

A peer-reviewed version of this preprint was published in PeerJ on 30 June 2016.

[View the peer-reviewed version](https://peerj.com/articles/2108) (peerj.com/articles/2108), which is the preferred citable publication unless you specifically need to cite this preprint.

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 <https://doi.org/10.7717/peerj.2108>

A New Method for Ecoacoustics? Toward the Extraction and Evaluation of Ecologically-Meaningful Sound Objects using Sparse Coding Methods

Efficient methods of biodiversity assessment and monitoring are central to ecological research and crucial in conservation management. Technological advances in remote acoustic sensing inspire new perspectives in ecology: environmental sound monitoring is emerging as a reliable non-invasive proxy for ecological complexity (Sueur and Farina, 2015). Rather than attempting to recognise species-specific calls, either manually or automatically, we are interested in monitoring the global acoustic environment, tackling the problem of diversity assessment at the community (rather than species) level. Preliminary work has attempted to make a case for community-level acoustic indices (e.g. Pieretti et al., 2011; Farina, 2014; Sueur et al., 2008b) which provide simple statistical summaries of the frequency or time domain signal. We suggest that under this approach, the opportunity to analyse spectro-temporal structural information is diminished, limiting their power both as monitoring and investigative tools. In this paper we consider sparse-coding and source separation algorithms (specifically, shift-invariant probabilistic latent component analysis in 2D) as a means to access and summarise ecologically-meaningful sound objects. In doing so we highlight a possible new approach for understanding and assessing ecologically relevant interactions within the conceptual framework of ecoacoustics.

A New Method for Ecoacoustics? Toward the Extraction and Evaluation of Ecologically-Meaningful Sound Objects using Sparse Coding Methods

Alice Eldridge¹, Michael Casey², Paola Moscoso³, and Mika Peck⁴

^{1,3,4}Dept. Evolution, Behaviour and Environment, University of Sussex, UK

²Depts. of Music and Computer Science, Dartmouth College, US

ABSTRACT

Efficient methods of biodiversity assessment and monitoring are central to ecological research and crucial in conservation management. Technological advances in remote acoustic sensing inspire new perspectives in ecology: environmental sound monitoring is emerging as a reliable non-invasive proxy for ecological complexity (Sueur and Farina, 2015). Rather than attempting to recognise species-specific calls, either manually or automatically, we are interested in monitoring the global acoustic environment, tackling the problem of diversity assessment at the community (rather than species) level. Preliminary work has attempted to make a case for community-level acoustic indices (e.g. Pieretti et al., 2011; Farina, 2014; Sueur et al., 2008b) which provide simple statistical summaries of the frequency *or* time domain signal. We suggest that under this approach, the opportunity to analyse spectro-temporal structural information is diminished, limiting their power both as monitoring and investigative tools. In this paper we consider sparse-coding and source separation algorithms (specifically, shift-invariant probabilistic latent component analysis in 2D) as a means to access and summarise ecologically-meaningful sound objects. In doing so we highlight a possible new approach for understanding and assessing ecologically relevant interactions within the conceptual framework of ecoacoustics.

Keywords: Ecoacoustics, Sparse Coding, Biodiversity, Probabilistic Latent Component Analysis, Unsupervised Learning, Acoustic Niche Hypothesis

1 INTRODUCTION

1 Biodiversity assessment is a central and urgent task, not only for research in the biological sciences, but
2 also in applied conservation biology, including major multi-lateral initiatives for promoting and protecting
3 biodiversity. At the governmental level biodiversity needs to be incorporated into national accounting by
4 2020 (Aichi Biodiversity targets A2)¹ and cost effective tools necessary to achieve this remain elusive.
5 Operating within the conceptual and methodological framework of the burgeoning field of Ecoacoustics,
6 (Sueur and Farina, 2015) we are interested in the potential for investigating the acoustic environment -
7 or soundscape - as a resource from which to infer ecological information. In this ideation paper we first
8 summarise the foundational premises for the automated ecoacoustics approach and provide an overview
9 of existing acoustic indices; a key constraint of current indices is identified and a potential new direction
10 for research is suggested. These ideas are illustrated with example analyses from a recent acoustic survey.

1.1 Current Methods for Biodiversity Assessment.

11 Biodiversity refers to the variability among living organisms and the ecological complexes of which
12 they are a part (Buckland et al., 2005) and is seen as an indicator of the 'health' of a habitat. It should
13 be noted that within traditional ecology, the question of how best to measure and describe biodiversity
14 is contentious. There are no absolute measures, rather indices reflecting difference or change across
15 or between time or space. A key part of biodiversity assessment then, is based on the estimation and
16 monitoring of changes in species composition and abundance in animal communities. There are a plethora
17

¹<http://www.cbd.int/sp/targets/>

18 of approaches to sampling and analysis (Magurran, 2004; Buckland et al., 2005; Pavoine and Bonsall,
19 2011) most of which are derived from observations of the richness and abundance of species encountered
20 in a given area at a specified time. In most strategies, there is a trade-off between data quality and quantity.
21 All Taxa Biodiversity Indices (ATBI) or multi-species field studies are desirable theoretically, but in
22 practice surveys are invariably subject to financial constraints that bind decision-makers (Lawton et al.,
23 1998). In ongoing work, we are exploring cost-effective solutions, including remote sensing (camera traps
24 and aerial photography of canopy) and identification of 'ecological-disturbance indicator species' (Caro,
25 2010). Remote sensors are an attractive choice for data collection in that they are noninvasive, scalable in
26 both space and time and remove the bias and cost associated with programs which require either experts
27 (ATBIs, Gewin, 2002) or even non-specialists (Rapid Biodiversity Assessment, Oliver and Beattie, 1993),
28 *in situ*.

29 Various forms of remote visual sensing technologies have been explored. Global satellite imaging
30 has been investigated to monitor biophysical characteristics of the earth's surface by assessing species
31 ranges and richness patterns indirectly (e.g. Wang et al., 2010). These methods are attractive, but rely
32 on expensive equipment, are difficult to adapt to small spatial scales and require a time-consuming
33 validation step. It is possible, for example, to infer valid species-level identification of canopy trees from
34 high-resolution aerial imagery, providing a means of remote sensing to assess forest status (Peck et al.,
35 2012). However, the principal weakness of this and other existing visual remote sensing methods is that
36 they cannot provide direct information on the status of taxa other than plants: they cannot detect 'silent
37 forests'. The need for innovative remote sensing methods to monitor the status of wildlife remains and
38 acoustic, rather than visual, sensors have many attractive characteristics.

39 1.2 Acoustic Approaches to Biodiversity Assessment

40 Acoustic surveys have most obvious relevance for the identification of vocal animals. Bird species
41 in particular are of interest as their importance as indicator species of environmental health has been
42 demonstrated in temperate (Gregory and Strien, 2010) and tropical (Peck et al., 2015) climates. One
43 approach is to focus on automatic species call identification, but current methods are far from reliable (e.g.
44 Skowronski and Harris, 2006, for bats), increasingly difficult in complex environments such as tropical
45 forest soundscapes, where tens of signals mix and many species still remain unknown (Riede, 1993) and
46 notoriously difficult to generalize across locations due to natural geographic variation in species' calls
47 (Towsey et al., 2013).

48 Rather than focusing on individual species, there is a growing interest in monitoring high-level
49 structure within the emerging field of *Soundscape Ecology* (Pijanowski et al., 2011) in which systematic
50 interactions between animals, humans and their environment are studied at the landscape level. From
51 this emerging perspective, the landscape's acoustic signature – the *soundscape* – is seen as a unique
52 component in the evaluation of its function, and therefore potential indicator of its status (Krause, 1987;
53 Schafer, 1977). We can consider similar processes occurring at the community level: vocalising species
54 establish an *acoustic community* when they sing at the same time at a particular place. The potential for
55 estimation of acoustic community dynamics as key to understanding what drives change in community
56 composition and species abundance is being recognised (Lellouch et al., 2014). The nascent discipline of
57 Ecoacoustics unites theoretical and practical research which aims to infer ecological information from the
58 acoustic environment at all levels (Sueur and Farina, 2015).

59 The motivations of this approach can be understood in evolutionary terms: the same competitive
60 forces which drive organisms to partition and therefore structure dimensions of their shared biophysical
61 environment (food supply, nesting locations etc.) apply in the shared sonic environment; the soundscape
62 is seen as a finite resource in which organisms (including humans) compete for spectro-temporal space.
63 These ideas were first explicitly captured in Krause's *Acoustic Niche Hypotheses* (ANH) (Krause, 1987)
64 which suggests that vocalising organisms have evolved to occupy unique spectro-temporal 'niches',
65 minimising competition and optimising intraspecific communication mechanisms. Formulated following
66 countless hours recording in pristine habitats, Krause goes so far as to posit that this spectro-temporal
67 partitioning structures the global soundscape, such that the global compositional structure is indicative
68 of the 'health' of a habitat. Crudely put, in ancient, stable ecosystems, the soundscape will comprise a
69 complex of non-overlapping signals well dispersed across spectro-temporal niches; a newly devastated
70 area would be characterised by gaps in the spectro-temporal structure; and an area of regrowth may
71 comprise competing, overlapping signals due to invasive species.

72 Krause's ANH can be understood in terms of several theories of the evolution of bird species, which
 73 are supported by field studies. Avian mating signals are thought to diverge via several processes: (1) as a
 74 by-product of morphological adaptation, the *Morphological Adaptation Hypothesis*; (2) through direct
 75 adaptation to physical features of the signalling environment, the *Acoustic Adaptation Hypothesis*; and
 76 (3) to facilitate species recognition, the *Species Recognition Hypothesis*. Field studies of the Neotropical
 77 suboscine antbird (Thamnophilidae) provide direct evidence that species recognition and ecological
 78 adaptation operate in tandem, and that the interplay between these factors drives the evolution of mating
 79 signals in suboscine birds (Seddon, 2005). The ANH is tenable in evolutionary terms, but to date we
 80 have lacked the tools for any serious experimental investigation of exactly which dimensions of acoustic
 81 ecospace niches may occupy.

82 1.3 Existing Acoustic Indices for Automated Ecoacoustics

83 This emerging framework, coupled with the technical feasibility of remote acoustic sensing and pressure to
 84 meet strategic biodiversity targets, fuels a growing research interest in ecological applications of acoustic
 85 indices; several dozen have been proposed over the last 6 years (see Sueur et al., 2014; Towsey et al., 2013;
 86 Lellouch et al., 2014, for good overviews). These are predominantly derived from statistical summaries of
 87 Fast Fourier Transforms (FFTs) or time wave representations of soundscape recordings and are motivated
 88 by slightly different approaches to measuring the 'health' of a habitat or species diversity or abundance.

89 The simplest indices provide summaries of the Sound Pressure Level (e.g. peaks, or specific times of
 90 day). In (Rodriguez et al., 2013), for example, root mean square values of raw signals from a network of
 91 recorders are used to create maps of amplitude variation to reveal spatiotemporal dynamics in a neotropical
 92 forest. The predominant approach, however, is to consider amplitude variation in time or magnitude
 93 differences between frequency bands of a spectrogram; a selection of these indices are described below.

94 Under the assumption that anthropogenic noise contribution is band-limited to a frequency range
 95 (anthrophony: 2 – 8 kHz) below that of the rest of the biological world (biophony: 0.2 – 2 kHz), the
 96 **Normalized Difference Soundscape Index** (NDSI) (Kasten et al., 2012) seeks to describe the 'health'
 97 of the habitat in terms of the level of anthropogenic disturbance by calculating the ratio (biophony -
 98 anthrophony) / (biophony + anthrophony). In long term studies, the NDSI has been shown to reflect
 99 assumed seasonal and diurnal variation in a landscape and may prove useful for observing high level, long
 100 term interactions between animals and human populations (Kasten et al., 2012). However, it does not
 101 give an estimation of local diversity *within* the range of biophony, or provide a means to investigate short
 102 term interactions in detail. Further, assumptions about frequency ranges may not generalize. For example
 103 in non-industrialized tropical climes (arguably the most precious in ecological terms) animals vocalize
 104 outside the 2-8 kHz range (Sueur et al., 2014), and industrial anthrophony is minimal.

105 A range of entropy indices are based on the assumption that the acoustic output of a community will
 106 increase in complexity with the number of singing individuals and species. A summary of the complexity
 107 of the sound is assumed to give a proxy of animal acoustic activity. Complexity here is used as a synonym
 108 of heterogeneity and many indices derive from classical ecological biodiversity indices. Shannon Entropy
 109 (Shannon and Weaver., 1949) (Equation 1) is favoured by ecologists as a measure of species diversity,
 110 where p_i is the proportion of individuals belonging to the i_{th} species in the data set of interest; it quantifies
 111 the uncertainty in predicting the species identity of an individual that is taken at random from the dataset.

$$H' = - \sum_{i=1}^R p_i \ln p_i \quad (1)$$

112 The **Acoustic Entropy Index**, H (Sueur et al., 2008b) is also calculated as the product of spectral
 113 (sh) and temporal (th) entropies, calculated on the mean spectrum and Hilbert amplitude envelope of a
 114 time wave respectively. H ranges from 0 for pure tones to 1 for high-energy, evenly distributed sound.
 115 The index was first tested against simulated choruses, generated by mixing together samples of avian
 116 vocalisations and systematically varying the number of species in each track. H values increased with
 117 species richness S following a logarithmic model. Field trials were carried out in pristine and degraded
 118 African coastal forests and H was shown to reflect assumed variation in species richness. (Sueur et al.,
 119 2008b). The study was in an area where animal acoustic activity was high and background noise low.
 120 When background noise (such as traffic) or broadband signals (such as rain, cicada or tropical cricket
 121 choruses) are higher, spectral entropy measures may give counter-intuitive results (values for white noise
 122 or motorway traffic, for example would approach 1).

123 The **Acoustic Diversity Index** (Villanueva-Rivera et al., 2011) (ADI) is a spectral entropy measure
124 which summarises the distribution of the proportion of signals across the spectrum. The FFT spectrogram
125 is divided into a number of bins (default 10) the proportion of the signals in each bin above a threshold
126 (default = 50 dBFS) is calculated. The Shannon Index (Eq.1) is then applied, where p_i is the fraction
127 of sound in each i_{th} of R frequency bands. An evenness metric, the **Acoustic Evenness Index** (AEI) is
128 similarly derived by calculating the Gini index (Gini, 1912) (commonly used by ecologists to estimate
129 species evenness) on the spectrum. These relatively simple indices are shown to effectively reflect
130 observed distinctions in gross acoustic activity, for example between dawn choruses and night activity, or
131 between diverse habitats (mature oak forest, secondary forest, wetland and agricultural land).

132 The spectral indices provide a statistical summary of the distribution of energy across the sample,
133 typically 1-10 mins are analysed at a time. These prove useful in long term studies or for observing gross
134 changes in time or space. Seeking to capture subtler changes in behaviour and composition of vocalising
135 communities, and to counter the noise-sensitivity of the entropy indices, the **Acoustic Complexity Index**
136 (ACI) was developed specifically to capture the dynamic changes in the soundscape: “many biotic sounds,
137 such as bird songs, are characterised by an intrinsic variability of intensities, while some types of human
138 generated noise (such as car passing or airplane transit) present very constant intensity values” (Pieretti
139 et al., 2011). The (ACI) is derived from measures of absolute difference in adjacent bins in a spectrogram
140 and was shown to correlate with the number of bird vocalisations in a small scale spatial study in an
141 Apennine National Park, Italy (Pieretti et al., 2011).

142 The **Bioacoustic Index** (Boelman et al., 2007) is presented as measure of avian abundance and is
143 calculated simply as the area under the mean frequency spectrum (minus the value of the lowest bin),
144 providing a measure of both the sound level and the number of frequency bands used by the avifauna. It
145 was used to investigate differences between exotic and native species in Hawaii and shown to be strongly
146 correlated with counts from direct ornithological survey when calculated for single samples taken across
147 a 6 week period.

148 These initial studies are encouraging: indices have been shown to correlate with aurally identified
149 changes in bird species richness (Depraetere et al., 2012) and reveal dynamic variation across landscape
150 [32], [28], however there are many open questions both methodologically and theoretically. Existing
151 indices are inherently likely to be affected by several factors including transitory or permanent background
152 noise, variation in distance of the animal to the microphone and relative intensity of particular species
153 call patterns. Theoretically, we are still far from understanding exactly what aspects of biodiversity these
154 indices might represent (Pijanowski et al., 2011; Sueur et al., 2008b; Servick, 2014). This is highlighted
155 in a recent temporal study of dissimilarity indices (Lellouch et al., 2014) in which indices were shown to
156 correlate well with simulated communities, but did not track community composition changes in the wild,
157 raising the question of what, if any, aspect of compositional diversity such indices represent.

158 By virtue of being based on either time-averaged spectrograms *or* amplitude changes in the time
159 domain, indices under this approach are fundamentally limited in their ability to detect *spectro-temporal*
160 patterns, which may be key to evaluating the acoustic dynamics of specific communities. Frequency-based
161 indices can pick up on crude differences in gross frequency range, but are inherently constrained in their
162 ability to detect global spectro-temporal patterns created by cohabiting species interacting in an acoustic
163 community. As the motivational premise of the community level approach assumes that spectro-temporal
164 partitioning is responsible for structuring the soundscape, this constraint may be relevant.

165 **1.4 Sparse Coding and Latent Component Analysis**

166 Time-frequency tradeoffs are an important issue in all signal processing tasks. Sparse coding is gaining
167 popularity in brain imaging, image analysis and audio classification tasks as an alternative to vector-based
168 feature representations. Depending on the dictionary used, sparse representations using overcomplete
169 dictionaries may have more time-frequency flexibility than standard Fourier transform representations.

170 The basic idea is fairly simple. For a given set of input signals, a number of atomic functions are
171 sought, such that each input signal can be approximated sparsely by a linear combination of a relatively
172 small number of atomic functions. This set of atoms is called a dictionary. In sparse coding, the number
173 of atomic functions is higher than the dimensionality of the signal such that a subset of them can span the
174 whole signal space – an overcomplete dictionary (Scholler and Purwins, 2011). Sparse approximations
175 of the signal area are then constructed by finding the “best matching” projections of multidimensional
176 data onto an over-complete dictionary, Matching Pursuit (Mallat and Zhang, 1993) (MP) being a popular

177 choice.

178 Sparse decomposition using dictionaries of atoms based on biologically informed time-frequency
 179 atoms such as Gabor and Gammatone functions – which are seen to resemble characteristics of cochlea
 180 filters – are intuitively attractive as they can provide an oriented feature set with which to approximate the
 181 original signal. This has been shown to be more efficient than Fourier or wavelet representations (Smith
 182 and Lewicki, 2005) and to provide effective and efficient input features in a range of audio discrimination
 183 tasks in everyday sounds (Adiloglu et al., 2012), drum samples (Scholler and Purwins, 2011) and similarity
 184 matching of bioacoustic data (Glotin et al., 2013).

185 Probabilistic Latent Component Analysis (PLCA) is one of a family of techniques used for source
 186 separation, which similarly provides a tool for extracting sound objects according to common frequency-
 187 amplitude statistics. PLCA is a probabilistic variant of non-negative matrix factorization (NMF) (Lee and
 188 Seung, 2001). It decomposes a non-negative matrix V into the product of two multinomial probability
 189 distributions, W and H , and a mixing weight, Z . In the auditory domain, V would be a matrix representing
 190 the time-frequency content of an audio signal:

$$V \approx WZH = \sum_{k=0}^{K-1} w_k z_k h_k^T \quad (2)$$

191 where each column of W can be thought of as a recurrent frequency template and each row of H as
 192 the excitations in time of the corresponding basis. $Z = \text{diag}(z)$ is a diagonal matrix of mixing weights
 193 z and K is the number of bases in W (Weiss and Bello, 2010). Each of V , w_k , z_k , and h_k correspond to
 194 probability distributions and are normalized to sum to 1.

195 Sparse and shift-invariant PLCA (SI-PLCA) extends PLCA to enable the extraction of multiple
 196 shift-invariant features from analysis of non-negative data of arbitrary dimensionality and was first
 197 demonstrated as an effective unsupervised tool for extracting shift-invariant features in images, audio
 198 and video (Smaragdis et al., 2008). The algorithm provides a very precise and perceptually meaningful
 199 description of content. A series of piano notes, for example, is automatically decomposed into a kernel
 200 distribution representing the harmonic series common to all notes, the peaks of the impulse distribution
 201 representing the fundamental frequency of each note and its location in time (Smaragdis et al., 2008). Weiss
 202 and Bello (2010) demonstrated application in segmentation task, showing SI-PLCA to be competitive with
 203 Hidden Markov Models and self-similarity matrices. More recently, Sarroff and Casey (2013) developed
 204 a shift and time-scale invariant PLCA which performed well against results of a human ‘groove-similarity’
 205 judgement task.

206 A common strategy used throughout the NMF literature is to favour sparse settings in order to
 207 learn parsimonious, parts-based decompositions of the data. Sparse solutions can be encouraged when
 208 estimating the parameters of the convolution matrix by imposing constraints using an appropriate prior
 209 distribution (Smaragdis et al., 2008). Under the Dirichlet distribution for example, hyper-parameters
 210 can be set to favour a sparse distribution. In these cases, the algorithm will attempt to use as few bases
 211 as possible, providing an ‘automatic relevance determination strategy’ (Weiss and Bello, 2010): The
 212 algorithm can be initialised to use many bases; the sparse prior then prunes out those that do not contribute
 213 significantly to the reconstruction of the original signal. In the context of pop song segmentation, this
 214 enables the algorithm to automatically learn the number and length of repeated patterns in a song. In
 215 soundscape analysis, might this provide an ecologically-relevant indicator of the compositional complexity
 216 of an acoustic community?

217 In Music Information Retrieval and composition tasks, SI-PLCA provides a tool for accessing
 218 perceptually relevant musical objects from time-frequency shifted patterns in a dynamic signal. From the
 219 perspective of community level ecoacoustics we are not necessarily concerned with the identification of
 220 specific species, so much as achieving a numerical description of the *qualitative* patterns of interaction
 221 between them. By way of musical analogy, we don’t care what the specific instruments of the orchestra
 222 are, rather we wish to assess characteristics of the arrangement and how the voices interact as an ensemble
 223 toward a coherent global composition through time, timbre and pitch space. Frequency-based indices
 224 may succeed in tracking species richness in simulated communities by measuring gross changes in
 225 frequency band occupancy. Perhaps their failure to track variation in species richness in the wild is
 226 because the defining feature of acoustic communities are global patterns of interaction across a more
 227 complex spectro-temporal space, rather than frequency band occupancy or amplitude variation alone.

228 Current indices based on frequency *or* amplitude statistics inherently throw away information crucial to
229 the analysis of spectro-temporal patterns: SI-PLCA provides a tool for extracting dynamic sound objects
230 grouped by common frequency-amplitude statistics, even when pitch or time shifted. That it has been
231 demonstrated to be effective in extracting the perceptually-meaningful but nebulous concept of ‘groove’
232 (Sarroff and Casey, 2013) suggests potential as a tool for beginning to interrogate the concept of the
233 acoustic niche.

234 In this paper we take a first look at how these methods might provide a complementary approach
235 to current acoustic indices for investigation of soundscape dynamics and ultimately for biodiversity
236 assessment. Taking a small sample of field recordings across different habitats in an Ecuadorian cloud
237 forest reserve we compare existing spectral and temporal indices with sample analyses of a number of
238 approaches to sparse approximation, including dictionaries built using mini-batch gradient descent, Gabor
239 functions and SI-PLCA2D. The potential value of this approach is illustrated with example reconstructions
240 from a new variant of SI-PLCA using dual dictionaries.

241 **2 METHODS AND MATERIALS**

242 **2.1 Data Collection**

243 **2.1.1 Study Area and Acoustic Survey Methods**

244 The data reported here was collected during an 8 week field survey (June - August 2014) in the Ecuadorian
245 Andean cloud forest at the Santa Lucia Cloud Forest Reserve (SLR). The SLR (0°07'30" N, 78°40'3" W)
246 is situated on the Western (Pacific) slopes of the Andes in northwestern Ecuador and spans an elevational
247 range of 1400 - 2560 m. The forest is lower montane rain forest (cloud forest). The area has a humid
248 subtropical climate and is composed of fragmented forest reserves surrounded by a matrix of cultivation
249 and pasture lands. It lies within the Tropical Andes biodiversity hotspot and exhibits high plant species
250 endemism and diversity. Topography is defined by steep-sloping valley systems of varying aspect.

251 The SLR was awarded reserve status 20 years ago, prior to which areas of Primary Forest had cleared
252 for fruit farming. The SLR therefore consists of a complex mosaic of habitat types: Ancient Primary
253 Forest (FP) punctuated by small areas of secondary regrowth of around 20 years (FS) and silvopasture (S),
254 typically elephant grass pastures used as grazing paddocks for the mules, which provide local transport.
255 These areas are less than 5 ha. In contrast to other studies where dramatically different sites have been
256 used to validate indices, this complex patchy habitat provides subtle habitat gradients.

257 Acoustic data was collected using nine digital audio field recorders Song Meter SM2+ (Wildlife
258 Acoustics), giving three replicates of each of the three habitat types. Minimum distance between recorders
259 was 300m to avoid pseudo sampling. Altitudinal range was minimised. Recording schedules captured the
260 full dawn (150 min), dusk choruses (90 min) plus midday (60 min) activity; throughout the rest of the
261 period 3min recordings were taken every 15 min and ran for a minimum of 14 days at each study site.

262 The SM2+ is a schedulable, off-line, weatherproof recorder, with two channels of omni-directional
263 microphone (flat frequency response between 20 Hz and 20 kHz). Gains were set experimentally at 36 dB
264 and recordings made at 16 bit with a sampling rate of 44.1 kHz. All recordings were pre-processed with a
265 high pass filter at 500 Hz (12 dB) to attenuate the impact of the occasional aircraft and local generator
266 noise.

267 **2.1.2 Species Identification**

268 Point counts were carried out by a local expert ornithologist and were made *in situ* in order to record
269 species seen as well as heard. A record was made for each individual, rather than individual vocalisations,
270 providing species presence-absence and abundance.

271 **2.2 Acoustic Indices**

272 For the purposes of this illustrative exercise, analyses were carried out on dawn chorus recordings from
273 just one day at one recording stations for three habitat types sampled. A range of indices described
274 in Section 1.3 were calculated: NDSI, H (including sh and th components), ADI, AEI, ACI and BI.
275 Indices were calculated for the same 10min periods during which point counts were made at each site.
276 Calculations were made using the seewave Sueur et al. (2008a)² and soundecology³ packages in R.

² Available at: <http://cran.r-project.org/web/packages/seewave/index.html>

³ Available at: <http://cran.r-project.org/web/packages/soundecology/index.html>

277 2.3 Audio Spectrum Approximation Methods

278 Three approaches to audio decomposition are illustrated using the Bregman Media Labs Audio Patch
279 Approximation Python package:⁴ Dictionary learning using mini-batch gradient learning, a Gabor
280 field dictionary and, shift invariant 2D Principle Latent Component Analysis (SI-PLCA2D). Each uses
281 Orthogonal Matching Pursuit (OMP) to build the component reconstructions. Samples were extracted
282 from analyses of 1 minute sections of the field recordings. These examples are aimed at illustrating the
283 potential of an atomic rather than vector approach in general, rather than experimental validation of any
284 particular algorithm. Default parameters were used in all cases.

285 A potential future direction is illustrated using a SI-PLCA variant (SI-PLCA2) using 2D dual dictio-
286 naries (Smaragdis and Raj., 2007; Weiss and Bello, 2010; Sarroff and Casey, 2013) based on frequency *
287 local time functions and frequency-shift * global time-activations. The expectation-maximisation (EM)
288 algorithm (Smaragdis and Raj., 2007) is used to build component reconstructions.

289 As described in Section 1.4, the algorithm returns a set of k from a K maximum components
290 ($K_{max} = 16$): independent component reconstructions, time-frequency kernels and shift-time activation
291 functions. Entropies of each are also returned. The technical details of this approach are not addressed
292 here, but example analyses are used to illustrate SI-PLCA as a potentially rich tool for future research in
293 investigating the complex quasi-periodic signals of wild soundscapes.

294 3 RESULTS AND DISCUSSION

295 3.1 Species Composition of Acoustic Communities

296 The species observations for each site, shown in Table 1, reveal little variation in overall abundance or
297 species number between the sites when seen and heard records are considered together. Several species
298 are observed in all sites; others are observed only in one habitat type. Discounting the seen-only counts,
299 the highest number of species, and individuals, was recorded at S, with least heard FP. The spectrograms
300 and mean spectrum profiles (Figure 1) for these recordings suggest that this information is present in the
301 soundscape. The number of shared species between sites results in acoustic communities with an overall
302 similar overlap, differentiated by calls of 'keynote' species. Each acoustic community occupies a broadly
303 similar frequency range, with variation in the peaks of spectral profiles according to the prevalence of
304 calls of habitat-specific species. FP appears to have lowest over-all activity, in line with the relatively
305 fewer number of species observed.

306 Despite occupying an overall similar frequency range and not differing dramatically in abundance,
307 each site is distinctly characterised by differing quasi-periodic patterns of calls. The same patterns
308 observed at the 1 minute shown continued for the full 10 minute sample⁵. The soundscape is structured,
309 not just by repetitions of specific species calls, but by turn taking, i.e. interactions *between* species. This
310 is most evident in listening, and can be observed visually as an interplay of periodic gestures in the
311 spectrogram. It is precisely this complex of interacting periodic structures which we wish to evaluate
312 under the soundscape approach, but which are impervious to analyses by current indices.

313 3.2 Acoustic Indices

314 Values for each of the acoustic indices calculated for the three habitats are given in Table 2 and shown as
315 bar plots in Figure 2. As we might expect given the minimal anthropogenic noise and broadly similar
316 spectral profile, the NDSI reports near maximum values for each site. The global complexity of each
317 scene is high; it is no surprise then that entropy indices approach 1 and differences between sites are
318 minimal. The ADI reports a small variation, following the rank-order pattern of species heard at each
319 site. Differences between Sueur's spectral, temporal and therefore overall, H entropy are minimal. ACI
320 similarly shows small variation between sites. This index in particular is very sensitive to the size of the
321 analysis window and requires further exploration to establish which aspects of community composition
322 may be being assessed. BI values report the differences in overall acoustic energy, observable in mean
323 spectrum plot (Figure 1, bottom), with the highest value at FS, FP being slightly higher than site S. These
324 basic features of the acoustic recordings are at odds with the field observations of abundance and species
325 numbers. An increase in overall energy could be due to certain individuals having intrinsically louder calls,
326 calling more frequently, or simply being closer to the microphone. In validation studies the latter could

⁴<https://github.com/bregmanstudio/audiospectrumpatchapproximation>

⁵1 min excerpts available via PeerJ linked data service

327 be countered by factoring in field-based point count distance measures (recorded, but not included here)
328 and call frequencies, as well as tallies of individuals, the latter being expedited by the use of automatic
329 segmentation software (as in Pieretti et al., 2011).

330 The key issue raised here, however is that in providing summaries of frequency *or* temporal amplitude
331 profile and magnitude differences, these current indices are not only sensitive to these largely irrelevant
332 variations in overall amplitude changes, but are all *insensitive* to the periodic structures which uniquely
333 characterise the three soundscapes.

334 3.3 Sparse-approximation Outputs

335 Dictionaries and sparse-approximations of recordings of each site using mini-batch gradient descent,
336 Gabor atoms and SI-PLCA2D are shown in Figure 3. The input for each is a log-frequency spectrogram
337 (constant-Q transform) of samples from the field recordings, as shown in Figure 1. Example dictionaries
338 (left) and sparse approximations of the input spectrogram (right) for site FP are shown for each method
339 (component reconstructions not shown). Comparing the sparse-approximation of the original spectrogram
340 for FP (see Figure 1 top image), the superior performance of SI-PLCA2D over the other two methods is
341 evident.

342 A detailed technical discussion of the methods is beyond the scope of the current paper, however of
343 key interest in this context are the qualitative differences in the dictionaries. The Gabor field dictionary
344 has an intuitive advantage over vector descriptors in representing features oriented in time-frequency
345 space. The dictionary learned under mini-batch gradient descent similarly exhibits time-frequency atoms
346 differing subtly in orientation. The SI-PLCA2D dictionary however, comprises a collection of spectrum
347 patches with a variety of micro-structures across a range of orientation and spread. In terms of the filter
348 model which motivates the use of Gabor atoms, the Gabor and mini-batch dictionaries could be described
349 as having relatively homogenous widths across the dictionary; the SI-PLCA2D dictionary by contrast
350 contains points not only differing in time-frequency orientation, but in spectral width, atoms 0, 2, 3 and 4
351 being considerably more focused than 1 and 5 (Figure 3(c), left). This is an appealing property for the
352 analysis of broad-spectrum versus pitched soundscape elements.

353 Full outputs for all three sites using the SI-PLCA2 algorithm with dual 2D dictionaries are shown in
354 Figures 4, 5 and 6. Each 10 min site recording is sampled, taking 16 time windows from across the file of
355 around 4 seconds each, arranged in order. The input is the log-frequency spectrogram of these samples, as
356 before. Extensive analysis of larger data sets across more diverse soundscapes is needed before we can
357 begin to evaluate the ecological significance or application of this approach, but a number of promising
358 observations can be made.

359 As can be seen in Figures 4(a), 5(a), 6(a), the component reconstructions appear faithful to the original
360 spectrogram. The individual component reconstructions (Figures 4(b), 5(b), 6(b)) pull out clearly distinct
361 components. This is clearest in S3 (Figure 6(b)) where the first component is broadband ambient noise,
362 and each of components 1 - 5 appear as distinct 'voices' grouped according to both spectral range and
363 spectro-temporal periodic gesture.

364 The time-frequency kernels provide a lower dimension representation of components with apparently
365 similar characteristics: compare each component in Figures 4(b) and (c), for example. The automatic
366 relevance determination feature deserves further investigation as a quick and dirty proxy for community
367 composition assessment. In this example in FP $k = 4$, FS $k = 7$ and S $k = 6$. Does K increase with the
368 number of vocalising species? Might it reflect the complexity or 'decomposability' of a scene in some
369 way?

370 The entropies of each distribution are given in the subfigure captions of Figures 4, 5 and 6. Whether
371 these can provide useful information as a difference measure either between components *within* a particular
372 reconstruction, or *between* kernels extracted from different soundscapes deserves further investigation.
373 No conclusions can be drawn from this illustrative analysis, but it raises a number of questions for future
374 research: (1) Are the component reconstructions meaningful soundscape objects in ecological terms?
375 Are the structurally distinct characteristics of geophony and biophony separated? Are vocalising species
376 separated in any meaningful ecological way (species, functionality etc.)? (2) Might the statistics generated
377 be meaningful? Does the number of components (k) returned reflect 'complexity' or 'decomposability' in
378 a way which may reflect the status of the acoustic community? Could the entropy summaries of each
379 component be used as a measure of diversity within or between communities?

380 The ability of PLCA to separate streams of distinct sound objects or voices from background noise

381 is well recognised. Within the conceptual framework of ecoacoustics, such techniques promise utility
382 at two levels. Within the acoustic community, as a possible means to investigate the composition of
383 the local soundscape in terms of dynamic interactions between spectro-temporal patterns of vocalising
384 component species, providing a new tool to begin to experimentally interrogate the concept of *acoustic*
385 *niche*. Within the conceptual framework of Soundscape Ecology, geophony (the noises created by wind,
386 rain etc. in interaction with the local landscape) is not noise to be removed from the signal, but a crucial
387 active component: the potential to separate geophony from biophony and anthrophony therefore provides
388 a means to further investigate the patterns of interaction between them.

389 4 SUMMARY AND FUTURE WORK

390 Monitoring subtle changes in complex ecosystems is crucial for ecological research and conservation but
391 far from straight forward. Acoustic indices hold promise as a rapid assessment tool, but are subject to
392 the same trade-offs as traditional ecological research of quality versus quantity: any metric necessarily
393 throws away some information. In this paper we have provided an overview of the motivational premises
394 of community-level ecoacoustics, including the concept that acoustic communities may be structured
395 according to competition across acoustic niches through spectro-temporal partitioning. We suggest that
396 existing indices operating in time *or* frequency domain may be insensitive to the dynamic patterns of
397 interaction in the soundscape which characterise specific acoustic communities and propose SI-PLCA2D
398 as a promising new tool for research. This was illustrated with example analyses of tropical dawn chorus
399 recordings along a gradient of habitat degradation.

400 It seems likely that if acoustic niches exist that they do not lie neatly along 1D vectors in the frequency
401 or time domain but dance dynamically across pitch-timbre-time space. SI-PLCA2D is computationally
402 expensive, but provides a tool for extracting shift-invariant patterns in a dynamic soundscape. In future
403 work we plan to investigate the potential for these tools to help us begin to experimentally investigate
404 the acoustic niche concept, seeking to understand what dimensions of acoustic ecospace it may occupy
405 toward the development of effective tools for rapid acoustic biodiversity assessment.

406 ACKNOWLEDGMENTS

407 Many thanks to Noé Morales of Santa Lucia for carrying out point count surveys.

408 REFERENCES

- 409 Adiloglu, K., Anniés, R., Wahlen, E., Purwins, H., and Obermayer, K. (2012). A graphical representation
410 and dissimilarity measure for basic everyday sound events. *Audio, Speech, and Language Processing,*
411 *IEEE Transactions on*, 20(5):1542–1552.
- 412 Boelman, N., Asner, G., Hart, P., and Martin, R. (2007). Multi-trophic invasion resistance in Hawaii:
413 Bioacoustics, field surveys, and airborne remote sensing. *Ecological Applications*, 17(8):2137–2144.
- 414 Buckland, S. T., Magurran, A., Green, R., and Fewster, R. M. (2005). Monitoring change in biodiversity
415 through composite indices. *Philosophical Transactions of the Royal Society of London B*, 360:243–254.
- 416 Caro, T. M. (2010). *Conservation by proxy: Indicator, umbrella, keystone, flagship, and other surrogate*
417 *species*. Island Press, Washington DC.
- 418 Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., and Sueur, J. (2012). Monitoring animal
419 diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*,
420 13(1):46–54.
- 421 Farina, A. (2014). *Soundscape Ecology: Principles, Patterns, Methods and Applications*. Springer.
- 422 Gewin, V. (2002). All living things, online. *Nature*, 418(18(6896)):362 – 363.
- 423 Gini, C. (1912). Variability and mutability: Contribution to the study of statistical distribution and
424 relations. *Studi Economico-Giuricici della R*.
- 425 Glotin, H., Sueur, J., Artières, T., Adam, O., and Razik, J. (2013). Sparse coding for scaled bioacoustics:
426 From humpback whale songs evolution to forest soundscape analyses. *The Journal of the Acoustical*
427 *Society of America*, 133(5):3311–3311.
- 428 Gregory, R. and Strien, A. (2010). Wild bird indicators: using composite population trends of birds as
429 measures of environmental health. *Ornithological Science*, 9:3 – 22.
- 430 Kasten, E., Gage, S., Fox, J., and Joo., W. (2012). The remote environmental assessment laboratory's
431 acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, 12:50–67.

- 432 Krause, B. (1987). Bioacoustics, habitat ambience in ecological balance. *Whole Earth Review*, 57:14 –
433 18.
- 434 Lawton, J., Bignell, D. E., Bolton, B., Bloemers, G. F., Eggleton, P., Hammond, P. M., Hodda, M.,
435 Holt, R. D., Larsen, T. B., Mawdsley, N. A., Stork, N. E., Srivastava, D. S., and Watt., A. D. (1998).
436 Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature*,
437 391:72–76.
- 438 Lee, D. D. and Seung, H. S. (2001). Algorithms for non-negative matrix factorization. In Leen, T.,
439 Dietterich, T., and Tresp, V., editors, *Advances in Neural Information Processing Systems 13*, pages
440 556–562. MIT Press.
- 441 Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H., and Sueur, J. (2014). Monitoring temporal change of bird
442 communities with dissimilarity acoustic indices. *Methods in Ecology and Evolution*, 4:495–505.
- 443 Magurran, A. E. (2004). *Measuring biological diversity*. Blackwell Science Ltd., Malden, USA.
- 444 Mallat, S. G. and Zhang, Z. (1993). Matching pursuits with time-frequency dictionaries. *Signal Processing*,
445 *IEEE Transactions on*, 41(12):3397–3415.
- 446 Oliver, I. and Beattie, A. J. (1993). A possible method for the rapid assessment of biodiversity. *Conserva-*
447 *tion Biology*.
- 448 Pavoine, S. and Bonsall, M. B. (2011). Measuring biodiversity to explain community assembly: a unified
449 approach. *Biol Rev Camb Philos Soc*, 86(4):792–812.
- 450 Peck, M., Maddock, S., Morales, J., Peñafiel, V., Torres-Carvajal, O., Pozo-Rivera, W., Cueva-Arroyo,
451 X., and Tolhurst., A. (2015). Cost-effectiveness of using small vertebrates as indicators of disturbance.
452 *Conservation Biology*.
- 453 Peck, M., Mariscal, A., Padbury, M., Cane, T., Kniveton, D., M., and A., C. (2012). Identifying tropical
454 ecuadorian andean trees from inter-crown pixel distributions in hyperspatial aerial imagery. *Applied*
455 *Vegetation Science*, 15(4):548 – 559.
- 456 Pieretti, N., Farina, A., and Morri., D. (2011). A new methodology to infer the singing activity of an avian
457 community: The acoustic complexity index (aci). *Ecological Indicators*, 11(3):868–873.
- 458 Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. S., Farina, A., Krause, B., Napoletano, B. M.,
459 Gage, S. H., and Pieretti, N. (2011). Soundscape ecology: the science of sound in the landscape.
460 *BioScience*, 61(3):203–216.
- 461 Riede, K. (1993). Monitoring biodiversity: Analysis of amazonian rainforest sounds. *Ambio*, pages
462 546–548.
- 463 Rodriguez, A., Gasc, A., Pavoine, S., Grandcolas, P., Gaucher, P., and Sueur., J. (2013). Temporal and
464 spatial variability of animal sound within a neotropical forest. *Ecological Informatics*.
- 465 Sarroff, A. M. and Casey, M. (2013). Groove kernels as rhythmic-acoustic motif descriptors. In
466 *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*,
467 pages 299–304.
- 468 Schafer, R. (1977). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Destiny Books,
469 Rochester, Vermont.
- 470 Scholler, S. and Purwins, H. (2011). Sparse approximations for drum sound classification sparse
471 approximations for drum sound classification sparse approximations for drum sound classification.
472 *Selected Topics in Signal Processing, IEEE Journal of*, 5(5):933 – 940.
- 473 Seddon, N. (2005). Ecological adaptation and species recognition drives vocal evolution in neotropical
474 suboscine birds. *Evolution*. 2005, 59(1):200–15.
- 475 Servick, K. (2014). Eavesdropping on ecosystems. *Science*, 343(6173):834–837.
- 476 Shannon, C. and Weaver., W. (1949). *The Mathematical Theory of Communication*. Illinois University
477 Press.
- 478 Skowronski, M. and Harris, J. (2006). Acoustic microchiroptera detection and classification using machine
479 learning: lessons learned from automatic speech recognition. *Journal of Acoustical Society of America*,
480 19(3):1817 – 1833.
- 481 Smaragdis, P. and Raj., B. (2007). Shift-invariant probabilistic latent component analysis. Technical
482 Report TR2007-009, MERL.
- 483 Smaragdis, P., Raj, B., and Shashanka., M. (2008). Sparse and shift-invariant feature extraction from
484 non-negative data. In *Proceedings of ICASSP*.
- 485 Smith, E. and Lewicki, M. S. (2005). Efficient coding of time-relative structure using spikes. *Neural*
486 *Computation*, 17(1):19–45.

- 487 Sueur, J., Aubin, T., and Simonis, C. (2008a). Seewave: a free modular tool for sound analysis and
488 synthesis. *Bioacoustics*, 18:213–226.
- 489 Sueur, J. and Farina, A. (2015). Ecoacoustics: the ecological investigation and interpretation of environ-
490 mental sound. *Biosemiotics*, pages 1–10.
- 491 Sueur, J., Farina, A., Gasc, A., Pieretti, N., and Pavoine., S. (2014). Acoustic indices for biodiversity
492 assessment and landscape investigation. *Acta Acoustica United with Acoustica*, 100:772–781.
- 493 Sueur, J., Pavoine, S., Hamerlynck, O., and Duvail, S. (2008b). Rapid acoustic survey for biodiversity
494 appraisal. *PLoS One*, 3(12):e4065.
- 495 Towsey, M., Wimmer, J., Williamson, I., I., and Roe., P. (2013). The use of acoustic indices to determine
496 avian species richness in audio-recordings of the environment. *Ecological Informatics*.
- 497 Villanueva-Rivera, L., Pijanowski, B., Doucette, J., and Pekin., B. (2011). A primer of acoustic analysis
498 for landscape ecologists. *Landscape Ecology*, 26:1233–1246.
- 499 Wang, K., F., S., Guo, X., and Cattet, M. (2010). Remote sensing of ecology, biodiversity and conservation:
500 a review from the perspective of remote sensing specialists. *Sensors*, 10:9647 – 9667.
- 501 Weiss, R. J. and Bello, J. P. (2010). Identifying repeated patterns in music using sparse convolutive
502 non-negative matrix factorization. In *In Proc. International Conference on Music Information Retrieval*
503 *(ISMIR)*.

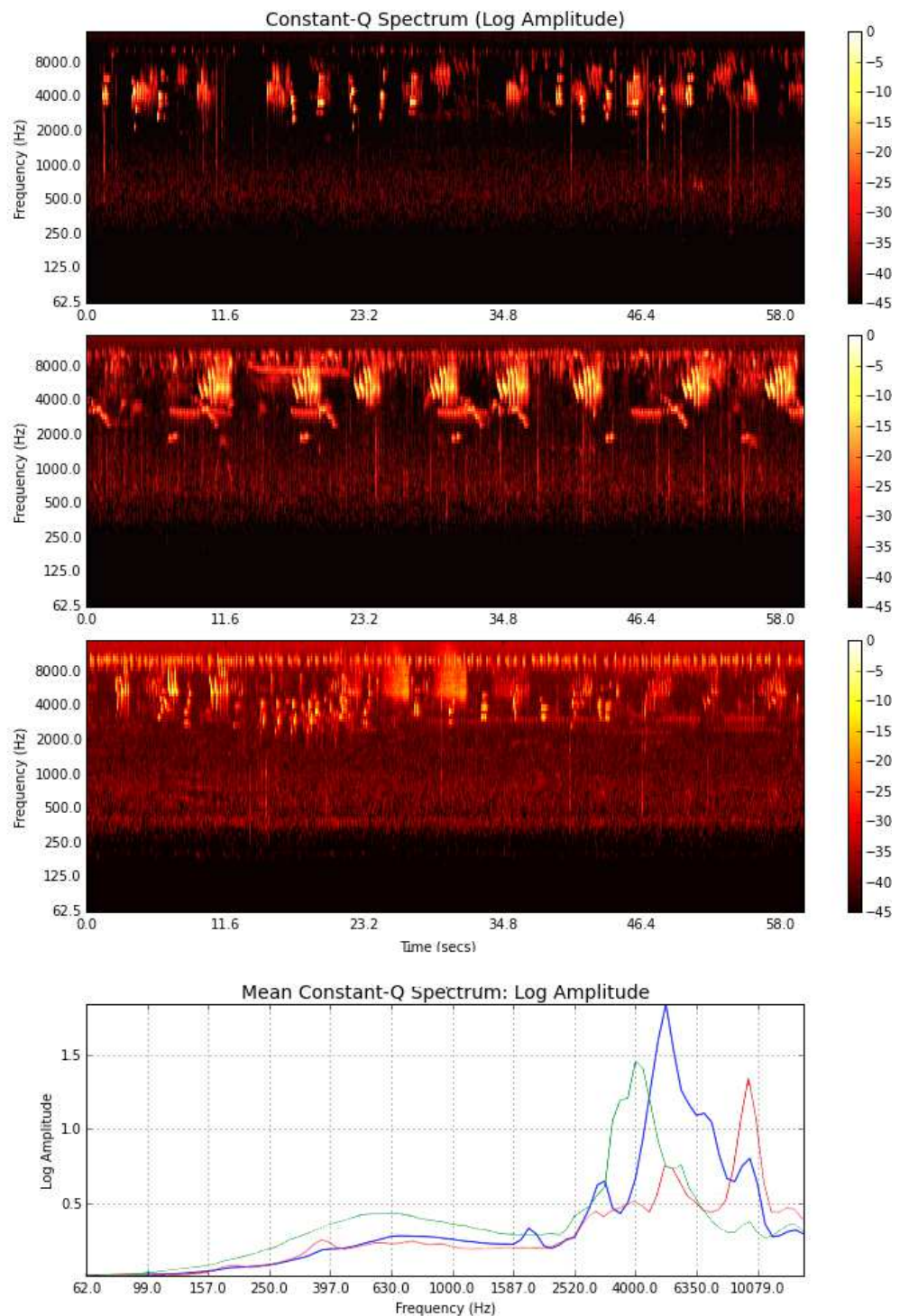


Figure 1. Top: Constant-Q spectrograms for 1 min excerpts for each site FP (top) FS (middle) S (bottom). 1 minute resolutions are presented to illustrate the periodic call patterns. These were consistent across the 10 min sampling time in each habitat. Bottom: Mean Constant-Q Spectrum (log amplitude) for FP (green), FS (blue) and S (red)

Common Name	Heard			Seen		
	FP	FS	S	FP	FS	S
Andean Solitaire	1	1	1	1	-	-
Azara's Spinetail	-	-	1	-	-	-
Beryl-spangled Tanager	-	-	-	1	1	-
Blue-winged Mountain-Tanager	-	1	-	-	-	1
Booted Racket-tail	-	1	-	1	-	-
Brown Inca	1	-	-	-	-	-
Brown-capped Vireo	1	-	-	-	-	-
Collared Forest-Falcon	1	-	-	-	-	-
Dusky Bush-Tanager	-	-	1	1	-	-
Empress Brilliant	-	-	1	-	-	1
Flame-faced Tanager	-	-	-	1	1	-
Golden-Crowned Flycatcher	-	-	1	-	-	-
Golden-Headed Quetzal	-	1	1	-	-	-
Gray-breasted Wood-wren	1	1	1	-	-	-
Immaculate Antbird	-	1	-	-	-	-
Lineated Foliage-Gleaner	-	1	1	1	-	-
Long-tailed Antbird	1	-	1	-	-	-
Masked Trogon	1	1	1	-	-	-
Metallic-Green Tanager	-	-	1	-	-	-
Nariño Tapaculo	-	1	-	-	-	-
Orange-bellied Euphonia	1	1	-	1	-	-
Plumbeous Pigeon	1	1	1	-	-	1
Red-faced Spinetail	-	1	1	-	-	-
Roadside Hawk	-	-	1	-	-	-
Rufous-breasted Antthrush	-	1	-	-	-	-
Russet-crowned Warbler	1	1	-	-	-	-
Scale-crested Pygmy-Tyrant	-	1	1	-	-	-
Slate-throated Whitestart	1	-	1	-	-	-
Smoke-colored Pewee	-	-	1	-	-	-
Three-striped Warbler	-	1	-	-	-	-
Toucan Barbet	-	-	1	-	-	-
Tricolored Brush-Finch	-	1	1	-	-	-
Tyrannine Woodcreeper	-	-	-	1	-	-
Uniform Antshrike	-	-	1	-	-	-
Wattled Guan	1	1	1	-	-	-
White-capped Parrot	1	-	-	-	-	-
Total	13	18	21	8	2	3

Table 1. Field observations for species heard and seen at FP (0600-0610), FS (0619-0629) and S(0639-0649) on June 15th 2014. As only 1 individual of each species was recorded, this data provides a record of species absence-presence as well as abundance for each site.

	NDSI	ADI	AEI	sh	th	H	ACI	BI
FP	0.9716	2.1919	0.2591	0.9567	0.9730	0.9309	18497.14	7.5393
FS	0.9727	2.2684	0.1418	0.9355	0.9729	0.9102	18315.61	11.1780
S	0.9809	2.2909	0.0749	0.9539	0.9825	0.9372	18686.78	6.5867

Table 2. Acoustic indices values for the three study sites: FP, FS and S shown to 4 decimal places.

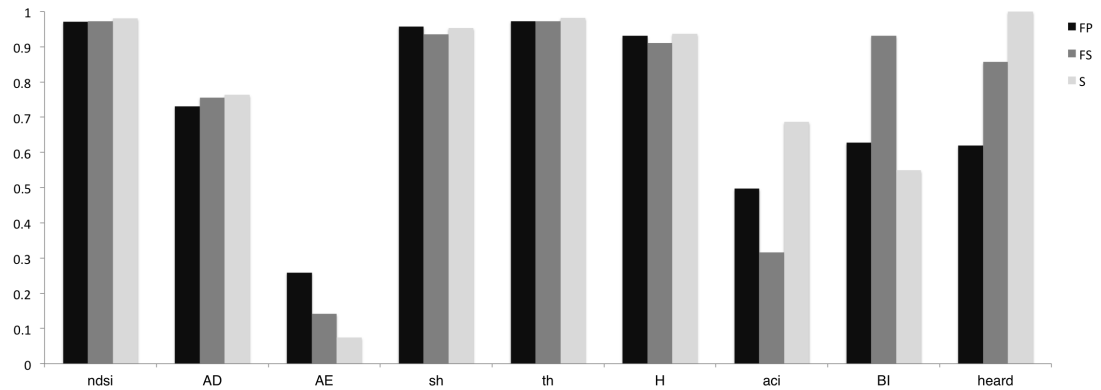
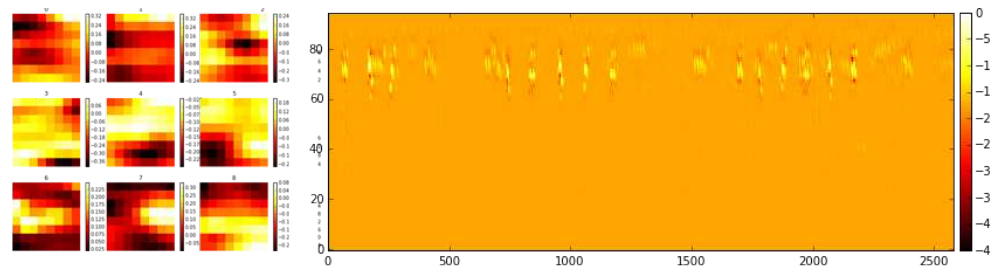
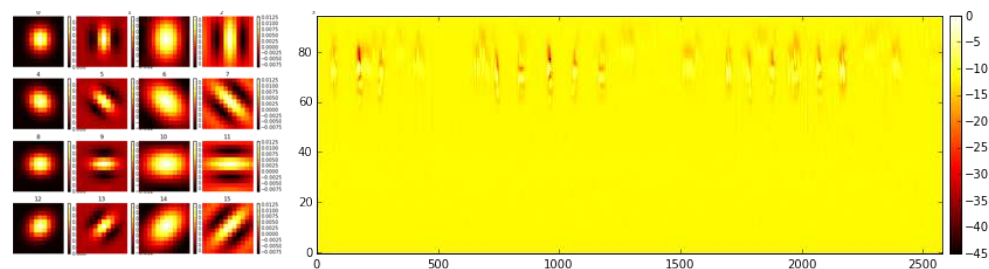


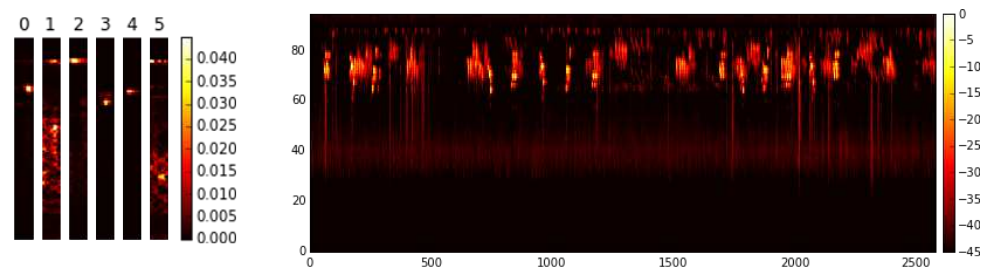
Figure 2. Bar plots of indices results for same recording showing indices values 10 minutes of dawn chorus at FP (black), FS (grey) and S (light grey) for NDSI, ADI, AEI, sh, th, H, ACI and BI. Values for AD, ACI and BI are scaled in the ranges 0:3, 1800:1900 and 0:12 respectively. Point count ‘heard’ data values for each site are given in the end column for comparison.



(a) Mini-batch Gradient Descent

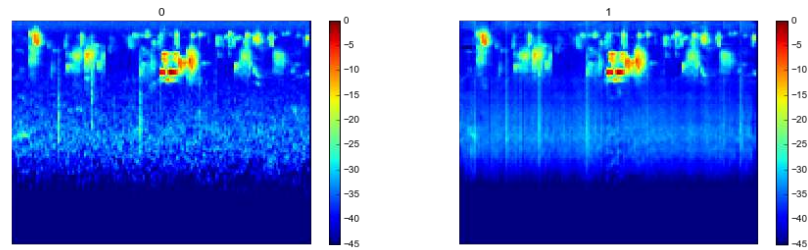


(b) Gabor Field

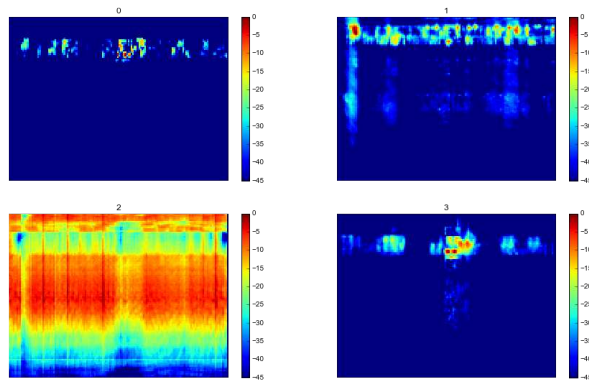


(c) SI-PLCA2D

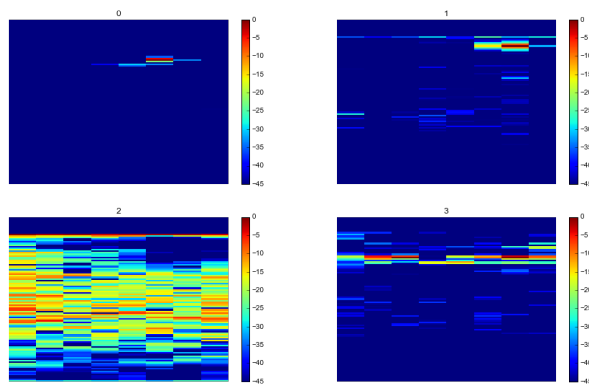
Figure 3. Over-complete Dictionaries (left) and sparse-approximations of original spectrogram shown in Figure 1 for (a) Mini-batch gradient learning, (b) Gabor Field Dictionary and (c) SI-PLCA2D Component Dictionary.



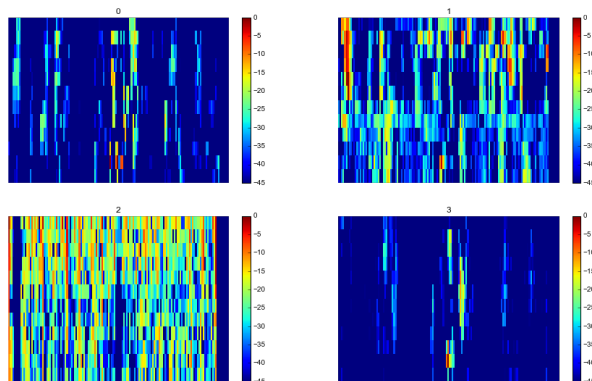
(a) FP: Original Spectrum (left) and Component Reconstructions (right) Original $S = 4.9439$; Reconstruction. $S = 4.9303$



(b) Individual Component Reconstructions (FP). $S = (5.3347, 5.5418, 0.0000, 5.4501)$

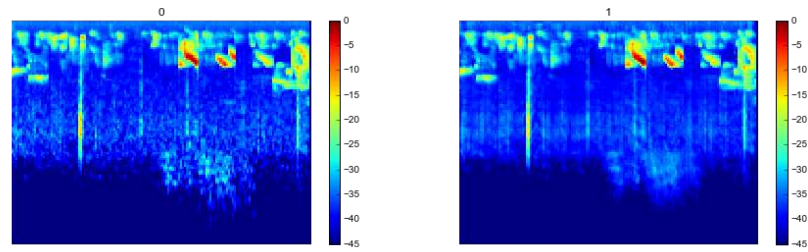


(c) Time-Frequency Kernels (FP). $S = (1.4749, 3.7474, 5.7675, 3.7552)$

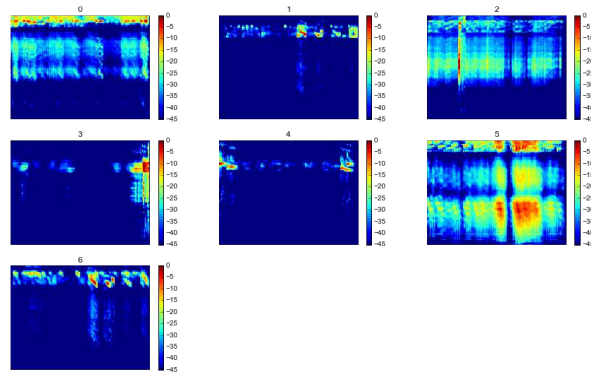


(d) Activation (shift-time) Functions (FP). $S = (5.0762, 5.7370, 6.2901, 4.6718)$

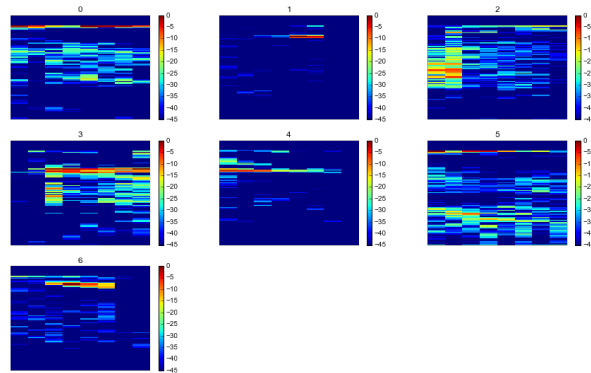
Figure 4. SIPLCA2 outputs for Primary Forest site dawn chorus. Entropy (S) values are shown in brackets



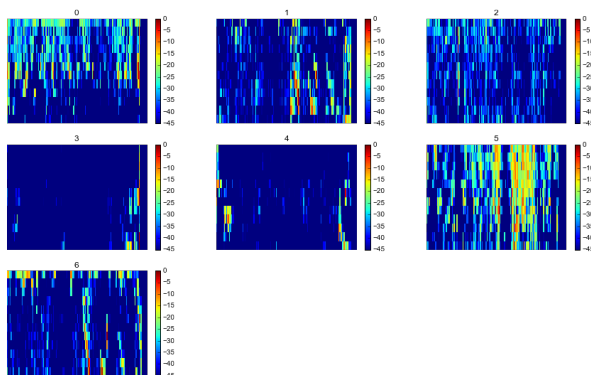
(a) FS: Original Spectrum (left) and Component Reconstructions (right) Original $S = 5.0291$; Reconstruction. $S = 4.9576$



(b) Individual Component Reconstructions (FS). $S = (5.4921, 5.4735, 4.3225, 5.5836, 5.1488, 4.7036, 5.6571)$

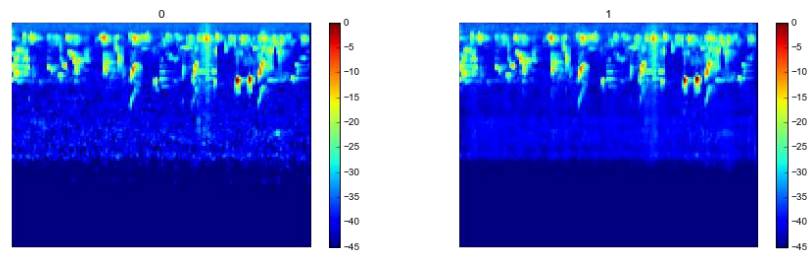


(c) Time-Frequency Kernels (FS). $S = (4.5131, 2.9622, 4.9541, 4.6132, 3.4991, 5.2196, 3.9248)$

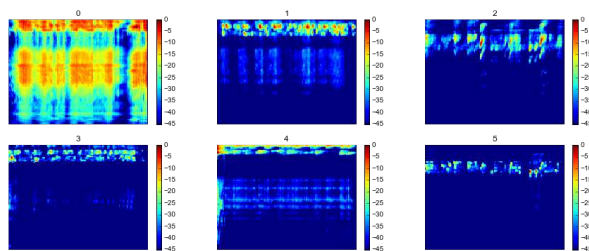


(d) Activation (shift-time) Functions (FS). $S = (5.7739, 5.1516, 5.9054, 3.6572, 4.2803, 6.0051, 4.9906)$

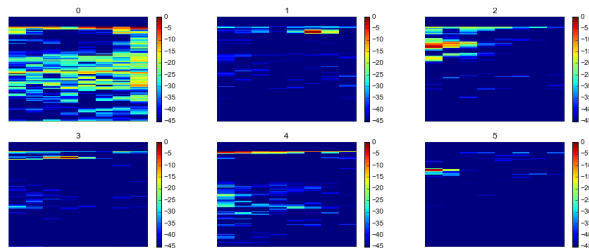
Figure 5. SI-PLCA2 outputs for Secondary Forest site dawn chorus. Entropy (S) values are shown in brackets.



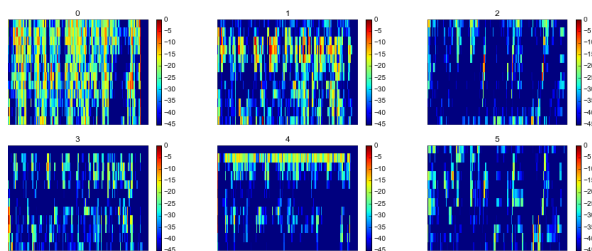
(a) S : Original Spectrum (left) and Component Reconstructions (right) Original. $S = 5.1797$; Reconstruction $S = 5.1715$



(b) Individual Component Reconstructions (S). $S = (0.6920, 5.8020, 5.5895, 5.6326, 5.9547, 5.4067)$



(c) Time-Frequency Kernels (S). $S = (5.4939, 3.8297, 3.9353, 3.7362, 4.0415, 3.1377)$



(d) Activation (shift-time) Functions (S). $S = (6.2345, 5.7911, 5.1608, 5.4627, 5.3665, 5.3127)$

Figure 6. SI-PLCA2 outputs for Silvopasture site dawn chorus. Entropy (S) values are shown in brackets.