nature portfolio

Corresponding author(s): Sylvain.Haupert
Last updated by author(s): May 7, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The datasets were downloaded from public repository such as Zenodo. No specific software was used for data collection
Data analysis	Annotated Python notebooks code, including the data needed to reproduce the statistical analyses and figures, is publicly available from github repository : https://github.com/ear-team/haupert_acoustic_indices_2025

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All datasets, including the one-minute audio files and associated species richness tables, will be publicly accessible via Zenodo. See the manuscript for the temporary link. Annotated Python notebooks code, including the data needed to reproduce the statistical analyses and figures, is publicly available from github repository : https://github.com/ear-team/haupert_acoustic_indices_2025

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

In this study conducted across sixteen terrestrial habitats, we highlight that while traditional indices often fail to correlate with species richness, new indices have robust and universal predictive species richness capabilities for terrestrial habitats. Additionally, we emphasize the importance of rigorous sampling protocols and the use of appropriate regression methods to ensure reliable and generalizable results.

Research sample

The dataset is composed of 2 dataset, a primary dataset used to build the regression model and a secondary dataset used to test the regression model to new data in order to test the generalization of the model.

The primary dataset contains 5,159 1-minute wav files recorded across 7 habitats and 8 countries. Each audio recordings is associated with the species richness.To build this dataset, we gathered ten different annotated soundscape open-access datasets, encompassing seven terrestrial habitats as defined by IUCN. We standardized the data into 1 min soundscape recordings, such that we obtained a total of 36,252 audio clips. To mitigate the significant yet often overlooked issue of temporal autocorrelation in PAM studies, we ensured a minimum 15-minute interval between each recordings, reducing the initial dataset to 5,159 minute-long soundscapes, distributed across “Forest - Temperate” (n=1,949, 65 sites), “Shrubland - Temperate” (n=675, 15 sites), “Arable land” (n=632, 15 sites), “Plantations” (n=562, 17 sites), “Forest - Subtropical-tropical moist montane” (n=580, 16 sites), “Forest - Subtropical-tropical moist lowland” (n=534, 22 sites) and “Shrubland - Subtropical-tropical high altitude” (n=227, 1 site), encompassing eight countries from North to South America and Europe, ranging from 53°N to 12°S. The soundscapes were recorded at different seasons, different periods of the day, with five types of autonomous recording unit (ARU) devices configured with three sampling rates. For all the datasets, most frequent taxa were manually labeled by experts directly on the audio recordings, except one dataset (i.e. Peru) where only avian was labeled while other taxa such as anuran and insect were also present. Species richness refers to the total number of manually identified singing species per audio.

The secondary dataset is fully based on the WABAD dataset and contains 1,678 1-minute wav files recorded across 16 habitats and 27 countries. Each audio recordings is associated with the species richness.To build this dataset,we ensured a minimum 15-minute interval between each recordings, reducing the initial dataset of 5,044 minute-long soundscapes to 1,678, distributed across “Forest - Temperate” (n=152, 4 sites), “Shrubland - Temperate” (n=77, 3 sites), "Forest - Boreal" (n=53, 5 sites), "Grassland - Temperate" (n=40, 3 sites), "Pastureland" (n=82, 3 sites), “Arable land” (n=64, 3 sites), “Plantations” (n=218, 3 sites), “Forest - Subtropical-tropical dry” (n=179, 11 sites), “Forest - Subtropical-tropical moist lowland” (n=177, 9 sites), "Forest - Subtropical-tropical moist montane" (n=213, 8 sites), "Shrubland - Subtropical-tropical dry" (n=31, site=1), “Shrubland - Subtropical-tropical high altitude” (n=37, 5 sites), "Shrubland - Mediterranean-type" (n=119, 3 sites), "Wetland (inland)" (n=172, 6 sites), "Savanna - Dry" (n=39, 2 sites) and "Desert - Hot" (n=25, 1 site) encompassing twenty seven countries. The soundscapes were recorded at different seasons, different periods of the day, with different types of ARU devices configured with different sampling rates. For all the datasets, only birdsongs were manually labeled by experts directly on the audio recordings. Species richness refers to the total number of manually identified singing species per audio.

References

Kahl, S., Charif, R. & Klinck, H. A collection of fully-annotated soundscape recordings from the Northeastern United States. Zenodo https://doi.org/10.5281/zenodo.7018484 (2022).

Budka, M., Sokołowska, E., Muszyńska, A. & Staniewicz, A. Data from: Acoustic indices estimate breeding bird species richness with daily and seasonally variable effectiveness in lowland temperate Białowieża forest. Zenodo https://doi.org/10.5061/dryad.zcrjdfnhc (2023).

Navine, A., Kahl, S., Tanimoto-Johnson, A., Klinck, H. & Hart, P. A collection of fully-annotated soundscape recordings from the Island of Hawai'i. Zenodo https://doi.org/10.5281/zenodo.7078499 (2022).

Vega-Hidalgo, Á. et al. A collection of fully-annotated soundscape recordings from neotropical coffee farms in Colombia and Costa Rica. Zenodo https://doi.org/10.5281/zenodo.7525349 (2023).

Kahl, S., Wood, C. M., Chaon, P., Peery, M. Z. & Klinck, H. A collection of fully-annotated soundscape recordings from the Western United States. Zenodo https://doi.org/10.5281/zenodo.7050014 (2022).

Eldridge, A., Moscoso, P., Guyot, P. & Peck, M. Data for 'Sounding out Ecoacoustic Metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats'. Zenodo https://doi.org/10.5281/zenodo.1255218 (2018).

Hopping, W. A., Kahl, S. & Klinck, H. A collection of fully-annotated soundscape recordings from the Southwestern Amazon Basin. Zenodo https://doi.org/10.5281/zenodo.7079124 (2022).

Pérez-Granados, C., & Sebastián-González, E. WABAD: A World Annotated Bird Acoustic Dataset for Passive Acoustic Monitoring (v 1.0). Zenodo. https://doi.org/10.5281/zenodo.14191524 (2024).

Sampling strategy	A minimum of five samples per site or per day par site were required for the site to be part of the analysis.
Data collection	10 datasets were collated from publicly accessible platforms that support open-access principles
Timing and spatial scale	To mitigate the temporal autocorrelation, we ensured a minimum 15-minute interval between each recordings. Then we aggregated measurements at the day level from each site or at the site level. The 221 sites were distributed across all continents such that no spatial autocorrelation is expected.
Data exclusions	No data were excluded
Reproducibility	We did not repeat the experiment but we rather test the robustness and generalization of the model. We selected the Huber linear regressor for its robustness to outliers, without intercept, to force the prediction to be 0 when no species are present in the soundscape. We estimate the prediction power of each acoustic indice alone as well as the combination of weakly correlated indices (i.e. correlation coefficient < 0.75). As the assumption that all samples stem from the same generative process is invalid because the soundscape's properties depend on the type of habitat, we cannot assume that the data is Independent and Identically Distributed (i.i.d.). The solution consists in using the group k-fold cross-validation technique. The model was trained iteratively on all habitats except one to assess its ability to generalize to the left-out habitat. The technique maintains the properties of k-fold cross-validation, thereby preventing overfitting and providing a more reliable assessment of the model's performance compared to a simple train-test split. We estimate the goodness of fit using several metrics: the coefficient of determination R^2, the mean absolute error (MAE) and the mean error (ME). R^2 informs about the proportion of variance that the two observed and predicted variables have in common. MAE is a measure of absolute errors between the observed and predicted values. ME is a measure of the bias between the observed and predicted values, a positive value meaning a systematic under-estimation of the ground truth. Finally, we selected the concordance correlation coefficient (CCC) score which tests the agreement and precision between the prediction and the ground truth to assess the capabilities of the model to generalize to a new dataset. CCC is specifically designed to evaluate agreement by considering both correlation and the closeness of the means and variances.
Randomization	n/a
Blinding	n/a

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a