Acoustic indices perform better when applied at ecologically meaningful time and frequency scales

Oliver C. Metcalf[1], Jos Barlow[2,3,4], Christian Devenish[1], Stuart Marsden[1], Erika Berenguer,[3,5], Alexander C. Lees[1,6]

1. Division of Biology and Conservation Ecology, Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK
2. Departamento de Biologia, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil
3. Lancaster Environment Centre, Lancaster University, Lancaster, Lancashire, UK
4. MCTIC/Museu Paraense Emílio Goeldi, Belém, Pará, Brazil
5. Environmental Change Institute, School of Geography and the Environment, University of Oxford, OX1 3QY, Oxford, UK
6. Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA

Contact author:

Oliver C. Metcalf[1]

o.metcalf@mmu.ac.uk

Division of Biology and Conservation Ecology, Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK

Running headline: Acoustic indices impacted by signal masking

Word count: 6736

Abstract:

1. Acoustic indices are increasingly employed in the analysis of soundscapes to ascertain biodiversity value. However, conflicting results and lack of consensus on best practices for their usage has hindered their application in conservation and land-use management contexts. Here we propose that the sensitivity of acoustic indices to ecological change and fidelity of acoustic indices to ecological communities are negatively impacted by signal masking. Signal masking can occur when acoustic responses of taxa sensitive to the effect of interest are masked by less sensitive acoustic groups, or target taxa sonification is masked by non-target noise. We argue that by calculating acoustic indices at ecologically appropriate time and frequency bins, masking effects can be reduced and the efficacy of indices increased.
2. We test this on a large acoustic dataset collected in Eastern Amazonia spanning a disturbance gradient of undisturbed, logged, burned, logged-and-burned, and secondary forests. We calculated values for two acoustic indices: the Acoustic Complexity Index and the Bioacoustic Index, across the entire frequency spectrum (0-22.1 kHz), and four narrower subsets of the frequency spectrum; at dawn, day, dusk and night.
3. We show that signal masking has a large impact on the sensitivity of acoustic indices to forest disturbance classes. Calculating acoustic indices at a range of narrower time-frequency bins substantially increases the classification accuracy of forest classes by random forest models. Furthermore, signal masking led to misleading correlations, including spurious inverse correlations, between biodiversity indicator metrics and acoustic index values compared to correlations derived from manual sampling of the audio data.
4. Consequently, we recommend that acoustic indices are calculated either at a range of time and frequency bins, or at a single narrow bin, predetermined by *a priori* ecological understanding of the soundscape.

1 – Índices acústicos são cada vez mais utilizados em análises de paisagens sonoras para entender padrões de biodiversidade. Entretanto, sua aplicação em biologia da conservação e em contextos de manejo do uso do solo têm sido atrasada devido a resultados conflitantes e a uma falta de consenso sobre as melhores práticas a serem empregadas. Aqui nós propomos que a sensibilidade de índices acústicos em capturar mudanças ecológicas, assim como a fidelidade com que índices acústicos capturam comunidades ecológicas, são severamente impactados por mascaramento do sinal. O mascaramento do sinal pode ocorrer quando respostas acústicas sensíveis aos efeitos que estão sendo monitorados são mascaradas por outros grupos menos sensíveis ou quando a vocalização do taxa alvo dos estudos é mascarado por barulho de outros taxa. Nós argumentamos que ao calcular índices acústicos em intervalos apropriados de tempo e frequência, efeitos mascaradores podem ser reduzidos e a eficácia dos índices acústicos aumentada.

2 – Nós testamos isso em um vasto grupo de dados acústicos coletados na Amazônia oriental, abrangendo um gradiente de distúrbios antrópicos, incluindo florestas primárias não perturbadas e aquelas afetadas por extração madeireira, incêndios florestais, extração madeireira e incêndios, assim como florestas secundárias. Nós calculamos os valores de dois índices acústicos, o Índice de Complexidade Acústica e o Índice Bioacústico. Para isso, empregamos todo o espectro de frequências (0-22.1kHz) e quatro subgrupos menores do espectro de frequências: o amanhecer, o dia, o anoitecer e a noite.

3 – Nós mostramos que o mascaramento do sinal tem um grande impacto na sensibilidade dos índices acústicos a distúrbios florestais. Calculando índices acústicos em um intervalo menor de tempo-frequência aumentou substancialmente a acurácia da classificação das classes florestais por modelos do tipo Random Forest. Além disso, o mascaramento do sinal levou a correlações errôneas, incluindo correlações negativas espúrias entre métricas de biodiversidade e valores de índices

acústicos, quando comparados com correlações geradas a partir de amostragem manual dos dados de áudio.

4 – Consequentemente, nós recomendamos que índices acústicos sejam calculados em intervalos de tempo e frequência menores, pré-determinados por conhecimento ecológico a priori da paisagem sonora.

Key words:

Acoustic indices, ecoacoustics, remote sensing, bioacoustics, biodiversity, Amazonia, soundscape, tropical ecology

Introduction

Acoustic monitoring is rapidly becoming a key tool to measure biodiversity, with strident calls for broader uptake (Burivalova et al., 2019; Deichmann et al., 2018; Wagner Ribeiro Jr et al., 2017). Despite increasing ease of data collection, there remain significant obstacles to the analysis of acoustic data, with species-level classification limited by the expertise and effort required to train machine-learning models, and the limited availability of both open source software and large audio libraries (Gibb et al., 2019; Priyadarshani et al., 2018). Consequentially, the use of acoustic indices has grown in popularity, often used as proxies for more traditional biodiversity metrics like species richness and composition, and presented as alternative effective tools for rapid biodiversity assessments (Sueur et al., 2008). There are a wide range of acoustic indices, but most involve calculating and comparing acoustic power within temporal and frequency bins (Farina, 2014; Sueur, Farina, Gasc, Pieretti, & Pavoine, 2014; Buxton et al., 2018; Gibb et al., 2019). These are, in turn, used to assess soundscape qualities such as evenness, entropy and complexity. Acoustic indices infer community level information from entire soundscapes, in contrast to species-level classification approaches that require time-consuming complex model-training techniques necessitating large training libraries, indices are relatively simple and readily available on a range of open-source platforms.

Despite their increasing popularity, acoustic indices are not always effective at answering key questions related to conservation or natural resource management. The first issue relates to their *sensitivity* to changes in environmental conditions. Acoustic indices have been shown to effectively distinguish between disparate land uses (Bradfer-Lawrence et al., 2019; Carruthers-Jones et al., 2019; Depraetere et al., 2012). However, they are less successful in distinguishing differences between similar land uses; for example between different types of forest (Bormpoudakis et al., 2013, Eldridge et al., 2018, Do Nascimento et al., 2020), or require a very large number of spatial replications to do so (Mitchell et al., 2020). The second issue relates to their *fidelity* as indicators of biodiversity, as they can be inconsistent predictors of traditionally-used biodiversity metrics such as species richness (Eldridge et al., 2018; Fuller et al., 2015; Jorge et al., 2018; Mammides et al., 2017).

The problems of low sensitivity and inconsistent fidelity are potentially caused by signal masking – whilst certain vocalising taxa or taxonomic groups may respond strongly to changes in environmental condition, others may not. By measuring acoustic indices at intervals that measure across multiple taxonomic groups, sensitivity to these varied responses is lost, which may not be the case if indices were measured with multiple intervals. Similarly, fidelity to a single taxonomic group is lost by the use of broad time and frequency intervals, which may be improved by the use of narrower, tailored intervals.

There are two key ways in which signal masking can occur in acoustic indices. The first, temporal masking, can occur when acoustic indices are measured over time periods that are too long, so that sounds from sensitive time periods may be confounded by a lack of change or contrasting responses in other time periods. For example, the vocal community at dawn may respond to a disturbance event very differently from the dusk community (Deichmann et al., 2017), so that measuring both together masks overall community responses. To avoid this, the analysis of acoustic indices often involves temporally limiting or splitting the data analysed into discrete periods, such as dawn and dusk (Bradfer-Lawrence et al. 2020, Deichmann et al., 2017; Fuller et al., 2015, Eldridge et al., 2018; Machado et al., 2017), selecting time periods that coincide with the peak communication time for certain groups.

The second form of signal masking, frequency masking, can occur when acoustic indices are measured at frequency bins that are too broad, so that sounds at sensitive frequencies are swamped by contrasting or null responses at other frequencies. Although the importance of frequency masking has not been explicitly considered in relation to acoustic index functioning, there is strong *a priori* reason to believe it may be important, and has been postulated by others (Eldridge et al., 2018). There is a broad negative relationship between body size and the frequency at which animals vocalize (Gillooly and Ophir, 2010; Ryan and Brenowitz, 1985; Seddon, 2005; Wilkins et al., 2013), meaning that the largest species, predominantly mammals, vocalize at the lowest frequencies, whilst orders composed of smaller species such as orthopterans predominate at higher frequencies. In addition, neotropical bird vocalisations exhibit both temporal and frequency partitioning to avoid signal masking from cicadas and other loud insects (Aide et al., 2017, Hart et al., 2015). At its simplest, this should result in different frequency bins being dominated by sounds from different broad taxonomic groupings.

Here, we use a data set from one of the world’s most speciose ecosystems - the Brazilian Amazon – to explore how the use of time and frequency bins (henceforth TFBs) can improve the sensitivity and fidelity of acoustic indices. By calculating acoustic index values within restricted frequency bands, the potential masking effect could be reduced, and correlations with specific taxonomic groups increased. Initially, we establish whether measuring indices at broad time and frequency scales, as is standard practice, masks variation in acoustic responses across narrower TFBs. Next, we look at the impact of signal masking on the efficacy of acoustic indices as a proxy for biodiversity and test the suitability of using TFBs as a solution by asking two questions of high relevance to practitioners and policy makers. First, do TFBs improve the sensitivity of acoustic indices to changes in forest condition (e.g. disturbance)? This is key to monitoring forest recovery following disturbances such as selective logging or wildfire, analyses which underpin many applied ecology questions. Second, do TFBs improve the fidelity of acoustic indices as proxies for traditional field surveys aimed at establishing species richness and composition? These field surveys can be expensive and inefficient for a range of taxonomic groups (Gardner et al., 2008), and if acoustic indices can be shown to be a reliable replacement for traditional survey methods, such as point count bird surveys, then they may offer a significant cost-saving.

Methods

Study area and data collection:

We collected acoustic data in the eastern Brazilian Amazon in the municipalities of Santarém-Belterra-Mojuí dos Campos (latitude -3.046, longitude -54.947, hereafter Santarém) in Pará state, between 12 June 2018 and 16 August 2018. We used the permanent transects of the Sustainable Amazon Network (Gardner et al. 2013) distributed in *terra firme* forest habitats. We sampled 28 300-m transects distributed into five forest classes: undisturbed primary forests (n = 4), logged primary forests (n = 4), burned primary forests (n= 5), logged-and-burned primary forests (n = 12), and secondary forests (n = 3); forests recovering after being completely felled). We installed Frontier Labs Bioacoustic Recording Units with a 16 bit 44.1 kHz sampling rate at points halfway along each transect. Recorders were placed in trees at a height of 7-10 m, with the microphone placed in a downward facing position, at a distance of 10-20 m from the transect to reduce the chance of recorder theft. Recording units were placed away from immediately overhanging dense vegetation to avoid sound being blocked and to limit geophony from leaves and branches. The microphones used have 80 dB signal to noise ratio and 14dBA self-noise, a fixed gain pre-amp of 20dB, a flat frequency response (±2dB) from 80Hz to 20kHz and an 80Hz high-pass filter to filter out low-frequency wind noise (Frontier Labs, 2015). All files were recorded in wav format. Recordings were made continuously (Frontier Labs software writes a new file every ~6 hrs) over multiple discrete time periods of differing length at each point with discrete time periods ranged in duration between 3 and 20 days. Total recording duration and first and last recording dates are included in SOM 1 Appendix 1. The inaccessibility of some transects used in previous studies meant that a balanced survey design was impossible across the disturbance categories (Table 1).

Table 1: Audio sampling by forest class after automated removal of recordings containing heavy rainfall

Forest Class	Sampled points (n)	Total sampling time (minutes)
Primary	4	90,600
Logged primary	4	89,540
Burned primary	5	139,720
Logged-and-burned primary	12	238,130
Secondary	3	60,970

<u>Data Analysis:</u>

We selected two acoustic indices, the Acoustic Complexity Index and the Biodiversity Index as they are two of the commonest indices used in ecoacoustic studies. However the Acoustic Complexity Index is commonly applied across broad frequency ranges, and the Bioacoustic Index is typically applied at restricted frequency ranges, making an ideal comparison for this study as in combination, they are likely to be representative of how many acoustic indices will be affected by the use of narrower time and frequency bins. The Acoustic Complexity Index is intended to quantify biotic sound whilst being robust to non-target noise (Duarte et al., 2015; Fairbrass et al., 2017; Pieretti et al., 2011), and is commonly applied across broad frequency bins. Acoustic Complexity Index measures the irregularity in amplitude across time samples by frequency bin, relative to the total amplitude of the frequency bin. The Acoustic Complexity Index has been found to significantly correlate with species richness for some taxa (Bertucci et al., 2016, Bradfer-Lawrence et al. 2020, Eldridge et al., 2018, Mitchell et al., 2020), whilst in others it showed little or no correlation (Fuller et al., 2015; Mammides et al., 2017; Moreno-Gomez et al., 2019) although this may be due to limitations in methodology and small sample sizes. In contrast, the Bioacoustic Index (BI) is generally applied to narrower frequency bins, and is intended to provide relative abundance of avian community within a frequency range that contains most bird sound (Boelman et al., 2007). It measures the disparity between the quietest and loudest 1 kHz frequency bins. Again, the Bioacoustic Index has been found to be a good predictor of diversity in some studies (Eldridge et al., 2018, Hilje et al., 2017, Gasc et al., 2017, Mitchell et al., 2020) whilst others have found it to be poor (Fuller et al.,2015, Moreno-Gomez et al., 2019), although concerns about the limitations of the methodologies used in these studies apply here too. We expect both indices to increase with increasing species richness and species abundance, and for correlations between both abundance and richness with the indices to be strongest in the frequency and time bins that are most dominated by the target taxa (Table 2), particularly diurnal bird species at dawn between 0.3-12 kHz and nocturnal taxa at night between 0.3-4 kHz.

We calculated the indices using the soundecology package (Villanueva-Rivera et al., 2011, v1.3.3) in R (R Core Team, 2019) which includes minimum and maximum frequency limits for both the Acoustic

Complexity Index and Bioacoustic Index, allowing easy and consistent index calculation at a range of frequency bins. To limit microphone self-noise the lowest frequency included in analysis was 300 Hz. We then calculated the mean index value per 10 minute interval of data collected for each acoustic index and each of the 20 TFBs (Fig.1a), having first screened out recording periods containing heavy rainfall (n = 527) using the hardRain package in R Studio (Metcalf et al., 2020, v0.1.1).

We selected TFBs with the objective of capturing periods of time and frequency bands that are most taxonomically homogenous. TFBs were not quantitatively optimised, but rather subjective approximations that aimed to effectively capture broad taxonomic groupings in tropical forest landscapes. Temporal limits were determined by patterns in animal communication in the diel cycle, to encapsulate dawn, dusk, daytime and night-time periods (Pieretti et al., 2015; Rodriguez et al., 2014); commonly used sampling periods in acoustic recording (Sugai et al., 2019). 'Dawn' was assigned to the period from 30 minutes prior to sunrise and for the following 2 hours, whilst 'Dusk' was the 2 hour period ending at 30 minutes after sunset. 'Day' and 'Night' are the respective intervening periods. Frequency limits were determined by a review of the literature and our own experience of manually analysing 100s of hours of acoustic data from the region. The taxonomic groupings we hypothesize dominate each TFB are illustrated in Figure 1a. Table 2 contains some of the TFBs likely to contain particularly high activity from particularly homogenous groupings. 0.3-22.1 kHz frequency bin (hereafter 'baseline') was used as baseline data, representative of how most terrestrial acoustic indices are currently calculated across the spectrum of human hearing or the common sampling rate of 44.1 kHz. However, it is worth noting that the Bioacoustic Index is commonly calculated with a narrower frequency bin than the baseline, typically from 2 to 8 or 11 Khz (Boelman et al., 2007, Bradfer-Lawrence et al., 2019, Villanueva-Rivera et al., 2011). We have used the same baseline as the Acoustic Complexity Index for ease of comparison, and because the mechanisms causing masking between ecologically relevant and non-relevant frequency bins is the same regardless of absolute frequency. Of course, macro frequency bands will never solely encompass single taxonomic groups, and boundaries will always be somewhat arbitrary due to variations in acoustic communication at species, temporal and even individual levels.

Table 2: Selected time-frequency bins and the taxonomic groups expected to dominate each sample.

Frequency Band (kHz)	Time Period	Taxonomic group	References
0.3-4	Night	Terrestrial/arboreal mammals, anuran and birds	Chek, Bogart, & Lougheed, 2003; Lima, Pederassi, Pineschi, & Barbosa, 2019
4-12	Day	Hemiptera/Orthoptera	Hart et al., 2015, Schmidt et al., 2013
4-12	Night	Hemiptera/Orthoptera	Hart et al., Schmidt et al., 2013
0.3-12	Dawn	Diurnal/crepuscular birds	Tobias, Planqué, Cram, & Seddon, 2014
12-22.1	Dusk	Insects, bats and frogs	Lima, Pederassi, Pineschi, & Barbosa, 2019, Schmidt et al., 2013
12-22.1	Night	Insects	Schmidt et al., 2013

Signal Masking

To investigate whether the soundscape responds differently to human-driven disturbance across time and frequency, we looked at the variation in response of each disturbance class for each TFB. Having removed periods with extreme outlying index values, we took a random sample of acoustic index values for each acoustic index and TFB from each forest class (n=500), giving a total sample size of n=2500 per TFB/index. For each TFB and acoustic index we conducted a Kruskal-Wallis (Kruskal and Wallis, 1952) test between the five forest classes, and calculated the effect size (ε^2). When significant differences between the classes were found, we used a Dunn's test (Dunn, 1964) to establish how many of the ten forest class pairs were significantly different from each other.

The sensitivity of acoustic indices to habitat

To assess whether the use of TFBs increased acoustic index sensitivity to forest classes, we built distributed random forest models from the 'h2o' R package (LeDell et al., 2020 v3.30.0.1), varying the number of TFBs used as predictors. Firstly, we tested if the use of TFBs improved classification accuracy between the two most ecologically distinct sampled habitats; undisturbed primary forest and secondary forest (Moura et al., 2013). To do so, we built two binomial random forest models, the first using training data only from the baseline frequency bin across all time periods, the second using training data from all frequency bins and time periods. Next, models were trained and tested on data from all five forest classes, which previous studies (e.g. Moura et al. 2013) suggest would provide a more challenging classification problem.

We used each combination of index and TFB as a separate predictor. The training datasets required subsampling to obtain predictors of equal length, as not all time periods were of the same duration, and forest classes had unequal survey effort. We used the same subsample as above (see Signal Masking), so that each TFB predictor had n=2,500 samples, with 500 samples from each forest class.

This resulted in a greatly reduced dataset for training the models with 100,000 acoustic indices values compared to 1,277,560 in the original dataset. Prior to model training, the dataset was split with 75% of observations used for training and 25% as a test dataset. Model parameters were kept constant across all models (SOM 1 Appendix 2). We used balanced accuracy (Fielding and Bell, 1997), F1 scores (Chinchor, 1993) and Matthew's Correlation Coefficient (Guilford., 1954) as accuracy metrics (Table 3), which were calculated per forest class based on predictions of the test dataset and are presented here as an unweighted mean across all forest classes included in the respective model.

Fidelity of acoustic indices to taxonomic measures of biodiversity

We assessed correlations between acoustic index scores and biodiversity indicator metrics, to see how representative the indices were of commonly used indicators of diversity. Data on the presence/absence of three sets of species were generated from two subsets of the audio data. Each audio subset consisted of 28 hrs of sound recordings, in the form of 240 15 s recordings from each point. The first data set was restricted to the dawn period (hereafter dawn birds), in which all identifiable avian vocalisations were assigned to species by an ornithologist (Nárgila Gomes De Moura) with extensive field experience of point counts in the same sites (e.g. Moura et al. 2013). This method of species detection is likely to produce comparable results to traditional point count surveys as several papers have shown that experienced observers reviewing recordings and spectrograms can be more or equally effective at detecting species than field-based surveys (Darras et al., 2019; Shonfield et al., 2018). The second set of data was restricted to the nocturnal period (hereafter nocturnal birds), and again all identifiable avian vocalisations were assigned to species by an experienced ornithologist (OCM). The third set was generated from the nocturnal data subset again (hereafter nocturnal taxa), but comprises all biophony below 4kHz, identified (by OCM) where possible or sonotyped if not. It is worth noting that all of the bird species identified at night vocalized below 4 kHz, so that the nocturnal bird set is wholly a subset of the nocturnal taxa set.

For each of these matrices (i.e. dawn birds, nocturnal birds and nocturnal taxa), five metrics were calculated; total number of encounters (the sum of the number of 15 s recordings each species was present in), species richness, Shannon diversity, Pielou's evenness, and the first axis from a nonmetric multidimensional scaling ordination (hereafter MDS1) using the Jaccard method from the vegan package (Oksanen et al., 2019). Total encounters was included as a proxy for the abundance of sounds, to test if indices responded more strongly to more sources of noise, regardless of composition. Estimated species richness from the dawn matrix was calculated for each point at 98.5% coverage based on rarefaction/extrapolation using the iNEXT package (Hsieh et al., 2020, v2.0.20), as some of the survey files were removed as they contained periods of heavy rain which affected the number of vocalizing species. Observed species richness was used for metrics from the nocturnal matrix, as the data were pre-screened for rain. Shannon diversity, Pielou's evenness and species richness were included as standard measures of ecological diversity (Oksanen et al., 2019). MDS1 was included to reflect turnover mediated by disturbance, as high values correspond with less disturbed habitats, whilst lower values have communities associated with more disturbed habitats. Correlations between these metrics are available in SOM 1 Appendix 3. Median values of each acoustic index were calculated for each point and TFB, and Spearman's rank order correlations ($\alpha = 0.05$) were calculated between these and the biodiversity metrics. Significant differences between each correlation and the respective baseline correlation were calculated using Zou's confidence interval test (Zou, 2007) in the 'cocor' package (Diedenhofen and Musch, 2015 v.1.1-3).

Results

Sensitivity: forest disturbance

The Kruskal-Wallis and Dunn's test revealed strong evidence that acoustic masking affects the sensitivity of acoustic indices, both temporally and by frequency. All of the Kruskal-Wallis tests were significant ($p<0.05$), showing that acoustic indices are sensitive to at least some disturbance events regardless of frequency band or time period. There were significant differences between all ten forest class pairs in every time period and with both indices when considering all frequency-restricted TFBs together. In contrast, there were no time periods with significant differences between all forest class pairs when using only the baseline TFBs, but Acoustic Complexity did have significant differences between nine forest class pairs in three time periods, and Bioacoustic Index once. Twelve TFBs showed significant difference ($p<0.05$) between more forest classes than the corresponding baseline, and 21 TFBs had higher effect sizes than the corresponding baseline, suggesting that in many cases stronger responses to disturbance events at narrower frequency bins are masked by the use of broad frequency bins (Fig. 2). No baseline TFB achieved perfect separation between all ten forest-class pairs but this was achieved by three of the non-baseline TFBs. Furthermore, Acoustic Complexity Index at dusk and the baseline frequency bin produced the lowest number of significantly different forest class pairs, just two, suggesting that using only the broadest frequency bin can result in relatively poor differentiation between forest disturbance classes. No one frequency bin or time period had a consistently larger effect size, or consistently differentiated between more forest classes. There were several occasions in which effect size increased in comparison to the baseline, whilst the number of different forest classes decreased (e.g. Acoustic Complexity Index at dawn, 0.3-4 kHz, Bioacoustic Index at night, 0.3-4 kHz). This suggests that the soundscape at this frequency bin is showing a particularly strong response to disturbance in one or more of the forest classes (in SOM 1 Appendix 4)."

Sensitivity: Forest class differentiation

The random forest models generated using all of the TFBs as predictors were able to classify forest classes with a high degree of accuracy, with 99.6% balanced accuracy between secondary and undisturbed forest and 88.2% between the five forest classes (Fig. 3). The models using all TFBs as predictors outperformed the corresponding baseline models in both tests, but as expected the baseline models performed particularly poorly when classifying between all five forest classes, achieving just 62.1%. The confusion matrix for the random forest model using all TFBs across all five classes suggest that acoustic indices do respond to soundscapes in ecologically meaningful ways, as both burned forest classes had comparatively high error between them, as did the two most disturbed classes, logged and burned vs secondary forest (SOM 1 Appendix 5).

Fidelity: Biodiversity Correlations

Correlations with traditional biodiversity metrics revealed complex patterns, underpinned by strong variation across index, frequency bins and time periods. For simplicity, we have focussed on time periods in which acoustic index values most directly reflect variation in manually reviewed datasets - dawn and day time for the dawn bird dataset, and night for the nocturnal datasets (Fig. 4), Correlations at other time periods are presented in SOM 2 Appendix 6. The two strongest correlations, were at night between Bioacoustic Index and diurnal avian MDS1 at 4-12 kHz (r_s =0.74) and dusk at 0.3-12 kHz (r_s = 0.72).

The correlation scores provide strong evidence that the use of TFBs increase the fidelity of correlations. We found that correlation directions of the frequency/ bins differed from the corresponding baseline frequency bin for at least one of the metrics in every time period in both indices. The strongest examples of this were between Bioacoustic Index at dawn with dawn birds, which saw predominantly significant positive correlations at frequency bands at which dawn birds vocalize, but negative correlations in the baseline and 12-22.1 kHz frequency bins. We also found 28 instances in which correlations were significantly different to the corresponding baseline frequency bin.

Overall, the Acoustic Complexity Index was inconsistently correlated with biodiversity indicator metrics with predominantly negative and significant correlations with diversity metrics at dawn, but mostly positive correlations during the day and at night for most frequency bands, with far fewer significant correlations. In contrast, the Bioacoustic Index showed predominantly positive correlations with most diversity metrics except MDS1 across all three time periods and all three frequency bins only including sound below 12 kHz. For the community metrics most likely to be useful to ecologists, species richness and Shannon diversity, there were 19 significant correlations. However, where we found significant correlations with these metrics, there were still strong reasons for doubting the fidelity of acoustic indices as proxies. At dawn, correlations for both indices were weaker than the respective correlations with total encounters or MDS1, suggesting the indices were more sensitive to the number of individual sounds or the overall community. At night, nocturnal birds and taxa correlations with the Bioacoustic Index showed conflicting patterns with correlations at frequency bins more likely to be relevant to the relevant taxonomic group, suggesting a strong masking effect by vocalisations of non-target taxa. The exception to this is the correlation between dawn birds and BI during the day at 0.3-4 kHz, which shows strong correlations with species richness and Shannon diversity (r_s = 0.44 and 0.40), a similar correlation with total encounters (r_s = 0.40) and no significant correlation with MDS1.

Discussion

We found that calculating acoustic indices at narrower TFBs results in large increases in the sensitivity of acoustic indices to the soundscape response of different forest classes. Calculating acoustic indices across a single broad frequency bin, as is commonplace in the ecoacoustic literature (Sueur et al., 2014; Buxton et al., 2018; Gibb et al., 2019) can mask varied responses across time periods and frequency bins, reducing the sensitivity of acoustic indices. Furthermore, when acoustic indices are used as proxies for biodiversity indicator metrics, masking can have a serious impact on the fidelity of the correlations. Correlating broad frequency bins with biodiversity metrics generated from taxa whose vocalisations do not occur across the entire frequency range is likely to be highly misleading. It not only misrepresents the magnitude of correlations, but potentially results in spurious inverse correlations caused by the responses of acoustically dominant species or patterns from acoustic space that are not biologically relevant.

The sensitivity gains of this new methodology are particularly apparent when using acoustic indices to differentiate between environments with classification accuracy greatly increased. Accuracy scores are as good or better than many in the literature (Bormpoudakis et al., 2013; Bradfer-Lawrence et al., 2019; Eldridge et al., 2018, Do Nascimento et al., 2020), especially considering that model hyperparameters are not optimised, fewer acoustic indices were used, and that the forest classes considered here are all of the same land cover (tropical forest) and within the same landscape. In addition, these results have been achieved whilst using only approximately one sixth of the training data compared to the models using only a single baseline frequency band, suggesting that by using TFBs, large efficiency savings can be made in terms of data collection. We therefore recommend that acoustic indices are calculated across a range of frequency bins and temporal periods in any study using acoustic index values to characterise and identify land use.

The impact of signal masking, and the benefits of using narrower time-frequency bins to avoid it, are equally apparent when correlating acoustic indices with biodiversity metrics. However, despite the increased fidelity of the correlations, the use of single acoustic indices as direct proxies for biodiversity indicator metrics is still problematic. Whilst we found the predicted positive correlations between diversity, richness and the Bioacoustic Index at the most relevant TFBs, we found a negative correlation between the Acoustic Complexity Index at dawn with bird species richness and abundance. This contrasts with other similar studies in comparable habitats that found positive relationships (Bradfer-Lawrence et al., 2020, Eldridge et al., 2018, Mitchell et al., 2020). However, Mitchell et al., (2020) found high Acoustic Complexity values in oil-palm plantations where diversity was low, and noted that the significant relationships they found were within habitat types, but not across different habitats. Furthermore, the complex mechanisms determining abundance and species richness in tropical forests remain poorly understood, particularly in relation to the impacts of disturbance (Barlow et al., 2016, Terborgh et al., 1990). It is possible that idiosyncratic responses of single or a few taxa to disturbance could create such a negative correlation (Moura et al., 2016), especially if the taxa are acoustically dominant. In general, the strongest correlations we found were with total encounters and MDS1 – metrics that would only be of ecological interest if the underpinning species were well understood, requiring extensive manual surveys and undermining the purpose of acoustic indices. Despite this, the strong positive correlation between Bioacoustic Index and dawn bird species richness and diversity in the day at 0.3-4 kHz is interesting. It is plausible that this time-frequency bin contains the least vocalization from non-target taxa, insects sonify predominantly around >4 kHz and after the dawn chorusing of acoustically dominant vertebrates, particularly red-handed howler monkey *Alouatta belzebul* (Sekulic, 1982). Furthermore, it is after the end of the bird dawn chorus, during which it is possible that intense vocal activity of a few species may mask underlying richness and diversity.

We have deliberately chosen to use subjective frequency bins determined by *a priori* knowledge of acoustic space use in our study system, to demonstrate both the wide applicability of this method, and that frequency bin selection need not be onerous to generate substantial benefits. However, choosing narrower or different frequency bins and time periods based on prior quantification of acoustic space use could provide substantial further benefits in understanding the effects of signal masking on correlations. Several existing methods exist to do so, either comprehensively through the multiscalar fractal approach (Monacchi and Farina, 2019), or more broadly using measures of acoustic space use or biophonic density (Aide et al., 2017,Eldridge et al., 2018) Quantifying TFB dominance by even broad acoustic clades could be highly informative and could provide quantifiable data on the relative effect size of the impact of disturbance types on those clades. Additionally, variation in the granularity of TFBs may well reveal further unknown ecological patterns. Whilst we have focused primarily on masking in the frequency domain, and across the diel cycle in the temporal domain, it is entirely plausible that analysis of acoustic indices at both greater and finer temporal scales, and broad frequency ranges, could reveal other patterns. For instance, within dawn choruses where we already know bird species can hold very specific temporal niches in the tropics (Fjeldså et al, 2020), or across seasonal scales such as the winter midday chorus in temperate forests (Farina and Gage, 2017)."

We found that acoustic indices are sensitive to soundscapes modified by habitat disturbance and can therefore be highly cost-effective tools for assessing forest condition and monitoring changes in conservation value in response to management interventions or other environmental changes. Acoustic indices are however highly susceptible to signal masking, where divergent responses across temporal and frequency spectrums are masked by calculating indices at inappropriate scales. We therefore recommend that acoustic indices are calculated either at a range of time and frequency bins when used to characterise a landscape, or a narrow bin predetermined by *a priori* ecological understanding of the soundscape when used as a proxy for the biodiversity of a specific taxonomic group.

Acknowledgements:

We thank the Large Scale Biosphere-Atmosphere Program (LBA) for logistical and infrastructure support during field measurements. We are very grateful for the hard work of Nárgila Gomes De Moura who undertook the identification of birds at dawn, as well as our field and laboratory assistants: Marcos Oliveira, Gilson Oliveira, Renílson Freitas, Josué Jesus de Oliveira, and Amanda Cardoso. We are also grateful to Joice Ferreira, Liana Rossi and Filipe. França for logistical field support in Brazil, as well as advice on data analysis and comments on the manuscript. Fieldwork in Brazil was supported by research grants ECOFOR (NE/K016431/1), and AFIRE (NE/P004512/1), PELD-RAS (CNPq/CAPES/PELD 441659/2016-0). We would like to thank both anonymous reviewers for their time and comments which substantially improved this manuscript. The authors have no conflicts of interest to declare.

Author’s contributions:

OCM designed the study, undertook data collection, analysed the data and drafted the article. ACL, JB and SM made significant contributions to the study conception and design, aided in data analysis and made substantial revisions to manuscript drafts. CD contributed significantly to the data analysis and made substantial revisions to manuscript drafts. EB contributed greatly to the study design, collection of the data, and made substantial revisions to manuscript drafts.

Data Availability:

The following datasets are available on the Dryad respository:

1. All of the indices values used in this research
2. The subset of indices values used for the sensitivity analysis
3. The full set of results from the sensitivity analysis
4. All of the diversity metrics used in the correlation analysis

Available at DOI:10.5061/dryad.dr7sqv9wt.

References:

Berenguer, E., Ferreira, J., Gardner, T.A., Aragão, L.E.O.C., De Camargo, P.B., Cerri, C.E.,…Barlow, J., 2014. A large-scale field assessment of carbon stocks in human-modified tropical forests. Glob. Chang. Biol. 20, 3713–3726. https://doi.org/10.1111/gcb.12627

Boelman, N.T., Asner, G.P., Hart, P.J., Martin, R.E., 2007. Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. Ecol. Appl. 17, 2137–2144. https://doi.org/10.1890/07-0004.1

Bormpoudakis, D., Sueur, J., Pantis, J.D., 2013. Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications. Landsc. Ecol. 28, 495–506. https://doi.org/10.1007/s10980-013-9849-1

Bradfer-Lawrence, T., Bunnefeld, N., Gardner, N., Willis, S.G., Dent, D.H., 2020. Rapid assessment of avian species richness and abundance using acoustic indices. Ecol. Indic. 115, 106400. https://doi.org/10.1016/j.ecolind.2020.106400

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S.G., Dent, D.H., 2019. Guidelines for the use of acoustic indices in environmental research. Methods Ecol. Evol. https://doi.org/10.1111/2041-210x.13254

Burivalova, Z., Game, E.T., Butler, R.A., 2019. The sound of a tropical forest. Science (80). 363, 28–29. https://doi.org/10.1126/science.aav1902

Burivalova, Z., Towsey, M., Boucher, T., Truskinger, A., Apelis, C., Roe, P., Game, E.T., 2018. Using soundscapes to detect variable degrees of human influence on tropical forests in Papua New Guinea. Conserv. Biol. 32, 205–215. https://doi.org/10.1111/cobi.12968

Buxton, R., McKenna, M.F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L.M.,..., Wittemyer, G., 2018. Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. Conserv. Biol. https://doi.org/10.1111/cobi.13119

Carruthers-Jones, J., Eldridge, A., Guyot, P., Hassall, C., Holmes, G., 2019. The call of the wild: Investigating the potential for ecoacoustic methods in mapping wilderness areas. Sci. Total Environ. 695. https://doi.org/10.1016/j.scitotenv.2019.133797

Chek, A.A., Bogart, J.P., Lougheed, S.C., 2003. Mating signal partitioning in multi-species assemblages: a null model test using frogs. Ecol. Lett. 6, 235–247. https://doi.org/10.1046/j.1461-0248.2003.00420.x

Chinchor N., 1992. MUC-4 evaluation metrics. In: Proceedings of MUC-4 – the 4th Conference on Message Understanding. McLean: Association for Computational Linguistics: p. 22–9.

Deichmann, J.L., Acevedo-Charry, O., Barclay, L., Burivalova, Z., Campos-Cerqueira, M., d'Horta, F.E., Mitchell Aide, T., 2018. It's time to listen: there is much to be learned from the sounds of tropical ecosystems. Biotropica 50, 713–718. https://doi.org/10.1111/btp.12593

Deichmann, J.L., Hernández-Serna, A., Delgado C., J.A., Campos-Cerqueira, M., Aide, T.M., 2017. Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest. Ecol. Indic. 74, 39–48. https://doi.org/10.1016/j.ecolind.2016.11.002

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., Sueur, J., 2012. Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. Ecol. Indic. 13, 46–54. https://doi.org/10.1016/j.ecolind.2011.05.006

Diedenhofen, B., Musch, J., 2015. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. PLoS One 10, e0121945. https://doi.org/10.1371/journal.pone.0121945

Do Nascimento, L.A., Campos-Cerqueira, M., Beard, K.H., 2020. Acoustic metrics predict habitat type and vegetation structure in the Amazon. Ecol. Indic. 117, 106679. https://doi.org/10.1016/j.ecolind.2020.106679

Dunn, O.J., 1964. Multiple Comparisons Using Rank Sums. Technometrics 6, 241–252. https://doi.org/10.1080/00401706.1964.10490181

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., Peck, M., 2018. Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. Ecol. Indic. 95, 939–952. https://doi.org/10.1016/j.ecolind.2018.06.012

Farina, A., 2014. Soundscape ecology: Principles, patterns, methods and applications, Soundscape Ecology: Principles, Patterns, Methods and Applications. https://doi.org/10.1007/978-94-007-7374-5

Farina, A., Ceraulo, M. 2017. The Acoustic Chorus and its Ecological Significance. In Ecoacoustics: The ecological role of sounds (Farina, A., Gage, S. Editors). Wiley, DOI:10.1002/9781119230724

Fjeldså, J., Marks, J. S., and Sharpe, C.J. (2020). Cryptic Forest-Falcon (*Micrastur mintoni*), version 1.0. In Birds of the World (del Hoyo, J. , Elliott, A., Sargatal, J., Christie, D. A., and de Juana, E., Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. https://doi.org/10.2173/bow.cryfof1.01

Frontier Labs, 2015. Bioacoustic Audio Recorder User Guide.

Fuller, S., Axel, A.C., Tucker, D., Gage, S.H., 2015. Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? Ecol. Indic. 58. https://doi.org/10.1016/j.ecolind.2015.05.057

Gardner, T.A., Barlow, J., Araujo, I.S., Cristina Vila-Pires, T., Bonaldo, A.B., Costa, J.E.,…Peres, C.A., 2008. The cost-effectiveness of biodiversity surveys in tropical forests. Ecol. Lett. 11, 139–150. https://doi.org/10.1111/j.1461-0248.2007.01133.x

Gardner, T.A., Ferreira, J., Barlow, J., Lees, A.C., Parry, L., Vieira, I.C.,…Zuanon, J., 2013. A social and ecological assessment of tropical land uses at multiple scales: the Sustainable Amazon Network. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368, 20120166. https://doi.org/10.1098/rstb.2012.0166

Gibb, R., Browning, E., Glover-Kapfer, P., Jones, K.E., 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods Ecol. Evol. https://doi.org/10.1111/2041-210X.13101

Gillooly, J.F., Ophir, A.G., 2010. The energetic basis of acoustic communication. Proc. R. Soc. B Biol. Sci. 277, 1325–1331. https://doi.org/10.1098/rspb.2009.2134

Guilford JP. Psychometric Methods. New York City: McGraw-Hill; 1954.

Hsieh, T. C., Ma K. H., Chao.A., 2020 iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.20 URL: http://chao.stat.nthu.edu.tw/wordpress/software-download/.

Jorge, F.C., Machado, C.G., da Cunha Nogueira, S.S., Nogueira-Filho, S.L.G., 2018. The effectiveness of acoustic indices for forest monitoring in Atlantic rainforest fragments. Ecol. Indic. https://doi.org/10.1016/j.ecolind.2018.04.001

Kruskal, W.H., Wallis, W.A., 1952. Use of Ranks in One-Criterion Variance Analysis. J. Am. Stat. Assoc. 47, 583–621. https://doi.org/10.1080/01621459.1952.10483441

LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T.,…M., Malohlava, M., (2020). h2o: R Interface for the 'H2O' Scalable Machine Learning Platform. R package version 3.28.0.4. https://CRAN.R-project.org/package=h2o

Lima, M.S.C.S., Pederassi, J., Pineschi, R.B., Barbosa, D.B.S., 2019. Acoustic niche partitioning in an anuran community from the municipality of Floriano, Piauí, Brazil. Brazilian J. Biol. 79, 566–576. https://doi.org/10.1590/1519-6984.180399

Mammides, C., Goodale, E., Dayananda, S.K., Kang, L., Chen, J., 2017. Do acoustic indices correlate with bird diversity? Insights from two biodiverse regions in Yunnan Province, south China. Ecol. Indic. 82, 470–477. https://doi.org/10.1016/j.ecolind.2017.07.017

Metcalf, O.C., Lees, A.C., Barlow, J., Marsden, S.J., Devenish, C., 2020. hardRain: An R package for quick, automated rainfall detection in ecoacoustic datasets using a threshold-based approach. Ecol. Indic. 109, 105793. https://doi.org/10.1016/j.ecolind.2019.105793

Metcalf, Oliver et al. (2020), Acoustic indices perform better when applied at ecologically meaningful time and frequency scales, Dryad, Dataset, https://doi.org/10.5061/dryad.dr7sqv9wt

Mitchell, S.L., Edwards, D.P., Bernard, H., Coomes, D., Jucker, T., Davies, Z.G., Struebig, M.J., 2018. Riparian reserves help protect forest bird communities in oil palm dominated landscapes. J. Appl. Ecol. 55, 2744–2755. https://doi.org/10.1111/1365-2664.13233

Monacchi, D., Farina, A. A Multiscale Approach to Investigate the Biosemiotic Complexity of Two Acoustic Communities in Primary Forests with High Ecosystem Integrity Recorded with 3D Sound Technologies. Biosemiotics 12, 329–347 (2019). https://doi.org/10.1007/s12304-019-09361-z

Moreno-Gómez, F.N., Bartheld, J., Silva-Escobar, A.A., Briones, R., Márquez, R., Penna, M., 2019. Evaluating acoustic indices in the Valdivian rainforest, a biodiversity hotspot in South America. Ecol. Indic. 103. https://doi.org/10.1016/j.ecolind.2019.03.024

Moura, N.G., Lees, A.C., Andretti, C.B., Davis, B.J.W., Solar, R.R.C., Aleixo, A.,…Gardner, T.A., 2013. Avian biodiversity in multiple-use landscapes of the Brazilian Amazon. Biol. Conserv. 167, 339–348. https://doi.org/10.1016/j.biocon.2013.08.023

Moura, N.G., Lees, A.C., Aleixo, A., Barlow, J., Berenguer, E., Ferreira, J., Mac Nally, R., Thomson, J.R., Gardner, T.A., 2016. Idiosyncratic responses of Amazonian birds to primary forest disturbance. Oecologia 180, 903–916. https://doi.org/10.1007/s00442-015-3495-z

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, Dan., Minchin, P.R., …Wagner, H., (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

Pieretti, N., Duarte, M.H.L., Sousa-Lima, R.S., Rodrigues, M., Young, R.J., Farina, A., 2015. Determining temporal sampling schemes for passive acoustic studies in different tropical ecosystems. Trop. Conserv. Sci. 8. https://doi.org/10.1177/194008291500800117

Pieretti, N., Farina, A., Morri, D., 2011. A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). Ecol. Indic. 11, 868–873. https://doi.org/10.1016/j.ecolind.2010.11.005

Priyadarshani, N., Marsland, S., Castro, I., 2018. Automated birdsong recognition in complex acoustic environments: a review. J. Avian Biol. https://doi.org/10.1111/jav.01447

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rodriguez, A., Gasc, A., Pavoine, S., Grandcolas, P., Gaucher, P., Sueur, J., 2014. Temporal and spatial variability of animal sound within a neotropical forest. Ecol. Inform. 21, 133–143. https://doi.org/10.1016/j.ecoinf.2013.12.006

Ryan, M.J., Brenowitz, E.A., 1985. The role of body size, phylogeny, and ambient noise in the evolution of bird song. Am. Nat. 126, 87–100. https://doi.org/10.1086/284398

Schmidt, A.K.D., Römer, H., Riede, K., 2013. Spectral niche segregation and community organization in a tropical cricket assemblage. Behav. Ecol. 24, 470–480. https://doi.org/10.1093/beheco/ars187

Seddon, N., 2005. Ecological adaptation and species recognition drives vocal evolution in Neotropical suboscine birds. Evolution (N. Y). 59, 200–215. https://doi.org/10.1111/j.0014-3820.2005.tb00906.x

Sedláček, O., Vokurková, J., Ferenc, M., Djomo, E.N., Albrecht, T., Hořák, D., 2015. A comparison of point counts with a new acoustic sampling method: a case study of a bird community from the montane forests of Mount Cameroon. Ostrich 6525. https://doi.org/10.2989/00306525.2015.1049669

Sekulic', R., 1982. The Function of Howling in Red Howler Monkeys (Alouatta seniculus). Behaviour,81(1), 38-54.

Sueur, J., Farina, A., Gasc, A., Pieretti, N., Pavoine, S., 2014. Acoustic indices for biodiversity assessment and landscape investigation. Acta Acust. united with Acust. 100. https://doi.org/10.3813/AAA.918757

Sueur, J., Pavoine, S., Hamerlynck, O., Duvail, S., 2008. Rapid acoustic survey for biodiversity appraisal. PLoS One 3. https://doi.org/10.1371/journal.pone.0004065

Sugai, L.S.M., Silva, T.S.F., Ribeiro, J.W., Llusia, D., 2019. Terrestrial Passive Acoustic Monitoring: Review and Perspectives. Bioscience. https://doi.org/10.1093/biosci/biy147

Tobias, J.A., Planqué, R., Cram, D.L., Seddon, A.N., 2014. Species interactions and the structure of complex communication networks. Proc. Natl. Acad. Sci. U. S. A. 111, 1020–1025. https://doi.org/10.1073/pnas.1314337111

Towsey, M., Wimmer, J., Williamson, I., Roe, P., 2014. The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecol. Inform. 21, 110–119. https://doi.org/10.1016/j.ecoinf.2013.11.007

Villanueva-Rivera, L.J., Pijanowski, B.C., Doucette, J., Pekin, B., 2011. A primer of acoustic analysis for landscape ecologists. Landsc. Ecol. 26, 1233–1246. https://doi.org/10.1007/s10980-011-9636-9

Wagner Ribeiro Jr, J., Larissa Sayuri Moreira Sugai, B., Marconi Campos-Cerqueira, B., by Dirk Sven Schmeller, C., 2017. Passive acoustic monitoring as a complementary strategy to assess biodiversity in the Brazilian Amazonia. Biodivers. Conserv. 26, 2999–3002. https://doi.org/10.1007/s10531-017-1390-0

Wilkins, M.R., Seddon, N., Safran, R.J., 2013. Evolutionary divergence in acoustic signals: Causes and consequences. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2012.10.002

Zou, G.Y., 2007. Toward Using Confidence Intervals to Compare Correlations. Psychol. Methods 12, 399–413. https://doi.org/10.1037/1082-989X.12.4.399