

Acoustic indices as proxies for biodiversity: a meta-analysis

Supporting information

Irene Alcocer, Herlander Lima, Larissa S. M. Sugai, Diego Llusia

A neat presentation of the supplementary material can be found in the online supplementary material at <https://irene-alcocer.github.io/Acoustic-Indices/>

Table S1. Data set used in the study. Due to the size of this table it is only available at <https://irene-alcocer.github.io/Acoustic-Indices/index.html#Dataset>

Table S2. Variable descriptions for Table S1. Available at <https://irene-alcocer.github.io/Acoustic-Indices/index.html#Dataset>

Table S3. List of the 34 features used to characterise studies that tested the relationship between acoustic indices and diversity metrics.

| Category | Features | Description |
|------------------------|--|--|
| Publication | Authors | |
| | Title | |
| | Journal | |
| | Year of publication | |
| | Peer reviewed | Whether the study was subjected to peer review (Yes or No) |
| Biological data | Environment | Ecosystem type where recordings were collected (aquatic or terrestrial) |
| | Taxonomic group | Primary studied group (invertebrates, fish, anurans, mammals, birds, or several) |
| | Diversity metric | Species abundance, species richness, species diversity, abundance of sounds, or diversity of sounds |
| | Diversity source | Method applied to obtain the diversity metric (acoustic or non-acoustic) |
| Acoustic data | Acoustic index | ACI, AEI, ADI, AR, BIO, H, H _i , H _f , M, NP, or NDSI. See Table 2 for definitions of acoustic indices |
| | Frequency range | Range (in Hz) used for index calculation |
| | FFT size | Window size of the Fast Fourier Transformation (FFT) |
| | Noise treatment | Audio pre-processing related to noise (noise filtering, noise addition, or exclusion of noisy recordings) |
| Recording | Sampling rate | Number of audio samples per second (in kHz) used for index calculation |
| | Audio format | Format of audio files (.wav, .mp3, etc.) |
| | Recording length | Length of each recording (in seconds) used for index calculation |
| | Recording method | Non-programmed (continuous), programmed (periodic) or manual (by an operator) |
| Sampling design | Study sites | Number of study sites (= spatial replicates) |
| | Distance between sites | Minimum distance between study sites (in meters) |
| | Recorders per site | Number of recording units per study site |
| | Recording days | Number of recording days per study site (= temporal replicates) |
| | Daily period | Period recorded within the day (dawn, morning, midday, evening, dusk, night, or all day) |
| | Daily sample | Number of recordings collected within a day per study site |
| Statistics | Statistical test | Statistical analysis used to test the relationship between acoustic indices and diversity metrics |
| | Independence | Whether the statistical test was considered independent from other tests of the same study |
| | R ² | Coefficient of determination (for regression analysis) |
| | r | Correlation coefficient (for Pearson or Spearman correlation) |
| | b | Regression coefficient (for linear regression analysis) |
| | t-statistic | Statistic value for Student's <i>t</i> -test |
| | Standard error | Standard error of the test coefficient |
| | Sample size | Number of observations included in the statistical test |
| | Pseudoreplication | Inadequate specification of the number of true replicates in the statistical test (Yes or No) |
| | Pseudoreplication type | Spatial, temporal, or spatial-temporal pseudoreplication |
| Adjusted sample size | Suitable specification of the number of true replicates (for pseudoreplicated studies) | |

Table S4. Number of effect sizes collected from each of the 34 studies included in the meta-analysis. ID corresponds to the study identification number in our data set.

| ID | Study | Number of effect sizes |
|------|--|------------------------|
| 2740 | Mammides <i>et al.</i> (2017) | 84 |
| 53 | Moreno-Gómez <i>et al.</i> (2019) | 42 |
| 87 | Eldridge <i>et al.</i> (2018) | 28 |
| 80 | Staaterman <i>et al.</i> (2017) | 24 |
| 90 | Ferreira <i>et al.</i> (2018) | 24 |
| 96 | Retamosa Izaguirre & Ramírez-Alán (2018) | 24 |
| 10 | Buscaino <i>et al.</i> (2016) | 22 |
| 2 | Desjonquères <i>et al.</i> (2015) | 12 |
| 11 | Bertucci <i>et al.</i> (2016) | 12 |
| 89 | Gage <i>et al.</i> (2017) | 12 |
| 2977 | Jorge <i>et al.</i> (2018) | 12 |
| 70 | Bolgan <i>et al.</i> (2018) | 11 |
| 9 | Harris <i>et al.</i> (2016) | 6 |
| 86 | Indraswari <i>et al.</i> (2018) | 6 |
| 427 | Fuller <i>et al.</i> (2015) | 6 |
| 14 | wa Maina <i>et al.</i> (2016) | 4 |
| 60 | Lyon <i>et al.</i> (2019) | 4 |
| 77 | Fairbrass <i>et al.</i> (2017) | 4 |
| 92 | Torti <i>et al.</i> (2018) | 4 |
| 15 | Roca & Proulx (2016) | 3 |
| 45 | McLaren & DeGroot (2012) | 3 |
| 13 | McWilliam & Hawkins (2013) | 2 |
| 41 | Paisley-Jones (2011) | 2 |
| 44 | Machado <i>et al.</i> (2017) | 2 |
| 251 | Buxton <i>et al.</i> (2016) | 2 |
| 4 | Parks <i>et al.</i> (2014) | 1 |
| 6 | Boelman <i>et al.</i> (2007) | 1 |
| 17 | Zhang <i>et al.</i> (2015) | 1 |
| 37 | Picciulin <i>et al.</i> (2016) | 1 |
| 1132 | Depraetere <i>et al.</i> (2012) | 1 |
| 1177 | Joo <i>et al.</i> (2011) | 1 |
| 1262 | Pieretti <i>et al.</i> (2011) | 1 |
| 2745 | Sueur <i>et al.</i> (2008) | 1 |
| 2986 | Raynor <i>et al.</i> (2017) | 1 |

Table S5. Number of effect sizes and studies per moderator levels. See Table 2 for definitions of acoustic indices.

| Moderator levels | Effect sizes | Studies |
|--------------------------|--------------|---------|
| Acoustic indices | | |
| ACI | 113 | 25 |
| ADI | 38 | 12 |
| AEI | 34 | 8 |
| AR | 18 | 5 |
| BIO | 36 | 10 |
| H | 55 | 16 |
| H _f | 12 | 3 |
| H _t | 15 | 4 |
| M | 5 | 2 |
| NDSI | 33 | 10 |
| NP | 5 | 2 |
| Diversity metrics | | |
| Species abundance | 27 | 6 |
| Species diversity | 49 | 9 |
| Species richness | 187 | 21 |
| Abundance of sounds | 66 | 11 |
| Diversity of sounds | 35 | 3 |
| Diversity source | | |
| Acoustic | 200 | 26 |
| Non-acoustic | 164 | 11 |
| Environment | | |
| Aquatic | 95 | 10 |
| Terrestrial | 269 | 24 |

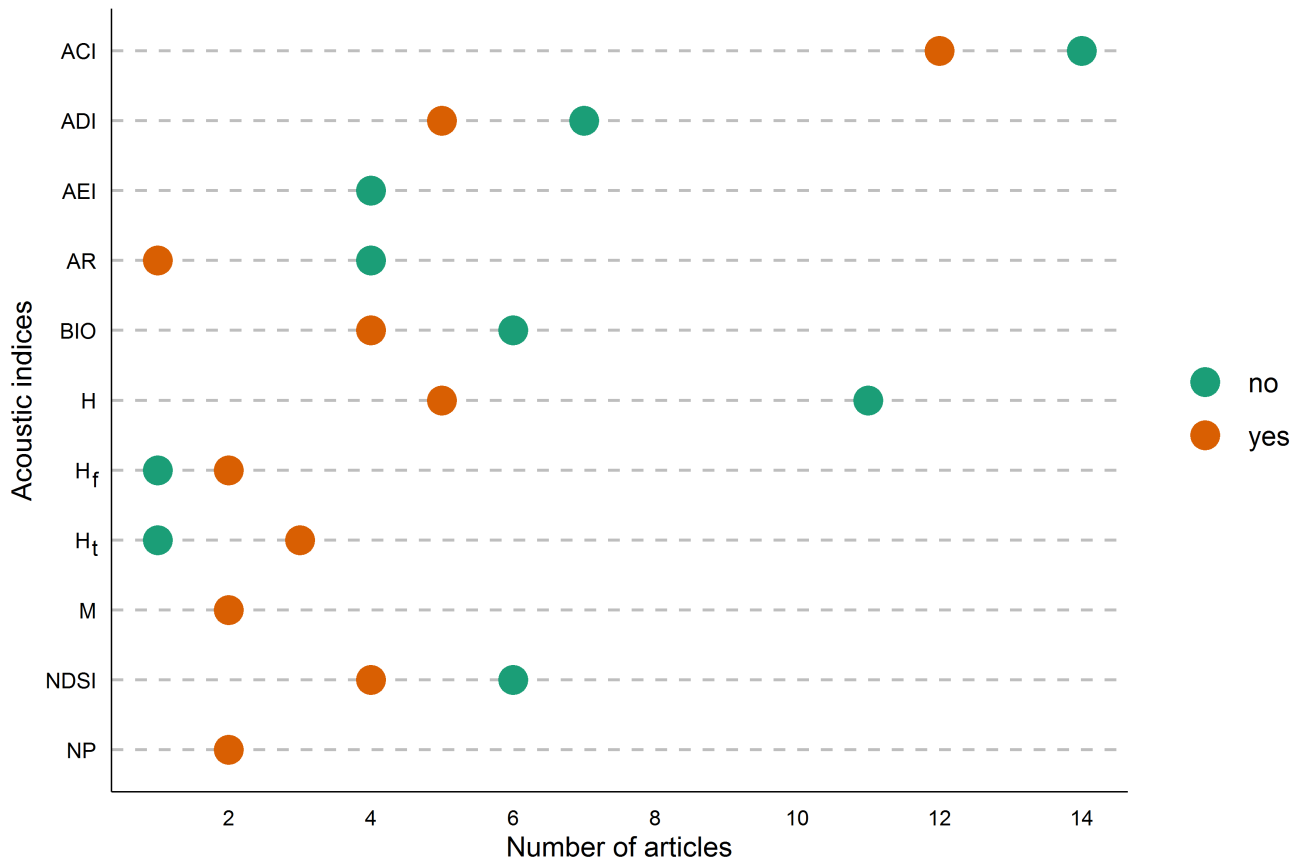


Fig. S1. Pseudoreplication summary. The data represent the total number of articles for each index. Orange dots represent the number of studies with pseudoreplication and green are without pseudoreplication. One article (Papin et al., 2019) was withdrawn from the pseudoreplication analysis as it was impossible to obtain data on pseudoreplication, thus leading to some variation in the total for each acoustic index. See Table 2 for definitions of acoustic indices.

Table S6. VIF (Variance Inflation Factor) values obtained for each moderator level. See Table 2 for definitions of acoustic indices.

| Moderators | VIF |
|---------------------|-------|
| ADI | 1.497 |
| AEI | 1.450 |
| AR | 1.250 |
| BIO | 1.478 |
| H | 1.472 |
| NDSI | 1.437 |
| Species abundance | 1.244 |
| Species diversity | 1.156 |
| Abundance of sounds | 1.454 |
| Aquatic | 1.393 |
| Acoustic | 1.623 |

Table S7. Model estimates for the intercept-only model. All values are Pearson correlation r . SE is the standard error and CI the confidence interval. Full model output can be found in the online supplementary material.

| | Estimate | SE | CI |
|----------------|----------|-------|----------------|
| Overall effect | 0.333 | 0.058 | [0.229, 0.429] |

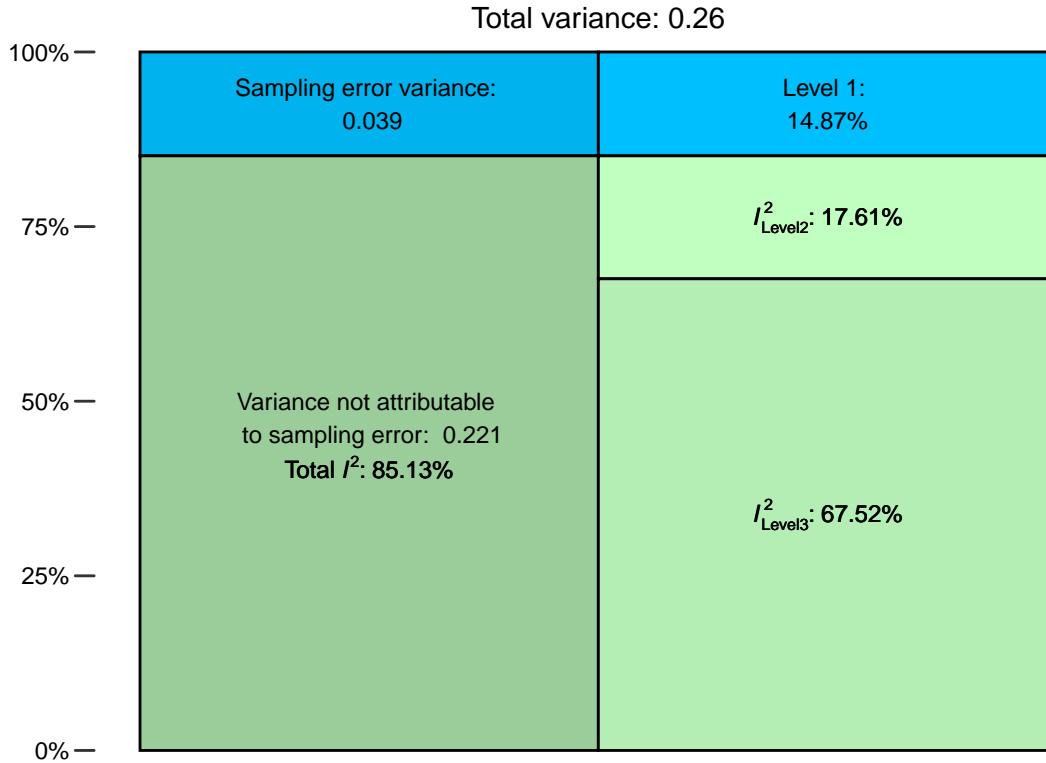


Fig. S2. Visual representation of the distribution of variance over the multilevel structure of the intercept-only model. Within-study heterogeneity (level 2) corresponds to the unaccounted for variation that is found on effect sizes within studies, and between-study heterogeneity corresponds to the unaccounted for variation between studies (level 3).

Table S8. Model estimates for the sub-group analysis. All values are Pearson correlation r . SE is the standard error and CI the confidence interval. See Table 2 for definitions of acoustic indices. Full model output can be found in the online supplementary material.

| Index | Estimate | SE | CI |
|-------|----------|-------|-----------------|
| ACI | 0.363 | 0.068 | [0.242, 0.474] |
| ADI | 0.244 | 0.097 | [0.058, 0.414] |
| AEI | 0.040 | 0.104 | [-0.164, 0.240] |
| AR | 0.078 | 0.135 | [-0.185, 0.331] |
| BIO | 0.193 | 0.101 | [-0.003, 0.374] |
| H | 0.501 | 0.090 | [0.358, 0.622] |
| NDSI | 0.427 | 0.103 | [0.247, 0.578] |

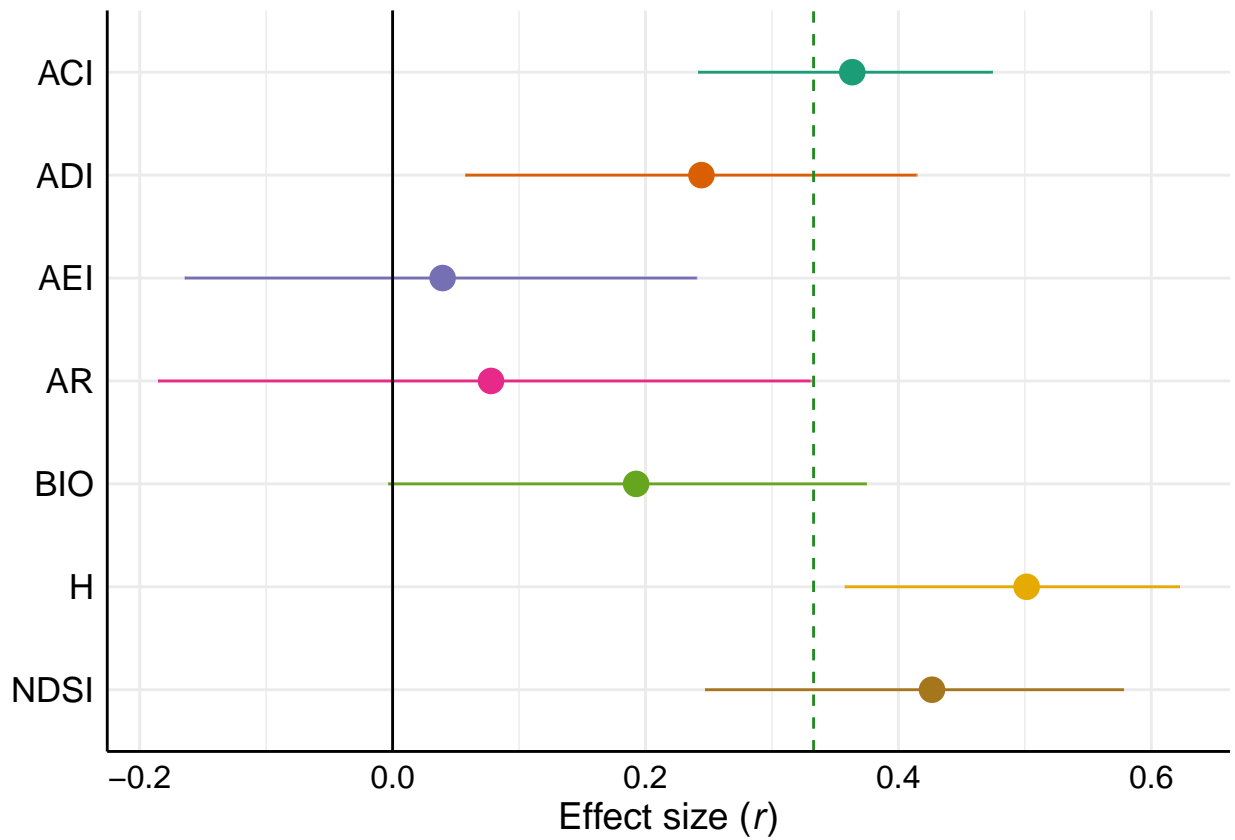


Fig. S3. Effect size mean estimates (circles) and corresponding 95% confidence intervals (horizontal lines) obtained from the sub-group meta-analysis with acoustic indices as the moderating factor. Estimated effect sizes whose 95% confidence intervals do not overlap zero (black vertical line) indicate a positive correlation between acoustic indices and diversity if they are to the right of zero, or a negative correlation if they are to the left of zero. The dashed green vertical line represents the summary effect size obtained from the intercept only meta-analysis.

Table S9. Results of Wald-type tests for all moderators (first row), and for each moderator separately (remaining rows). “Q” is the Wald statistic; “df” are the degrees of freedom; and “p” is the probability that moderator estimates came from a chi-square distribution, where all estimates are equal to zero. Thus $p < 0.05$ provides support against the null hypothesis that moderator levels estimates are equal to zero (i.e. they do not explain variation in effect sizes).

| Moderator | Q | df | p |
|-------------------|--------|----|-------|
| All moderators | 27.428 | 11 | 0.004 |
| Acoustic indices | 22.353 | 6 | 0.001 |
| Diversity metrics | 3.561 | 3 | 0.313 |
| Environment | 0.196 | 1 | 0.658 |
| Diversity source | 0.004 | 1 | 0.950 |

Table S10. Results of Wald-type tests for the contrasts between acoustic index H and all other acoustic indices. The column “Compared” expresses the difference between the estimate for H and the estimate of each of the other acoustic indices. The column “Estimate” is the estimate obtained from the difference shown in the previous column. “SE” is the standard error of the difference, and CI.lb, CI.ub the confidence interval lower and upper bound, respectively. “ Q_m ” is the Wald statistic; “p” is the probability that the difference between estimates is equal to zero. Thus, $p < 0.05$ gives support against the null hypothesis of no difference between the estimate of the H index and the estimate of the other index.

| Compared | Estimate | SE | CI.lb | CI.ub | Q_M | p |
|----------|----------|-------|--------|-------|--------|-------|
| H - ADI | 0.327 | 0.121 | 0.088 | 0.565 | 7.223 | 0.007 |
| H - AEI | 0.489 | 0.126 | 0.240 | 0.737 | 14.901 | 0.000 |
| H - AR | 0.471 | 0.151 | 0.173 | 0.769 | 9.622 | 0.001 |
| H - BIO | 0.342 | 0.124 | 0.098 | 0.586 | 7.590 | 0.005 |
| H - ACI | 0.197 | 0.109 | -0.016 | 0.411 | 3.280 | 0.070 |
| H - NDSI | 0.113 | 0.130 | -0.141 | 0.368 | 0.765 | 0.381 |

Table S11. Results of Wald-type tests for the contrasts between acoustic index NDSI and all other acoustic indices. The column “Compared” expresses the difference between the estimate NDSI and the estimate of each of the other acoustic indices. The column “Estimate” is the estimate obtained from the difference expressed in the previous column; “SE” is the standard error of the difference, and CI.lb, CI.ub the confidence interval lower and upper bound, respectively; “ Q_m ” is the Wald statistic; “p” is the probability that the difference between estimates is equal to zero. Thus, a $p < 0.05$ gives support against the null hypothesis of no difference between the estimate of the NDSI index and the estimate of the other index.

| Compared | Estimate | SE | CI.lb | CI.ub | Q_M | p |
|------------|----------|-------|--------|-------|-------|-------|
| NDSI - ADI | 0.213 | 0.132 | -0.046 | 0.473 | 2.585 | 0.107 |
| NDSI - AEI | 0.375 | 0.137 | 0.105 | 0.645 | 7.442 | 0.006 |
| NDSI - AR | 0.357 | 0.161 | 0.041 | 0.673 | 4.913 | 0.026 |
| NDSI - BIO | 0.228 | 0.135 | -0.035 | 0.493 | 2.869 | 0.090 |
| NDSI - H | -0.113 | 0.130 | -0.368 | 0.141 | 0.765 | 0.381 |
| NDSI - ACI | 0.084 | 0.126 | -0.162 | 0.331 | 0.444 | 0.504 |

Table S12. Results of Wald-type tests for the contrasts between the diversity metric “abundance of sounds” and all other diversity metrics. The column “Compared” expresses the difference between the estimate “abundance of sounds” and the estimate of each of the other diversity metrics. The column “Estimate” is the estimate obtained from the difference shown in the previous column; “SE” is the standard error of the difference, and CI.lb, CI.ub the confidence interval lower and upper bound, respectively; “ Q_m ” is the Wald statistic; “p” is the probability that the difference between estimates is equal to zero. Thus, a $p < 0.05$ gives support against the null hypothesis of no difference between the estimate of the abundance of sounds metric and the estimate of the other metric.

| Compared | Estimate | SE | CI.lb | CI.ub | Q_M | p |
|---|----------|-------|--------|-------|-------|-------|
| Abundance of sounds - Species abundance | 0.341 | 0.214 | -0.079 | 0.762 | 2.530 | 0.111 |
| Abundance of sounds - Species diversity | 0.301 | 0.171 | -0.034 | 0.638 | 3.087 | 0.078 |
| Abundance of sounds - Species richness | 0.259 | 0.146 | -0.028 | 0.547 | 3.129 | 0.076 |

Table S13. Model estimates for the meta-analysis inspecting differences between pseudoreplicated and non-pseudoreplicated studies. The intercept correspond to the non-pseudoreplicated effect sizes. All values are Pearson correlation r . SE is the standard error, CI the confidence interval. Full model output can be found in the online supplementary material

| Coefficients | Estimate | SE | CI |
|------------------|----------|-------|-----------------|
| Intercept | 0.359 | 0.071 | [0.232, 0.474] |
| Pseudoreplicated | -0.078 | 0.123 | [-0.310, 0.161] |

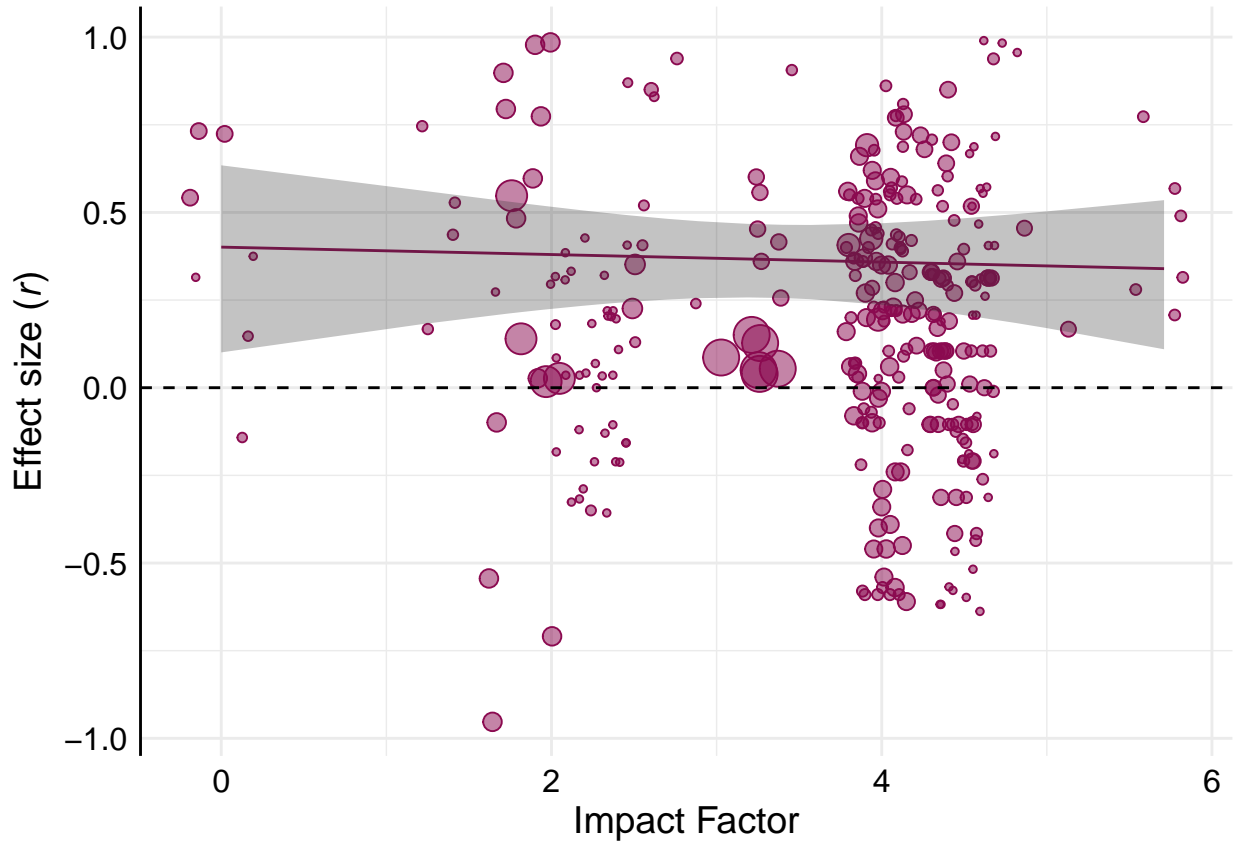


Fig. S4. Relationship between reported mean effect sizes and journal impact factor. Circle size indicates the relative sample size for each effect size. The fitted line is a meta-regression over the journal impact factor with the corresponding 95% confidence interval region shaded grey. The dashed horizontal line represents an effect size of 0. Effect size mean values are positioned along the impact factor axis with minor random noise to reduce overlapping. Model statistics in Pearson correlation r estimate [CI], intercept 0.40 [0.10, 0.63], slope -0.01 [-0.10, 0.08]. Full model output can be found in the online supplementary material

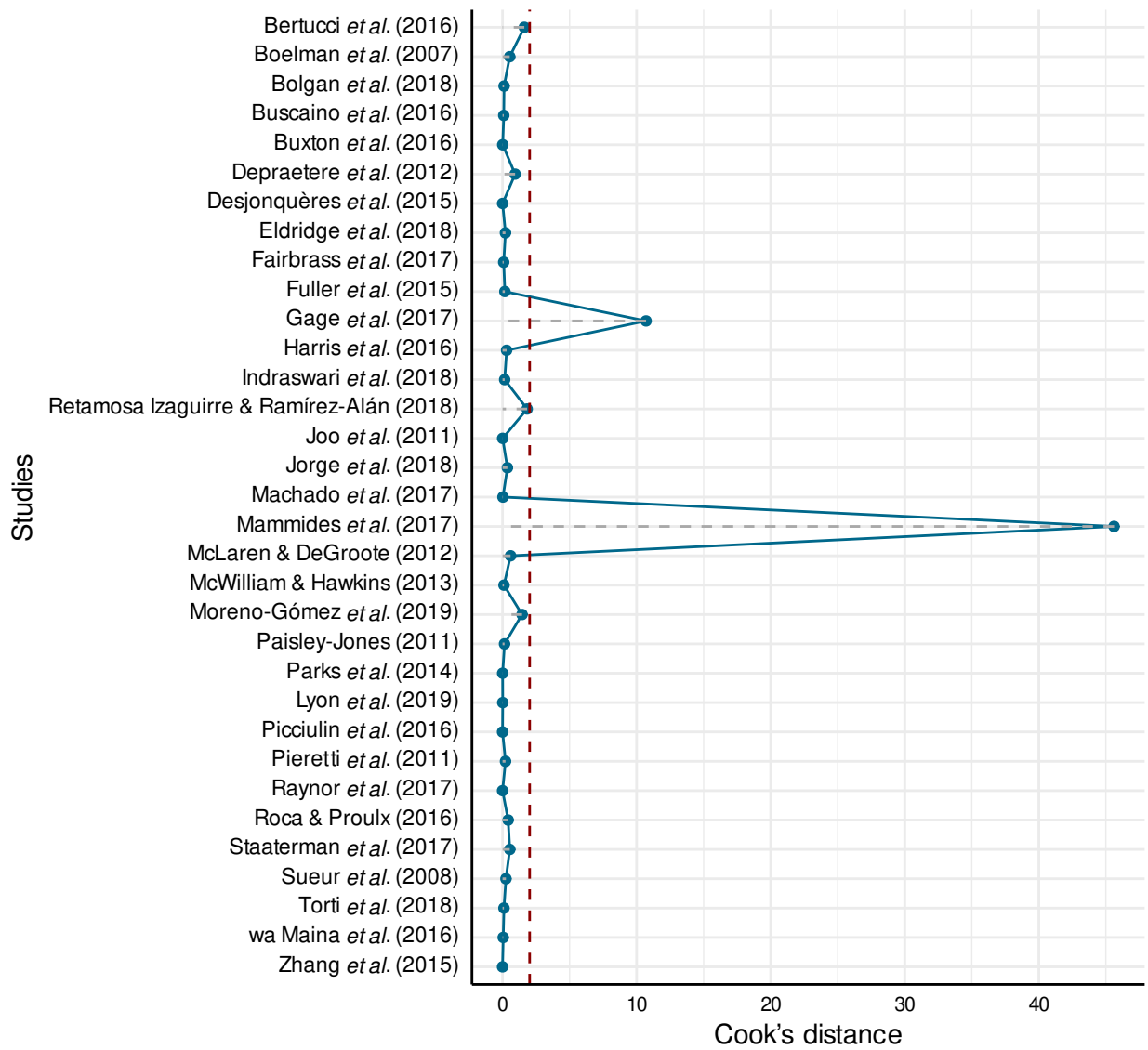


Fig. S5. Cook's distance values for each study (blue dots) and average Cook's distance over all studies (dashed vertical red line). The Cook's distance for a given study can be interpreted as the distance between the entire set of predicted values with this study included and when this study is excluded from the model fitting procedure.

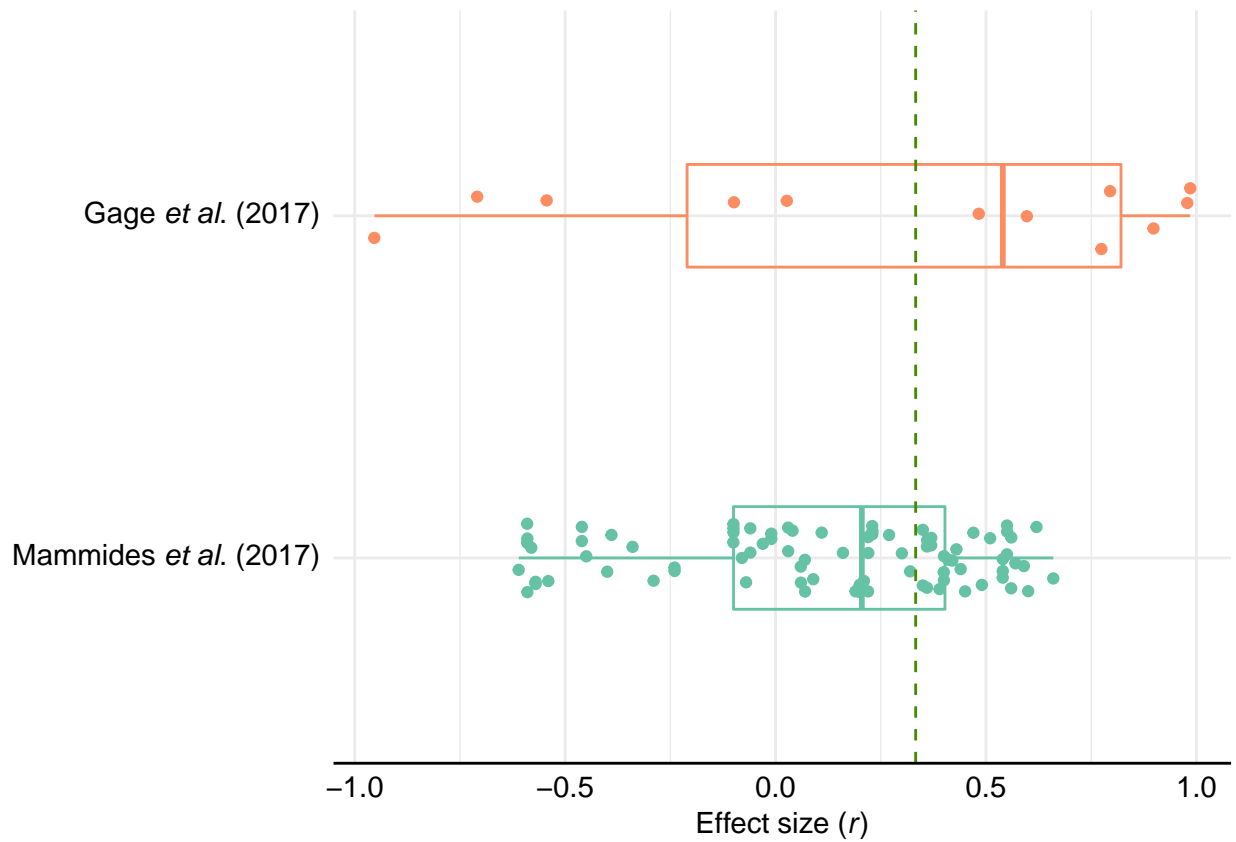


Fig. S6. Boxplot and distribution of effect size values (dots) of the two studies identified as outliers. The x-axis shown is the Pearson r effect size. The green vertical dashed line is the summary effect obtained in the intercept-only model.

Table S14. Model estimates for the meta-regression over the data set without outliers. All values are Pearson correlation r . SE is the standard error, CI the confidence interval. Full model output can be found in the online supplementary material.

| Coefficients | Estimate | SE | CI |
|---------------------|----------|-------|-----------------|
| Intercept | 0.569 | 0.185 | [0.272, 0.767] |
| ADI | -0.203 | 0.111 | [-0.401, 0.013] |
| AEI | -0.075 | 0.125 | [-0.312, 0.170] |
| AR | -0.077 | 0.205 | [-0.450, 0.319] |
| BIO | -0.099 | 0.119 | [-0.321, 0.134] |
| H | -0.036 | 0.105 | [-0.238, 0.170] |
| NDSI | -0.016 | 0.140 | [-0.284, 0.253] |
| Species abundance | -0.155 | 0.150 | [-0.424, 0.139] |
| Species diversity | -0.106 | 0.129 | [-0.346, 0.147] |
| Abundance of sounds | 0.158 | 0.207 | [-0.248, 0.517] |
| Aquatic | -0.188 | 0.203 | [-0.533, 0.209] |
| Acoustic | -0.137 | 0.189 | [-0.472, 0.232] |

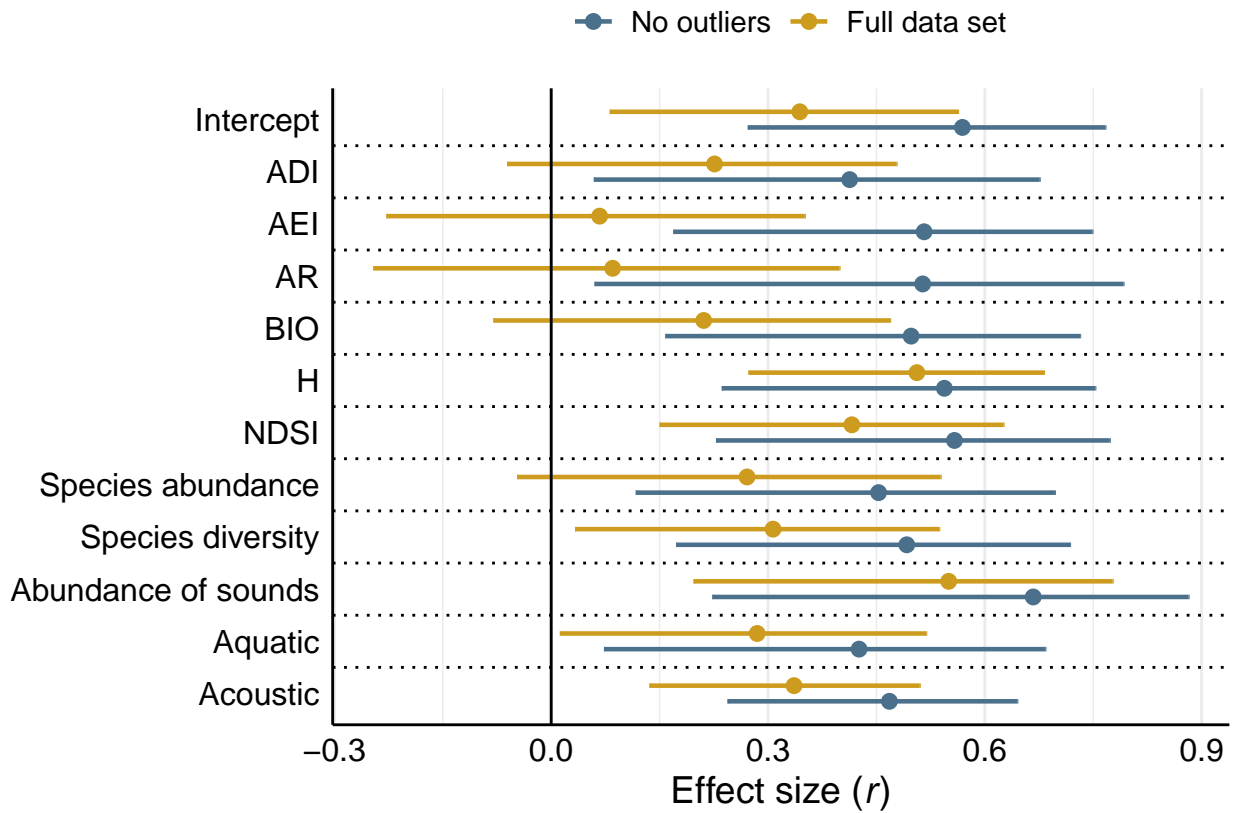


Fig. S7. Contrast of model estimates obtained with meta-regression analysis using the full data set (yellow) and over the data set with outliers removed (blue). Estimates are represented by circles and corresponding 95% confidence intervals by horizontal lines. Each estimate (except the intercept) corresponds to the additive effect of each coefficient as obtained with the *predict_rma* function from *metafor* R package. Estimated effect sizes whose 95% confidence intervals do not overlap zero (black vertical line) indicate a positive correlation between acoustic indices and diversity if they are to the right of zero, or a negative correlation if they are to the left of zero. We considered as outliers every study that had a Cook's distance value higher than the mean of all Cook distances (see Fig.S5). Model moderators were acoustic indices (ADI, AEI, AR, BIO, H, NDSI, with ACI as intercept), diversity metric (Species abundance, Species diversity, Abundance of sounds, with Species richness as intercept), environment (Aquatic, with Terrestrial as intercept), diversity source (Acoustic, with Non-Acoustic as intercept). The solid vertical black line represents a null effect size. See Table 2 for definitions of acoustic indices.