Willis, SG, Dent, DH. Guidelines for the use of acoustic indices in environmental research. *Methods Ecol Evol.* 2019; 10: 1796– 1807, which has been published in final form at https://doi.org/10.1111/2041-210X.13254. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for self-archiving.

Methods in Ecology and Evolution

Handling editor: Dr Veronica Zamora-Gutierrez

Guidelines for the use of acoustic indices in environmental research

Tom Bradfer-Lawrence[1]*, Nick Gardner[2], Lynsey Bunnefeld[1], Nils Bunnefeld[1], Stephen G. Willis[3], Daisy H. Dent[1,2].

1. Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK
2. Smithsonian Tropical Research Institute, Panama City, Republic of Panama
3. Conservation Ecology Group, Department of Biosciences, Durham University, South Road, Durham, DH1 3LE, UK

*email for correspondence: tom.bradfer-lawrence@stir.ac.uk

Running headline: “Guidelines for the use of acoustic indices”

Abstract

1. Ecoacoustics, the study of environmental sound, is a growing field with great potential for biodiversity monitoring. Audio recordings could provide a rapid, cost-effective monitoring tool offering novel insights into ecosystem dynamics. More than 60 acoustic indices have been developed to date, which reflect distinct attributes of the soundscape, (i.e. the total acoustic energy at a given location, including noise produced by animals, machinery, wind and rain). However, reported

patterns in acoustic indices have been contradictory, possibly because there is no accepted best practice for the collection and analysis of audio recordings.

2. Here, we propose: (1) guidelines for designing studies using audio recordings for the rapid assessment of multiple sites, and (2) a workflow for comparing recordings with seven of the most commonly used indices, permitting discrimination among habitat-specific soundscapes. We collected and analysed over 26,000 hours of recordings from 117 sites across a range of habitats in a human-modified tropical landscape in central Panama; an order of magnitude more recordings than used in previously published studies.

3. We demonstrate that: (1) Standard error variance of indices stabilises within 120 hours of recordings from a single location. (2) Continuous recording should be used rather than sub-sample recording on a schedule; sub sampling is a common practice but delays capture of site variability and maximising total duration of recording should be prioritised. (3) Use of multiple indices to describe soundscape patterns reveals distinct diel and seasonal soundscape patterns among habitats.

4. We advocate collecting at least 120 hours of continuous recordings per site, and using a range of acoustic indices to categorise the soundscape, including the Acoustic Complexity Index, Acoustic Evenness Index, Acoustic Entropy Index and the Normalised Difference Soundscape Index. Differences among habitat types can be captured if multiple indices are used, and magnitude of variance is often more important than mean values. The workflow we provide will enable successful use of ecoacoustic techniques for environmental monitoring.

Key-words: Acoustic index, Bioacoustics, Biodiversity, Ecoacoustics, Landscape, Sound recording, Soundscape

Introduction

Ecoacoustics, the study of environmental sound, is a rapidly evolving field (Sueur et al. 2014). Recent developments in automated sound collection and processing offer enormous potential for rapid and cost-effective monitoring of biodiversity, an essential task in the face of global land-use change (Burivalova et al. 2019; Laiolo 2010; Ribeiro et al. 2017). By identifying temporal shifts in soundscapes, this monitoring can be used to assess how species are affected by anthropogenic disturbance (Burivalova et al. 2019). However, the relative novelty of this field and the pace of innovation mean there are currently no accepted standards regarding the quantity of data (i.e. the length of recordings) or sampling intensity necessary for characterising the soundscape of a given habitat. Similarly, guidance on how such data can be used for effective, yet simple, biodiversity monitoring is lacking (Priyadarshani et al. 2018).

Thousands of hours of sound recordings have been collected from a multitude of habitats around the world, but methods for translating these data into a rapid monitoring process are not keeping pace (Gibb et al 2018; Priyadarshani et al 2018). Data on faunal presence can be extracted from audio recordings using either manual or automated methods. However, both manual and automated approaches are time-consuming, necessitate expert knowledge and, in the case of automated recognisers, are still subject to high error rates (Furnas & Callas 2015; Sevilla & Glotin 2017). Rather than focus on individual species, alternative approaches are required that summarise the huge quantities of sound recordings now available. To this end, over 60 indices have been developed to rapidly classify soundscapes based on their acoustic properties, providing metrics for habitat assessment and monitoring (Sueur et al. 2014; Buxton et al. 2018).

The soundscape is comprised of the total acoustic energy at a given location, incorporating biophony (noise produced by animals), anthrophony (noise produced by humans and machines), and geophony (noise from natural processes such as wind and rain) (Pijanowski et al. 2011). Each acoustic index utilises different characteristics of the soundscape, such as pitch, saturation, and amplitude. Often these involve contrasting short time steps or frequency bands within a recording. For example, the widely used Acoustic Complexity Index (ACI) contrasts the amplitude difference between one short time step (e.g. 0.03 secs) and the next, within a narrow frequency band (e.g. 62 Hz). The ACI is sensitive to the inherent irregularity of biophony, particularly from bird song, while it is relatively impervious to persistent sound of a constant intensity. Audio indices such as ACI reduce the enormous complexity of the soundscape to a single number, greatly simplifying extraction of information from recordings.

Acoustic indices are now used in a range of ecological research. Recently, a promising method using false colour spectrograms constructed with acoustic indices has been developed as a means of detecting particular species or taxon choruses (Towsey et al. 2014b, 2018). However, research focus has generally concentrated on investigating overall soundscape patterns. For example, Rodriguez et al. (2014) used acoustic indices to describe clear diel cycles in tropical forest soundscapes, and differences between the canopy and understory strata. Seasonal shifts in soundscapes have been examined in both temperate and tropical habitats (Farina et al. 2011; Pieretti et al. 2015; Rankin & Axel 2017). There are also clear distinctions in the soundscapes of different habitat types (Villanueva-Rivera et al. 2011; Depraetere et al. 2012; Bormpoudakis et al. 2013), with habitat disturbance or conversion reflected in changes in the soundscape, likely triggered by shifts in faunal assemblages (Burivalova et al 2017; Deichmann et al. 2017; Tucker et al. 2014). From these studies it is clear that ecoacoustics has enormous potential for environmental research.

Despite the promising results described above, studies have reported contradictory patterns, even when using the same acoustic indices. For example, some have found higher biophony and lower soundscape variability to be associated with lower levels of disturbance (Fuller et al. 2015; Machado et al. 2017), while others have found no differences among habitat types (Mammides et al. 2017; Ng et al. 2018). These disagreements may have arisen because to date there has been no consistency in data collection, and little agreement on the best indices for soundscape assessment. Although guidelines on the use of ecoacoustics for biodiversity monitoring have been published, these focus on assessing faunal presence rather than soundscape analysis (Browning et al. 2017; Llusia et al. 2011). Thus, inconsistent methodologies may underlie the inconsistent patterns.

Ecoacoustic studies have used a wide range of recording schedules (e.g. from continuous to <1 minute per hour) and data volumes (e.g. from >200 hours to <5 minutes per site; see supporting information). Inter-soundscape comparisons are common, without consideration of whether intra-soundscape variation has been accurately captured. While geophony is a key constituent of natural soundscapes, recordings with “high” levels of geophony are often removed from analyses, without a common definition of what “high” might be. Moreover, studies often present just one or two indices, with little justification for their selection. We argue that these inconsistencies are limiting the efficacy of acoustic indices in biodiversity monitoring. Given each index reflects different spatio-temporal features (Eldridge et al. 2016), considering several indices in concert may give a much better representation of the soundscape rather than any one individual index. Here, we use seven commonly employed acoustic indices derived from recordings collected across a human-modified landscape in central Panama to ask:

1. What duration of recordings is necessary to quantify the soundscape of a site, and does this vary among habitat type or index?
2. Should recordings be continuous, or can they be limited to temporal sub-samples to minimise storage volumes and subsequent analysis?
3. Which indices best reflect temporal variation over the course of the day, and between seasons, and are there different patterns among habitats?

Materials and Methods

Study Landscape

Acoustic data were collected in the Emparador landscape, located in the central region of the Republic of Panama, to the west of the Panama Canal (Fig. 1). The landscape covers approximately 700 km^2 and is highly heterogeneous, with tracts of extensive continuous forest, agricultural pasture, remnant forest fragments, non-native tree plantations, regenerating scrub and small urban centres. The landscape is bordered by the Panama Canal to the north and east, and the Interamericana highway to the south. The human population is distributed throughout the landscape, with sizeable areas of new urban development close to the Interamericana Highway. Rainfall varies from 2334mm in the north to 1969mm in the south (Pyke et al. 2001). There is a pronounced dry season between late December and late April when the mean daytime temperature is 31C. The remainder of the year is wet, with a mean daytime temperature of 28C (Robinson et al. 2004).

Data Collection

One hundred and seventeen deployment sites (hereafter “sites”) were selected for this study, with 14-24 sites in each of the six main habitats present in the Emparador landscape: continuous forest, fragmented forest, riparian forest, scrub, teak plantation and pasture (Fig. 1). See Table S1 for a detailed description of the six habitats, their typical features, and numbers of sites in each habitat. Sites were positioned in patches of uniform habitat of at least one hectare, and were separated by a minimum distance of 500m from sites in other habitat types, and 1000m from those in the same habitat.

Data were collected between January and September 2017. Recorders were deployed for one week at each site. There were a total of 154 deployments, 90 in the dry season and 64 in the wet season. Eighty sites were only visited once during the study, the remaining 37 sites had two separate deployments of a week each, one in the dry season and one in the wet season, to facilitate study of intra-site seasonal patterns (mean 160 days +/- 26 between deployments). After each deployment, recorders were rotated between habitat types to minimise any bias that might arise from hardware variability. Sound recordings were collected using Solo recorders with omnidirectional microphones positioned between 1 and 2m above the ground (Whytock & Christie 2017). Most recordings were collected with Primo EM172 microphones (Primo, Singapore), however logistical issues necessitated switching to Snowflake microphones (Blue, USA) during the wet season in some cases. Testing suggested no systematic disparity in recordings collected with the different microphone models (see supporting information), so we did not distinguish between the two sets of recordings in the main analyses. Solo recorders collect audio continuously, but for ease of analysis recordings are automatically divided into 10-minute files. A sampling rate of 32,000 Hz was used as a balance between capturing the majority of the human-audible soundscape against storage volume requirements.

Pre-processing was limited to a 500 Hz low-stop filter prior to analyses to reduce microphone self-noise (Pieretti et al. 2015). This will have removed some genuine sources of low-frequency noise, but microphone self-noise may bias indices values. Several studies have screened recordings to exclude those with high levels of geophony (Pieretti et al. 2015; Gasc et al. 2013; Depraetere et al. 2012), or anthrophony (Bormpoudakis et al. 2013), as some indices can be strongly influenced by these elements. However, we consider these to be key components of the soundscape, so no recordings were excluded.

Data Analyses

Calculation of Acoustic Indices

157,476 10-minute files were included in this study, equivalent to three years of continuous audio. Index calculation and all analyses were conducted with the software R (ver 3.5.1; R Core Team

2018). We calculated soundscape indices values for each 10-minute recording, although for one section of the analysis, values were calculated for individual minutes (see below). Using the packages *seewave* (ver 2.1.0; Sueur et al. 2008a) and *soundecology* (ver 1.3.3; Villanueva-Rivera & Pijanowski 2018), the following seven indices were calculated using the default values of each function: Acoustic Complexity Index (ACI), Acoustic Diversity Index (ADI), Acoustic Evenness (AEve), Bioacoustic Index (Bio), Acoustic Entropy (H), Median of the amplitude envelope (M), and the Normalised Difference Soundscape Index (NSDI). A description of each index and the patterns they reflect is in Table 1, with additional details in Table S2 and example sonograms in Fig. S1. We selected these indices as they are the most frequently used in ecoacoustic research and have been compared in other multi-index studies (Fuller et al. 2015; Machado et al. 2017b; Mammides et al. 2017; Ng et al. 2018 but see Buxton et al. 2016, 2018). Selected indices also had to meet the following criteria: simple to calculate (i.e. existing functions in R packages), their values reflect soundscape patterns with links to ecological dynamics, and they are supported by peer-reviewed publication.

Minimum quantity of recordings and recording schedules

To determine the minimum number of recordings required to describe a site's soundscape, we randomly assigned the 10-minute recordings from each site into groups of six to create "pseudo-hours". Randomisation of the entire recording set from each site removed any diel effects, permitting focus on overall soundscape variability. We calculated mean index value and standard error for each pseudo-hour. Mean value was determined from the six recordings within each pseudo-hour, but standard errors were cumulative over time, i.e. error was estimated using all pseudo-hours up to and including the latest to simulate successively longer deployments. For example, standard error for the fourth hour was calculated using the indices values from the first four pseudo-hours, for the fifth hour standard error was calculated with the first five pseudo-hours and so on, (Fig. S2). As simulated deployments became longer, the inclusion of more data led to a decline in standard errors. Standard errors stabilised when natural variability rather than data paucity was determining the index variance. Reduction in this variance over time was modelled using nonlinear regression, to quantify the effect of increasing deployment lengths. For this analysis we treated all deployments as separate, even though some were revisits to the same site. A global model with index as a random effect would not converge, and so each index was modelled separately using the same distribution. We selected the Weibull distribution as it is both relatively simple and versatile, with a range of potential shapes from exponential to humped (Bolker, 2008). We explored potential habitat and seasonal differences in variance reduction but found no support for separate models, implying similar patterns across all habitats and seasons.

Recording on a temporal schedule, rather than continuously, is common practice in acoustic monitoring to improve battery performance and reduce data storage (Pieretti et al. 2015). To examine the effect of scheduled recording, we divided recordings into single minutes and calculated acoustic indices for each. We then simulated a range of schedules used in previous acoustic indices studies: continuous, one minute in every two, one in five, one in 10, one in 30 and one in 60 minutes. All schedules were treated as if they came from a one-week deployment, so resulting datasets spanned the same length of time but those from sparser schedules contained fewer data. Cumulative standard errors were calculated for each schedule as described above. Reduction in variance of standard error as a percentage of the maximum was modelled over deployment length using nonlinear regression. Again, a global model with schedule as a random effect would not converge and so separate models with a Weibull distribution were used for each schedule.

Indices for characterising temporal and spatial patterns

Acoustic indices from the 10-minute recordings were used to generate mean and standard deviation values per hour for each habitat in dry and wet seasons (Pieretti et al. 2015). To test for diel patterns, each hour was classed as either day (06:00 – 17:00) or night (18:00 – 05:00). Finer scale temporal trends were explored using the mean value per 10-minute recording within habitat and season.

Temporal and spatial soundscape patterns among habitats were explored in four ways. First, we performed non-metric multidimensional scaling (NMDS) to investigate habitat-specific diel and seasonal patterns, using the mean and standard deviation of acoustic indices values per hour. We used two axes and the Horn-Morisita dissimilarity index (Horn 1966), and checked the output met minimum stress requirements (Kruskal 1964). Second, the ordination was extended with permutational multivariate analysis of variance (PERMANOVA) to quantitatively test the effect of diel phase, season and habitat type on mean hourly indices values (Anderson 2001). These two analyses were conducted with the package *vegan* (ver 2.5.2; Oksanen et al. 2018). Third, to illustrate finer-scale temporal patterns over 24 hours, we considered trends in mean index value from each 10-minute recording block. This was undertaken for each habitat for both dry and wet seasons, and curves were fitted to these patterns with Generalised Additive Models (GAMs), using the package *mgcv* (ver 1.8.26; Wood 2004).

Finally, to determine which indices were most important in separating habitat-specific soundscapes, we undertook a random forest (RF) classification (Breiman 2001) using the *randomForest* package (ver 4.6.14; Liaw & Wiener 2002). We built a RF using mean hourly indices values and standard deviations, plus the factors “dry” or “wet” season, diel phase “day” or “night”. 75% of the data were used for forest construction and the remaining 25% reserved for testing.

Results

Minimum quantity of recordings and recording schedules

Standard errors rapidly shrank with increasing deployment time; all indices showed a common pattern of exponential decline as standard errors converging on the mean (Fig. 2). These patterns were consistent for all indices, across all sites and habitats, and between seasons, as evidenced by fitting the same Weibull distribution to all datasets (Table S3). After 120 hours of recordings, variance stabilised to 8.9 - 12.1 %.

Similar patterns of exponential decline in standard errors were evident when exploring the importance of scheduling (Fig. 3). Sparser schedules were associated with greater variability, a pattern consistent across all indices. Extrapolation of the one-minute-in-10 model suggested that more than 26 weeks of recording would be required to reduce index variance to a similar level achieved with seven days of continuous recordings (one-in-10 = 2.1% after 4368 hours, continuous = 1.56% after 168 hours). Convergence was a product of total recording length irrespective of the schedule used.

Indices for characterising spatial and temporal patterns

The NMDS ordination showed clear patterns (Fig. 4); dry and wet season recordings separated along axis 1 suggesting distinct soundscapes at different times of the year, while axis 2 illustrated a clear division between day and night soundscapes. The effect of habitat was less clear, with no obvious pattern in habitat type driving separation among the points. These results were reflected in the PERMANOVA; diel phase, season, habitat type and the diel phase-season interaction were all significant (Table S5).

Six of the seven indices exhibited distinct patterns over the 24 hour period (Fig. 5, Table S6), often with marked shifts between the day and night soundscapes as implied by the NMDS ordination and PERMANOVA. Curves from the GAM fitted to 10-minute mean index values showed ADI and H values were high across all habitats during the day, but 6% and 9% lower at night respectively. Conversely, the AEve, Bio, M and NSDI values were lower during the day but 50 – 200% higher at night. Habitat-specific diel patterns were also apparent; the rise to daytime H values in pasture was an hour behind the other habitats, a lag mirrored in the Bio and NDSI indices values. AEve and Bio values in both fragmented forest and scrub were higher at night and lower during the day compared with other habitats. Diel patterns in ACI were not apparent, but values in pasture were more variable with a standard deviation of 17, compared to <10 for other habitats. Seasonal differences in diel patterns of all habitats were also evident. The switch from diurnal to nocturnal values in the ADI, AEve and Bio indices was much more gradual in the wet season, beginning around 90 minutes earlier than in the dry season across all habitats. In the wet season, diel variation was reduced in the AEve index (standard deviation 50% lower) but magnified in NDSI (standard deviation 50% higher).

The RF classifier built with mean hourly values per habitat was able to readily separate the data into the six habitat classes (Fig. 6). When applied to the testing dataset, the RF was 84.7% accurate in assigning soundscapes to the correct habitat type. Both mean and standard deviation of ACI were more important than any other variables for distinguishing among habitat types (with removal accounting for a proportional drop in accuracy of 0.17 and 0.08 respectively). Mean AEve values, and the mean and standard deviations of H and NDSI were also important. Season and diel phase were of least importance.

Discussion

Minimum quantity of recordings and recording schedules

We found a consistent pattern in variance reduction across all indices, habitats and seasons. Index variability was reduced to a mean of 10.9% of its maximum after 120 recording hours. We selected this cut-off as a balance between deployment length and capturing the majority of site variability. Beyond 120 hours variance decreased so slowly with increasing sampling duration that it was not worth the increased input of time and resources. Convergence of indices values was a product of the total amount of recordings used rather than the length of the deployment or recording schedule. Thereafter continued variability was likely due to inherent soundscape features of the site rather than insufficient length of recording.

Previous acoustic studies have used a median of 24 hours (range 0.1 – 1436 hours) of recordings per site in terrestrial systems (see supporting information); at which point variance in index standard error will still be high. Our analysis suggests that studies using <120 recording hours may not have fully described the soundscapes, limiting the power of their conclusions. Similar analyses with recordings from other terrestrial landscapes would be required to determine if this rate of variance reduction is typical. Tropical soundscapes are often more complex than those of temperate systems, therefore convergence might be achieved more rapidly in simpler environments. However, all habitats in this study shared common convergence patterns despite considerable variation in vegetation structure and faunal communities, so the time required for other locations may prove similar.

To our knowledge there is only one other study of temporal sampling, which advocated a one-minute-in-5 schedule as retaining the majority of information found in continuous recordings (Pieretti et al 2015). Our results from simulated datasets suggest that sparser sampling schedules (even 1-in-5) delays capture of inherent soundscape variability; and that continuous recordings are more effective for reliably capturing a soundscape. Sparse sampling schedules also require longer deployment times, so that site patterns might be complicated by seasonal shifts. Where monitoring seeks to describe patterns over longer temporal scales, it might be difficult to distinguish between short-term stochasticity and longer-term variability such as seasonal changes.

Indices for characterising temporal and spatial patterns in soundscapes

We found clear diel, seasonal and habitat-specific patterns among the soundscapes. Diel patterns were particularly pronounced, with all soundscapes showing a consistent distinction between day and night. Such findings are intuitive; almost every habitat in the world has discreet diurnal and nocturnal faunal assemblages. The diel division is evidenced by the common trends shown in most of the indices, regardless of habitat type, and reinforced by the clear division along axis 2 of the NMDS and the significant effect in the PERMANOVA.

Diel patterns in the soundscape were particularly marked with the AEve, Bio, H and NDSI indices. Overall, the indices imply that nocturnal soundscapes were more uneven; with fewer occupied frequency bands (ADI, AEve), a greater disparity between loudest and quietest bands (Bio, H) and lower levels of anthrophony (NDSI). This is consistent with insect and anuran communities dominating a limited range of frequencies (Villaneuva-Rivera et al. 2011). Conversely, diurnal soundscapes were typically more even. Greater levels of anthrophony and more variable biophony lead to an increase in the number of occupied frequency bands (ADI, AEve, NDSI), and with a more even amplitude (Bio, H). The only index without a clear diel pattern was ACI, perhaps because it effectively filters out consistent sounds such as insect choruses that are likely to underlie the diel differences in the other indices.

The division of soundscapes into distinct diel phases has been widely reported. Equivalent patterns in NDSI values have been found previously, presumably because there is generally more anthrophony during daylight hours (Fuller et al. 2015). However, for some indices, specific patterns appear strongly dependent on region. Studies of Australian woodland sites report a diel split with the reverse of our results; high ADI and H values at night, and high Bio during the day (Fuller et al. 2015, Gage et al. 2017). This would be consistent with insects in nocturnal soundscapes in Australia occupying a broader range of frequency bands than in Panama. Trends in ACI are also inconsistent: either no clear pattern (this study; Fuller et al. 2015), or marked diel differences arising from nocturnal insect biophony (Pieretti et al. 2015). Villanueva-Rivera et al. (2011) showed that most of their temperate sites had distinct diel patterns in ADI, with strong peaks corresponding to dawn and dusk choruses, a pattern not evident in our recordings. These patterns may reflect genuine differences among soundscapes but, as noted earlier, such contradictory results may arise from the variable amounts of recordings analysed in these studies (Table S7).

Soundscapes differed between seasons, with different diel patterns. Seasonal variation might be driven by changes in vegetation structure, or follow behavioural shifts in faunal communities, such as the onset of territorial birdsong (Buxton et al. 2016; Rankin & Axel 2017). In our recordings the most important seasonal influence was the frequency of storms, which had a notable impact on wet season soundscapes. Wet season diel variation was weaker in Bio and H index values, but stronger in NDSI, suggesting a smaller disparity between loudest and quietest frequency bands. Furthermore, the diel switch in ADI, AEve , Bio and H indices values was more gradual in the wet season, implying a less abrupt transition to nocturnal dominance of a reduced range of frequencies. Storm geophony likely underpinned this reduced diel shift, as storm events are less temporally restricted than biophony.

While temporal influences drove overall patterns, the RF implied consistent finer-scale differences among habitat soundscapes. Mean ACI and standard deviation were the most important variables for distinguishing among habitats, matching previous findings of habitat-specific patterns in ACI values (Fuller et al 2015; Pieretti et al 2015). The influence of geophony on ACI values likely permitted effective discrimination between open habitats (scrub and pasture), and those habitats with trees. The low vegetation characteristic of pasture and scrub make these habitats exposed to wind and the associated sound. Conversely, rainstorms in forested habitats have a much greater influence on the soundscape, as water continues to drip from vegetation long after the rain has ceased.

We did not include urban sites in this study, and distinct patterns in acoustic indices have been found in urban habitats (Fairbrass et al 2017; Joo et al 2011). It would be interesting to ascertain how an acoustically rich anthrophony is reflected in acoustic indices values and whether urban sites might exhibit the same patterns in variance reduction and the effects of temporal subsampling reported here. Although our seven indices describe a range of soundscape features they are only a fraction of those available; other less commonly used indices may well contain additional important information. Further testing would be required to determine whether variance reduction in other acoustic indices follows similar patterns to those we report for the seven in this study.

Recommendations for acoustic monitoring

Recent reviews have highlighted the critical need for standardised protocols in ecoacoustic data collection and processing (Gibb et al 2018; Priyadarshani et al. 2018). We provide the following workflow to guide future ecoacoustic studies;

1. *Collect 120 hours of audio recordings per site*. This balances deployment length and capture of soundscape variability, although the time required to improve precision might vary in other biomes or ecosystems. Repeated short deployments during distinct seasons may be as suitable as a single long deployment.

2. *Avoid temporal sub-sampling*. Recording on a schedule only delays the capture of soundscape variability; sparse sampling schedules will require longer deployment times.

3. *Use multiple indices to describe soundscape patterns*. No single acoustic index can describe the entire soundscape; capturing inter-habitat differences requires multiple indices, as there are often competing explanations for a particular index value. For example, low Bio values could indicate either an impoverished soundscape with little noise or an acoustically rich environment; if the soundscape also has low H and high AEve values it would support the latter interpretation. The seven indices used in this study will not necessarily suit all situations and systems, and identifying the most appropriate indices to use will depend on study aims. Using a suite of indices will offer complimentary impressions of different aspects of the soundscape. Selection should be based on a solid understanding of the soundscape patterns underlying index values, and hence the ecological patterns they may reflect.

4. *Use mean values and standard deviations rather than raw values*. This draws out patterns that might otherwise be obscured by short-term variability. The magnitude of variability provides additional information, and in many cases standard deviations of indices were more important than mean values for distinguishing among the habitats.

5. *Consider more than just a single portion of the day*. Diel patterns are important for extracting differences between habitat types. Dry season values for Bio and NDSI were near uniform among habitats between 12:00 and 17:00, but differed widely outside these hours. Conversely, the greatest variation in ACI and AEve values was during afternoon.

Traditional approaches to biodiversity assessment are time-consuming, expensive and often limited to a small geographic area. Automated recording and analysis of soundscapes can be conducted at far greater spatial and temporal scales, potentially at lower costs. Soundscape analysis has been used as a tool for rapid biodiversity assessment; acoustic indices have been linked with measures of bird species richness, compositional shifts in bird communities, and songbird phenology (Buxton et al. 2016; Fuller et al. 2015; Lellouch et al. 2014; Towsey et al. 2014a). Increased forest disturbance has been associated with lower acoustic diversity in Tanzania (Sueur et al. 2008b), and lower acoustic saturation in Papua New Guinea (Burivalova et al. 2017). Yet it is unclear whether such results are representative of more general relationships between soundscapes and habitat integrity (Burivalova et al. 2019; Gibb et al. 2018; Merchant et al. 2015). There are inconsistent patterns in the literature, which have led some to question the efficacy of acoustic indices for biodiversity monitoring (Browning et al. 2017; Eldridge et al. 2016; Servick 2014). We argue that variations in collection and processing methodologies probably underlie some of these uncertainties. Further research is needed to elucidate the complementarity of standard biodiversity monitoring methods and ecoacoustics, but a key aspect of integrating these approaches will be consistency in both data collection and analysis.

Acknowledgements

We thank two anonymous reviewers for their helpful and constructive comments. Our thanks to the Smithsonian Tropical Research Institute, the Autoridad de Canal de Panama, and all of the private landowners who granted permission to deploy recorders on their land. We thank the Autoridad Nacional del Ambiente for providing research permits to work in the Republic of Panama (permits SE/A-134-16 and SE/A-67-18). Robbie Whytock and James Christie offered helpful advice and support regarding the Solo audio recorders. Scott Jackson provided invaluable IT resources and support during indices calculations. TBL was funded by the Natural Environment Research Council UK, and the IAPETUS Doctoral Training Partnership.

Author Contributions

TBL, LB, NB, SGW and DD conceived the study; TBL and NG collected the data; TBL, NB and DD designed the analysis; TBL led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data Accessibility

Data have been archived at DataSTORRE, the University of Stirling's online repository for research data, available at http://hdl.handle.net/11667/132

References

Anderson, M. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*, 32–46.

Boelman, N., Asner, G., Hart, P., & Martin, R. (2007). Multi-Trophic Invasion Resistance in Hawaii:

Bioacoustics, Field Surveys, and Airborne Remote Sensing. *Ecological Applications*, *17*, 2137–2144.

Bolker, B. (2008). *Ecological Models and Data in R*, Princeton, Princeton University Press.

Bormpoudakis, D., Sueur, J., & Pantis, J. (2013). Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications. *Landscape Ecology*, *28*, 495–506.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Browning, E., Gibb, R., Glover-Kapfer, P., & Jones, K. (2017). Passive acoustic monitoring in ecology and conservation. *WWF Conservation Technology Series*, *2*, 1–75.

Burivalova, Z., Game, E., & Butler, R. (2019). The sound of a tropical forest. *Science*, *363* (*6422*), 28–29.

Burivalova, Z., Towsey, M., Boucher, T., Truskinger, A., Apelis, C., Roe, P., & Game, E. (2017). Using soundscapes to detect variable degrees of human influence on tropical forests in Papua New Guinea. *Conservation Biology*, *32*, 205–215.

Buxton, R., McKenna, M., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L., Crooks, K., & Wittemyer, G. (2018). Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*, 32, 1174-1184.

Buxton, R., Brown, E., Sharman, L., Gabriele, C., & McKenna, M. (2016). Using bioacoustics to examine shifts in songbird phenology. *Ecology and Evolution*, *6*, 4697–4710.

Deichmann, J., Hernández-Serna, A., Delgado, A., Campos-Cerqueira, M., & Aide, M. (2017). Soundscape analysis and acoustic monitoring document impacts of natural gas exploration on biodiversity in a tropical forest. *Ecological Indicators*, *74*, 39–48.

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., & Sueur, J. (2012). Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*, *13*, 46–54.

Eldridge, A., Casey, M., Moscoso, P., & Peck, M. (2016). A new method for ecoacoustics? Toward the extraction and evaluation of ecologically-meaningful soundscape components using sparse coding methods. *PeerJ*, *4*, e2108.

Fairbrass, A.J., Rennett, P., Williams, C., Titheridge, H. & Jones, K.E. (2017). Biases of acoustic indices measuring biodiversity in urban areas. *Ecological Indicators*, *83*, 169-177.

Farina, A., Pieretti, N., & Piccioli, L. (2011). The soundscape methodology for long-term bird monitoring: A Mediterranean Europe case-study. *Ecological Informatics*, *6*, 354–363.

Fuller, S., Axel, A., Tucker, D., & Gage, S. (2015). Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? *Ecological Indicators*, *58*, 207–215.

Furnas, B., & Callas, R. (2015). Using Automated Recorders and Occupancy Models to Monitor Common Forest Birds Across a Large Geographic Region. *Journal of Wildlife Management*, *79*, 325–337.

Gage, S., Wimmer, J, Tarrant, T. & Grace, P. (2017). Acoustic patterns at the Samford Ecological Research Facility in South East Queensland, Australia: The Peri-Urban SuperSite of the Terrestrial Ecosystem Research Network. *Ecological Informatics*, 38, 62-75.

Gasc, A., Sueur, J., Jiguet, F., Devictor, V., Grandcolas, P., Burrow, C., ... Pavoine, S. (2013). Assessing biodiversity with sound: Do acoustic diversity indices reflect phylogenetic and functional diversities of bird communities? *Ecological Indicators*, *25*, 279–287.

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. (2018). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, *2018*, 1–17.

Horn, H. (1966). Measurement of “overlap” in comparative ecological studies. *The American Naturalist*, *100*, 419–424.

Joo, W., Gage, S.H. & Kasten, E.P. (2011). Analysis and interpretation of variability in soundscapes along an urban–rural gradient. *Landscape and Urban Planning*, *103*, 259-276.

Kasten, E., Gage, S., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, *12*, 50–67.

Kruskal, J. (1964). Nonmetric Multidimensional Scaling: a numerical method. *Psychometrika*, *29*, 115–129.

Laiolo, P. (2010). The emerging significance of bioacoustics in animal species conservation. *Biological Conservation*, *143*, 1635–1645.

Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H., & Sueur, J. (2014). Monitoring temporal change of bird communities with dissimilarity acoustic indices. *Methods in Ecology and Evolution*, *5*, 495–505.

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*, 18–22.

Llusia, D., Márquez, R., & Bowker, R. (2011). Terrestrial sound monitoring systems, a methodology for quantitative calibration. *Bioacoustics*, *20*, 277–286.

Machado, R., Aguiar, L., & Jones, G. (2017). Do acoustic indices reflect the characteristics of bird communities in the savannas of Central Brazil? *Landscape and Urban Planning*, *162*, 36–43.

Mammides, C., Goodale, E., Dayananda, S., Kang, L., & Chen, J. (2017). Do acoustic indices correlate with bird diversity? Insights from two biodiverse regions in Yunnan Province, south China. *Ecological Indicators*, *82*, 470–477.

Merchant, N., Fristrup, K. M., Johnson, M., Tyack, P., Witt, M., Blondel, P., & Parks, S. (2015). Measuring acoustic habitats. *Methods in Ecology and Evolution*, *6*, 257–265.

Ng, M., Butler, N., & Woods, N. (2018). Soundscapes as a surrogate measure of vegetation condition for biodiversity values: A pilot study. *Ecological Indicators*, *93*, 1070–1080.

Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2018). vegan: Community Ecology Package. R package version 2.5-2.

Pieretti, N., Duarte, M., Sousa-Lima, R., Rodrigues, M., Young, R., & Farina, A. (2015). Determining temporal sampling schemes for passive acoustic studies in different tropical ecosystems. *Tropical Conservation Science*, *8*, 215–234.

Pieretti, N., Farina, A., & Morri, D. (2011). A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). *Ecological Indicators*, *11*, 868–873.

Pijanowski, B., Villanueva-Rivera, L., Dumyahn, S., Farina, A., Krause, B. L., Napoletano, B., … Pieretti, N. (2011). Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*, *61*, 203–216.

Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. *Journal of Avian Biology*, *49,* e01447.

Pyke, C., Condit, R., Aguilar, S., & Lao, S. (2001). Floristic composition across a climatic gradient in a neotropical lowland forest. *Journal of Vegetation Science*, *12*, 553–566.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rankin, L., & Axel, A. (2017). Biodiversity Assessment in Tropical Biomes using Ecoacoustics: Linking Soundscape to Forest Structure in Human-dominated Tropical Dry Forest in Southern Madagascar. In A. Farina & S. Gage (Eds.), *Ecoacoutics: The ecological role of sounds* (pp. 129–144). Hoboken, NJ: John Wiley & Sons Ltd.

Ribeiro, J., Sugai, L., & Campos-Cerqueira, M. (2017). Passive acoustic monitoring as a complementary strategy to assess biodiversity in the Brazilian Amazonia. *Biodiversity and Conservation*, *26*, 2999–3002.

Robinson, W., Angehr, G., Robinson, T., Petit, L., Petit, D., & Brawn, J. (2004). Distribution of Bird Diversity in a Vulnerable Neotropical Landscape. *Conservation Biology, 18*, 510–518.

Rodriguez, A., Gasc, A., Pavoine, S., Grandcolas, P., Gaucher, P., & Sueur, J. (2014). Temporal and spatial variability of animal sound within a neotropical forest. *Ecological Informatics*, *21*, 133–143.

Servick, K. (2014). Eavesdropping on Ecosystems. *Science*, *343*, 834–838.

Sevilla, A., & Glotin, H. (2017). Audio bird classification with inception-v4 extended with time and time-frequency attention mechanisms. *CEUR Workshop Proceedings*, *1866*.

Sueur, J., Aubin, T., & Simonis, C. (2008a). Seewave, a Free Modular Tool for Sound Analysis and Synthesis. *Bioacoustics*, *18*, 213–226.

Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica United with Acustica*, *100*, 772–781.

Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008b). Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, *3*, e4065.

Towsey, M., Wimmer, J., Williamson, I., & Roe, P. (2014a). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*, 21 (100), 110–119.

Towsey, M., Zhang, L., Cottman-Fields, M., Wimmer, J., Zhang, J., & Roe, P. (2014b). Visualization of long-duration acoustic recordings of the environment. *Procedia Computer Science*, 29, 703–712.

Towsey, M., Znidersic, E., Broken-Brow, J., Indraswari, K., Watson, D., Phillips, Y., Truskinger, A. & Roe, P. (2018). Long-duration, false-colour spectrograms for detecting species in large audio data-sets. *Journal of Ecoacoustics*, 2, doi.org/10.22261/jea.iuswui

Tucker, D., Gage, S., Williamson, I., & Fuller, S. (2014). Linking ecological condition and the soundscape in fragmented Australian forests. *Landscape Ecology*, *29*, 745–758.

Villanueva-Rivera, L., & Pijanowski, B. (2018). soundecology: Soundscape Ecology. R package version

1.3.3.

Villanueva-Rivera, L., Pijanowski, B., Doucette, J., & Pekin, B. (2011). A primer of acoustic analysis for landscape ecologists. *Landscape Ecology*, *26*, 1233–1246.

Whytock, R., & Christie, J. (2017). Solo: an open source, customizable and inexpensive audio recorder for bioacoustic research. *Methods in Ecology and Evolution*, *8*, 308–312.

Wood, S. (2004). Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. *Journal of the American Statistical Association*, *99*, 673–686.

Figures

Fig. 1. Map of the study region in central Republic of Panama showing the 117 sites where audio recordings were collected.

Fig. 2. Reduction in variance of standard errors for seven acoustic indices, from a total of 154 recordings sets. Curves for each index show predicted values from nonlinear regression models with a Weibull distribution +/- 1 standard deviation.

Fig. 3. Effect of collecting audio recordings on a schedule. Reduction in variance of standard errors for six temporal recording schedules with increasing lengths of recording. Total dataset includes seven acoustic indices from 154 recording sets. Curves show predicted values from nonlinear regression models with a Weibull distribution +/- 1 standard deviation.

Fig. 4. Ordination plot showing the strong diel and seasonal divisions between soundscapes. Ordination was performed using non-metric multidimensional scaling (NMDS) with the Horn-Morisita dissimilarity index (Horn 1966). Stress value 0.002. Each point represents the soundscape of a habitat during one hour. This soundscape is composed of the mean hourly values and standard deviations for each of the seven acoustic indices. Circles show dry season soundscapes, triangles wet season soundscapes.

Fig. 5. Diel patterns in mean acoustic indices, with predicted values and standard errors from GAM output for each habitat. Solid line shows dry season values, dashed line wet season. Values calculated for each 10-minute recording window over 24 hours, from 154 recording sets.

Fig. 6. Variables ranked by importance for classifying habitat type in a random forest model, showing decline in predictive accuracy if a predictor is removed. Random forest constructed with the hourly mean and standard deviation values of seven acoustic indices, season (wet or dry), and diel phase (day or night). Internally estimated error rate was 20.8%, while testing with an independent dataset showed the classifier to be 84.7% accurate.

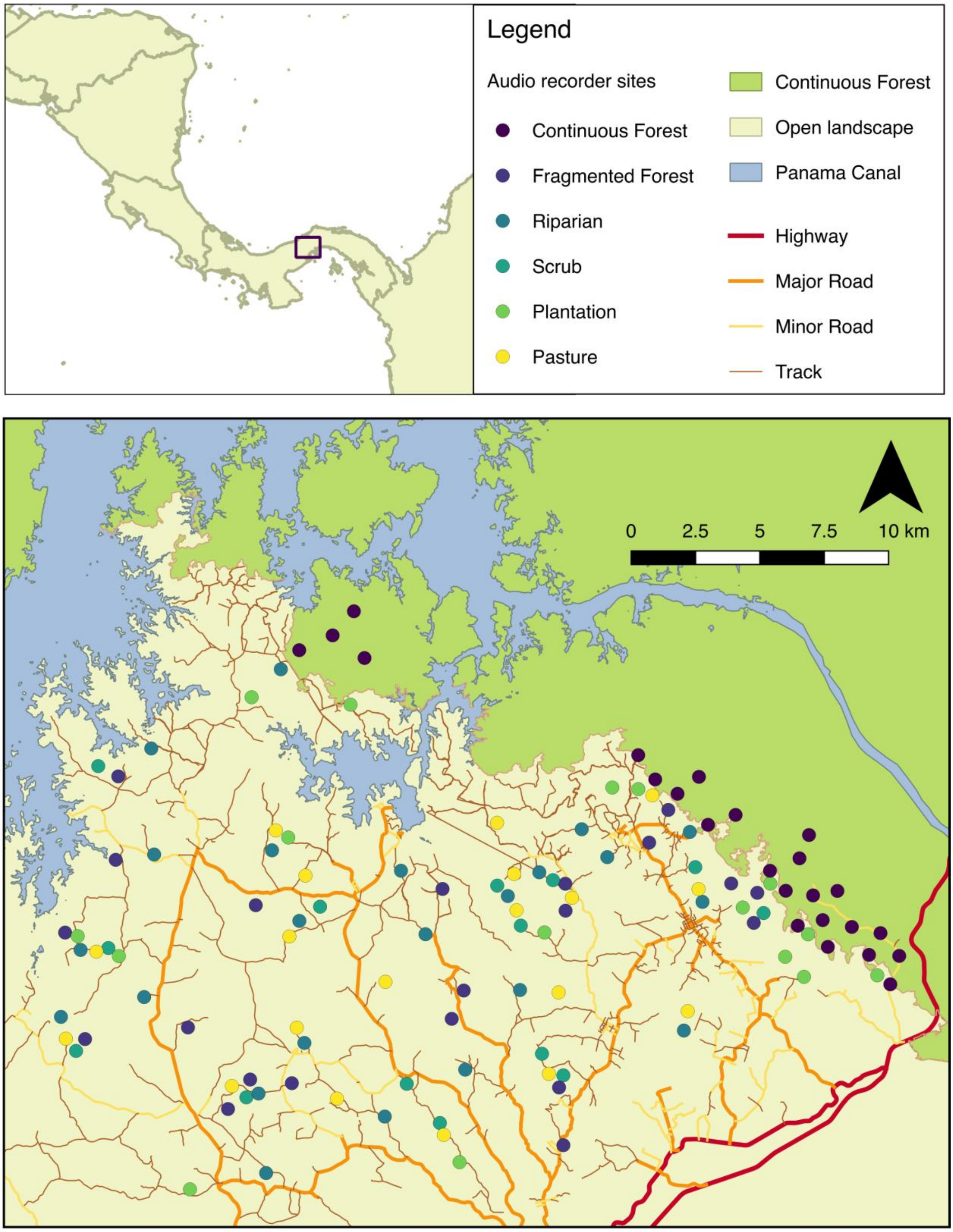
Tables

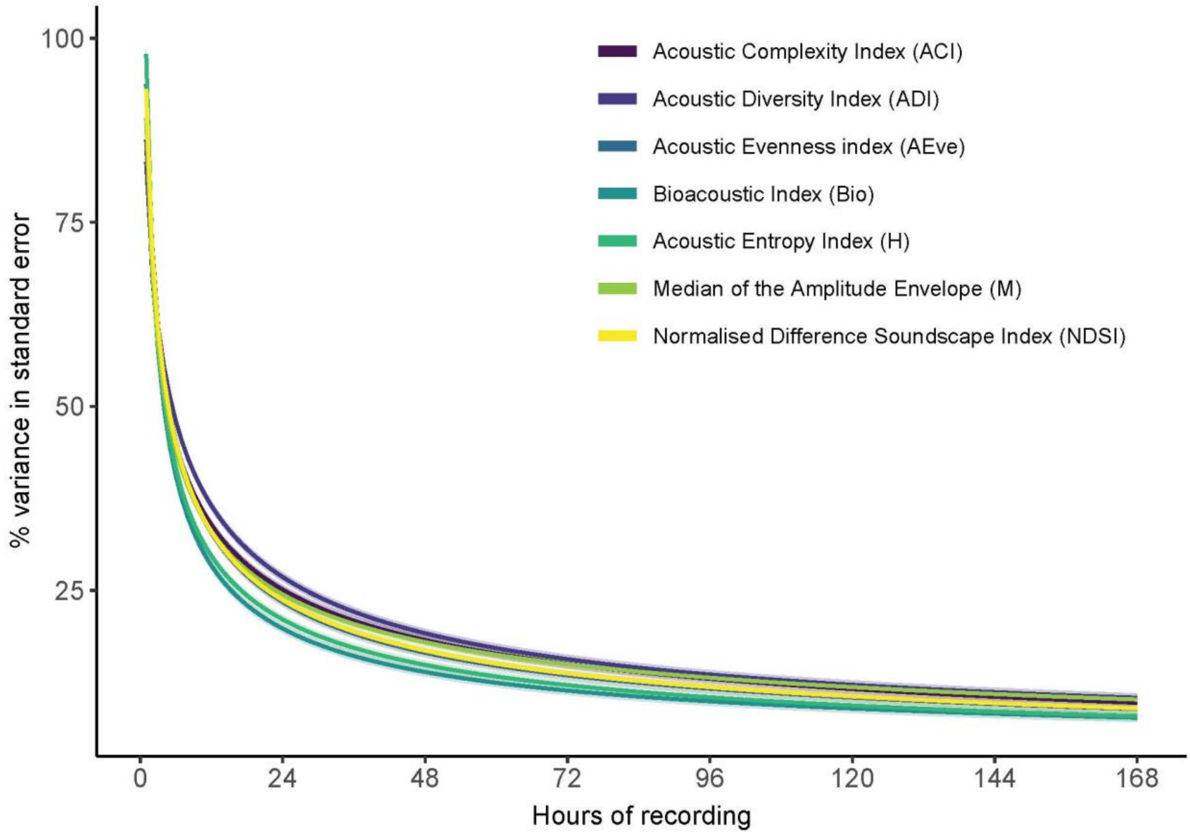
Table 1. Summary of the indices used in this study, the general soundscape patterns they reflect, and examples from this study. Further information including how the indices are calculated is detailed in Table S2.

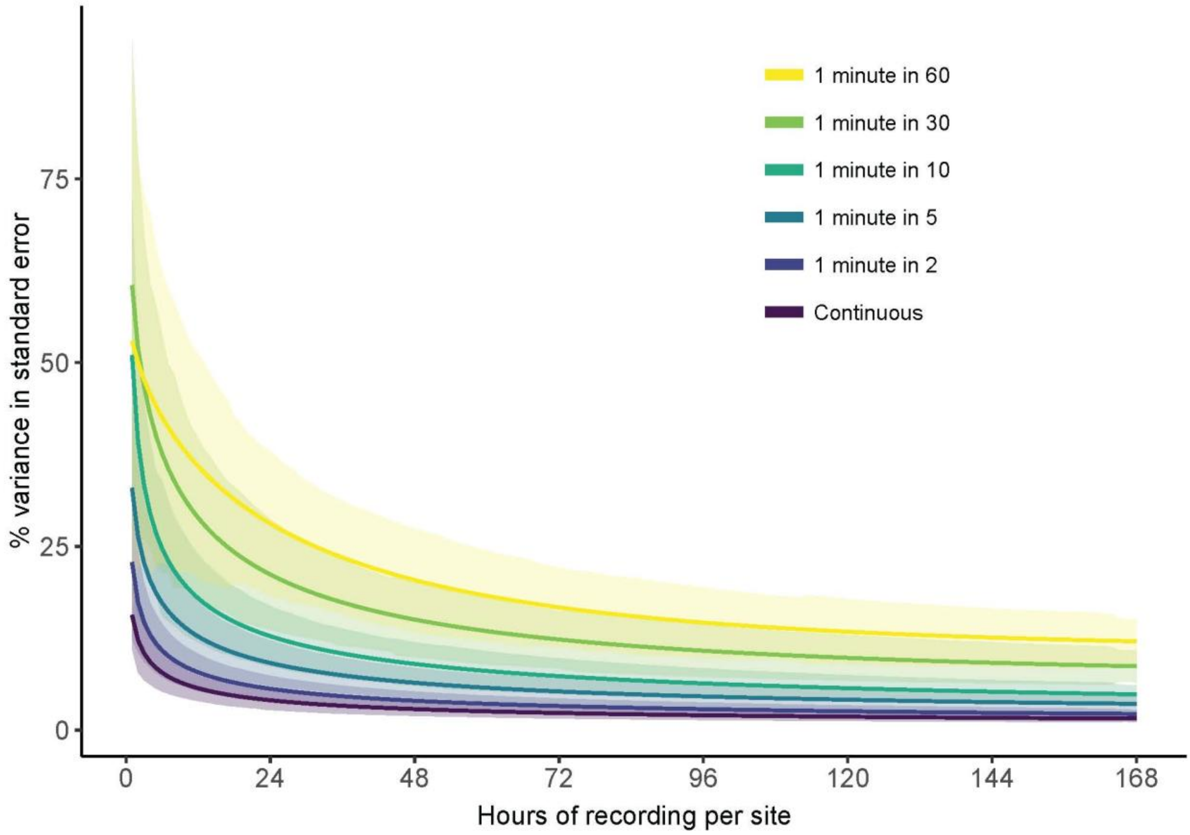
Index and reference	Soundscape patterns	Patterns in this study
Acoustic Complexity Index (ACI) (Pieretti et al. 2011)	Based on difference in amplitude between one time sample and the next within a frequency band, relative to the total amplitude within that band. Designed to quantify the inherent irregularity in biophony, while being relatively impervious to persistent sound of a constant intensity.	High values indicate storms, intermittent rain drops falling from vegetation, stridulating insects, or high levels of bird activity. Lowest values came from recordings with consistent cicada noise that fills the whole spectrogram.
Acoustic Diversity Index (ADI) (Villanueva-Rivera et al. 2011)	Increases with greater evenness across frequency bands. An even signal (either noisy across all frequency bands or completely silent) will give a high value, while a pure tone (i.e. all energy in one frequency band) will be closer to 0.	Highest values were from recordings with high levels of geophony or anthrophony (wind, helicopters or trucks) blanketing the spectrogram with noise, or from very quiet recordings with little variation among frequency bands. Lowest values reflect dominance by a narrow frequency band, usually by nocturnal insect noise.
Acoustic Evenness (AEve)	Higher values indicating greater unevenness among frequency bands, i.e. most of the sound intensity appears in a restricted range	Reverse of ADI patterns. High values identify recordings with dominance by a narrow frequency band of insect noise.

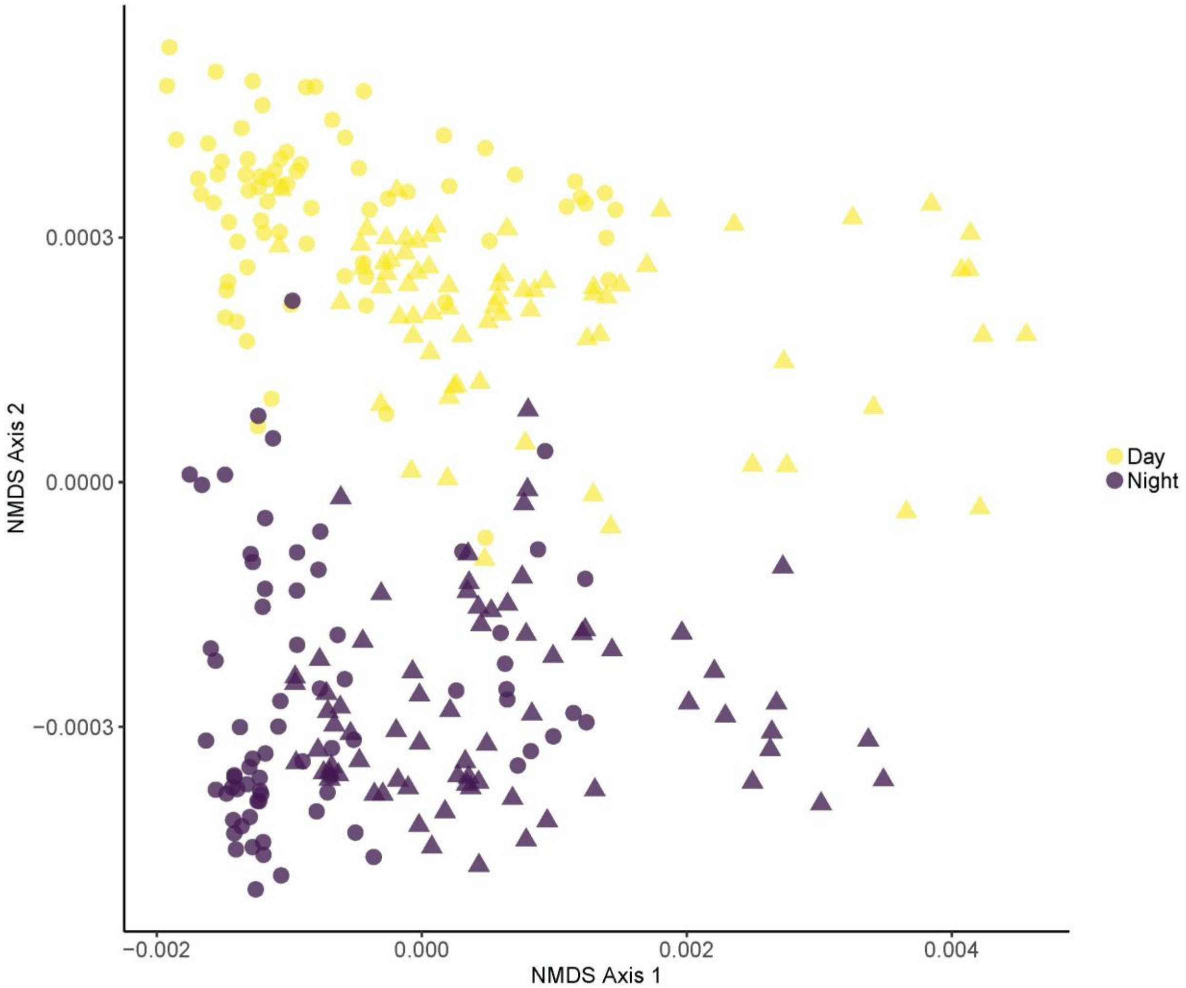
(Villanueva-Rivera et al. 2011)	of frequencies. Acoustically rich habitats may produce low values because there is little variation in intensity among frequency bands in saturated soundscapes.	Low values are associated with windy recordings with many occupied frequency bands, or near silent recordings with no acoustic activity.
Bioacoustic Index (Bio) (Boelman et al. 2007)	A function of both amplitude and number of occupied frequency bands between 2 – 11 kHz. Value is relative to the quietest 1 kHz frequency band; higher values indicate greater disparity between loudest and quietest bands.	Highest values produced by blanket cicada noise, with high amplitude and minimal variation among frequency bands. Low values arise when there is no sound between 2 and 11 kHz, although there is sometimes insect biophony outside these bounds.
Acoustic entropy (H) (Sueur et al. 2008b)	Increases with greater evenness of amplitude among frequency bands and/or time steps. Returns a value between 1 (an even signal, either noisy across frequency bands or completely silent) and 0 (a pure tone with all energy in one frequency band).	Highest values from near-silent recordings, with no wind, and only faint bird calls. Lowest values produced when insect noise dominated a single frequency band.
Median of the amplitude envelope (M) (Depraetere et al. 2012)	Reflects the amplitude of a recording. Louder recordings will give higher values, reflecting noisier soundscapes.	Highest values associated with high levels of geophony, particularly storms. Low levels of M produced by very quiet recordings, little biophony or geophony.
Normalised	Relies on a theoretical frequency split	High values reflect high levels of insect

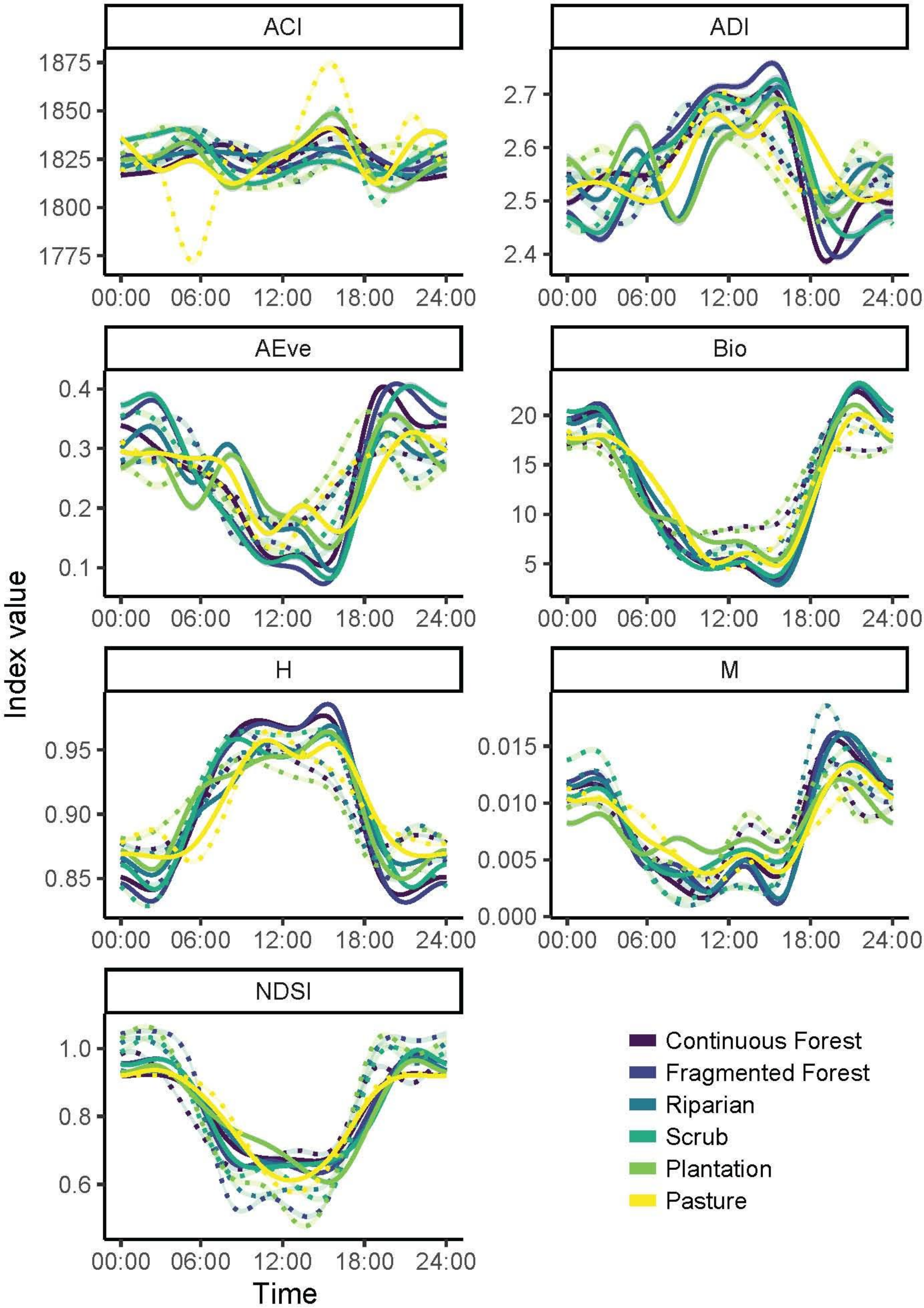
Difference Soundscape Index (NDSI) (Kasten et al. 2012)	between anthrophony (1-2 kHz) and biophony (2-11 kHz). The ratio of the two components give values of -1 to +1, with +1 indicating no anthrophony in the soundscape.	biophony, with minimal noise in the 1 – 2 kHz range. Low values arise when insect biophony dominates the 1 – 2 kHz band.

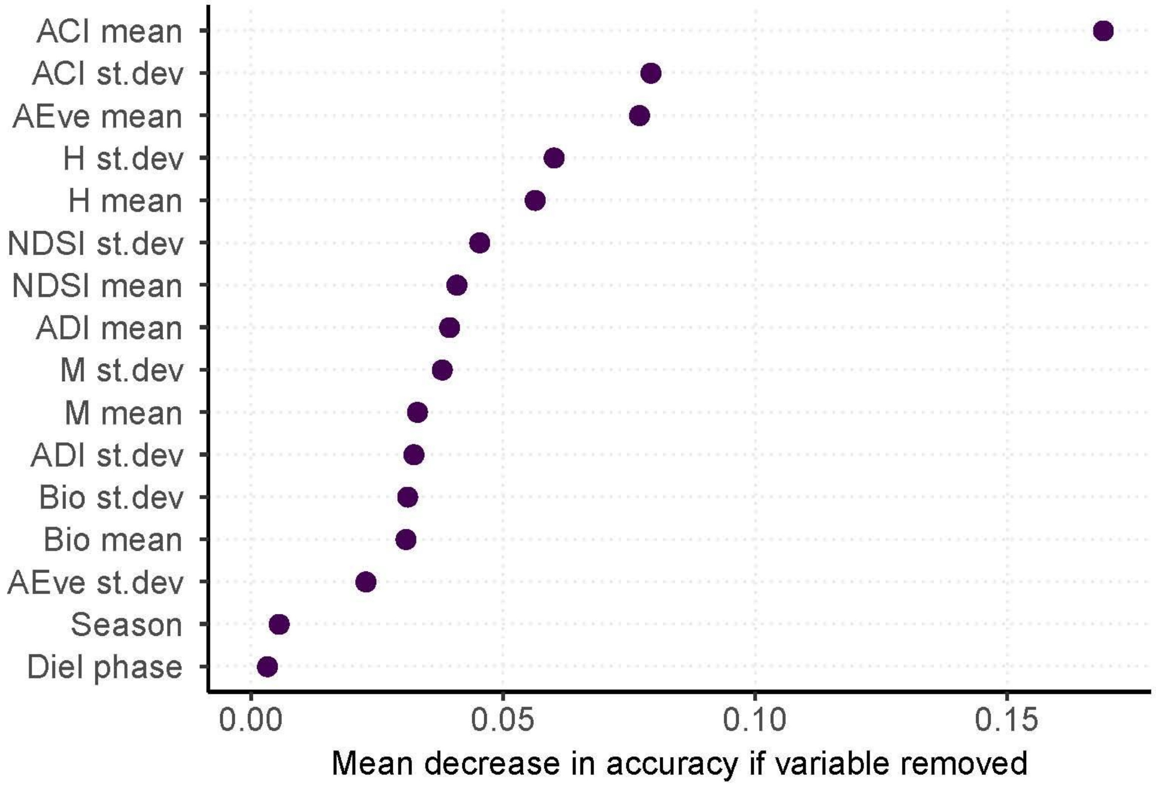

Legend
Audio recorder sites
Continuous Forest
Fragmented Forest
Riparian
Scrub
Plantation
Pasture
Continuous Forest
Open landscape
Panama Canal
Highway
Major Road
Minor Road
Track
0
2.5
5
7.5
10 km


Acoustic Complexity Index (ACI)
Acoustic Diversity Index (ADI)
Acoustic Evenness index (AEve)
Bioacoustic Index (Bio)
Acoustic Entropy Index (H)
Median of the Amplitude Envelope (M)
Normalised Difference Soundscape Index (NDSI)
% variance in standard error
100
75
50
25
0
24
48
72
96
120
144
168
Hours of recording


1 minute in 60
1 minute in 30
1 minute in 10
1 minute in 5
1 minute in 2
Continuous
% variance in standard error
75
50
25
0
0
24
48
72
96
120
144
168
Hours of recording per site


0.0003
0.0000
−0.0003
NMDS Axis 2
−0.002
0.000
0.002
0.004
NMDS Axis 1
Day
Night


ACI
ADI
AEve
Bio
H
M
NDSI
Index value
Time
00:00 06:00 12:00 18:00 24:00
1875 1850 1825 1800 1775
2.7 2.6 2.5 2.4
0.4 0.3 0.2 0.1
20 15 10 5
0.95 0.90 0.85
0.015 0.010 0.005 0.000
1.0 0.8 0.6
Continuous Forest
Fragmented Forest
Riparian
Scrub
Plantation
Pasture


ACI mean
ACI st.dev
AEve mean
H st.dev
H mean
NDSI st.dev
NDSI mean
ADI mean
M st.dev
M mean
ADI st.dev
Bio st.dev
Bio mean
AEve st.dev
Season
Diel phase
0.00
0.05
0.10
0.15
Mean decrease in accuracy if variable removed