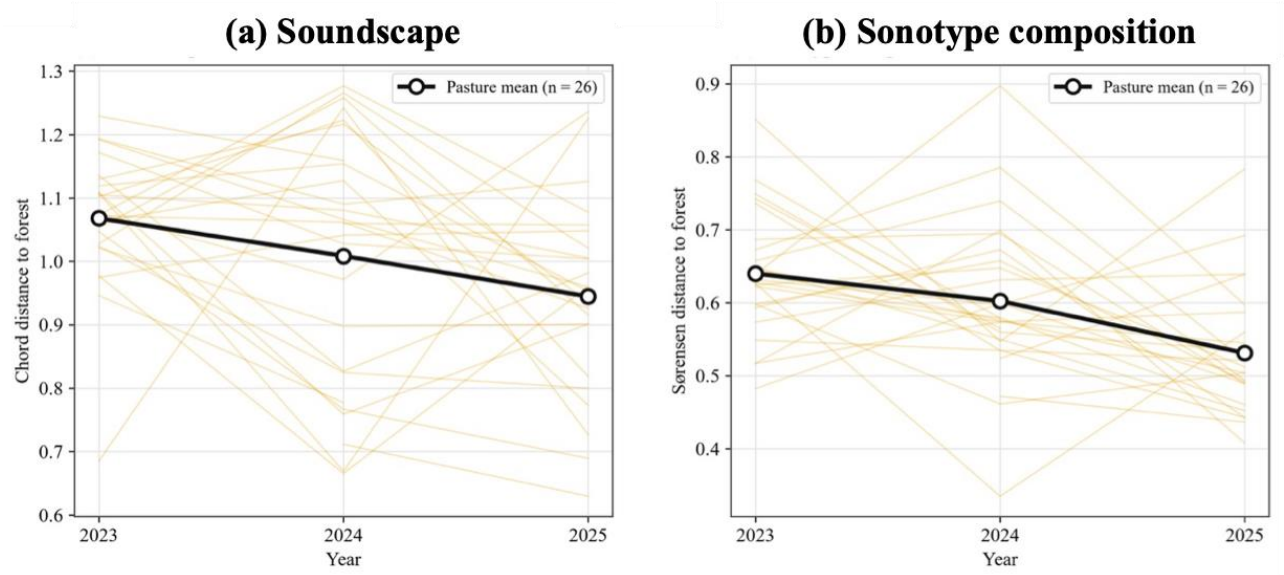


Figure 6. Distance from each pasture site-year to the fixed forest reference, under both distance measures. The reference is the centroid of the three forest sites, pooled across sites and years. (a) Soundscape shows the chord distance between each pasture site-year embedding centroid and the forest reference centroid. (b) Sonotype composition distance is the Sørensen dissimilarity to the forest reference, computed from the 15-sonotype presence-absence matrix. Faint orange lines are individual sites and the bold black line is the pasture mean, across 74 pasture site-years at 26 sites. Lower values are more forest-like. Year effects were estimated with a Gamma GLMM in (a) and a beta GLMM in (b), each taking 2023 as the reference year and carrying a site-level random intercept. Neither distance differed significantly from 2023 in 2024, but both had significantly diverged from baseline by 2025 whilst becoming more similar to the forest reference (Appendix B, Table B.7).

Discussion

We show that an audio embeddings pipeline can track frog and insect community change at a reforestation site without a local labelled sound repository. Using Perch embeddings and the classifiers for the 15 sonotypes identified, we built two measures of distance to forest of the restored pasture, one from sonotype composition and one from the whole nocturnal soundscape. Both distances fell by 2025. Two forest-associated sonotypes were detected at a rising share of sites over time while one pasture-associated sonotype declined. Using reference databases, five sonotypes were assigned candidate frog identification, while the remaining ten were assessed to be insects. None of this process required any fine-tuning of the foundation model or confirmed species identification, so the pipeline can run without machine learning experts or taxonomists.

By 2025, both distances to forest had fallen significantly below their 2023 baseline, though neither had shifted by 2024 (Figure 6). Severe disturbances from active planting and associated machinery in the months leading into 2024 may account for this lag. Two years is a short assessment window when Neotropical compositional recovery across taxa has been claimed to take decades or centuries (Rozendaal *et al*., 2019; Elsy *et al*., 2024; Metz *et al*., 2026). However, those estimates rest on chronosequences that compare plots of different ages sampled once, which can miss the directional recovery that repeated surveys of the same plots detect (Kreyling, 2025). Our results suggest that tracking the same sites against a fixed reference may reveal short-term changes by disentangling restoration time from the landscape-level context that chronosequences confound with it. The two distance to forest measures only correlated moderately, which may be because the 15 sonotypes are a subset of the whole nocturnal community that the embedding index represents. Sethi *et al*. (2023) encountered the same disconnect between avian community change and learned features, attributing it to insect and amphibian calls and environmental noise possibly contributing more to the soundscape than birds. Our result is consistent with that finding, as sampling only part of the non-avian community is still not enough to close the gap. The other calls and noise still missing from the sonotype composition may influence the embedding index enough to create the disconnect we observed. Additionally, acoustic measures have been shown to explain the composition of a largely silent nocturnal insect assemblage (Müller *et al*., 2023), so an embedding index can encompass changes beyond the vocalising community. Nonetheless, over a longer recording period, a continued trajectory of pasture soundscapes towards those of forest is expected.

Neotropical frogs often show strong habitat preferences due to physiological (Titon Jr. & Gomes, 2015) and diet specialisation (Born *et al*., 2010). Therefore, they can be sensitive to changes in their habitat (Faggioni *et al*., 2021) which may come from reforestation. The five sonotypes identified as frogs were largely a pasture-associated assemblage (Figure 5) which reflects the opportunistic sonotype subset chosen from a highly pasture-skewed acoustic recording design. The habitat associations found were ecologically consistent. For example, *L. fuscus* is found almost exclusively in pasture (Bernarde & Macedo, 2008) matching its pasture-association from our results and *A. hylaedactyla* was widespread across all sites, consistent with a clade of open formations (de Carvalho *et al*., 2019) whose species group reaches its highest abundances in heavily deforested plots (Torralvo *et al*., 2022). *D. microcephalus* was absent from all forest sites aligning with its known range being from open savanna to pasture (Barrio-Amorós, 1998) and *D. nanus* was strongly pasture-associated also consistent with preferring open habitat (Santos *et al*., 2007). The only frog which showed any significant change in presence across sites with reforestation was *A. andreae*, known as a forest specialist frog (de Carvalho *et al*., 2019), which became more prevalent. The presence of pasture or widespread species did not change over time. This may be because the environment is still open, since Amazonian pasture regrowth keeps canopy open until six to eight years post-abandonment (Feldpausch *et al*., 2005). These results suggest that recolonisation of forest specialists may be happening faster than extirpation of pasture species. Frog composition on a pasture legacy is projected to take 37 years to return (Metz *et al*., 2026), though a Costa Rican leaf-litter frog reached mature-forest abundance in under 17 years (Thompson & Donnelly, 2023). Two years sits far below both timelines, so one forest frog's increase is still a meaningful signal of recovery from possibly a small sample of the site's overall frog community.

Insects can predict habitat change as their calls are affected by microhabitat temperature, moisture and vegetation structure (Riede & Balakrishnan, 2025). However, with reforestation, insect communities have been shown to change very slowly, over one to two decades in Amazonian forest emptied by logging (Rappaport *et al*., 2022) and over eight to 15 years in replanted or restored ground elsewhere (Hugel, 2012; Cole *et al*., 2016). Our results after two years reflect this lag, with eight of the ten insect sonotypes remaining relatively stable. Without species-level identification, inferring why one forest-associated insect becomes more present across sites and one pasture-associated declines remains speculative. Nonetheless, this insect sonotype community may have only shifted slightly as the canopy is still undeveloped, as it was shown that

crickets across Amazonian sites tracked canopy cover rather than habitat (Do Nascimento *et al*., 2024). Our small subset of the insect community is still a valuable measure of composition, since according to Anso *et al*. (2022), even when the cricket community was narrowed to only the species that call, recovery stage still explained more than half of compositional variation. Thus, it can be extrapolated that the vocalising pasture insect assemblage is still dominant after recent reforestation.

The manual effort in this pipeline sits in sonotype validation. Embedding, retrieval and scoring are automated, and the manual review is the only human-in-the-loop stage. Dumoulin *et al*. (2025) developed their classifiers in less than one hour each while in our pipeline, roughly six hours of manual review carried one sonotype from a single exemplar clip to a finished classifier. This is due to our choice of reviewing a greater number of clips for robustness. Comparable classifiers have been built elsewhere from far fewer labels per class over BirdNET embeddings (Jacuzzi & Olden, 2025; Weldy *et al*., 2025), so the several hundred confirmed positives behind an average classifier here is an upper bound rather than a requirement. Therefore, what limits the sonotype library's size is listener time rather than the volume of audio. Unsupervised clustering would also be able to automate finding sonotypes, needing no exemplar or labelled training set (Guerrero *et al*., 2023; Turlington *et al*., 2025). Under tighter time constraints, vector search alone may serve for presence-absence work (Hamer *et al*., 2023; Allen-Ankins *et al*., 2025), which has not been tested against the classifier matrix but performed well at finding the call of interest (Appendix B, Table B.4). Sonotype libraries and their classifiers built by this process can potentially keep pace with fast changing soundscapes.

Several considerations bear on how the classifier metrics should be read (Figure 4; Appendix B, Table B.5). They are best-case estimates rather than field performance, because label leakage may have been built into the evaluation, as each classifier was scored on windows from the recordings and sites it was trained on. Perch is trained in part to predict which recording a window came from, so similarity can arise from a shared session as readily as a shared species (van Merriënboer *et al*., 2025; Peralta *et al*., 2025). Time of day is recoverable from tropical soundscape audio alone (Loo *et al*., 2025), so a classifier fitted to a narrow calling window may be separating hours as much as calls. However, neither claim can be confirmed. A site-year split from the outset would show how much performance survives them (Rasmussen *et al*., 2025). Non-detection is also read as absence throughout, with no occupancy model to separate the two (MacKenzie *et*

al., 2002). Having the second highest published call descriptions among Brazilian anurans (Guerra *et al*., 2018), *Leptodactylus fuscus* drew 34 of its 50 nearest recordings against one to six for the rest of anurans. That count measures archive coverage rather than identity, since *L. fuscus* may denote a species complex spanning several genetic lineages (Wynn & Heyer, 2001), which also inflates the match. No insect sonotype drew a matching candidate species, which further highlights how limited acoustic repositories are for insects. Future work with agile modelling should reduce the embedding dimensionality to focus on the call of interest. This would address the label leakage at its source and improve retrieval of that call from public archives and on-site recordings alike.

This study converted a taxonomically unresolved nocturnal assemblage into 15 validated sonotype classifiers, and the directional change they recorded was reproduced by a soundscape-level measure of the same recordings. Most species of the sound-producing community at a tropical restoration site have no names, reference recordings, or classifiers. Building these classifiers took one exemplar clip per sonotype and no model fine-tuning, removing the technical expertise requirement that has kept automated acoustic monitoring out of the tropical systems that need it most (Zwerts *et al*., 2021). The output was a candidate species list and two dissimilarity measures stemming from the same embeddings, computed from audio that a restoration project often already holds and can read against its own baseline. Standardising this protocol may enable cross-site and cross-region comparison, provided the foundation model stays the same. Even without ground-truthing each sonotype, a restoration project can adapt this scalable pipeline to quantify soundscape changes of its own uncatalogued vocalising communities.

Acknowledgements

We thank Jasmine Nelson for her initial library of sonotypes and for helping to identify frog and insect sonotypes, and all members of the CaLE lab for their valuable feedback and suggestions on the project.

Code and Data Availability

All code and metadata used in the study can be found via the following link: https://github.com/jareddesilva1/perch_embeddings_acoustic_pipeline. Any additional data or information used for this paper may be available by request from the CaLE lab.

Generative AI Use Statement

Claude Opus 5 was used to assist formatting, spelling and clarity, and for programming support and debugging R code. All original content is the authors' own; all text was reviewed and edited by the authors, and no content generated by AI has been presented as the authors' own work.

References

Alain, G. & Bengio, Y. (2016) Understanding intermediate layers using linear classifier probes. *arXiv*. [Preprint] https://doi.org/10.48550/arXiv.1610.01644

Alcocer, I., Lima, H., Sugai, L.S.M. & Llusia, D. (2022) Acoustic indices as proxies for biodiversity: a meta-analysis. *Biological Reviews*. 97 (6), 2209-2236. https://doi.org/10.1111/brv.12890

Allen-Ankins, S., Hoefer, S., Bartholomew, J., Brodie, S. & Schwarzkopf, L. (2025) The use of BirdNET embeddings as a fast solution to find novel sound classes in audio recordings. *Frontiers in Ecology and Evolution*. 12, 1409407. https://doi.org/10.3389/fevo.2024.1409407

Anso, J., Gasc, A., Bourguet, E., Desutter-Grandcolas, L. & Jourdan, H. (2022) Crickets as indicators of ecological succession in tropical systems, New Caledonia. *Biotropica*. 54 (5), 1270-1284. https://doi.org/10.1111/btp.13151

Barrio-Amorós, C.L. (1998) Sistemática y biogeografía de los anfibios (Amphibia) de Venezuela. *Acta Biológica Venezuelica*. 18 (2), 1–93.

Bernarde, P.S. & Macedo, L.C. (2008) Impacto do desmatamento e formação de pastagens sobre a anurofauna de serapilheira em Rondônia. *Iheringia, Série Zoologia*. 98 (4), 454-459. https://doi.org/10.1590/S0073-47212008000400006

Born, M., Bongers, F., Poelman, E.H. & Sterck, F.J. (2010) Dry-season retreat and dietary shift of the dart-poison frog *Dendrobates tinctorius* (Anura: Dendrobatidae). *Phyllomedusa*. 9 (1), 37–52. https://doi.org/10.11606/issn.2316-9079.v9i1p37-52

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M. & Bolker, B.M. (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modelling. *The R Journal*. 9 (2), 378-400. https://doi.org/10.32614/RJ-2017-066

Burns, A., Harrell, L., van Merriënboer, B., Dumoulin, V., Hamer, J. & Denton, T. (2025) Perch 2.0 transfers "whale" to underwater tasks. *arXiv*. [Preprint] https://doi.org/10.48550/arXiv.2512.03219

de Carvalho, T.R., Giaretta, A.A., Maciel, N.M., Barrera, D.A., Aguilar-Puntriano, C., Haddad, C.F.B., Kokubum, M.N.C., Menin, M. & Angulo, A. (2019) On the uncertain taxonomic identity of *Adenomera hylaedactyla* (Cope, 1868) and the composite type series of *A. andreae* (Müller, 1923) (Anura, Leptodactylidae). *Copeia*. 107 (4), 708-723. https://doi.org/10.1643/CH-19-237

Climate-Data.org. (n.d.) *Clima Pará (Brasil): temperatura, tempo e dados climatológicos*. https://en.climate-data.org/south-america/brazil/para-189/ [Accessed 5th July 2026]

Cole, R.J., Holl, K.D., Zahawi, R.A., Wickey, P. & Townsend, A.R. (2016) Leaf litter arthropod responses to tropical forest restoration. *Ecology and Evolution*. 6 (15), 5158-5168. https://doi.org/10.1002/ece3.2220

Cromberg, M., Naime, J., Chervier, C., Angelsen, A. & Martius, C. (2026) *Forest policy diagnostics for Brazil*. Working Paper 71. Bogor, Indonesia, CIFOR; Nairobi, Kenya, ICRAF. https://doi.org/10.17528/cifor-icraf/009462

da Cruz, D.C., Rey Benayas, J.M., Costa Ferreira, G. & Santos Ribeiro, S. (2020) Tree communities in three-year-old post-mining sites under different forest restoration techniques in the Brazilian Amazon. *Forests*. 11 (5), 527. https://doi.org/10.3390/f11050527

De Cáceres, M. & Legendre, P. (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology*. 90 (12), 3566-3574. https://doi.org/10.1890/08-1823.1

De Cáceres, M., Legendre, P. & Moretti, M. (2010) Improving indicator species analysis by combining groups of sites. *Oikos*. 119 (10), 1674-1684. https://doi.org/10.1111/j.1600-0706.2010.18334.x

Do Nascimento, L.A., Pérez-Granados, C., Alencar, J.B.R. & Beard, K.H. (2024) Time and habitat structure shape insect acoustic activity in the Amazon. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 379 (1904), 20230112. https://doi.org/10.1098/rstb.2023.0112

Dumoulin, V., Stretcu, O., Hamer, J., Harrell, L., Laber, R., Larochelle, H., van Merriënboer, B., Navine, A., Hart, P., Williams, B. & Lamont, T.A. (2025) The search for squawk: agile modelling in bioacoustics. *arXiv*. [Preprint] https://doi.org/10.48550/arXiv.2505.03071

Elsy, A.D., Pfeifer, M., Jones, I.L., DeWalt, S.J., Lopez, O.R. & Dent, D.H. (2024) Incomplete recovery of tree community composition and rare species after 120 years of tropical forest succession in Panama. *Biotropica*. 56 (1), 36-49. https://doi.org/10.1111/btp.13275

Esri. (2026) *World Imagery (Wayback), ArcGIS Living Atlas of the World.* [Map service] Redlands, CA, Esri. https://livingatlas.arcgis.com/wayback [Accessed 19th August 2026]

Faggioni, G.P., Souza, F.L., Paranhos Filho, A.C., Gamarra, R.M. & Prado, C.P.A. (2021) Amount and spatial distribution of habitats influence occupancy and dispersal of frogs at multiple scales in an agricultural landscape. *Austral Ecology*. 46 (1), 126-138. https://doi.org/10.1111/aec.12966

Feldpausch, T.R., Riha, S.J., Fernandes, E.C.M. & Wandelli, E.V. (2005) Development of forest structure and leaf area in secondary forests regenerating on abandoned pastures in central Amazônia. *Earth Interactions*. 9 (6), 1-22. https://doi.org/10.1175/EI140.1

Ferreira, L.M., Oliveira, E.G., Lopes, L.C., Brito, M.R., Baumgarten, J., Rodrigues, F.H. & Sousa-Lima, R.S. (2018) What do insects, anurans, birds, and mammals have to say about soundscape indices in a tropical savanna. *Journal of Ecoacoustics*. https://doi.org/10.22261/JEA.PVH6YZ

GBIF (Global Biodiversity Information Facility). (2026) *GBIF occurrence download*. https://www.gbif.org [Accessed 7th July 2026]

Ghani, B., Denton, T., Kahl, S. & Klinck, H. (2023) Global birdsong embeddings enable superior transfer learning for bioacoustic classification. *Scientific Reports*. 13 (1), 22876. https://doi.org/10.1038/s41598-023-49989-z

Gibb, R., Browning, E., Glover-Kapfer, P. & Jones, K.E. (2019) Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*. 10 (2), 169-185. https://doi.org/10.1111/2041-210X.13101

Grau, J., Grosse, I. & Keilwagen, J. (2015) PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. *Bioinformatics*. 31 (15), 2595-2597. https://doi.org/10.1093/bioinformatics/btv153

Guayasamin, J.M., Ribas, C.C., Carnaval, A.C., Carrillo, J.D., Hoorn, C., Lohmann, L.G., Riff, D., Ulloa Ulloa, C. & Albert, J.S. (2024) Evolution of Amazonian biodiversity: a review. *Acta Amazonica*. 54, e54bc21360. https://doi.org/10.1590/1809-4392202103601

Guerra, V., Llusia, D., Gambale, P.G., Morais, A.R., Márquez, R. & Bastos, R.P. (2018) The advertisement calls of Brazilian anurans: historical review, current knowledge and future directions. *PLoS ONE*. 13 (1), e0191691. https://doi.org/10.1371/journal.pone.0191691

Guerrero, M.J., Bedoya, C.L., López, J.D., Daza, J.M. & Isaza, C. (2023) Acoustic animal identification using unsupervised learning. *Methods in Ecology and Evolution*. 14, 1500-1514. https://doi.org/10.1111/2041-210X.14103

Hamer, J., Triantafillou, E., van Merriënboer, B., Kahl, S., Klinck, H., Denton, T. & Dumoulin, V. (2023) BIRB: a generalization benchmark for information retrieval in bioacoustics. *arXiv*. [Preprint] https://doi.org/10.48550/arXiv.2312.07439

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J. & Inger, R. (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*. 6, e4794. https://doi.org/10.7717/peerj.4794

Hill, A.P., Prince, P., Piña Covarrubias, E., Doncaster, C.P., Snaddon, J.L. & Rogers, A. (2018) AudioMoth: evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*. 9 (5), 1199-1211. https://doi.org/10.1111/2041-210X.12955

Howells, N., Lees, A.C., Barlow, J., Berenguer, E., Rossi, L.C., Sueur, J., Sullivan, M.J., Gabriel Ramos, Y. & Metcalf, O.C. (2025) Temporal patterns of Amazonian insect acoustic activity. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 380 (1928), 20240337. https://doi.org/10.1098/rstb.2024.0337

Hugel, S. (2012) Impact of native forest restoration on endemic crickets and katydids density in Rodrigues island. *Journal of Insect Conservation*. 16 (3), 473-477. https://doi.org/10.1007/s10841-012-9476-1

Hutschenreiter, A., Aureli, F., Torres-Araneda, A. & Andresen, E. (2026) Performance of four ecoacoustic indices as proxies of bird species richness in a Neotropical forest. *Biodiversity and Conservation*. 35, 13. https://doi.org/10.1007/s10531-025-03208-5

iNaturalist. (2026) *iNaturalist research-grade observations*. https://www.inaturalist.org [Accessed 7th July 2026]

Isometric. (2026) *Project prj_1KA95092G1S093EM, Isometric Registry.* https://registry.isometric.com/project/prj_1KA95092G1S093EM [Accessed 19th August 2026]

Jacuzzi, G. & Olden, J.D. (2025) Few-shot transfer learning enables robust acoustic monitoring of wildlife communities at the landscape scale. *Ecological Informatics.* 90, 103294. https://doi.org/10.1016/j.ecoinf.2025.103294

Johnson, J.M. & Khoshgoftaar, T.M. (2019) Survey on deep learning with class imbalance. *Journal of Big Data*. 6 (1), 27. https://doi.org/10.1186/s40537-019-0192-5

K. Lisa Yang Center for Conservation Bioacoustics. (2019) *Raven Pro: interactive sound analysis software* (Version 1.6) [Software] Ithaca, NY, The Cornell Lab of Ornithology. https://www.ravensoundsoftware.com/software/raven-pro/ [Accessed 7th July 2026]

Kath, H., Serafini, P.P., Campos, I.B., Gouvêa, T.S. & Sonntag, D. (2024) Leveraging transfer learning and active learning for data annotation in passive acoustic monitoring of wildlife. *Ecological Informatics*. 82, 102710. https://doi.org/10.1016/j.ecoinf.2024.102710

King, R.A. & Halpern, B.S. (2025) Implementation of automated biodiversity monitoring lags behind its potential. *Environmental Research Letters*. 20 (6), 064022. https://doi.org/10.1088/1748-9326/add02d

Kingma, D.P. & Ba, J. (2015) Adam: a method for stochastic optimization. In: *3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, 7-9 May. https://doi.org/10.48550/arXiv.1412.6980

Kreyling, J. (2025) Space-for-time substitution misleads projections of plant community and stand-structure development after disturbance in a slow-growing environment. *Journal of Ecology*. 113 (1), 68-80. https://doi.org/10.1111/1365-2745.14438

Kümmet, S., Müller, J., Buřivalová, Z., Schaefer, H.M., Gelis, R., Freile, J., Busse, A., Püls, M., Kriegel, P., Seibold, S., Blüthgen, N., de la Hoz, M., Schleuning, M., Neuschulz, E.L., Mitesser, O. & Kortmann, M. (2025) Acoustic indices predict recovery of tropical bird communities for taxonomic and functional composition. *Conservation Letters*. 18 (4), e13131. https://doi.org/10.1111/conl.13131

Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*. 16 (8), 951-963. https://doi.org/10.1111/ele.12141

Lennox, G.D., Gardner, T.A., Thomson, J.R., Ferreira, J., Berenguer, E., Lees, A.C., Mac Nally, R., Aragão, L.E.O.C., Ferraz, S.F.B., Louzada, J., Moura, N.G., Oliveira, V.H.F., Pardini, R., Solar, R.R.C., Vaz-de Mello, F.Z., Vieira, I.C.G. & Barlow, J. (2018) Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. *Global Change Biology*. 24 (12), 5680-5694. https://doi.org/10.1111/gcb.14443

Ligges, U., Krey, S., Mersmann, O. & Schnackenberg, S. (2023) tuneR: analysis of music and speech. [R package] https://doi.org/10.32614/CRAN.package.tuneR

Loo, Y.Y., Lee, M.Y., Shaheed, S., Maul, T. & Clink, D.J. (2025) Temporal patterns in Malaysian rainforest soundscapes demonstrated using acoustic indices and deep embeddings trained on time-of-day estimation. *The Journal of the Acoustical Society of America*. 157 (1), 1-16. https://doi.org/10.1121/10.0034638

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*. 83 (8), 2248-2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

McCarthy, R.A., Zhang, Y., Verburg, S.A., Jenkins, W.F. & Gerstoft, P. (2025) Machine learning in acoustics: a review and open-source repository. *npj Acoustics*. 1 (1), 18. https://doi.org/10.1038/s44384-025-00021-w

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

Metz, T., Farwig, N., Dormann, C.F., Schaefer, H.M., Guevara Andino, J.E., Brehm, G., Burneo, S., Chao, A., Chazdon, R.L., Colwell, R.K., Diniz, U.M., Donoso, D.A., Endara, M.-J., Erazo, S., Escobar, S., Falconí-López, A., Feldhaar, H., Garcia Villamarin, M., Grella, N., Heer, K., Heethoff, M., Keller, A., Landim, A.R., Leonhardt, S.D., Tamargo Lopez, E., Marín-Armijos, D., Müller, J., Neira-Salamea, K., Neuschulz, E.L., Pedersen, K.M., Rödel, M.-O., Schleuning, M., Schmitt, T., Staab, M., Tartara, A., Tinoco, B.A., Tremlett, C.J., Tschapka, M., Unsicker, S., Villa-Galaviz, E. & Blüthgen, N. (2026) Biodiversity resilience in a tropical rainforest. *Nature*. https://doi.org/10.1038/s41586-026-10365-2

Müller, J., Mitesser, O., Schaefer, H.M., Seibold, S., Busse, A., Kriegel, P., Rabl, D., Gelis, R., Arteaga, A., Freile, J., Leite, G.A., de Melo, T.N., LeBien, J., Campos-Cerqueira, M., Blüthgen, N., Tremlett, C.J., Böttger, D., Feldhaar, H., Grella, N., Falconí-López, A., Donoso, D.A., Moriniere, J. & Buřivalová, Z. (2023) Soundscapes and deep learning enable tracking biodiversity recovery in tropical forests. *Nature Communications*. 14 (1), 6191. https://doi.org/10.1038/s41467-023-41693-w

Napier, T., Ahn, E., Allen-Ankins, S., Schwarzkopf, L. & Lee, I. (2024) Advancements in preprocessing, detection and classification techniques for ecoacoustic data: a comprehensive review for large-scale passive acoustic monitoring. *Expert Systems with Applications*. 252, 124220. https://doi.org/10.1016/j.eswa.2024.124220

Navine, A.K., Denton, T., Weldy, M.J. & Hart, P.J. (2024) All thresholds barred: direct estimation of call density in bioacoustic data. *Frontiers in Bird Science*. 3, 1380636. https://doi.org/10.3389/fbirs.2024.1380636

Open Acoustic Devices. (n.d.) *AudioMoth*. https://www.openacousticdevices.info/audiomoth [Accessed 31st May 2024]

Pan, W., Goodale, E., Jiang, A. & Mammides, C. (2024) The effect of latitude on the efficacy of acoustic indices to predict biodiversity: a meta-analysis. *Ecological Indicators*. 159, 111747. https://doi.org/10.1016/j.ecolind.2024.111747

Peralta, A.D., dos Santos, E.M., Gordo, M., Xie, J. & Colonna, J.G. (2025) Improving soundscape retrieval for bioacoustic monitoring: an analysis of fusion techniques with pre-trained embeddings. In: *Proceedings of the 31st Brazilian Symposium on Multimedia and the Web (WebMedia 2025*). Sociedade Brasileira de Computação, pp. 312-320. https://doi.org/10.5753/webmedia.2025.15196

Pickering, A., Martinez Balvanera, S., Jones, K.E. & Hedwig, D. (2025) A scalable transfer learning workflow for extracting biological and behavioural insights from forest elephant vocalizations. *Remote Sensing in Ecology and Conservation*. https://doi.org/10.1002/rse2.70008

Prechelt, L. (1998) Early stopping — but when? In: Orr, G.B. & Müller, K.-R. (eds.) *Neural networks: tricks of the trade*. Lecture Notes in Computer Science, vol. 1524. Berlin, Heidelberg, Springer, pp. 55-69. https://doi.org/10.1007/3-540-49430-8_3

R Core Team. (2026) R: a language and environment for statistical computing. Version 4.5.3 [Software] Vienna, Austria, R Foundation for Statistical Computing. https://www.R-project.org/

Ramesh, V., Hariharan, P., Akshay, V.A., Choksi, P., Khanwilkar, S., DeFries, R. & Robin, V.V. (2023) Using passive acoustic monitoring to examine the impacts of ecological restoration on faunal biodiversity in the Western Ghats. *Biological Conservation.* 282, 110071. https://doi.org/10.1016/j.biocon.2023.110071

Rappaport, D.I., Swain, A., Fagan, W.F., Dubayah, R. & Morton, D.C. (2022) Animal soundscapes reveal key markers of Amazon forest degradation from fire and logging. *Proceedings of the National Academy of Sciences*. 119 (18), e2102878119. https://doi.org/10.1073/pnas.2102878119

Rasmussen, N.R., Rizk, R., Wang, L. & Santosh, K.C. (2025) Ecologically valid benchmarking and adaptive attention: scalable marine bioacoustic monitoring. *arXiv*. [Preprint] https://doi.org/10.48550/arXiv.2509.04682

Rauch, L., Huseljic, D., Wirth, M., Decke, J., Sick, B. & Scholz, C. (2024) Towards deep active learning in avian bioacoustics. In: *Proceedings of the 8th International Workshop and Tutorial on Interactive Adaptive Learning (IAL@ECML-PKDD 2024), Vilnius, Lithuania, 9th September.* CEUR Workshop Proceedings, vol. 3770. Aachen, CEUR-WS.org, pp. 12-17. https://doi.org/10.48550/arXiv.2406.18621

Riede, K. (2018) Acoustic profiling of Orthoptera: present state and future needs. *Journal of Orthoptera Research*. 27 (2), 203-215. https://doi.org/10.3897/jor.27.23700

Riede, K. & Balakrishnan, R. (2025) Acoustic monitoring for tropical insect conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 380 (1928), 20240046. https://doi.org/10.1098/rstb.2024.0046

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C. & Müller, M. (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 12, 77. https://doi.org/10.1186/1471-2105-12-77

Rozendaal, D.M.A., Bongers, F., Aide, T.M. *et al*. (2019) Biodiversity recovery of Neotropical secondary forests. *Science Advances*. 5 (3), eaau3114. https://doi.org/10.1126/sciadv.aau3114

Santos, T.G., Rossa-Feres, D.C. & Casatti, L. (2007) Diversidade e distribuição espaço-temporal de anuros em região com pronunciada estação seca no sudeste do Brasil. *Iheringia, Série Zoologia*. 97 (1), 37-49. https://doi.org/10.1590/S0073-47212007000100007

Schwinger, R., Vali Zadeh, P., Rauch, L., Kurz, M., Hauschild, T., Lapp, S. & Tomforde, S. (2026) Foundation models for bioacoustics — a comparative review. *Ecological Informatics*. 96, 103765. https://doi.org/10.1016/j.ecoinf.2026.103765

Sethi, S.S., Jones, N.S., Fulcher, B.D., Picinali, L., Clink, D.J., Klinck, H., Orme, C.D.L., Wrege, P.H. & Ewers, R.M. (2020) Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. *Proceedings of the National Academy of Sciences*. 117 (29), 17049-17055. https://doi.org/10.1073/pnas.2004702117

Sethi, S.S., Bick, A., Ewers, R.M., Klinck, H., Ramesh, V., Tuanmu, M.-N. & Coomes, D.A. (2023) Limits to the accurate and generalizable use of soundscapes to monitor biodiversity. *Nature Ecology & Evolution*. 7 (9), 1373-1378. https://doi.org/10.1038/s41559-023-02148-z

Simpson, G.L. (2022) permute: functions for generating restricted permutations of data. [R package] https://doi.org/10.32614/CRAN.package.permute

Smithson, M. & Verkuilen, J. (2006) A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*. 11 (1), 54-71. https://doi.org/10.1037/1082-989X.11.1.54

Somervuo, P. *et al*. (2025) Human contributions to global soundscapes are less predictable than the acoustic rhythms of wildlife. *Nature Ecology & Evolution*. 9 (9), 1585-1598. https://doi.org/10.1038/s41559-025-02786-5

Steck, H., Ekanadham, C. & Kallus, N. (2024) Is cosine-similarity of embeddings really about similarity? In: *Companion Proceedings of the ACM Web Conference 2024*. New York, Association for Computing Machinery, pp. 887-890. https://doi.org/10.1145/3589335.3651526

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

690 Sueur, J., Aubin, T. & Simonis, C. (2008) Seewave: a free modular tool for sound analysis and synthesis.
691 *Bioacoustics*. 18 (2), 213-226. https://doi.org/10.1080/09524622.2008.9753600

692 Thompson, M.E. & Donnelly, M.A. (2023) Leaf-litter frog abundance increases during succession of
693 regenerating pastures. *Journal of Tropical Ecology*. 39, e41.
694 https://doi.org/10.1017/S0266467423000305

695 Titon Jr., B. & Gomes, F.R. (2015) Relation between water balance and climatic variables associated with
696 the geographical distribution of anurans. *PLoS ONE*. 10 (10), e0140761.
697 https://doi.org/10.1371/journal.pone.0140761

698 Torralvo, K., de Fraga, R., Lima, A.P., Dayrell, J. & Magnusson, W.E. (2022) Environmental filtering and
699 deforestation shape frog assemblages in Amazonia: an empirical approach assessing species abundances
700 and functional traits. *Biotropica*. 54 (1), 226-238. https://doi.org/10.1111/btp.13053

701 Turlington, K., Suárez-Castro, A.F., Teixeira, D., Linke, S. & Sheldon, F. (2025) A novel protocol for
702 exploratory analysis of unknown sound-types in large acoustic datasets. *Methods in Ecology and*
703 *Evolution*. 16 (10), 2487-2499. https://doi.org/10.1111/2041-210X.70134

704 Ushey, K., Allaire, J.J. & Tang, Y. (2024) reticulate: interface to Python. [R package]
705 https://doi.org/10.32614/CRAN.package.reticulate

706 Weldy, M.J., Lesmeister, D.B., Denton, T., Duarte, A., Vernasco, B.J., Gasc, A., Rowe, J.C., Adams, M.J. &
707 Betts, M.G. (2025) Simulated soundscapes and transfer learning boost the performance of acoustic
708 classifiers under data scarcity. *Methods in Ecology and Evolution*. 17 (2), 322-338.
709 https://doi.org/10.1111/2041-210X.70089

710 Wickramasingha, B., West, J., Bellanthudawa, B., Graziano, M.P. & Surasinghe, T.D. (2026) The
711 multifaceted importance of amphibians: ecological, biomedical, and socio-economic perspectives.
712 *Biology*. 15 (1), 98. https://doi.org/10.3390/biology15010098

713 Wildlife Acoustics, Inc. (n.d.) *Song Meter Mini 2 Bat AA*. https://www.wildlifeacoustics.com/products/song-
714 meter-mini-2-bat-aa [Accessed 8th July 2026]

715 Williams, B., van Merriënboer, B., Dumoulin, V., Hamer, J., Fleishman, A.B., McKown, M., Munger, J.,
716 Rice, A.N., Lillis, A., White, C., Hobbs, C., Razak, T., Curnick, D., Jones, K.E. & Denton, T. (2025a)

Using tropical reef, bird and unrelated sounds for superior transfer learning in marine bioacoustics. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 380 (1928), 20240280. https://doi.org/10.1098/rstb.2024.0280

Williams, B., Balvanera, S.M., Sethi, S.S., Lamont, T.A.C., Jompa, J., Prasetya, M., Richardson, L., Chapuis, L., Weschke, E., Hoey, A., Beldade, R., Mills, S.C., Haguenauer, A., Zuberer, F., Simpson, S.D., Curnick, D. & Jones, K.E. (2025b) Unlocking the soundscape of coral reefs with artificial intelligence: pretrained networks and unsupervised learning win out. *PLOS Computational Biology*. 21 (4), e1013029. https://doi.org/10.1371/journal.pcbi.1013029

Wynn, A. & Heyer, W.R. (2001) Do geographically widespread species of tropical amphibians exist? An estimate of genetic relatedness within the neotropical frog *Leptodactylus fuscus* (Schneider 1799) (Anura Leptodactylidae). *Tropical Zoology*. 14 (2), 255-285. https://doi.org/10.1080/03946975.2001.10531157

Xeno-canto Foundation. (2026) *Xeno-canto: sharing animal sounds from around the world*. https://xeno-canto.org [Accessed 7th July 2026]

Zwerts, J.A., Stephenson, P.J., Maisels, F., Rowcliffe, M., Astaras, C., Jansen, P.A., van der Waarde, J., Sterck, L.E.H.M., Verweij, P.A., Bruce, T., Brittain, S. & van Kuijk, M. (2021) Methods for wildlife monitoring in tropical forests: comparing human observations, camera traps, and passive acoustic sensors. *Conservation Science and Practice*. 3 (12), e568. https://doi.org/10.1111/csp2.568

Supplementary Information

Appendix A

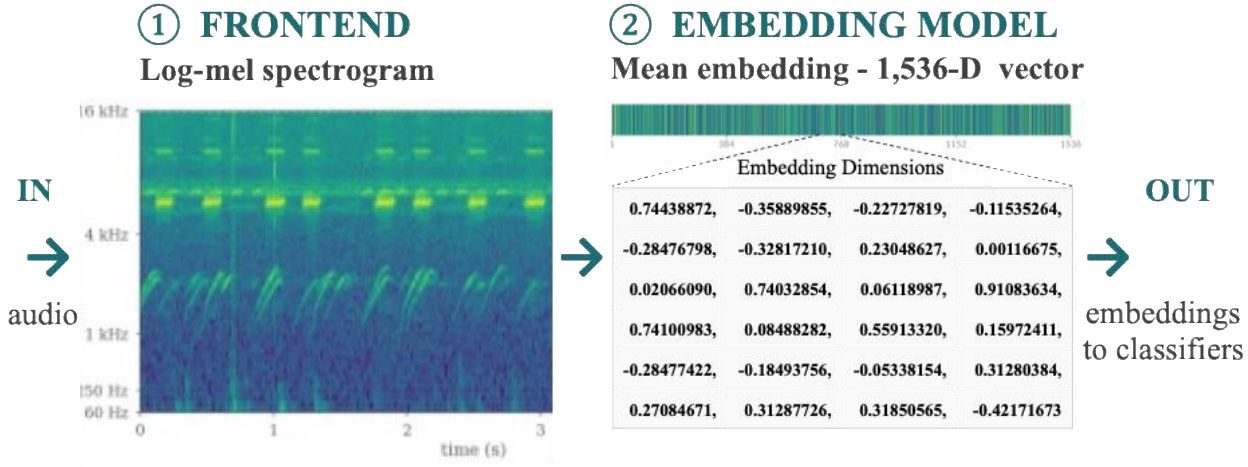


Figure A.1. Two of the three Perch v2 components used to generate embeddings in this study, as defined in the published architecture. (1) The frontend converts five seconds of mono 32 kHz audio into a log mel-spectrogram of 500 frames by 128 mel bands, covering 60 Hz to 16 kHz. (2) The embedding model is an EfficientNet-B3 backbone of roughly 12 million parameters that returns a 5 × 3 time-frequency grid in which each cell carries a 1,536-dimensional feature vector. Averaging that grid over its time and frequency axes yields the 1,536-dimensional mean embedding. (3) Three output heads drive training and produce the released Perch v2 model. They are not shown because that step sits outside the pipeline used in this study. Architecture and parameters after van Merriënboer *et al*. (2025).

Appendix B

Table B.1. The 15 candidate sonotypes taken forward for classifier training, selected by manual review of spectrograms for distinctiveness and ease of recognition. Each spectrogram is five-seconds long but differs in frequency based on how much of the frequency spectrum the sonotype call covers. This ranges from 0-8 kHz to 0-24 kHz. Boxes highlight the calls of interest that represent the sonotype. Sonotype names are arbitrary identifiers and carry no taxonomic information.

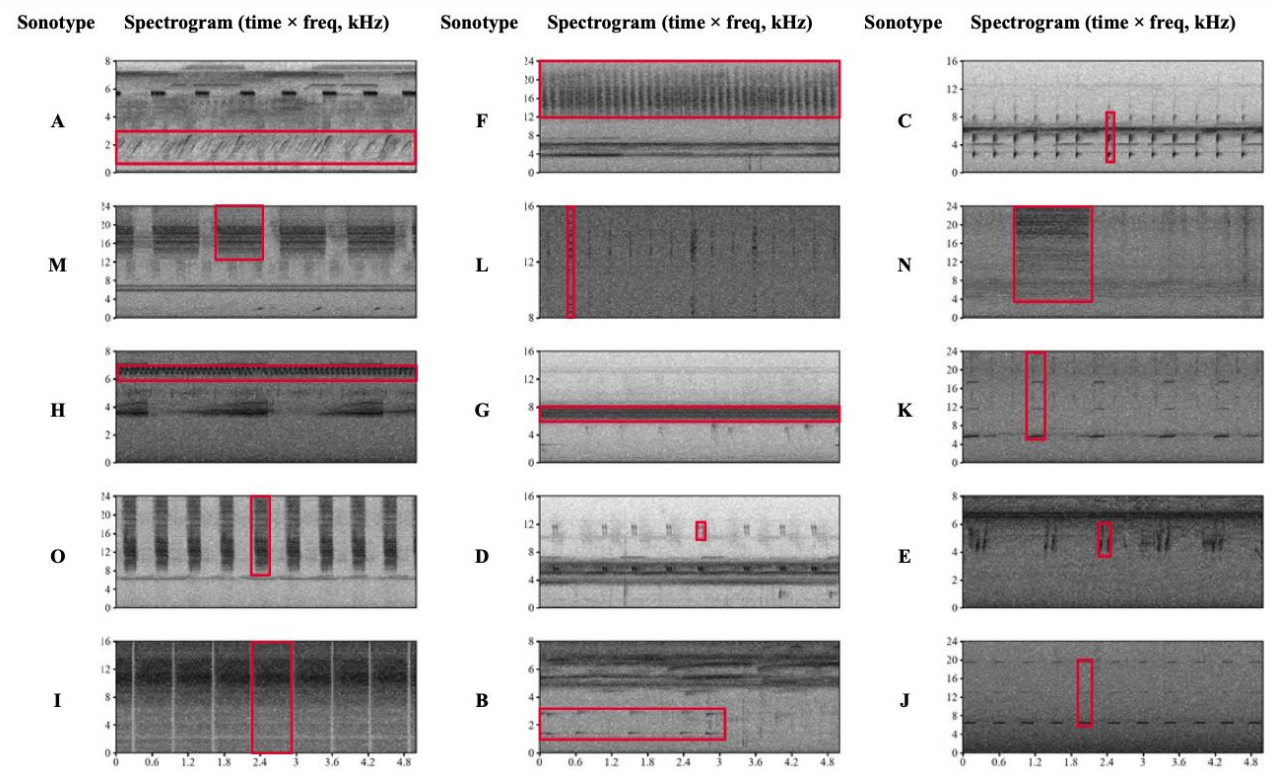


Table B.2. Spectrogram comparison of each sonotype against its best-matching reference recording, for the five sonotypes whose candidate identification is an anuran. Matches come from cosine-similarity retrieval in Perch v2 embedding space across the three reference databases. Each row gives the sonotype identifier (Sonotype ID) and the candidate species name (Species), followed by two spectrograms. Spectrogram of Sonotype is the five-second reference clip that is a good example of the call made by that sonotype, and Spectrogram of Matching Species is the best cosine-similarity match drawn from the three databases. Source Database of Match names the database that supplied that best match, and Number of Matches is how many of the sonotype's 50 best matches carry the same species-level name as the candidate species across the three databases.

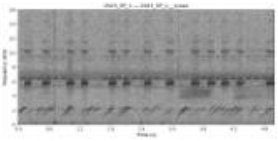
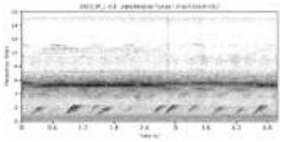
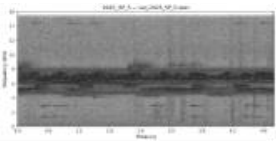
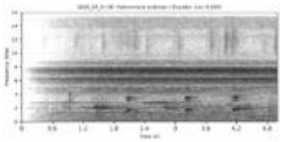
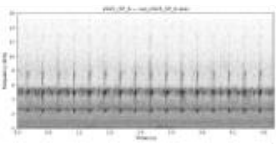
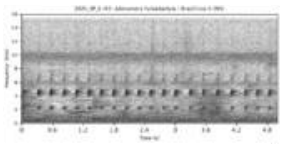
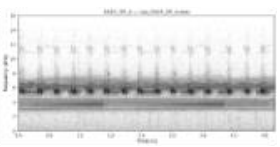
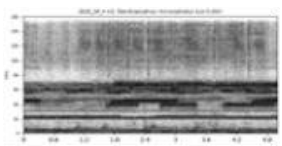
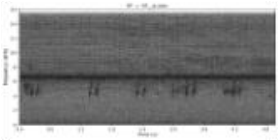
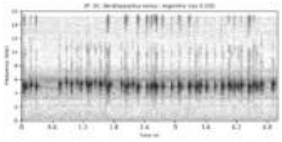
Sonotype ID	Species	Spectrogram of Sonotype	Spectrogram of Matching Species	Source Database of Match	Number of Matches
A	*Leptodactylus fuscus*			iNaturalist	34
B	*Adenomera andreae*			Xeno-canto	2
C	*Adenomera hylaedactyla*			Xeno-canto	6
D	*Dendropsophus microcephalus*			iNaturalist	1
E	*Dendropsophus nanus*			iNaturalist	1

Table B.3. Construction of the training data for the 15 sonotype classifiers, summarised from §2.5. The negative pool was shared across all 15 classifiers. Positive counts per sonotype are given in Table B.4.

Step	Detail
Negative pool, automated draw	Manually selected 'forest noise' and 'rain' clips were pooled with cross-sonotype slices taken across years from each sonotype's confirmed positives, yielding 3,077 clips (2,074 forest-noise, 79 rain, 924 cross-sonotype).
Negative pool, manual curation	The cross-sonotype slice was dropped, leaving 3,000 negatives shared by all 15 classifiers. Because curation was manual, the pool was supplied to the code as an input rather than rebuilt at run time.
Initial positives	77 to about 1,500 confirmed windows per sonotype from the year-stratified vector search of §2.5, giving positive-to-negative ratios of 1:2 to 1:39 (Table B.4).
Active-learning sampling	Three cycles. In each, all ~2,790,000 windows were scored and 200 detections per sonotype were reviewed: the top 50 after 30-sec non-maximum suppression, plus 150 spread across 8 log-score quantile bins. k was set to 50 to raise early-cycle throughput on a 2.79 M-window archive.
Retraining	Cumulative, each cycle refitting on every label collected so far, reaching roughly 600 labels per sonotype (Dumoulin *et al*., 2025).
Guards against label leakage	Cross-negative pairs with centroid cosine > 0.85 were excluded, and any negative-pool window scoring > 0.8 was removed as a likely contaminant positive.

Table B.4. Confirmed positives obtained at each labelling stage for the 15 sonotypes. The vector search column gives the number of retrieved windows manually confirmed as the target sonotype, from the retrieval of §2.5 in which the top 500 windows per year were retained, giving up to 1,500 candidates per sonotype for review. Each of the three active-learning cycles drew 200 detections per sonotype by the top-k plus quantile strategy, so the cycle columns are counts out of 200. Retrieval yield varied by more than an order of magnitude, from 77 windows for sonotype F to the full 1,500 for sonotypes A and C, whereas the per-cycle active-learning yield was comparatively stable, with a median of 51.

Sonotype	Vector search	Cycle 1	Cycle 2	Cycle 3
A	1,500	61	65	67
B	1,486	51	50	53
C	1,500	52	56	54
D	987	57	50	53
E	562	38	36	21
F	77	55	56	52
G	815	22	41	38
H	341	45	47	49
I	401	48	51	48
J	304	54	53	59
K	239	52	52	50
L	90	54	55	51
M	260	16	24	38
N	1,292	51	61	45
O	1,497	54	54	51

Table B.5. Validation performance of the 15 sonotype classifiers after three active-learning cycles summarised in Figure 4. ROC-AUC, PR-AUC and F1-max (with the score threshold at which F1 is maximised) are from the 80:20 stratified validation split. The dataset-corrected in-context metrics are the plain retrieval-AUC (rank separation of classifier score against the manual yes/no review labels), the stratified P(b)-weighted retrieval-AUC, and the call-density P(+) with a 95% confidence interval from 10,000 bootstrap resamples. 'n/a (rare)' marks a stratified retrieval-AUC that cannot be estimated because the sonotype is too rare for any score bucket to contain both a positive and a negative.

Sonotype	ROC-AUC	PR-AUC	F1-max @ thr	Retr-AUC (plain)	Retr-AUC (strat.)	Call-density P(+) [95% CI]
A	0.999	0.999	0.987 @0.79	0.985	0.751	0.156 [0.11, 0.21]
B	1.000	1.000	1.000 @0.59	0.999	0.857	0.068 [0.04, 0.11]
C	1.000	1.000	0.996 @0.74	1.000	0.962	0.075 [0.04, 0.12]
D	1.000	0.999	0.998 @0.88	0.985	0.513	0.073 [0.04, 0.12]
E	0.997	0.993	0.981 @0.84	0.957	0.028	0.054 [0.03, 0.09]
F	0.999	0.992	0.979 @0.68	0.994	0.031	0.061 [0.03, 0.10]
G	0.990	0.958	0.932 @0.83	0.912	0.630	0.066 [0.03, 0.11]
H	0.996	0.981	0.960 @0.63	0.980	0.506	0.068 [0.03, 0.11]
I	1.000	1.000	1.000 @0.80	0.996	1.000	0.055 [0.03, 0.09]
J	1.000	0.999	0.995 @0.76	0.996	0.794	0.106 [0.07, 0.15]
K	1.000	1.000	1.000 @0.80	1.000	n/a (rare)	0.048 [0.02, 0.08]
L	0.999	0.985	0.980 @0.89	1.000	0.875	0.055 [0.03, 0.09]
M	0.997	0.962	0.948 @0.96	0.979	n/a (rare)	0.048 [0.02, 0.08]
N	0.997	0.996	0.988 @0.81	0.973	0.433	0.061 [0.03, 0.10]
O	1.000	1.000	0.997 @0.89	1.000	1.000	0.055 [0.03, 0.09]

Table B.6. Habitat association and temporal trend for each of the 15 sonotypes. Habitat association is the group-size-corrected φ point-biserial coefficient, signed positive toward forest, and the two-group forest-versus-pasture IndVal, both across 3 forest and 26 pasture sites. The trend is the slope in log-odds per year from a per-sonotype binomial GLMM of presence against centred year with a site-level random intercept, fitted on 74 pasture site-years across 26 sites. All P values are false-discovery-rate corrected within their own family and given as Padj. Bold marks Padj < 0.05. Sonotypes are ordered by φ, most forest-associated first. O was present in 69 of the 74 pasture site-years, so its trend sits near complete separation and the slope is not interpretable.

	Habitat association				Trend in proportion of sites present			
Sonotype	φ	Padj	IndVal	Padj	Slope yr^{-1}	SE	P	Padj
B	+0.834	**0.003**	0.921	**0.004**	+1.382	0.491	0.005	**0.037**
F	+0.508	**0.023**	0.793	**0.021**	+0.926	0.349	0.008	**0.040**
C	+0.215	0.358	0.701	0.361	+0.227	0.377	0.547	0.632
G	+0.159	0.498	0.690	0.501	+0.662	0.418	0.113	0.166
H	−0.043	1.000	0.505	1.000	−0.772	0.407	0.058	0.145
I	−0.338	0.194	0.453	0.441	−0.169	0.345	0.623	0.668
D	−0.361	0.175	0.480	0.292	+0.290	0.344	0.399	0.499
J	−0.479	**0.023**	0.686	0.176	−1.152	0.368	0.002	**0.026**
K	−0.518	**0.023**	0.650	0.142	+0.668	0.332	0.044	0.133
E	−0.547	**0.005**	0.773	**0.013**	+2.065	1.208	0.087	0.164
L	−0.558	**0.011**	0.689	0.064	−0.080	0.283	0.777	0.777
M	−0.677	**0.003**	0.793	**0.013**	−0.801	0.371	0.031	0.115
N	−0.680	**0.003**	0.848	**0.004**	−1.534	0.992	0.122	0.166
A	−0.774	**0.003**	0.887	**0.004**	+3.117	1.937	0.108	0.166
O	−0.891	**0.003**	0.941	**0.004**	n/a	n/a	n/a	n/a

Table B.7. Distance to the fixed forest reference across years, and the concordance between the two distance measures. Both models take 2023 as the reference year and carry a site-level random intercept, and both were fitted on 74 pasture site-years across 26 sites. Estimates are on the link scale. The final column back-transforms each year term to the response scale, as a ratio to 2023 for the Gamma model and as a predicted mean for the beta model. Bold marks $p < 0.05$. The concordance row is Spearman's rank correlation between the two distances across the 72 pasture site-years at which both were defined, tested against a site-blocked permutation null of 999 permutations. Predicted mean chord distance was 1.066 (95% CI 0.998 to 1.138) in 2023, 1.006 (95% CI 0.944 to 1.072) in 2024 and 0.943 (95% CI 0.883 to 1.007) in 2025. Predicted mean Sørensen distance was 0.637 (95% CI 0.597 to 0.675) in 2023, 0.604 (95% CI 0.564 to 0.643) in 2024 and 0.536 (95% CI 0.496 to 0.575) in 2025.

Term	Estimate	SE	z	P	Response scale (95% CI)
Embedding chord distance, Gamma GLMM (log link)					
2024 vs 2023	−0.058	0.044	−1.32	0.186	0.944 (0.866, 1.028)
2025 vs 2023	−0.123	0.045	−2.75	**0.006**	0.885 (0.811, 0.965)
Sonotype Sørensen distance, beta GLMM (logit link)					
2024 vs 2023	−0.139	0.115	−1.21	0.227	0.604 (0.564, 0.643)
2025 vs 2023	−0.419	0.112	−3.74	**0.0002**	0.536 (0.496, 0.575)
Concordance between the two distances					
Spearman ρ, n = 72	0.350			**0.033**	