UNIVERSIDAD
DE ANTIOQUIA
1 8 0 3

METHODOLOGY FOR ECOSYSTEM CHANGE ASSESSING USING ECOACOUSTICS ANALYSIS

Diana Carolina Duque Montoya

Universidad de Antioquia UDEA
Facultad de Ingeniería
Medellín, Colombia
2019

METHODOLOGY FOR ECOSYSTEM CHANGE ASSESSING USING ECOACOUSTICS ANALYSIS

Diana Carolina Duque Montoya

In partial fulfillment of the requirements for the degree of

Master's in engineering

Advisor

Claudia Victoria Isaza Narváez

Research Group

Sistemas Embebidos e Inteligencia Computacional

-SISTEMIC-

Universidad de Antioquia UDEA

Facultad de Ingeniería

Medellín, Colombia

2019

This thesis is dedicated to my family

Thank you for supporting me in all the decisions I made up to this point

Acknowledgments

First, I would like to thank God for guiding me through His will, I would have never dare to take all these steps without His help. I would also like to thank my advisor, Claudia Isaza, who patiently leaded me to overcome my failures, advised me for every decision I made on this journey and helped me to improve the quality of my work. Thanks also to the people in SISTEMIC for providing me with everything I needed to accomplish all the goals during my master, specially to the professor Sebastián Isaza, who mentored me for developing my teaching skills and accepted me as his assistant in the informatics I class; in this job I learned new tools for the development of computer algorithms and could have a financial support during my studies.

Special thanks to the interdisciplinary group formed by members of Grupo Herpetológico de Antioquia and SISTEMIC: Juan Manuel Daza, Claudia Isaza, Estefany Cano, William Gómez, Camilo Sánchez and Juan Parra, who helped me to understand the main definitions for ecology, ecoacoustics, artificial intelligence and the importance of an accurate ecosystem state assessment. They were also responsible for setting and executing the guidelines for recording the locations around ISAGEN facilities; company that I would also like to thank. Thanks also to the Alexander von Humboldt institute, for providing another set of recordings for the validation of my work.

Finally, I would like to thank the friends I made during the past couple of years, who made this experience much more pleasant and to my boyfriend and family, who supported me with their unconditional love during this time of my life and will continue to do so for the rest of it.

Abstract

Ecoacoustics has become a field of growing interest for ecosystem monitoring. Its main advantages over traditional methods include cost effectiveness, non-invasiveness and simplicity; besides the distribution of many recorder units makes possible the recollection of more information. However, for long term studies, the quantity of collected data makes the manual inspection of recordings a cumbersome task, leading to reduced analysis. As an alternative to manual inspection, a series of indices have been proposed to summarize the acoustical information in recordings. Nonetheless, these indices have been applied mainly to biodiversity studies and their connection to ecosystem state is still not clear.

In this work we trusted ANOVA robustness for non-normal data for proposing a methodology that selected the best acoustical indices or features for a specific application and used them to model the ecosystem soundscape patterns with hidden Markov models and Gaussian mixture emissions (GMMHMM). Additionally, the set of input features included by default a biodiversity indicator per 1kHz band.

This methodology was applied to two Colombian cases with defined ecosystem types. In the first case, a series of forest, stubble and pasture recordings were collected for over a year in the east of Antioquia. The second application aimed to find the soundscape patterns of dry forests transformations in two regions of the Colombian Caribbean. The model identified six and three soundscape patterns for the first and second dataset respectively. In the first application, continuous sounds, high biophonic intensity and multiple occupied frequency bands were found in the patterns associated to forest sites; on the other hand, stubble sites presented more general entropy, which we related to high geophonic presence, preventing biophonic activity. Lastly, pasture soundscapes alternated between periods of high geophony and high frequency complexity, making it an intermediate ecosystem in the acoustical sense. The adaptation of the model for classification resulted in the identification of 81% of the forest samples, 96.6% of the stubble samples and 51.2% of the pasture samples.

The classification results for the second application were not as high, with 68% for the low transformation samples, 58.8% for the medium transformation and 31.8% for the high transformation. Nonetheless, the confusion matrices indicated that the training samples were not enough, and more sampling should be provided for attaining better results.

Given that GMMHMM is a sequential model, it also presented the temporal configuration of the acoustical patterns by their transition probabilities. This feature allowed us to emphasize the importance of conservation, when we found that the most stable and inaccessible states were associated to the most acoustically diverse ecosystems.

Resumen

La ecoacústica se ha convertido en un área de creciente interés para el monitoreo de ecosistemas. Entre las principales ventajas que presenta sobre las técnicas tradicionales se encuentran su bajo costo, poca afectación al entorno y simplicidad; además de que la distribución de varias grabadoras hace posible la recolección de más información. Sin embargo, para estudios de largo plazo, la cantidad de datos hace que la inspección manual de las grabaciones sea una tarea tediosa y por consiguiente el análisis sea limitado. Como alternativa a la inspección manual, una serie de índices han sido propuestos para resumir la información acústica de las grabaciones. No obstante, estos índices han sido aplicados principalmente a estudios de biodiversidad y su relación con el estado del ecosistema no es claro aún.

En este trabajo se confió en la robustez del ANOVA frente a datos que no se distribuyen normalmente para proponer una metodología de selección de los mejores índices o descriptores acústicos para una aplicación específica y usarlos para modelar los patrones del paisaje acústico del ecosistema con modelos ocultos de Markov y emisiones por mezclas Gaussianas (GMMHMM). Además, el conjunto de descriptores que entran al modelo incluye por defecto un indicador de biodiversidad para cada banda de 1kHz.

Esta metodología fue aplicada a dos casos colombianos con tipos de ecosistema definidos. En el primer caso, una serie de grabaciones de bosque, rastrojo y pastizal fueron colectados por más de un año en el este de Antioquia. La segunda aplicación buscaba encontrar patrones de paisaje acústico de las transformaciones de bosque seco en dos regiones del caribe colombiano. El modelo identificó seis y tres patrones acústicos para la primera y segunda base de datos respectivamente. En la primera aplicación, se encontraron sonidos continuos, alta intensidad biofónica y ocupación de varias bandas en los patrones asociados a bosque, mientras que en los rastrojos se presentó más entropía, que se relaciona con alta presencia geofónica, lo que limita la actividad biofónica. Finalmente, los paisajes acústicos de pastizal alternaron entre periodos de alta geofonía y alta complejidad frecuencial, haciéndolo un ecosistema intermedio en el sentido acústico. La adaptación del modelo para clasificación resultó en la identificación del 81% de las muestras de bosque, 96,6 % de las muestras de rastrojo y 51,2 % de las muestras de pastizal.

Los resultados de clasificación para la segunda aplicación no fueron altos, con 68% para las muestras de baja transformación, 58,9% para la transformación media y 31,8% para la transformación alta. No obstante, las matrices de confusión indicaron que las muestras de entrenamiento no fueron suficientes, y que debería proporcionarse mayor muestreo para obtener mejores resultados.

Dado que GMMHMM es un modelo secuencial, también presentó la configuración temporal de los patrones acústicos dadas sus probabilidades de transición. Esta característica nos permitió destacar la importancia de la conservación, cuando encontramos que los estados más estables e inaccesibles fueron asociados a los ecosistemas más diversos acústicamente.

Contents

List of Figures

Lists of Tables

Conventions Index

$\widehat{\ }$	New estimated value for a variable
$'$	Transpose of an array variable
$\bar{\ }$	Mean value for a variable
$\lceil\ \rceil$	Nearest integer greater than a variable
$\lfloor\ \rfloor$	Nearest integer smaller than a variable
α - $\alpha_i(t)$	Technophony level - Joint probability for a partial starting sequence finishing in the state i and time t
a_{ij}	Probability of transition from state i to state j
A	Transition matrix of λ
$ACIf_t$	Acoustic Complexity Index (frequency-time)
$ACIf_{te}$	$ACIf_t$ evenness
$ACIt_f$	Acoustic Complexity Index (time-frequency)
$ACIt_{fe}$	$ACIt_f$ evenness
ADI	Acoustic Diversity Index
ADI_b	Modified acoustic diversity index for band b
AEI	Acoustic Evenness Index
ANOVA	Analysis of Variance
β - $\beta_i(t)$	Bioacoustics Index, biophony level - Joint probability for the partial ending sequence starting at $t+1$ and finishing at T, given state i and time t
$b_{(\#)} - b_i(y_t)$	A specific frequency band – the emission probability of state i to observation y_t
B	Emission Probability Distribution of λ
$BN_{f(j)}$	Frequency Background Noise (in the bin j)
BN_t	Temporal Background Noise
$\gamma_i(t)$	Probability of being in state i at time t
$\gamma_{im}^{e}(t)$	Probability for the e^{th} sequence that the m^{th} component of the i^{th} state generated observation y_t
c	Number of temporal clusters in the recording
C	Number of likely values for N_m
CF	Crest Factor
$\delta_t(i)$	Highest probability for the first t observations in the sequence ending in state i
d	Audio signal duration – if preceding another variable, it represents the variable's derivative
D	Signal digitalization depth
e	A specific sequence
$\varepsilon_{(i)}$	Hilbert transform (evaluated in time step i)
E	Number of sequences
E_{peak}	Peak energy value of the signal
$\phi(y_t \mid \mu_{im}, \Sigma_{im})$	The probability of y_t to be generated by a multivariate gaussian distribution with mean vector μ_{im} and covariance matrix Σ_{im}

$f_{(j)}$	A specific frequency value (correspondent to frequency bin j)
F_f	P_f for P_f values higher than a certain threshold, 0 otherwise
FFT	Fast Fourier Transform
FM	Frequency Modulation
Fs	Frequency of sampling for an audio signal
ζ	A specific recording session
g	A specific group (label)
G	Number of likely values for N_c
GHA	Grupo Herpetológico de Antioquia
GMMHMM	Hidden Markov Models with Gaussian mixture emissions
H	Acoustic Entropy Index
H_m	Entropy of spectral maxima
HMM	Hidden Markov Model(s)
H_s	Spectral Entropy
H_t	Temporal Entropy
H_v	Entropy of spectral variance
i	A specific time step or state in the HMM context
IDEAM	Colombian climate authority
$I_{i,j}$	Normalized amplitude of each pulse in the time step i and frequency bin j.
j	A specific frequency bin or state in the HMM context
k	A specific temporal cluster for $ACIt_f$ calculation
λ	A certain HMM
Λ_j	Denotes major peaks in the mean spectrum
$\mu_{i(m)}$	Mean vector (of the m^{th} gaussian distribution) for the state i
m	Frequency bins at the left of j – A certain gaussian component
l	Frequency bins surrounding j
M	Median of amplitude envelope
$MAPC_{(Y)}$	Mean absolute Pearson coefficient for feature Y
MD	Musicality Degree
MFCC	Mel Frequency Cepstral Coefficients
MID	Mid-band activity
MSE	Mean Squared Error
nan	Not a number
N_b	Total number of frequency bands to be considered
N_c	Number of states in the model
NDSI	Normalized Difference Soundscape Index
$nfft$	Number of points for the FFT
N_f	Total number of frequency bins
N_g	Number of groups
N_m	Number of gaussian components in the GMMHMM
NP	Number of peaks
N_t	Total number of temporal steps

N_{tr}	Total number of recordings in the training set
$\xi_{ij}(t)$	Probability of being in state i and j in time t and $t+1$ respectively
$\pi_{(i)}$	Initial probability distribution of λ (initial probability for state i)
$p_{b(d)} - p_{im}$	Occupation of the frequency band b (or d) – the weight for the m^{th} gaussian distribution in the i^{th} state
$\hat{P}$	Estimated probability for the given sequence
$PCC_{(YZ)}$	Pearson Correlation Coefficient (for Y and Z)
P_f	Temporal spectrogram average for frequency f
$PSD(n_{r\zeta})$	Power Spectral Density (averaged over the band between 600 Hz and 1200 Hz for recording r in session ζ)
ρ	Ratio of biophony to anthrophony
r	A specific recording
R_ζ	Total number of recordings in the session ζ
R_G	Number of times that the Baum-Welch algorithm is run
RMS	Root mean square
ROI	Region of Interest
σ_{eq}	Energy-equivalent sound pressure level standard deviation
σ_L	Instantaneous sound pressure level standard deviation
$\sigma_{Y(Z)}$	Variance for index Y (Z)
$\sigma_{Y,Z}$	Covariance for indices Y and Z
Σ_{im}	Covariance matrix of the m^{th} gaussian distribution in the state i
S_j	Mean spectrum value for frequency bin j
$s_{i,w}$	Spectrogram cell in the time step i and frequency bin w
SF	Spectral Flatness
SPL	Sound Pressure Level signal
SSE	Sum of Squared Errors
STFT	Short-Time Fourier Transform
t	A specific time in the sequence
T	Time for the last element in the sequence
T_e	Time duration for sequence e
thr_ζ	Threshold for heavy rain detection in the session ζ
TLSV	Temporal Sound Level Variance Descriptor
U_j	Maximum value of the spectrogram in the frequency bin j
v_t	Assigned state to the observation y_t
V_j	Spectrogram variance along frequency bin j
ω	Window size for the spectrogram computation
w	A specific frequency bin (different to j)
W_j	Welch's power spectral density estimate (for frequency bin j)
x	Raw audio signal
$\psi_t(i)$	Temporary vector keeping track of the state sequence that maximizes $\delta_t(i)$
Y (Z)	Denotes a specific index
$\bar{y}$	Index mean for all samples

y_g	Index mean for the group g
y_{gr}	Index value for the recording r in the group g
y_t	Observation vector in time t

Chapter 1. Introduction

1.1. Problem Description and Justification

Sound, as an important indicator for ecosystem activity, can be of great help to monitor the natural periodicity in the ecosystem and its degradation due to industrial pollution or climate change (Joo, Gage, & Kasten, 2011; B. L. Krause & Farina, 2016). Currently, there are many tools to automatize the collection of recordings and analyze the data (Aide et al., 2013; Kasten, Gage, Fox, & Joo, 2012; Merchant et al., 2015; Sueur, Farina, Gasc, Pieretti, & Pavoine, 2014) but given the diversity of applications and the different approaches to address them, there is still not a clear consensus about the best methodology for extracting the ecosystem acoustical patterns.

Summarizing the soundscape behavior for specific sites generally relies on the analysis of certain acoustic metrics or indices. These analysis have been carried on visually (Gage & Axel, 2014; Gaitán Forero, 2017; Job, Myers, Naghshineh, & Gill, 2016; Razali et al., 2015; Mangalam Sankupellay, Towsey, Truskinger, & Roe, 2015) or applying statistical tools like correlation or linear models (Fuller, Axel, Tucker, & Gage, 2015; Gage, Wimmer, Tarrant, & Grace, 2017; Jorge, Machado, da Cunha Nogueira, & Nogueira-Filho, 2018; Bernie Krause, Gage, & Joo, 2011; Pekin, Jung, Villanueva-Rivera, Pijanowski, & Ahumada, 2012; Tucker, Gage, Williamson, & Fuller, 2014). Although some of the studies have employed machine learning techniques for grouping the acoustic measurements (Mangalam Sankupellay et al., 2015; Tucker et al., 2014); these techniques have been used rather rarely for soundscape pattern recognition (Bormpoudakis, Sueur, & Pantis, 2013; Mullet, Gage, Morton, & Huettmann, 2016).

Few of the studies cited above analyze temporal variation of soundscapes beyond daily fluctuations. Those that do, seemed to describe accurately the respective seasonal soundscape patterns, but some presented analysis lacking predictive power; then, for these methodologies to detect early anomalies in the ecosystem acoustic signature is more difficult (Gage & Axel, 2014; Bernie Krause et al., 2011). Only one study proposed predictive models for each of the soundscape components, instead of building a model for the soundscape as a whole (Mullet et al., 2016). In this work, we proposed such a model by the application of predictive machine learning algorithms.

The region where the recordings are collected is another factor to consider. The acoustic richness is expected to increase as we approach to the Equator, as this region favors the conditions for life and diversity of species (Gaston, 2000). Nonetheless, few ecoacoustic studies are made on equatorial ecosystems.

In this work, we describe a new methodology for assessing the temporal and structural changes in tropical soundscapes. For this purpose, two Colombian databases were used for building the model; the first one was collected from the western mountain chain of the Andes; the second one contained recordings from the Caribbean dry forests. Firstly, the acoustic metrics were selected according to the ecosystem types of each database; then a model based on hidden Markov models (HMM) (Rabiner, 1989) was trained. HMM is a statistical model that can be fit as a machine learning model; it also groups the samples by acoustical similarity

and stablishes relations between groups (components or states), in order to predict the most likely future state for a given soundscape pattern.

1.2. Background

Sound is an important indicator in ecological studies. Through sound, many species can interact with the environment, mark their territory, start mating rituals, and detect preys or predators (Mcwilliam & Hawkins, 2013). The study of animal communication by analyzing the sounds they produce has become into a science called Bioacoustics (Fletcher, 2007).Thanks to bioacoustics applications, it has been possible to discover new aspects of animal behavior, such as the differences between female and male vocalizations (Langbauer, 2000) or preferred position for making alarm calls (Collier, Blumstein, Girod, & Taylor, 2010), etc.

The development of sophisticated tools for efficient recording has allowed the enhancement of bioacoustics studies, leading to the analysis of higher organizational levels in ecology. In the past, the studies focused mainly on species behavior (Erbe, King, Yedlin, & Farmer, 1999; Leong, Ortolani, Burks K. D, Mellen, & Savage, 2003; Marcellini, 1974), now there are multiple studies at the population and community level (Baker & Logue, 2003; De Solla, Fernie, Barrett, & Bishop, 2006; Hobson, Rempel, Greenwood, Turnbull, & Van Wilgenburg, 2002; Tobias et al., 2010). Currently, the goal is to take the next step and extract information from the ecosystem level.

To study the ecosystem acoustic dynamics, it is necessary to include in the analysis sounds from non-biological sources, which are proven to modify the communication patterns in living organisms (B. Krause, 1987) and even help them choose their habitat (Slabbekoorn & Bouton, 2008). In addition, the seasonality of ecosystem soundscape must be characterized, to understand its natural cycles and recognize atypical behavior. For the first purpose, three types of sounds are to be analyzed: biological sounds, or biophony; sounds produced by physical phenomena, such as wind, rain, etc. or geophony, and sounds produced by human-made machinery, which are called technophony or anthrophony. The collection of these sounds is the definition of soundscape for ecological studies (Pijanowski, Villanueva-Rivera, et al., 2011).

The study of soundscape surpasses the bioacoustics domain, because this area is focused mainly on biological sound. On the other hand, the area of soundscape ecology is specifically addressed to landscape ecology studies (Pijanowski, Villanueva-Rivera, et al., 2011). Hence, a new term was needed to enclose the analysis of soundscape for ecosystem studies, and it has now been defined as Ecoacoustics (Sueur & Farina, 2015). Consequently, ecoacoustics concerns both bioacoustics and soundscape ecology and other spatiotemporal, large-scaled studies intended to respond questions about the ecosystem through sound.

For ecoacoustics studies to include spatiotemporal ecosystem change, a big amount of data is needed, because the perturbation gradient is only visible through long periods of time (Magurran et al., 2010). Additionally, big quantities of data demand more accuracy for acoustical measurements. Over the past few years, many acoustic metrics have been proposed to describe the ecological components in sound and summarize acoustical data (Sueur et al., 2014). Most of them were design to estimate biodiversity (Pieretti, Farina, & Morri, 2011; Sueur, Pavoine, Hamerlynck, & Duvail, 2008; Villanueva-Rivera, Pijanowski, Doucette, & Pekin, 2011), while only a few

consider other soundscape elements in the analysis (Kasten et al., 2012; Qi, Gage, Joo, Napoletano, & Biswas, 2008). Although these indices convey important information about the ecosystem dynamics, their relation to ecosystem state and communities' distribution is not clear yet.

In other acoustical sciences such as psychoacoustics and acoustic design, the relationship between the listeners and the acoustical surroundings is explored (Everest & Shaw, 2001; Truax, 1998). Particularly, in psychoacoustics, three components of sound are considered to be the most important for analyzing its effects on people: intensity, frequency changes and temporal structure (De Coensel, Botteldooren, Debacq, Nilsson, & Berglund, 2007). Although these studies focus on human perception of sound, other studies have adapted some psychoacoustical metrics to measure other species hearing (Bedoya, Isaza, Daza, & López, 2014; Blumstein et al., 2011; Delaney, Grubb, Beier, Pater, & Hildegard Reiser, 1999; Fuller et al., 2015). Given that most of ecoacoustics indices are based on spectrograms, which is a representation of sound intensity in time and frequency, the combination of these three components should be a good criterion to find the right model for soundscape characterization.

Several types of techniques have been applied to build an ecoacoustic model for ecosystem assessment. Some are based on statistical models (Fuller et al., 2015; Bernie Krause et al., 2011; Tucker et al., 2014) and others are based on machine learning techniques (De Coensel, Botteldooren, Debacq, Nilsson, & Berglund, 2008; Mullet et al., 2016; Rychtáriková & Vermeir, 2013). Machine learning techniques are being preferred recently for their versatility, practicality and adaptability to any type of problem. A subgroup of these techniques, known as unsupervised learning or clustering, are especially useful for problems where the data patterns are not clear, like soundscape analysis, because they are able to automatically identify clusters by grouping similar data (Bishop, 2006).

However, in ecoacoustic modeling studies, few machine learning approaches have taken into account the temporal dimension of data, which is key for seasonal soundscape characterizing or degradation detection. A robust model should include time series analysis for interpreting correctly the dynamism of ecosystem soundscape through time. One of the most popular method for time series analysis is hidden Markov models (HMM) (Rabiner, 1989). Hidden Markov models (HMM) are based on Markov chains, a stochastic model that stablishes that the current state depends on the previous one. They have been used in many bioacoustics applications, leading to high performance in species vocalizations recognition (Agranat, 2009; Aide et al., 2013; Erbe et al., 1999; Kirschel et al., 2009; Kogan & Margoliash, 1998). In this work, it is explored for first time to assess the ecosystem change through space and time.

1.3. Objectives

1.3.1. General Objective

Develop a methodology to analyze ecosystem changes by their soundscape, considering environmental factors and the distribution of its communities.

1.3.2. Specific Objectives

- Analyze current ecoacoustic indices to assess their main strengths and drawbacks.

- Propose a new methodology based on ecoacoustics for a structured analysis of the ecosystem soundscape.

- Test the proposed methodology on a large dataset with known ecosystem changes and compare the results to other approaches.

Chapter 2. Proposed Methodology

The proposed methodology for ecosystem soundscape characterization consists of 3 stages: in the first stage, or preprocessing stage, the noisiest recordings are discarded. Then, in the features stage, soundscape features are extracted and selected, and the acoustic diversity indices are calculated for biodiversity estimation. Finally, in the modeling stage, the parameters of the hidden Markov model with Gaussian mixture emissions (GMMHMM) are set up and the model is trained and tested with the data (Figure 2.1). In this chapter, each of the stages is introduced and justified and the details of the implementation are explained.

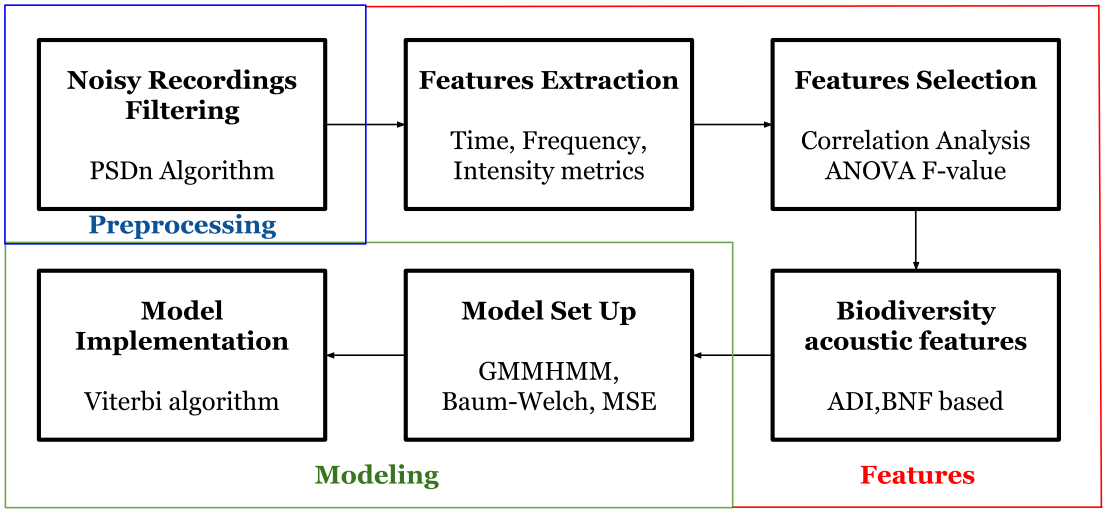

Figure 2.1 Block Diagram of the Proposed Methodology to Detect Ecosystem Acoustical Patterns

2.1. Preprocessing

2.1.1. Recordings Filtering

Loud geophonic and technophonic sounds can modify animal communication patterns (Slabbekoorn, 2004; Towsey, Wimmer, Williamson, & Roe, 2014). It has been proven that these sounds can prevent individuals of making calls (Lengagne & Slater, 2002) or provoke the increase of callings to overcome the acoustical obstacles (Lengagne, Aubin, Lauga, & Jouventin, 1999). Wind and rain are specially problematic, because they occupy a large range of the spectrum, masking the calls of many species (Qi et al., 2008). Therefore, in order to identify ecosystem change by acoustic analysis, the common practice is to discard the recordings with high levels of these sounds (Mangalam Sankupellay et al., 2015).

Manual inspection of recordings for this purpose could be tedious for large datasets. In this work, we estimated and implemented an automatic threshold to the algorithm proposed by Bedoya et al. (Bedoya, Isaza, Daza, &

López, 2017) for detecting the recordings containing heavy rain or wind. As mentioned in their work, we found that these recordings had the highest levels of power spectral density (PSD) in the band between 600 and 1200 Hz (not only for rain, but also for wind, see Figure 2.2). For automating this process, we tested the geometric and arithmetic mean as thresholds for filtering out the noisiest recordings per session and discovered that the best threshold was the average of the two means (see Eq. (2.1)).

$$thr_\zeta = \frac{\frac{\sum_r^{R_\zeta} PSDn_{r\zeta}}{R_\zeta} + \sqrt[R_\zeta]{\prod_r^{R_\zeta} PSDn_{r\zeta}}}{2} \tag{2.1}$$

Where thr_ζ is the threshold per session, $PSDn_{r\zeta}$ is the mean PSD in the band between 600 and 1200 Hz for each recording r in the session ζ and R_ζ is the number of recordings per session. A recording session is a set of consecutive recordings made by the same recorder, in the same location (during this period the recorder was not reconfigured or recharged). It was necessary to compute the threshold per session, because we noticed that, even though recorders were supposed to be set with the same specifications, there were considerable differences in intensity, which could be due to different models of recorders, microphones, or misfunctioning.

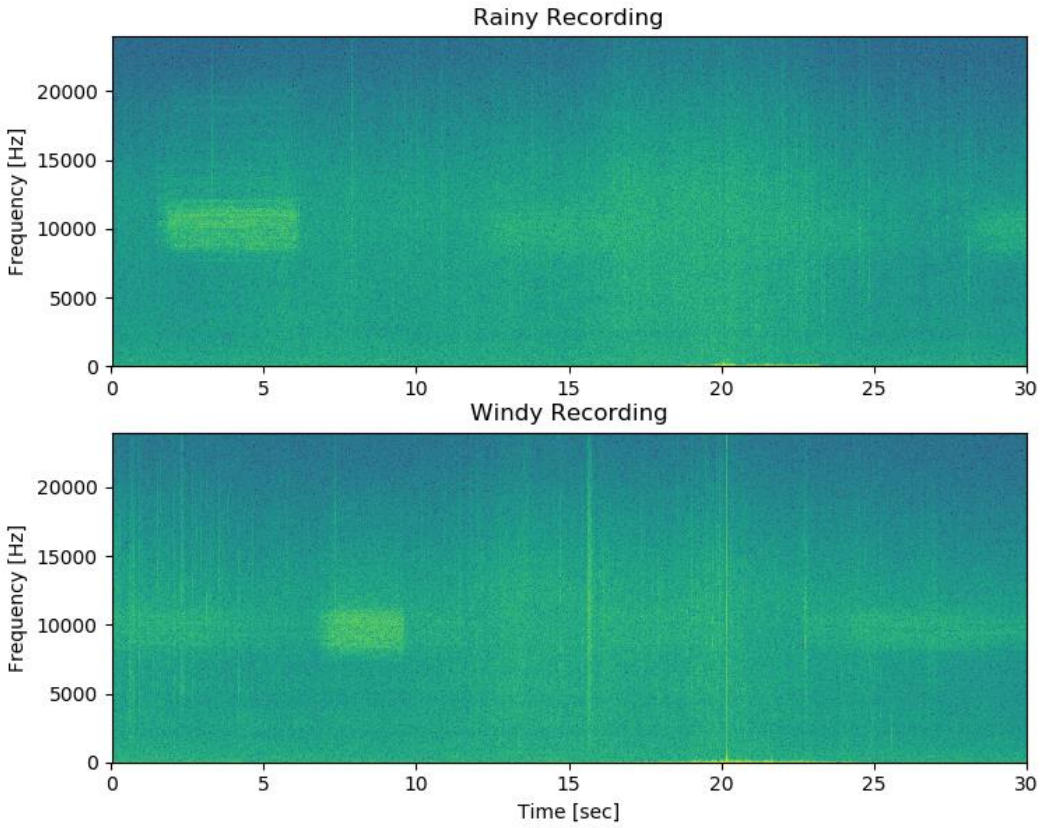

Figure 2.2 Two spectrogram segments of discarded recordings in the Humboldt case. The upper segment presented rain noise, while the lower segment presented wind noise. The y axis represents frequency, the x axis represents time and the coloring represents the intensity of the signal

The recordings with high level noise detected by the algorithm were discarded from the subsequent stages of the analysis.

2.2. Features

In this stage, different sets of features are extracted from the recordings. For acoustic signal characterization, there are a wide variety of proposed features which are to be considered according to the application. Maybe the most popular type of acoustical features is the Mel Frequency Cepstral Coefficients (MFCC) for automatic speech recognition. This technique has been used not only for human speech (Davis & Mermelstein, 1980; Hermansky, 1990; Milner, 2002; Wang & Paliwal, 2003), but also for birds (Kirschel et al., 2009; Kogan & Margoliash, 1998; Lellouch, Pavoine, Jiguet, Glotin, & Sueur, 2014), anurans (Bedoya et al., 2014; Duque-Montoya, Isaza, & Cano Rojas, 2017), and other species calls classification (Mazhar, Ura, & Bahl, 2007; Soltis, Leong, & Savage, 2005). However, the main use of MFCC and other automatic feature extraction techniques (like those used in deep learning) is for supervised applications, where the final clusters are commonly known. For the goal of characterizing acoustical dynamics in the ecosystem, in which we look for identifying patterns that have not been fully described, we chose to use features with physical interpretation, so the resulting clusters could be analyzed using their features profile by the final users of our model.

The users of our model will be the people responsible to study the ecosystem soundscape and set the guidelines for protecting it, as an important part of ecological conservation. These experts traditionally base their studies on identifying the species in an ecosystem and summarizing the findings using biodiversity indices (Magurran et al., 2010). To interpret the acoustical information, like in traditional methods, a series of ecoacoustic indices based on the biodiversity indices have been proposed (Sueur et al., 2014). In this work we used the alpha indices, which describe the diversity within a group by its sound intensity, acoustic complexity and soundscape elements.

Some of these indices have become key elements to many ecoacoustic applications (e.g. Depraetere et al., 2012; Harris, Shears, Radford, & Reynolds, 2016; Sankupellay et al., 2015). However, they can be easily affected by factors such as background noise, distance of sound source to the microphone, number of individuals singing or relative amplitude difference between same kind calls (Sueur et al., 2014). Furthermore, their theoretical interpretation is still unclear (Eldridge, Casey, Moscoso, & Peck, 2015). For this reason, besides these indices, we included in our work acoustic metrics for more general applications, looking to compensate indices bias and complement the soundscape description.

2.2.1. Features Extraction

For selecting the additional features to our problem, we focused on the first source of information for ecoacoustical analysis: the spectrogram. Before the indices were proposed, the spectrogram was the main tool to visualize the acoustic dynamics of a community (Blumstein & Munos, 2005; Kirschel et al., 2009; Leong et al., 2003; Urazghildiiev & Clark, 2006) or an ecosystem (Job et al., 2016; Joo et al., 2011) and it is still the first step for the calculation of many indices (Sueur et al., 2014). A spectrogram is an intensity plot of the Short-Time

Fourier Transform (STFT) magnitude (Smith, 2007), it allows to see the intensity distribution of the audio signal in both frequency and time (see Figure 2.2).

If we think intuitively on the components of the spectrogram, we identify three elements: intensity, frequency and time. Thence, we looked for acoustic features that could retrieve information about these three dimensions and describe accurately the spectrogram. The selected acoustic features are mathematically explained, described and referenced in Table 2.1.

Table 2.1 Calculation, description and reference of the extracted acoustical features.

Acoustic Indices			
Acoustic Feature	**Calculation**	**Description**	**Reference**
Acoustic Complexity Index (frequency-time) ($ACIf_t$)	$$\sum_{i=1}^{N_t}\sum_{j=1}^{N_f-1}\frac{\lvert I_{i,j}-I_{i,j+1}\rvert}{(I_{i,j}+I_{i,j+1})} \quad (2.2)$$	Where $I_{i,j}$ is the normalized amplitude of each pulse in the time step i and frequency bin j, N_t is the number of temporal steps and N_f is the number of frequency bins in the spectrogram. $ACIf_t$ measures the information in two successive frequency bins along each time step.	(Farina, Pieretti, Salutari, Tognari, & Lombardi, 2016)
ACIft evenness ($ACIf_{te}$)	$$\frac{1}{\sum_{i=1}^{N_t}\left(\sum_{j=1}^{N_f-1}\frac{\lvert I_{i,j}-I_{i,j+1}\rvert}{(I_{i,j}+I_{i,j+1})}\right)^2} \quad (2.3)$$	Levins evenness applied to the $ACIf_t$. $ACIf_{te}$ measures the distribution of intensity along each time step.	(Farina et al., 2016)
Acoustic Complexity Index (time-frequency) ($ACIt_f$)	$$\sum_{k=1}^{c}\frac{\sum_{i=1}^{\frac{N_t}{c}}\lvert I_{i,j}-I_{i+1,j}\rvert}{\sum_{i-1}^{\frac{N_t}{c}} I_{i,j}} \quad (2.4)$$	Originally named ACI, it is used to assess the acoustic complexity information in an audio signal. In this expression, c stands for the number of temporal clusters in the recording (usually, $c=5$ for one-minute recordings)	(Pieretti et al., 2011)
ACItf evenness ($ACIt_{fe}$)	$$\frac{1}{\sum_{k=1}^{c}\left(\frac{\sum_{i=1}^{\frac{N_t}{c}}\lvert I_{i,j}-I_{i+1,j}\rvert}{\sum_{i-1}^{\frac{N_t}{c}} I_{i,j}}\right)^2} \quad (2.5)$$	Levins evenness applied to the $ACIt_f$. $ACIt_{fe}$ measures the distribution of intensity along each frequency bin.	(Farina et al., 2016)
Acoustic Diversity Index (ADI)	$$-\sum_{b=1}^{N_b} p_b \log_2 p_b \quad (2.6)$$	The Shannon index is used to estimate diversity by using 1 kHz frequency bands as acoustic categories. p_b is the occupation of the frequency band b, measured as the number of cells in the band	(Pekin et al., 2012)

		higher than a given threshold. N_b is the number of frequency bands.	
Acoustic Evenness Index (AEI)	$\frac{\sum_{b_1}^{N_b}\sum_{b_2}^{N_b}\lvert p_b - p_d\rvert}{2N_b^2\bar{p}}$ (2.7)	The Gini coefficient is applied to obtain acoustic evenness across frequency bands of 1kHz.	(Villanueva-Rivera et al., 2011)
Bioacoustics Index (β)	$\int_{f=2kHz}^{f=8kHz} P_f$ (2.8)	Area below the P_f curve in the biophony frecuencies (between 2kHz and 8kHz) in dB. P_f is the mean of the spectrogram along the time dimension. This index gives an estimate of the biophonic energy in the recording.	(Boelman, Asner, Hart, & Martin, 2007)
Acoustic Entropy Index (H)	$H_t H_s$ (2.9)	It is the product of H_s and H_t (see below). It ranges between 0 and 1; 0 for pure tones and 1 for white noise.	(Sueur et al., 2008)
Entropy of spectral maxima (H_m)	$U_j = \max_{\forall w=j}(s_{i,w})$ (2.10) $H_m = -\sum_{j=1}^{N_f} U_j \log_2 U_j$ (2.11)	The Shannon index is applied to the maximum values of each frequency bin in the spectrogram, but only in the band between 482 Hz and 8820 Hz (expanding the biophony band). $s_{i,w}$stands for a cell in the spectrogram in the time step i and frequency bin w and U_j is the maximum value in the frequency bin j.	(Towsey et al., 2014)
Spectral Entropy (H_s)	$-\frac{\sum_{j=1}^{N_f} S_j \log_2 S_j}{\log_2 N_f}$ (2.12)	The Shannon evenness Index is applied to the average energy (S_j) in each frequency bin.	(Sueur et al., 2008)
Temporal Entropy (H_t)	$-\frac{\sum_{i=1}^{N_t} \varepsilon_i \log_2 \varepsilon_i}{\log_2 N_t}$ (2.13)	Similarly, to H_s, the Shannon evenness Index is applied to the amplitude envelope of the signal, obtained by applying the Hilbert transform (ε).	(Sueur et al., 2008)
Entropy of spectral variance (H_v)	$V_j = \operatorname{var}_{\forall w=j}(s_{i,w})$ (2.14) $H_v = -\sum_{j=1}^{N_f} V_j \log_2 V_j$ (2.15)	The Shannon index is applied to the variance value for each frequency bin. V_j stands for the variance value in the frequency bin j.	(Towsey et al., 2014)

Median of amplitude envelope (M)	$2^{D-1}\,\mathrm{median}(\varepsilon)$ (2.16)	The median of the amplitude envelope (obtained by using the Hilbert transform), where D is the signal digitalization depth.	(Depraetere et al., 2012)
Mid-band activity (MID)	$F_f = \begin{cases} P_f \; if \; P_f > \overline{P_f} \\ 0 \; if \; P_f < \overline{P_f} \end{cases}$ (2.17) $\int_{f=482z}^{f=3500Hz} F_f$ (2.18)	The area below P_f in the mid band (482 Hz – 3500 Hz) for values of P_f that surpass a certain threshold. F_f corresponds to the values of P_f over that threshold (in this case the mean of P_f)	(Towsey et al., 2014)
Normalized Difference Soundscape Index (NDSI)	$\frac{\beta - \alpha}{\beta + \alpha}$ (2.19)	A normalized measure of the ratio of biophony to technophony. For pure technophony NDSI = -1 and for pure biophony NDSI = 1. β is the bioacoustics index and α is the estimated technophony, which is measured similarly to β for the frequency band between 200 Hz and 1500 Hz	(Kasten et al., 2012)
Number of peaks (NP)	$\Lambda_j = \begin{cases} 1 \; if \; S_j > \; S_m \; \cap \; \overline{\lvert dS_l \rvert} > 0.01 \\ 0 \qquad\qquad otherwise \end{cases}$ (2.20) $NP = \sum_{j=1}^{N_f} \Lambda_j$ (2.21)	Number of major frequency peaks in the mean spectrum. It assumes that the more peaks, the higher spectral complexity and richness. In this equation m represents the frequency bins lower than j; and l represents the surrounding bins of the peak. Λ_j is then 1 if j is a major peak and 0 otherwise.	(Gasc, Sueur, Pavoine, Pellens, & Grandcolas, 2013)
Ratio of biophony to anthrophony (ρ)	$\frac{\beta}{\alpha}$ (2.22)	The first index to estimate the effect of technophony to biophony.	(Qi et al., 2008)
Complementary Acoustic Features			
Frequency Background Noise (BN_f)	$\overline{\underset{\forall w=j}{\mathrm{mode}}}(s_{i,w})$ (2.23)	Mean background noise of the spectrogram frequency bins, based on the mode value for each frequency bin j.	(Towsey, 2013)
Temporal Background Noise (BN_t)	$\mathrm{mode}(SPL)$ (2.24)	Temporal background noise in dB. Based on the mode value for the immediate sound pressure level signal (SPL)	(Towsey, 2013)

Crest Factor (CF)	$\frac{E_{peak}}{RMS}$ (2.25)	The ratio of the peak value of the energy signal (E_{peak}) to its root mean square value (RMS, explained below). The higher the value, the more expected peaks in the energy signal.	(Torija, Ruiz, & Ramos-Ridao, 2013)
Frequency Modulation (FM)	$\frac{\sum_{i=1}^{N_t-1}\sum_{j=1}^{N_f-1}\frac{180°}{\pi}\left\lvert\arctan\frac{I_{i,j+1}-I_{i,j}}{I_{i+1,j}-I_{i,j}}\right\rvert}{(N_f-1)(N_t-1)}$ (2.26)	The mean angle of the directional derivatives. The higher the value, the more abrupt changes in intensity.	(Tchernichovski, Nottebohm, Ho, Pesaran, & Mitra, 2000)
Musicality Degree (MD)	$\frac{\sum_{j}^{N_f-1}\frac{\log(W_{j+1}^2-W_j^2)}{\log(f_{j+1}-f_j)}}{N_f-1}$ (2.27)	The mean slope of the 1/f curve. W is the Welch's power spectral density estimate. MD is a measure of the temporal complexity of the signal.	(De Coensel, 2007)
Root mean square (RMS)	$\sqrt{x^2}$ (2.28)	The root mean square of the raw signal (x). It is a measure of the amplitude of the signal.	(Rodriguez et al., 2014)
Spectral Flatness (SF)	$\frac{\sqrt[N_f]{\prod_{j=1}^{N_f}W_j}}{\bar{W}}$ (2.29)	The ratio of the geometric mean of the power spectrum to its arithmetic mean. The higher the value, the more frequency complexity.	(Mitrović, Zeppelzauer, & Breiteneder, 2010)
Temporal Sound Level Variance Descriptor (TLSV)	$\sigma_L\sigma_{eq}$ (2.30)	The product of the instantaneous sound pressure level standard deviation (σ_L) and the energy-equivalent sound pressure level standard deviation (σ_{eq}). TLSV is a measure of temporal complexity.	(Torija et al., 2013)

All the acoustic indices were computed on the selected recordings, then each site daily averages were calculated and standardized using (2.31).

$$\hat{y} = \frac{y - \bar{y}}{\sigma_Y} \quad (2.31)$$

Where $\hat{y}$ is the new standardized value, y is the original value, $\bar{y}$ and σ_Y are the index mean and the index standard deviation calculated over all samples respectively.

2.2.2. Features Selection

In this stage, the most significant indices for our application are selected as features.

2.2.2.1. Data Division

At this point, the data was divided in two subsets: the training data, containing 80 percent of the total data; and the testing data, with the remaining 20 percent. Given the requirements of the clustering method and the purpose of this study, the data division was made in such a way that the chronological order was maintained, and each subset contained samples from each recording sequence as possible. A recording sequence is a set of recordings in the same session with possible gaps not larger than 6 days. The features extraction and selection methods and the model set up (section 2.3.1) were applied only to the training data, and after the model is configured, it is implemented to both the training and testing data. In other words, the complete methodology is executed using only the training data, except for the model implementation stage (section 2.3.2), in which both training and testing data were classified by the resulting model. The data division is suggested to avoid overfitting and obtain a model capable of generalization.

2.2.2.2. Correlation Analysis

A first stage of feature selection was based on correlation analysis. The Pearson Correlation coefficient (PCC) was computed among every pair of indices to identify those which were strong correlated, and therefore redundant for the analysis (2.32). For each pair of strongly correlated indices (Absolute value of PCC was 0.9 or higher), only one of the indices were selected for the following stage.

$$PCC_{Y,Z} = \frac{\sigma_{Y,Z}}{\sigma_Y \sigma_Z} \tag{2.32}$$

Where $\sigma_{Y,Z}$ is the covariance for indices Y and Z, σ_Y is the variance for index Y and σ_Z is the variance for index Z.

2.2.2.3. F-value

For selecting the most representative indices for each case, the Analysis of Variance (ANOVA) F-value was calculated for each index. Computing this value requires knowing the prior classification of the data; which for our case, corresponded to the ecosystem type of the site where the audio files were recorded. Then, the F-value is computed as shown in (2.33).

$$\frac{N_{tr}(N_{tr} - N_g)\sum_g (y_g - \bar{y})^2}{(N_g - 1)\sum_g^{N_g}\sum_r^{N_{tr}} (y_{gr} - y_g)^2} \tag{2.33}$$

Where N_{tr} is the total number of recordings in the training set, N_g is the number of groups, y_g is the index mean for the group g, $\bar{y}$ is the index mean for all samples and y_{gr} is the index value for the recording r in the group g.

The F-value assesses the discrimination power of the index, i.e. the higher the F-value, the more uniform index values for same group samples, and the more different index values for different group samples. Consequently, the indices with highest F-values were selected as soundscape features, for our hypothesis poses that soundscape differs with ecosystem type.

For this selection we excluded the ADI and BN_f indices, which we used for computing the biodiversity features in the following stage. Then, based on the assumption that each soundscape feature would differentiate at least two groups, the number of soundscape features was determined by (2.34)

$$\binom{N_g}{2} \tag{2.34}$$

Which is the number of possible pairs between groups, or ecosystem types.

Although the ANOVA poses the normality of the data as a requirement, it is deemed to be robust for non-normal data too (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010), therefore the data normality test was omitted.

2.2.3. Biodiversity Acoustic Features

Biophony remains still the most important soundscape element (Gage & Axel, 2014). Whereas these sounds have been traditionally demarcated in the frequency band between 2 kHz and 8 kHz (Boelman et al., 2007; Kasten et al., 2012), this convention has been less applicable to tropical habitats (Sueur et al., 2014) and has already been extended to the 2-11 kHz band for some other studies (Eldridge, Guyot, Moscoso, Johnston, & Eyre-walker, 2018; Gage & Axel, 2014). The band under 2 kHz is by no means lacking biological sounds, especially in the tropic, but it is often dominated by technophony (Fuller et al., 2015; Joo et al., 2011). To encompass biodiversity in the soundscape level completely, in this work we included as features a modified ADI vector for the 0-11 kHz band as descriptors of acoustic diversity.

As other acoustic indices (AEI, H, H_s, H_v, H_m), ADI is based on the Shannon index, which is used frequently in ecology for richness estimation (Spellerberg & Fedor, 2003; Sueur et al., 2014). This index was proposed originally for measuring the information content in a signal by summing the information of its components (Shannon, 1948). In ADI, these components correspond to 1 kHz frequency bands, which we proposed to analyze separately because it is easier to interpret (An unique value would not say much about the calls distribution in the signal, on the other hand, a value per band would help identify which are the bands with higher information or complexity, hence the most active acoustical species by mapping their calls in the spectrum).

The main advantage for ADI over other Shannon based indices is that it considers background noise as a part of the calculation (Pekin et al., 2012). Therefore, this index should be more robust to technophony effect in the corresponding band. Nonetheless, the main drawback of ADI is that it uses a fixed background noise value for

the whole spectrum. Assuming that this value differs between frequency bins, we calculated a different background noise value per bin, using the original proposed BN_f vector (Towsey, 2013) [1] (2.35).

$$BN_{fj} = \underset{\forall j=w}{\text{mode}}(s_{i,w}) \tag{2.35}$$

Where BN_{fj} is the background noise value in the frequency bin j, and $s_{i,w}$ is the spectrogram value for time step i and frequency bin w.

Using BN_{fj} the modified ADI is calculated per 1 kHz band as indicated in (2.36)

$$ADI_b = -p_b \log p_b \tag{2.36}$$

Where ADI_b is the modified ADI for the band b and p_b is the occupation in the band b, which was measured as the number of spectrogram cells in the band higher than their respective BN_{fj}.

As a result, we obtained 11 biodiversity metrics, which join the soundscape features for the complete feature set per day.

2.3. Modeling

In this stage, the selected features served as inputs to a GMMHMM for modeling ecoacoustical patterns. HMM is a stochastical model in which a sequence of random observations is generated by a set of hidden states (or clusters) given a Markov chain and a emission probability distribution (Bilmes, 1998); this process is illustrated in Figure 2.3. This model assumes Markov property, i.e. the probability of a given state depends uniquely on the previous one (Bilmes, 1998). For denoting this model, (2.37) is used

$$\lambda = (A, B, \pi) \tag{2.37}$$

Where λ is the HMM, A is the transition matrix, i.e. the element a_{ij} in A denotes the probability of passing to state j being previously in the state i; B is the emission probability distribution, or the distribution that determines the probability that a given state produces a given observation; and π is the initial state distribution, or the distribution that sets the probability for each state to be the first in the sequence (Rabiner, 1989).

GMMHMM is a type of HMM in which B follows a gaussian mixture distribution. Therefore, the probability for the state i to emit the observation vector y_t at the time t ($b_i(y_t)$) is given by a weighted sum of normal functions (see (2.38)).

$$b_i(y_t) = \sum_{m=1}^{N_m} p_{im}\, \phi(y_t|\mu_{im}, \Sigma_{im}), \quad \sum_{m=1}^{N_m} p_{im} = 1 \tag{2.38}$$

[1] This is a simplification to the algorithm, to see the complete algorithm, please refer to (Towsey, 2013)

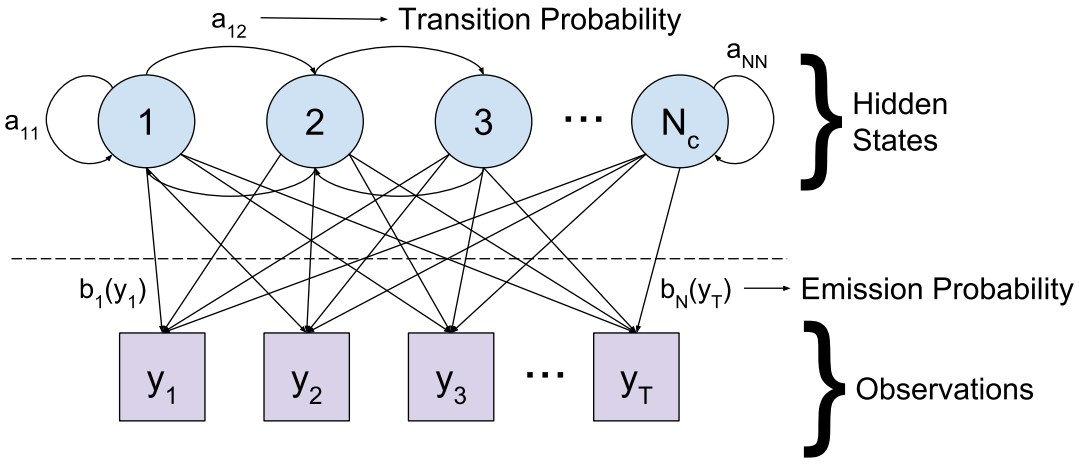

Figure 2.3 HMM Diagram. The Markov chain is represented in the upper half of the figure, with its hidden states (denoted by the blue circles) and transition probabilities (denoted by the curved arrows). In the lower part of the figure, the observed sequence is produced by the Markov chain given the emission probabilities (the observed samples are denoted by purple squares and the emission probabilities are denoted by the arrows connecting the states to the observations).

Where p_{im} is the weight for the m^{th} distribution in the state i, $\phi(y_t|\mu_{im},\Sigma_{im})$is the probability for the observation y_t to be generated by the multivariate gaussian distribution described by the mean vector μ_{im} and covariance matrix Σ_{im}, and N_m is the number of components in the mixture. This model has the advantage of fitting complex data but requires the calculation of more parameters for the model. In Figure 2.4, an example with a univariate 3-component-gaussian mixture model is plotted showing that the distribution is shaped by its components and, unlike the normal distribution, in which the data is grouped around a unique value, this model presents three different data groups, allowing each state to be associated to different values of the observed data.

Lastly, this model was chosen because it does not only serve to characterize the hidden states (patterns or clusters) which describe the ecosystem soundscape, but it also recognizes the relation between them, making it a predictable model, suitable for both spatial and temporal analysis.

2.3.1. Model Set Up

In this stage, the optimal GMMHMM parameters A, B and π are estimated using the Baum-Welch Algorithm and the Mean Squared Error (MSE) applied to the training dataset. Firstly, a set of likely values for N_c (number of states) and N_m (number of mixture components) are established, then the Baum-Welch Algorithm is executed and the sum of squared errors (SSE) is calculated several times for each combination of these values. Using the SSE resultant values, the MSE is computed per combination and the optimal values for N_c and N_m are chosen according to the lowest MSE. Finally, the Baum-Welch algorithm is executed again several times using the selected N_c and N_m values and the final selected model is that with the lowest SSE. The flowchart for the described process is shown in Figure 2.5. The resulting model performance is displayed in Chapter 3.

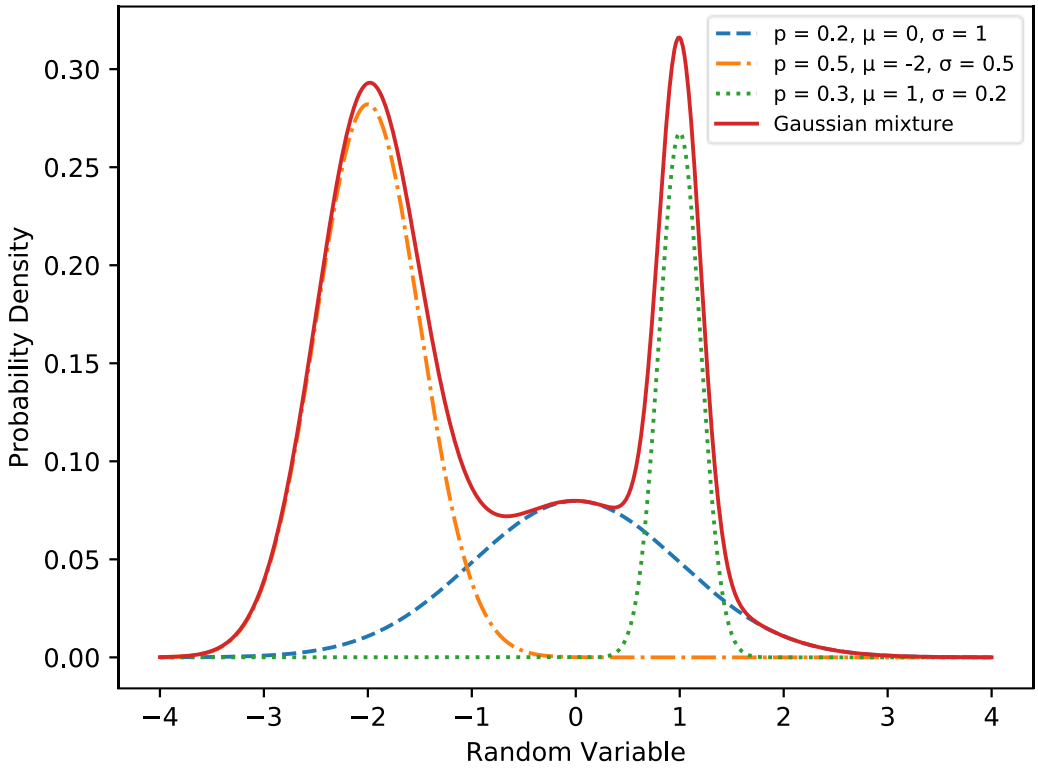

Figure 2.4 Gaussian mixture model example with three components. It shows that three different distributions of the observed samples could be assigned to the same hidden state.

In the following subsections the model set up process is detailed.

2.3.1.1. Baum-Welch Algorithm

For finding the GMMHMM parameters that best fit to the problem, we applied the commonly used Baum-Welch algorithm to the training data (Bilmes, 1998). This algorithm finds the HMM parameters that maximize the probability for the observed sequence of feature vectors.

Firstly, the parameters A, B and π are initialized with random values. Then, let π_i be the initial probability for state i, $b_i(y_t)$ the probability for state i to emit the observation y_t, N_c the number of states, a_{ij} the transition probability from state i to state j, and T the time for the last element of the sequence. Next, the algorithm is recursively executed using forward, backward and update procedures:

For the forward procedure, we denote $\alpha_i(t)$ as the joint probability for a partial starting sequence finishing in the state i and time t. $\alpha_i(t)$ can be obtained as shown in (2.39) and (2.40).

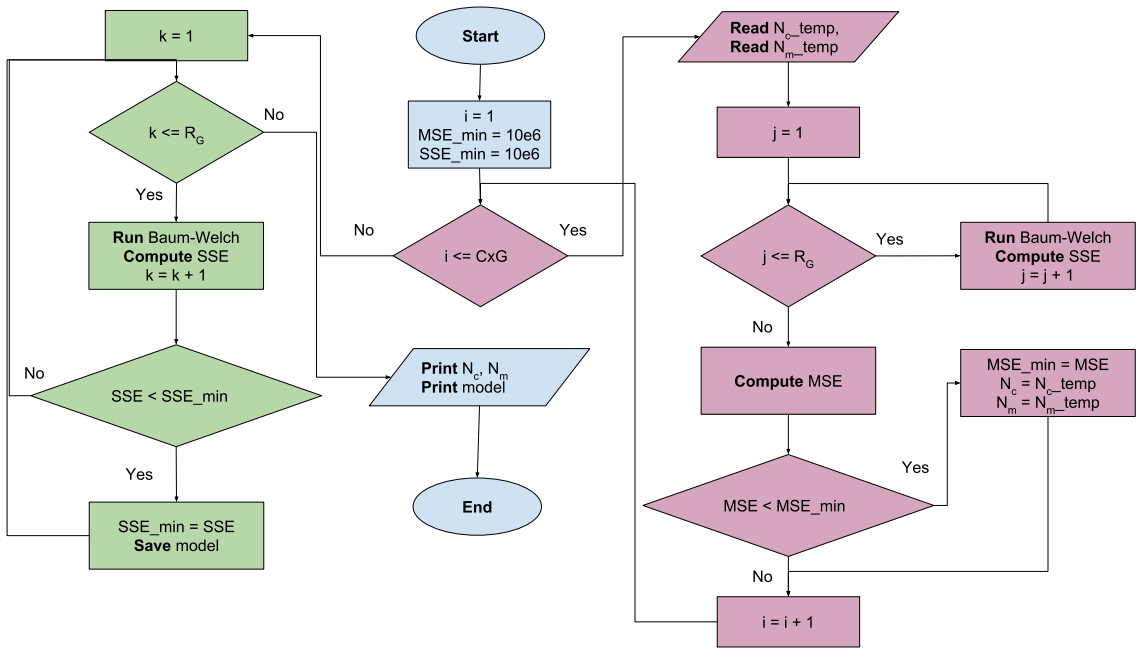

Figure 2.5 Model set up diagram. This flowchart describes the process for the estimation of the model parameters. The purple items describe the first part of the process, where N_c and N_m are selected. The green items describe the second part of the process where the remaining model parameters (A, B and π) are obtained. Note that C is the number of tested values for N_m, G is the number of tested values for N_c and R_G is the number of times that the Baum-Welch algorithm is executed, and the SSE is computed.

$$\alpha_i(1) = \pi_i b_i(y_1) \tag{2.39}$$

$$\alpha_j(t+1) = \left[\sum_{i=1}^{N_c} \alpha_i(t) a_{ij}\right] b_j(y_{t+1}) \tag{2.40}$$

Whereas for the backward procedure we denote $\beta_i(t)$ as the joint probability for the partial ending sequence starting at $t+1$ and finishing at T given state i and time t. $\beta_i(t)$ is calculated using (2.41) and (2.42).

$$\beta_i(T) = 1 \tag{2.41}$$

$$\beta_i(t) = \sum_{j=1}^{N_c} \beta_j(t+1) a_{ij}\, b_j(y_{t+1}) \tag{2.42}$$

Lastly, in the update step, we denote $\gamma_i(t)$ as the probability of being in state i at time t, and it is estimated by (2.43).

$$\gamma_i(t) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{j=1}^{N_c} \alpha_j(t)\beta_j(t)} \tag{2.43}$$

In this step it is also defined $\xi_{ij}(t)$ as the probability of being in state i and j in time t and $t+1$ respectively. This probability is estimated by (2.44).

$$\xi_{ij}(t) = \frac{\alpha_i(t)a_{ij}\beta_j(t+1)b_j(y_{t+1})}{\sum_{i=1}^{N_c}\sum_{j=1}^{N_c}\alpha_i(t)a_{ij}\beta_j(t+1)b_j(y_{t+1})} \tag{2.44}$$

For several sequences, $\alpha_i(t)$, $\beta_i(t)$, $\gamma_i(t)$ and $\xi_{ij}(t)$ are calculated for each sequence with different initialization parameters.

Now, if there are E sequences, and the e^{th} sequence has a length of T_e, the parameters π_i and a_{ij} are updated using (2.45) and (2.46)

$$\hat{\pi}_i = \frac{\sum_{e=1}^{E}\gamma_i^e(1)}{E} \tag{2.45}$$

$$\hat{a}_{ij} = \frac{\sum_{e=1}^{E}\sum_{t=1}^{Te}\xi_{ij}^e(t)}{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_i^e(t)} \tag{2.46}$$

To estimate $b_i(y_t)$, let $\gamma_{im}^e(t)$ be the probability for the e^{th} sequence that the m^{th}component of the i^{th}state generated observation y_t, and it is defined by (2.47)

$$\gamma_{im}^e(t) = \gamma_i^e(t)\frac{p_{im}\phi(y_t|\mu_{im},\Sigma_{im})}{b_i(y_t)} \tag{2.47}$$

Finally, B is then updated by replacing (2.48), (2.49) and (2.50) in (2.38)

$$\hat{p}_{im} = \frac{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_{im}^e(t)}{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_i^e(t)} \tag{2.48}$$

$$\hat{\mu}_{im} = \frac{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_{im}^e(t)}{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_i^e(t)} \tag{2.49}$$

$$\hat{\Sigma}_{im} = \frac{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_{im}^e(t)(y_t^e-\mu_{im})(y_t^e-\mu_{im})'}{\sum_{e=1}^{E}\sum_{t=1}^{Te}\gamma_i^e(t)} \tag{2.50}$$

Baum-Welch algorithm is repeated for all values of t, then executed again until a desired level of convergence. Since this algorithm is gradient-based, it can stop in a local optimum. Therefore, it is suggested to be restarted and executed several times, to generate several models and choose the one with the best score.

2.3.1.2. Mean Squared Error (MSE)

Baum-Welch algorithm finds the optimal values for most of GMMHMM parameters, except for the number of clusters (N_c) and the number of mixture components (N_m). For finding such values, a 2D-grid with the combinations of likely quantities for these parameters was arranged, then the Baum-Welch algorithm was run for several times for each of the combinations in the grid.

Every time the Baum-Welch algorithm was executed, the SSE was computed for the resulting model. But firstly, v_t was computed for each observation using Viterbi algorithm (explained in the following section), then the mean vector μ_i for the model state i was calculated using (2.51).

$$\mu_i = \sum_{m=1}^{N_m} p_{im}\mu_{im} \tag{2.51}$$

Let us remember that μ_{im} is the mean vector for the m^{th} distribution that composes the emission probability for state i, and p_{im} is its mixture weight. Using μ_i and v_t, SSE can be computed using (2.52)

$$\sum_{t=1}^{T}(y_t - \mu_i)(y_t - \mu_i)', \quad \forall i = v_t \tag{2.52}$$

Finally, the MSE value was obtained as an average of all SSE values within the same N_c and N_m combination (Rasmussen & Williams, 2004). MSE is a measure of cluster compactness, i.e. the higher the MSE the more variance within the same cluster samples. Accordingly, the chosen values for N_c and N_m were the ones that led to the lowest MSE value.

MSE was preferred over maximum likelihood estimation (MLE), Bayesian information criterion (BIC) and Akaike information criterion (AIC) (Bishop, 2006), because the first showed convergence with low-complexity models, whereas the other methods kept increasing their scores with model complexity (i.e. greater values for N_c and N_m, which would lead to overfitting).

Once the optimal parameters for the model have been found, it is not necessary to repeat the model set up for new similar data, but only classify these data using the Viterbi algorithm (explained below). This is the reason why the data division (section 2.2.2.1) was executed in the first place, to ensure that the model responds correctly to new samples from the same study.

2.3.2. Model Implementation

In this stage we apply the resulting GMMHMM model to the complete dataset (both training and testing data) using the Viterbi algorithm (Forney, 1973; Rabiner, 1989).

For the implementation of the algorithm, let $\delta_t(i)$ be the highest probability for the first t observations in the sequence ending in state i, and $\psi_t(i)$ the vector keeping track of the state sequence that maximized said probability. Now, the procedure is developed by the following steps:

1) Initialization:

$$\delta_1(i) = \pi_i b_i(y_1) \tag{2.53}$$

$$\psi_1(i) = 0 \tag{2.54}$$

2) Recursion:

$$\delta_t(j) = \max_i[\delta_{t-1}(i)a_{ij}]\, b_j(y_t) \tag{2.55}$$

$$\psi_1(i) = \operatorname{argmax}_i[\delta_{t-1}(i)a_{ij}] \tag{2.56}$$

3) Termination:

$$\hat{P} = \max_i[\delta_T(i)] \tag{2.57}$$

$$\upsilon_T = \operatorname*{argmax}_i[\delta_T(i)] \tag{2.58}$$

4) State sequence backtracking:

$$\upsilon_t = \psi_{t+1}(\upsilon_{t+1}) \tag{2.59}$$

Where $\hat{P}$ is the estimated probability for the given sequence and υ_t is the probable state that emitted observation y_t. This process is repeated for each of the observed sequences.

As a result, the data is grouped in clusters and their patterns can be now described and analyzed. Notice that the obtained clusters might not correspond to the ecosystem types, but they are meant to describe the acoustical processes that configure the observed soundscape dynamics.

Chapter 3. Results

In this chapter the methodology is applied to two Colombian datasets, and the results for each of the processing stages are described. The final interpretation of the resulting clusters was reviewed by a biologist.

3.1. Jaguas Application

The first dataset was provided by Grupo Herpetológico de Antioquia (GHA)[2] and ISAGEN[3], an energy generation company. The recordings were collected for monitoring the ecosystems surrounding San Lorenzo dam and assess the influence of the Jaguas energy plant in the natural communities and environment.

3.1.1. Database Description

Six sites surrounding the Jaguas hydroelectric power station in Alejandría, Antioquia, were acoustically monitored for this study. A total of 124,989 recordings were collected from April 22nd of 2016 to May 6th of 2017 using Wildlife Acoustics[4] Songmeter SM2 recorders (Wildlife Acoustics, 2011). The proprietary recording format WAC could be converted to WAV using Wildlife's Kaleidoscope software. The resulting recordings had a digital format of 16-bit PCB ($D = 16$), a sampling frequency of 22,050 Hz ($Fs = 22{,}050\ Hz$), one minute of duration ($d = 60\ sec$) and were programmed to be recorded every 20 minutes. Even though some recorders had two functioning channels during this period, only one channel per recorder was considered.

The sites were chosen and labeled by experts along a perturbation gradient. Each of the sites was assigned with a specific recorder during this monitoring period. In Table 3.1 the location of the monitored sites, their classification and the number of resulting recordings are described.

Table 3.1 Description of the recording sessions for the Jaguas application. The rows are colored according to their ecosystem type.

Recorder	Location	Ecosystem Type	Number of recordings
5253	6.356 N, 74.997 W	Stubble	20,256
5255	6.367 N, 74.997 W	Stubble	24,018
5256	6.367 N, 75.027 W	Forest	10,119
5260	6.361 N, 74.998 W	Forest	25,756
5257	6.375 N, 75 W	Pasture	19,416
7125	6.375 N, 74.998 W	Pasture	25,424
		Total	124,989

[2] http://grupoherpetologicodeantioquia.org/
[3] https://www.isagen.com.co/SitioWeb/es/
[4] https://www.wildlifeacoustics.com/

In conclusion, three ecosystem types were monitored: forest, stubble and pasture, with two sites per ecosystem type (see Figure 3.1).

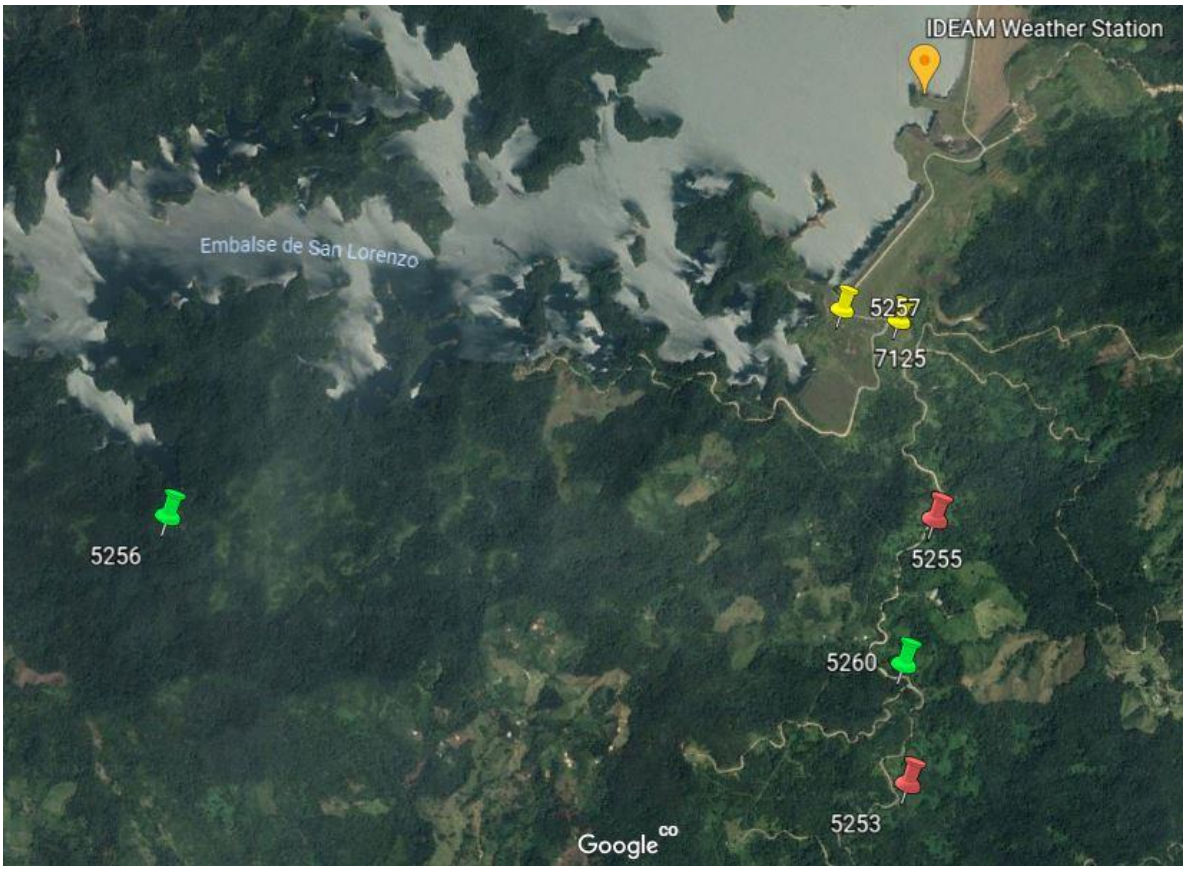

Figure 3.1 Google Earth screenshot of the monitored area. Yellow pin color indicates a pasture site, green pin color indicates a forest site and red pin color indicates a stubble site. The closest weather station is also marked.

3.1.2. Preprocessing Results

In this stage, a total of 25,560 recordings were automatically discarded due to probable heavy rain. This quantity corresponded to 20.4 % of the total recordings. In Table 3.2. the number of original recordings, the number of discarded recordings and the percentage of excluded recordings per site are presented.

It is interesting to notice that the quantity of discarded recordings per site were very different. This seems to imply that it did not rain the same for all six sites, even though they are in the same locality. To test this hypothesis, we requested the local climate authority (IDEAM) to send the area daily rain records for the same dates (see in Fig. 3.1. the relative location of the weather station). We found that there was no strong correlation between any of the monitored sites $PSDn$ and the weather station measurements, reinforcing our hypothesis (see Table 3.3). Notice that the longest distance between measuring points is approximately 3.9 km

for the weather station and recorder 5256, and the shortest distance is approximately 260 m for recorder 5257 and 7125.

Table 3.2 Results for the preprocessing stage in the Jaguas application.

Recorder	Number of Original Recordings	Number of Discarded Recordings	Percentage of noisy recordings
5253	20,526	5,679	27,67 %
5255	24,018	3,590	14,95 %
5256	10,119	2,281	22,54 %
5257	19,416	1,030	5,3 %
5260	25,756	9,561	37,12 %
7125	25,424	3,419	13,45 %

Table 3.3 PCC matrix for the recorders and weather station (IDEAM). The cells are colored according to their value, Greener cells indicate high PCC values; Redder cells indicate low PCC values.

	IDEAM	5253	5255	5256	5257	5260
5253	0.31	-	-	-	-	-
5255	-0.034	0.21	-	-	-	-
5256	-0.063	-0.0098	-0.052	-	-	-
5257	0.2	0.18	-0.0047	-0.005	-	-
5260	0.17	0.17	-0.017	0.015	0.04	-
7125	0.11	0.088	0.052	-0.037	0.23	0.086

Nonetheless, IDEAM measurements seem to be the more correlated to the remaining sites measurements, slightly proving the rain algorithm reliability. On the other hand, the site 5256 shows the least correlation to the other sites, probably because it is the most remote point. Finally, relatively high PCC can be observed for same ecosystem type sites (except for the forest type), indicating that, on this band at least, there is some evidence for similar acoustical properties.

3.1.3. Feature Extraction Results

For the ecoacoustic indices calculation, MATLAB (Mathworks, 2013) home-made routines were executed. For $ACIf_t$, $ACIf_{te}$, $ACIt_f$, $ACIt_{fe}$, β, H_s, NP and MID, the spectrogram was obtained using a Hann Window with no overlap (N_{ov} = 0), size of 512 points ($\omega = 512$) and 512 points for the Fast Fourier Transform (FFT) ($nfft = 512$). This configuration produced a spectrogram of dimension 257 x 2,584 (see (3.1) and (3.2)). For NP, the smoothed mean spectrum (S_j) was inspected for 10-point zones to find possible major peaks, and then if two peaks were found with a separation of less than 200 Hz, only the highest one was considered for the count.

$$N_f = \left\lceil \frac{nfft}{2} + 1 \right\rceil \tag{3.1}$$

$$N_t = \left\lfloor \frac{d \times Fs - N_{ov}}{\omega - N_{ov}} \right\rfloor \tag{3.2}$$

For ADI, AEI, H_m and H_v, $nfft$ and ω were set as a tenth of Fs ($nfft = \omega = 2{,}205$) leading to a spectrogram of 1,104 x 600 cells. Lastly, for NDSI and ρ the spectrogram was calculated using a Hamming window, and $nfft$ and ω were set to Fs ($nfft = \omega = 22{,}050$) leading to a spectrogram dimension of 11,026 x 60. Additionally, for ADI and AEI, the threshold for computing the band occupation was set at -50 dBFS.

For the remaining features (complementary soundscape features and biodiversity features), the spectrogram was computed using a Hamming window with $nfft = \omega = 512$ and 256 points of overlapping ($N_{ov} = 256$) producing a spectrogram dimension of 257 x 5,161. On the other hand, the Welch's method for SF and MD was computed using a Hann window with $nfft = \omega = 256$ and 128 points of overlapping ($N_{ov} = 128$), leading to a signal of 129 elements (this value is computed using (3.1)). The calculation of this last group of features was executed using Scipy library for Python (Jones, Oliphant, & Peterson, 2001).

Before the calculation of the daily averages, we noticed that some recordings presented "nan" values (not a number) for the $ACIf_t$ index. Inspecting these recordings, we found that they presented clicks and moments of silence, as if the microphone stopped working for those instants (see Figure 3.2). We opted for discarding such recordings to avoid introducing noise to the analysis. In total, only 34 recordings were discarded for this reason.

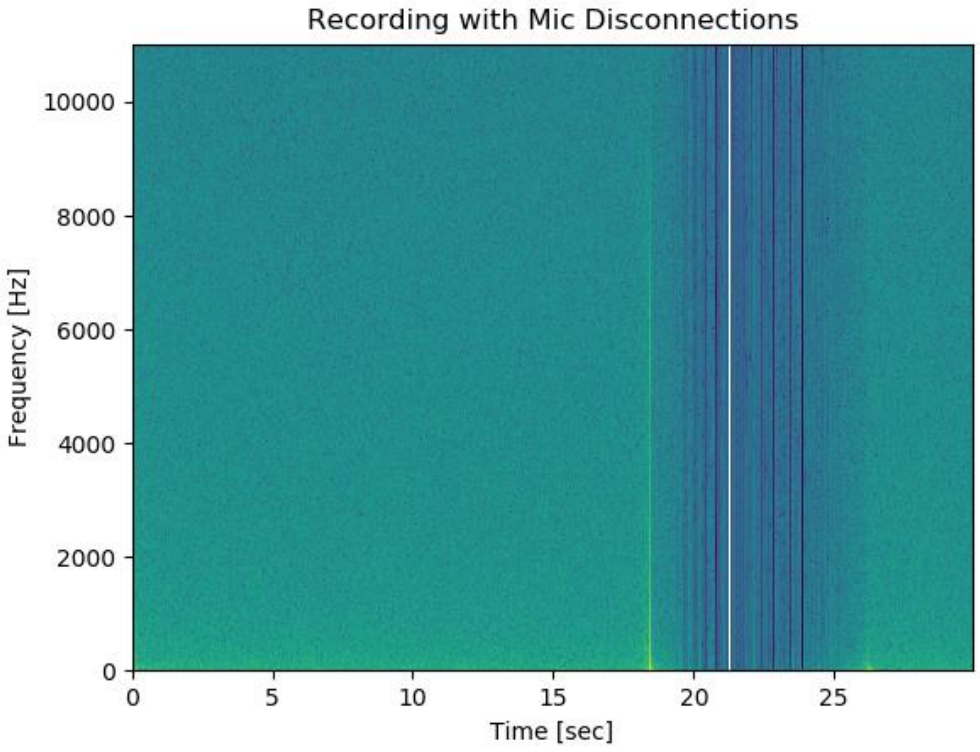

Figure 3.2 Spectrogram segment of a discarded recording. Observe the vertical black lines towards the end of the segment indicating a microphone disconnection.

Given the discarded recordings in the previous stage and in this one, a minimum number of recordings per day and recorder was established for computing the daily average. Considering that a recorder worked every 20 minutes, a total of 72 recordings should be obtained per day for that recorder. Thence, we established that at least 60 recordings per day and recorder should have passed the previous filters for computing the daily average and including that day in the subsequent analysis.

In conclusion, a total of 1172 daily samples were obtained: 142 for recorder 5253, 229 for recorder 5255, 84 for recorder 5256, 259 for recorder 5257, 195 for recorder 5260 and 262 for recorder 7125.

3.1.4. Feature Selection Results

Once the daily averages were obtained and standardized, the correlation analysis was realized. For every pair of features with absolute value of PCC higher than 0.9, only one feature was selected for the following stages. The criterion for selecting the chosen feature was its mean absolute Pearson coefficient ($MAPC$) calculated as the average of the feature's absolute $PCCs$ over the rest of the features (3.3). The feature with the highest $MAPC$ within the pair, was removed from the study. The discarded features, the corresponded correlated feature, their $MAPCs$ and $PCCs$ are displayed in Table 3.4.

$$MAPC_Y = \overline{|PPC_{YZ}|}, \qquad \forall Z \neq Y \tag{3.3}$$

Where $MAPC_Y$ is the $MAPC$ value for feature Y, PPC_{YZ} is the Pearson correlation coefficient for features Y and Z and Z is any feature different than Y.

Table 3.4 Discarded features with their correspondent correlated feature, their $MAPCs$ and their $PPCs$

Discarded Features		Correlated Features		PPC
Feature	$MAPC$	Feature	$MAPC$	
AEI	0.418	ADI	0.3875	-0.97
H_v	0.4071	H_m	0.3993	0.939
$ACIt_{fe}$	0.3354	$ACIt_f$	0.3281	-0.983
$ACIf_{te}$	0.4328	$ACIf_t$	0.4282	-0.999
H_s	0.3879	ADI	0.3875	0.938
H	0.3887	ADI	0.3875	0.941
TLSV	0.3902	ADI	0.3875	-0.946

Notice that, though ADI is preferred over several other features, the original ADI is not used for training the model, but the modified ADI vector is used instead. Nonetheless, including the ADI for the feature selection stage helped us identify the features that would have been redundant for the analysis.

Next, the training samples are grouped according to the ecosystem type classification and the F-value is computed. The obtained F-value is depicted in Figure 3.3. Considering that there are 3 ecosystem types and applying (2.34), the selected soundscape features are the three features with the highest F-value: H_t, β and NP. The grouped boxplots for these features are displayed in Figure 3.4.

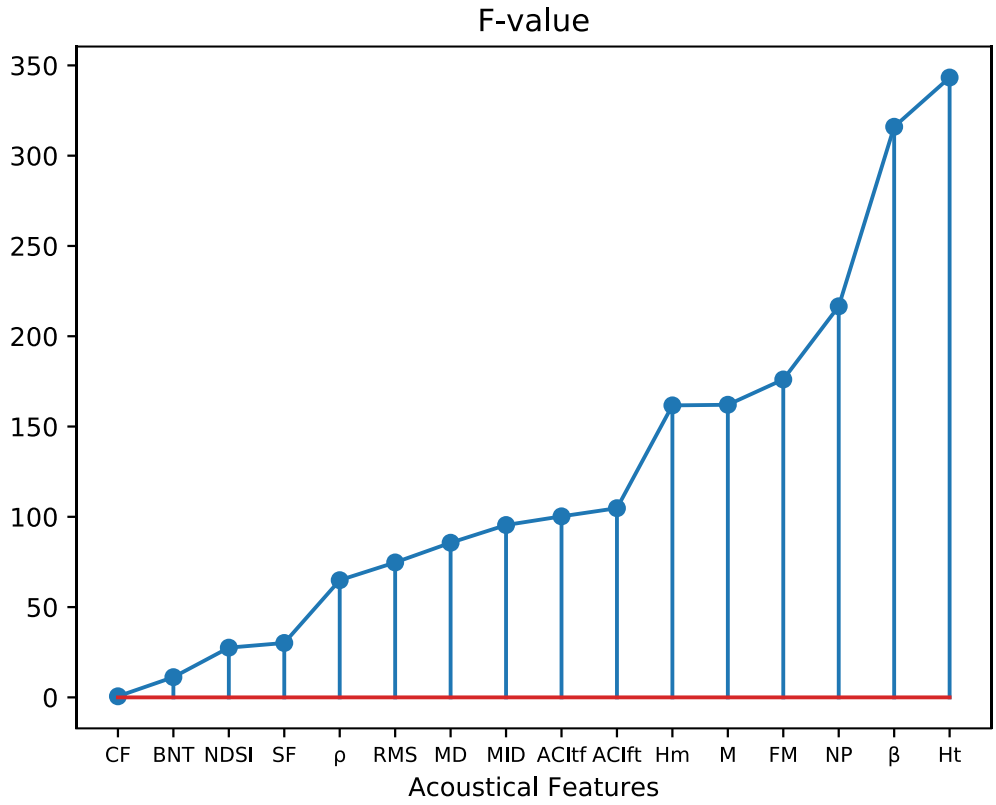

Figure 3.3 F-value for the selected features in the Jaguas application.

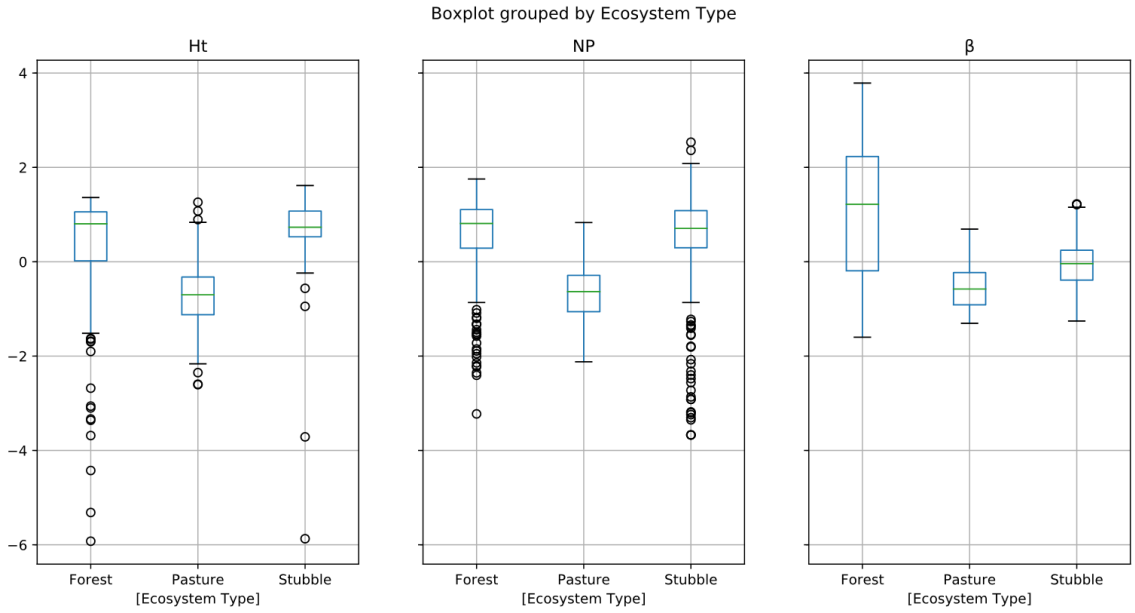

Figure 3.4 Grouped boxplots by ecosystem type for the three soundscape features.

It is interesting to notice that each of the soundscape features are related to one of the spectrogram dimensions: time (H_t), frequency (NP) and intensity (β). Also, notice in Figure 3.4 that the three ecosystem types are not completely distinguished using the soundscape features, even though their F-values are very high (Figure 3.3). In general, we can infer that the forest values seem to be more heterogeneous, while the pasture values are the most homogeneous. However, since there is not a clear distinction between the classes, we expect to find common underlying acoustical patterns for the three ecosystems soundscapes.

Finally, the biodiversity acoustic features are extracted as explained in section 2.2.3 and the set of features is completed (see Table 3.5).

Table 3.5 Final feature vector for characterizing a daily sample in the Jaguas application.

Soundscape feat.			**Biodiversity features**										
β	NP	H_t	ADI_1	ADI_2	ADI_3	ADI_4	ADI_5	ADI_6	ADI_7	ADI_8	ADI_9	ADI_{10}	ADI_{11}

3.1.5. Model Set Up Results

For the model set up and implementation stage, the Hmmlearn library for Python was used (Hmmlearn developers, 2016). Firstly, N_m was tested in the range between one and five (C = 5) and Nc was tested in the range between 3 and 20 (G = 18). Then, the Baum-Welch algorithm was executed 20 times (R_G = 20) for each of the Nc and N_m combinations; and 20 times again when the optimal Nc and N_m were found. The MSE values for each N_m x Nc combination are shown in Table 3.6.

In the table we find the lowest values in the column for $N_m = 3$, and the lowest value in this column for N_c = 16. However, when we inspected the resulting model, we found that there were seven repeated states; therefore, we maintained the value for N_m, changed N_c to 9 and ran the Baum-Welch algorithm 20 more times, finding this time only 3 repeated states. Lastly, we retrained the model for N_c = 6 and obtained our final model for Jaguas.

Figure 3.5 and Table 3.7 summarize the resulting model. The graphs in the left column of Figure 3.5 show the mean vectors for each cluster, which were computed using (2.51); the graphs at the right correspond to the clusters' prototypes, i.e. the recordings that best fit to the corresponding mean vectors. Table 3.7 shows the transition matrix A and initial probability π of the final model.

For cluster one, the prototype recording presented four occupied frequency bands by mainly constant calls. This type of occupation presents low biodiversity features because the entropy is low. On the other hand, β is high, because the constant calls sum to the biophonic intensity and NP is also high for the number of occupied bands.

Table 3.6 MSE values for all the possible combinations of N_c and N_m in the Jaguas application. The table cells are colored according to their values. Greener cells indicate lower MSE and redder cells indicate higher MSE.

Values for N_c	Values for N_m				
	1	2	3	4	5
3	13,582.69	13,558.85	11,874.24	12,603.17	12,741.83
4	13,582.69	13,557.71	11,784.15	12,433.33	12,427.01
5	13,582.69	13,533.03	11,781.18	12,367.13	12,260.71
6	13,582.69	13,483.86	11,607.86	12,179.76	12,049.71
7	13,582.69	13,495.78	11,574.52	11,965.19	12,013.15
8	13,582.69	13,507.9	11,292.71	12,030.13	11,813.66
9	13,582.69	13,532.99	11,403.33	11,991.37	11,747.56
10	13,582.69	13,519.78	11,174.39	11,819.75	11,869.43
11	13,582.69	13,519.98	11,094.87	11,846.88	11,543.39
12	13,582.69	13,481.7	11,098.45	11,735.96	11,640.71
13	13,582.69	13,507.62	11,178.56	11,831.09	11,693.77
14	13,582.69	13,494.92	11,028.94	11,579.31	11,596.84
15	13,582.69	13,481.28	11,075.92	11,688.5	11,458.84
16	13,582.69	13,480.98	**10,881.52**	11,538.69	11,511.13
17	13,582.69	13,492.81	10,895.7	11,449.9	11,443.27
18	13,582.69	13,467.54	10,960.35	11,633.1	11,484.58
19	13,582.69	13,441.88	10,950.44	11,461.99	11,479.11
20	13,582.69	13,455.77	10,896.47	11,488.65	11,374.74

Table 3.7 Transition matrix (A) of the resulting model in the Jaguas case. The initial probability for each state (π) is represented by row 0. Each cell represents the transition probability of the row state (i) to the column state (j). The greener the cell color, the higher probability for that transition, and the red cells indicate low transition probability. Notice that the row's sum equals one.

Current State (i)	Next state (j)					
	1	2	3	4	5	6
0	0.0456	0.1364	0.2723	0.1807	0.137	0.2273
1	0.8535	0.0275	0.0070	0.0770	0.0350	0.0000
2	0.0292	0.7772	0.0818	0.0665	0.0453	0.0000
3	0.0062	0.0784	0.7838	0.0096	0.0858	0.0362
4	0.0845	0.0437	0.0388	0.7820	0.0361	0.0150
5	0.0296	0.0269	0.0618	0.0163	0.8654	0.0000
6	0.0000	0.0169	0.0374	0.0152	0.0000	0.9305

Cluster two prototype had low values for β and NP because there were not many species calls. However, this recording had a high value for H_t and entropy for the medium band because it presented rain, which produced a series of temporal pulses (increasing temporal entropy) and occupied most of the spectrum.

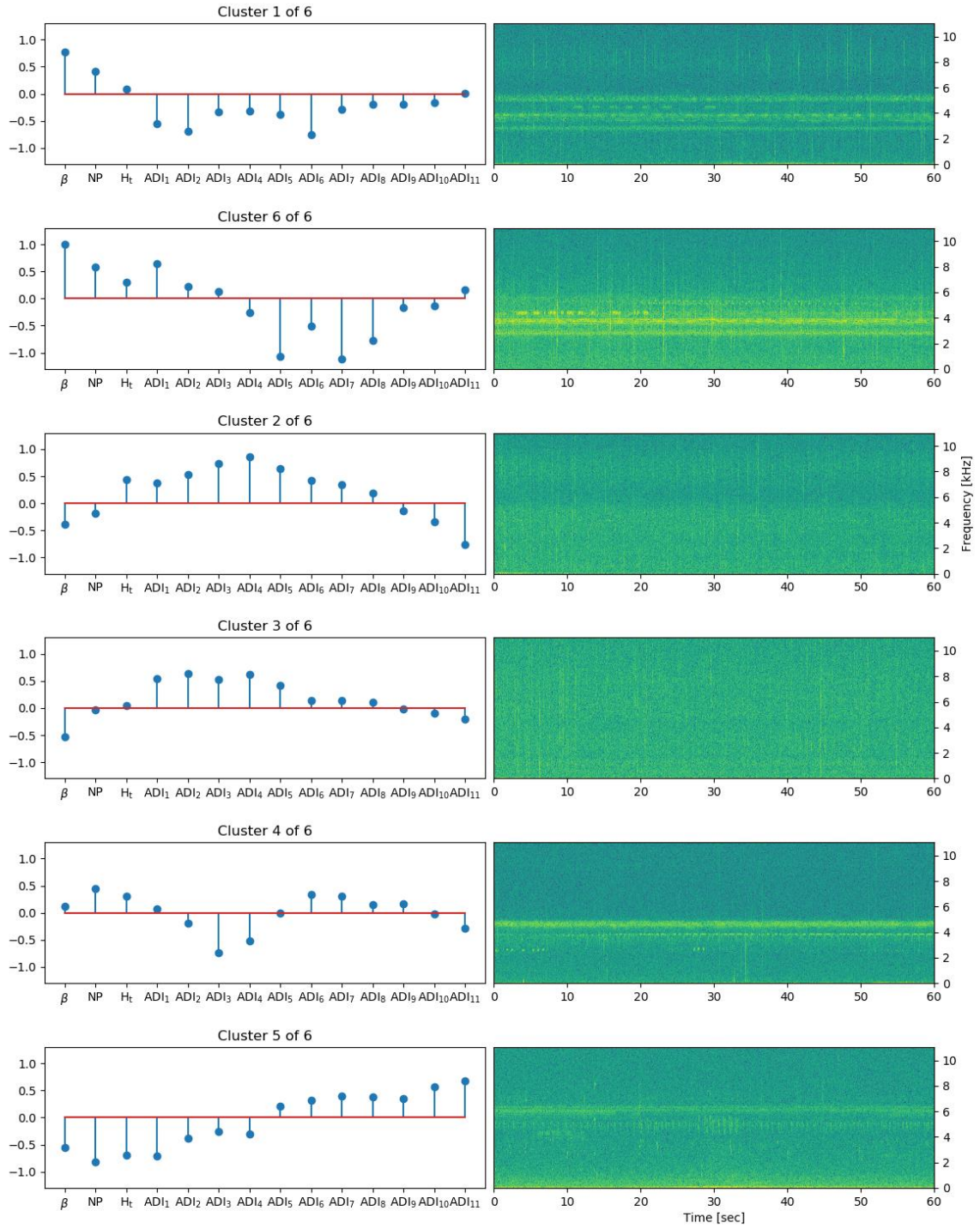

Figure 3.5 Mean vectors for the resulting clusters and the clusters' prototypes in the Jaguas application.

Cluster three prototype also presented rain, but this time no biophonic sound, which explained the low value for β.

Cluster four prototype had a couple of totally occupied frequency bands and some others partially occupied by periodical calls.

Although cluster five seemed to be the most interesting for high frequency analysis, we only found one peculiar compound call occupying the band between 7 kHz and 11 kHz. This recording also had constant, periodical and sporadic calls but with low intensity.

Finally, cluster six prototype was very similar to cluster one's, but with higher intensity and variety of calls.

In conclusion, cluster 1 and 6 are the most acoustically complex, because they presented more occupied bands, but the found calls tended to be periodical or constant. Clusters 5 and 4 had singing species but in less intensity and continuity. Lastly, clusters 2 and 3 had the least number of singing species but presented geophonic sound which produced high temporal entropy (for cluster 2) or high values for the $ADIs$, given that biophony was not constant and therefore their presence meant more information.

Certainly, higher values of ADI do not mean high occupation for the respective frequency bands, but higher randomicity of sounds, which can be due to rain or to the presence of different species in the band; whereby these values must be interpreted carefully.

Finally, looking at the transition matrix, we conclude that all clusters are stable, meaning that the probability of transition to a different cluster is relatively low (around 8% for the highest values). Cluster six seems to be the most stable and the most isolated cluster, given that from the other states is practically impossible to pass to this state (see column six in Table 3.7). However, this is one of the most likely starting points with a probability of 22.73 %.

The physical interpretation of the obtained clusters is made more clearly in the next section.

3.1.6. Model Implementation Results

In this stage, the Viterbi algorithm is applied to the dataset and the results for the cluster assignation to the observed samples are presented. Firstly, we inspected the cluster composition using the classification of the training samples, or how the ecosystem types of the training samples were distributed into the final clusters in Figure 3.6. It shows that cluster 1 and 6 were mostly assigned to forest samples, but both had a high percentage of pasture observations; cluster 2 had the highest percentage of stubble samples, but also had an important share of forest. On the other hand, cluster 3 and 4 had a majority of stubble samples but were also assigned to a big share of pasture samples. Lastly, cluster 5 had the highest percentage of assigned pasture observations and probably is the cluster that best represent the typical pasture soundscape.

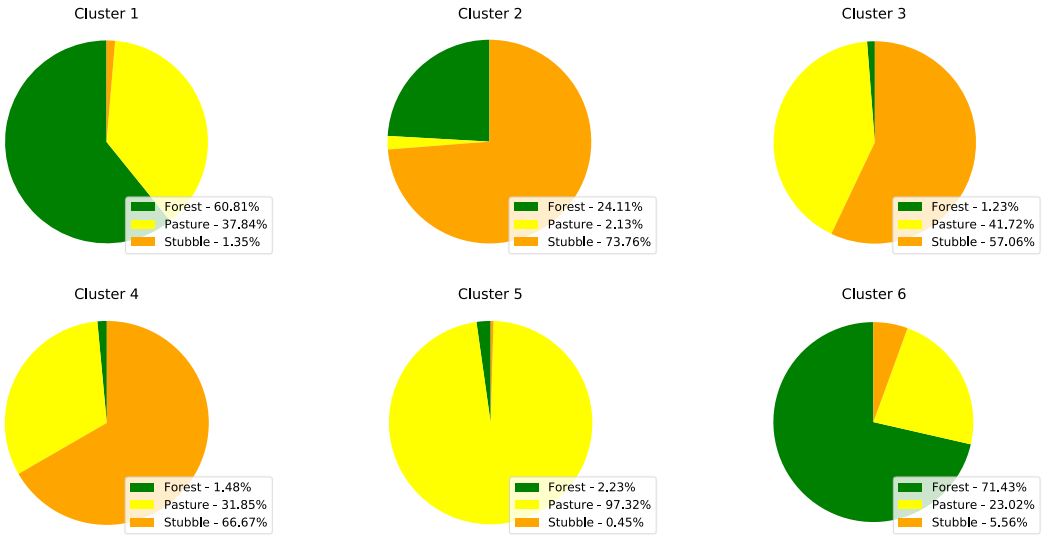

Figure 3.6 Training distribution of ecosystem types in the resulting clusters.

Notice that there is a high percentage of pasture samples in most clusters (in all of them, except for cluster 2), meaning that the pasture habitat presents many soundscape patterns, probably because its acoustical components are constantly changing through time or the local physical structure allows the propagation of many sound sources.

Next, each of the clusters was labeled with the ecosystem type presenting the highest proportion in it (the cluster's label is shown in

Table 3.8). Then, using these labels and the cluster's assignation to the complete dataset (training and testing), a compound confusion matrix was constructed for assessing the generalization capacity of the model. The matrix can be observed in Table 3.9.

Table 3.8 Cluster's label according to their training data distribution in the Jaguas case.

Label	Forest	Stubble	Pasture
Clusters number	1, 6	2, 3, 4	5

Table 3.9 Normalized confusion matrix for the training and testing Jaguas datasets. Notice that the sum of the rows equals one. Greener cells indicate higher values, redder cells indicate lower values.

True Label	**Dataset**	**Predicted Label**					
		1 (Forest)	6 (Forest)	5 (Pasture)	2 (Stubble)	3 (Stubble)	4 (Stubble)
Forest	Training	0.39	0.4	0.02	0.15	0.02	0.01
	Testing	0.46	0.43	0	0.07	0.04	0
Pasture	Training	0.13	0.07	0.53	0	0.16	0.1
	Testing	0.2	0	0.44	0.01	0.3	0.05
Stubble	Training	0.01	0.02	0	0.35	0.31	0.3
	Testing	0	0.01	0	0.28	0.41	0.3

A well generalized model would have similar values for training and testing data, which seems to be the case, although there is some variation for pasture samples. If the model is intended for classification, a diagonal confusion matrix is ideal. For our case, we found that 81% of the forest samples were classified in the forest clusters (1 and 6), 51.2% of the pasture samples were classified in the pasture cluster (5) and 96.6% of the stubble samples were classified in the stubble clusters (2, 3 and 4). In conclusion, forest and stubble were the easiest ecosystem types to be classified given the proposed methodology, but pasture was the hardest one to identify. Considering that ecosystem type classification is a complex task even for ecologists, because they can share multiple characteristics, we deemed these results very promising.

Finally, the temporal change in the sites is explained by Figure 3.7. In it, the classification for the complete dataset (training and testing) is displayed through time.

We can observe that site 5256 was the most stable site, but also had less samples than the others. This site was the furthest monitored point, and was in the depths of the forest, therefore it corresponded to the least disturbed habitat. Most of its samples were classified in cluster 6, proving that the better preserved the forest, the more different communities and more stability through time.

Cluster six was also assigned to the last samples of site 5260 during the first months of 2017. This indicates increase of calls intensity. For the past samples in this site, the associated cluster was number one, with a few exceptions in April and November, when cluster two had some apparitions, maybe due to rain.

Cluster two was mostly present in site 5255, which is a stubble site, but then disappeared in November of 2016 and made way for cluster 4 from that moment. This could indicate that for the first six months of monitoring, there was not much biophonic activity, but the following six months more species started to sing, probably because of a climatic factor.

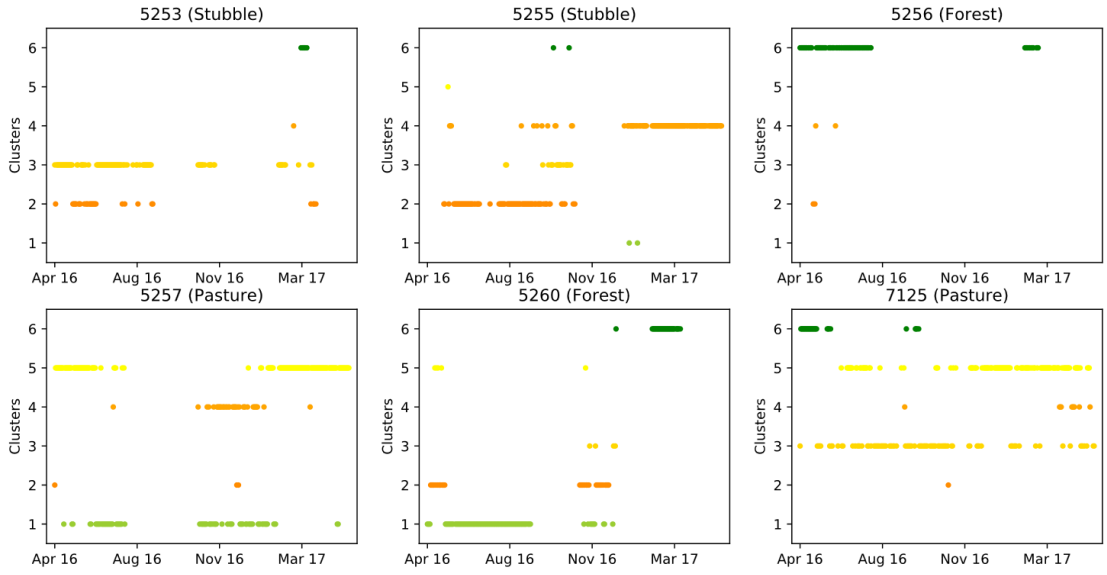

Figure 3.7 Temporal Classification of samples per site in the Jaguas case.

Cluster four and one alternated from November of 2016 to February of 2017 in the pasture site with recorder 5257. However, in the first and last monitored months, cluster five was more dominant. This site seemed to vary its acoustical activity by seasons, but apparently there were always some singing species in this place.

The same cannot be said about site 5253, which seemed to be a rather lacking space of biophonic activity, but with more presence of geophony. Finally, pasture site 7125 could have sometimes dominance of geophony (explained by cluster 3) or some biophonic activity (indicated by cluster 5).

Summing up, forest presented more occupied bands, indicating more species singing constant and periodical calls. Pasture seemed to be more of a passage area, where geophony dominated for some seasons, but species calls were still noticeable. Lastly, stubble sites did not present continuous and periodical sounds, but rather random sounds, probably due to geophony influence; also, did not seem a preferred space for making calls, given that not many frequency bands were occupied.

3.2. Dry Tropical Forest Transformation Application

This database was provided by the institute Alexander von Humboldt[5]. The original project was aimed to understand the biodiversity changes due to landscape transformation and succession in the Colombian tropical dry forests (Rodríguez-Buritica, 2017). Given the dimension and the heterogeneity of the dataset, in this work we only selected a subgroup of recordings from the Cañas river basin in the Guajira and Arroyo river basin in Bolivar.

3.2.1. Database Description

From December of 2015 to March of 2017, a series of recording sessions were realized in different points of the Cañas river basin and the Arroyo river basin in the Colombian Caribbean region. The sessions were established to record continuously for five days, then stop for another five days and restart this process during approximately one month. This process was repeated for other months during this period. However, in practice we found that this protocol was not strictly followed, or some recordings were missing, which negatively affected the model performance. A complete description of the selected recording sessions is presented in Table 3.10.

The recorders were programmed to register acoustical activity every 10 minutes. The resulting audio files were monoaural, presented a digital format of 16-bit PCB ($D = 16$), a sampling frequency of 48,000 Hz ($Fs = 48{,}000$ Hz) and lasted five minutes ($d = 300\ sec$). In general, the exact monitored sites were not repeated between sessions, except for sessions 3 and 7, sessions 4 and 5, sessions 10 and 12 and session 11 and 13.

For this application, we chose the recording sessions made by the same model of recorders: Wildlife Acoustics SM3. We found that SM2 recorders were also present in the original dataset and configured with the same settings as SM3 recorders, which actually produced significant differences in intensity (Eldridge et al., 2018).

[5] http://www.humboldt.org.co/es

The recording sessions were also selected to obtain a balanced model respecting the ecosystem transformation, which is the ecosystem type for this case.

Table 3.10 Description of the selected recording sessions for the Humboldt Application. The rows are colored according to their transformation value.

Session Number	Region	Transf.	Recorder	Location	Initial Date (AAAAMMDD)	Final Date (AAAAMMDD)	Number of Recordings
1	Guajira	High	5067	11.213 N, 73.434 W	20151212	20160123	3,239
2	Guajira	High	5071	11.214 N, 73.44 W	20161221	20161231	786
3	Bolívar	High	5071	9.896 N, 75.162 W	20160305	20160326	1,332
4	Bolívar	High	5069	9.893 N, 75.139 W	20160224	20160228	574
5	Bolívar	High	5071	9.893 N, 75.139 W	20170123	20170204	1,087
6	Bolívar	High	5071	9.888 N, 75.141 W	20170213	20170314	628
7	Bolívar	High	302143	9.896 N, 75.162 W	20170123	20170224	2,476
8	Guajira	Low	JSC5069	11.197 N, 73.436 W	20151212	20160123	3,237
9	Guajira	Low	302143	11.198 N, 73.435 W	20161221	20170114	1,103
10	Bolívar	Low	5071	9.908 N, 75.187 W	20160224	20160227	576
11	Bolívar	Low	5071	9.899 N, 75.192 W	20160214	20160217	576
12	Bolívar	Low	302151	9.908 N, 75.187 W	20170202	20170313	2,597
13	Bolívar	Low	302298	9.899 N, 75.192 W	20170123	20170206	1,440
14	Bolívar	Low	302298	9.903 N, 75.185 W	20170212	20170306	1,571
15	Guajira	Medium	5072	11.167 N, 73.444 W	20151212	20160102	2,010
16	Guajira	Medium	302151	11.168 N, 73.431 W	20161221	20170120	2,219
17	Bolívar	Medium	5070	9.94 N, 75.170 W	20160214	20160316	1,865
18	Bolívar	Medium	5067	9.936 N, 75.159 W	20160214	20160326	2,535
						Total	29,851

For the classification of the sites, the institute Alexander von Humboldt collected forest/non-forest layers from satellite imagery in the period between 1990 and 2012. Then, forest patches (of 30m x 30m dimension) were classified as “retained” if they were old forests, “lost” if they were younger forest than 4 years old, or “new” if they were older forest than 4 but younger than 22 years old. Consequently, high transformation zones corresponded to more proportion of lost forest than retained or new, low transformation zones presented higher proportion for retained or new forest and medium transformation zones were an approximate balance between forest and non-forest (see the third column of Table 3.10).

3.2.2. Preprocessing Results

Using the $PSDn$ algorithm, a total of 2,861 recordings were removed from the analysis due to heavy wind or rain, for 9.58% of the total recordings.

3.2.3. Features Extraction Results

Similar parameters to Jaguas’ were chosen for computing the set of features. Given that this is a comparative study, notice that the details for feature calculation are not relevant as long as the main concept is respected, and it is uniformly applied to the data samples.

For this case, the discarded metrics from Table 3.4 were not computed to speed up the process. This omission was made on the assumption that the correlation converges as the number of samples increase. Then, since Jaguas case had a considerable number of samples, the obtained correlation represented a generalized measure that could be applied to other applications. Also, H_v was replaced by H_m because in a previous result, the former had presented a lower $MAPC$. This change should not affect the results, because these features were highly correlated.

In this application the $ACIf_t$ index presented nan values as well. For this reason, 7 more recordings were discarded. In total 9.6 % of the initial database recordings was rejected in the previous and current stage.

Finally, the averages were computed over the days with at least 120 recordings (the proportion of minimum required recordings was maintained; remind that in this study the number of recordings per day doubled Jaguas’). A total of 143 daily samples were obtained. These samples were also standardized using (2.31).

3.2.4. Feature Selection Results

An additional correlation analysis between the remaining features was executed. Although Jaguas correlation analysis showed a general relation between features, the small number of samples for this case could lead to an increase of PCC for some features. Indeed, we found that the PCC for M and RMS was higher than 0,9 as

well as the PCC for $ACIf_t$ and FM (the respective Jaguas' $PCCs$ were 0.72 and 0.81). Consequently, we removed M and $ACIf_t$ from the analysis.

Next, we calculated the F-value for the obtained features given the transformation of the training samples. Figure 3.8 shows the resulting F-values.

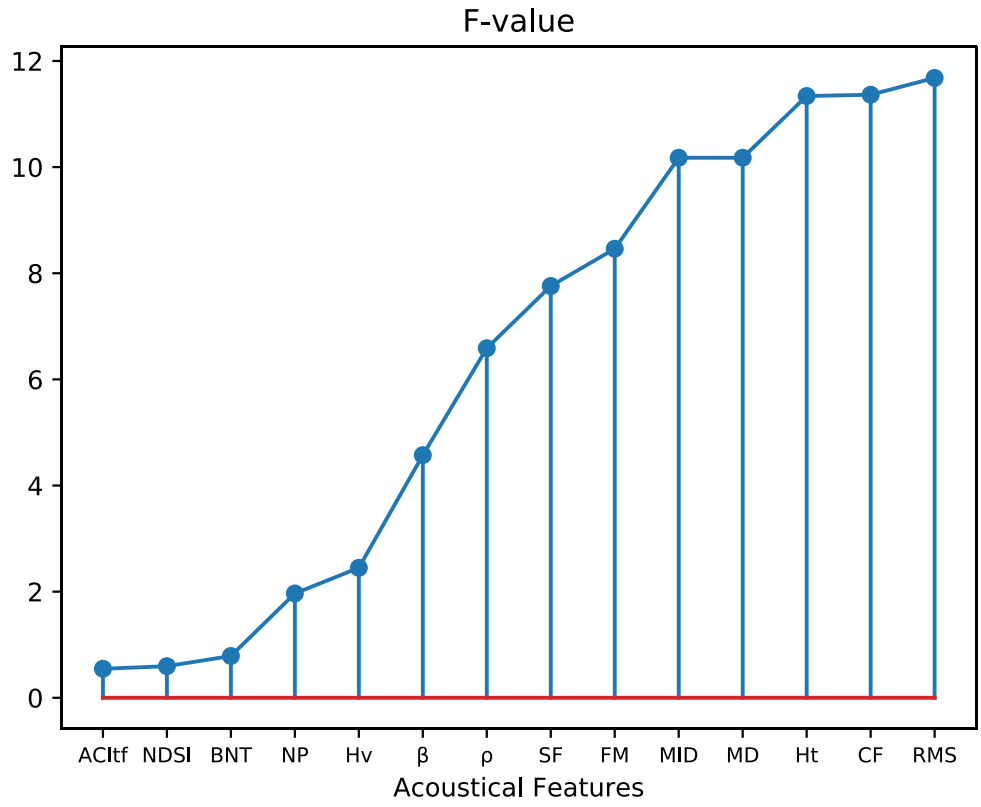

Figure 3.8 F-value for the selected features in the Humboldt application.

Observe that the features with highest F-value correspond to temporal complexity or intensity metrics; conversely, frequency complexity measures (as NP and $ACIt_f$) presented the lowest values, indicating that the number of singing species did not differ much between transformations. Note also that the highest F-values are much lower than Jaguas', which could be due to the dataset size difference or the difficulty in this case for identifying ecosystem types from acoustical measures. In Figure 3.9 we inspected the boxplots of the three highest scored features.

The boxplots confirmed that it was very difficult to separate transformation types given the soundscape features. Low and medium transformations seemed to present similar values for CF and H_t but differed a little in their RMS values. Nonetheless, transformation high and low were very similar for this last feature. Apparently, the low transformation could be easily confused with the other transformations on the acoustic level.

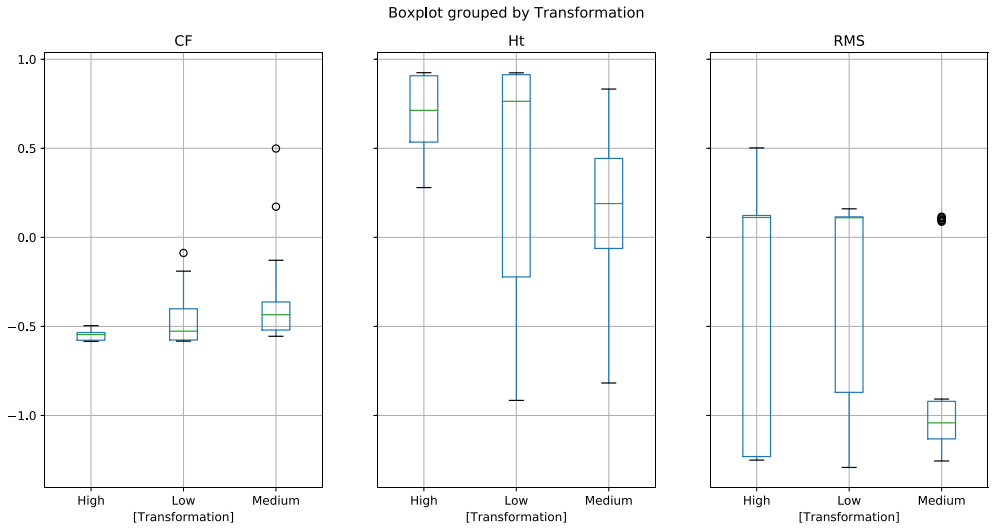

Figure 3.9 Grouped boxplots by ecosystem type for the three soundscape features in the Humboldt application

Lastly, we computed the $ADIs$ values maintaining the analysis up to 11 kHz. Soundscape studies usually limit to this frequency band, because higher frequencies present little information (Gage & Axel, 2014; Pijanowski, Farina, Gage, Dumyahn, & Krause, 2011; M. Sankupellay et al., 2016). The final feature vector is described in Table 3.11.

Table 3.11 Final feature vector for characterizing a daily sample in the Humboldt application.

Soundscape feat.			**Biodiversity features**										
RMS	H_t	CF	ADI_1	ADI_2	ADI_3	ADI_4	ADI_5	ADI_6	ADI_7	ADI_8	ADI_9	ADI_{10}	ADI_{11}

3.2.5. Model Set Up Results

In this case the N_m values were tested in the range between one and five, and the N_c values were tested in the range between three and fifteen, because higher N_c values were slowing down the Baum-Welch algorithm. The resulting MSEs are displayed in Table 3.12.

As shown in Table 3.12, the best N_c x N_m combination was 10 x 4. In this case, five states were replications of the others, then the model was reconfigured for having only five states and the Baum-Welch was repeated 20 more times. This process was repeated until the states were unique. At the end, only three states remained. Figure 3.10 and Table 3.13 summarize the final model. Figure 3.10 shows the mean vectors for each cluster and Table 3.13 presents the transition matrix A and initial probability π of the resulting model.

Although there were three final states, states two and three were very similar; the main difference between them was that state two had lower ADI values for the lowest frequency bands, whereas state three had lower ADI values for the highest frequency bands. These values indicate that state two presented a little more

information in the upper bands and state three had more information at the lowest bands. Besides, state two presented a lower RMS value, which is related to intensity, meaning that either the sound sources were further than in state three, or the soundscape was quieter. In general, these two states should present monotonous soundscapes, either for continuous sounds or silence.

Table 3.12 MSE values for all the possible combinations of N_g and N_m in the Humboldt application. The table cells are colored according to their values. Greener cells indicate lower MSE and redder cells indicate higher MSE.

Values for N_c	Values for N_m				
	1	2	3	4	5
3	1,192.40	1,192.40	1,150.22	1,065.53	1,114.81
4	1,192.40	1,192.40	1,117.79	1,077.29	1,071.45
5	1,192.40	1,192.40	1,126.38	1,043.69	1,077.86
6	1,192.40	1,192.40	1,111.73	1,011.58	1,079.30
7	1,192.40	1,192.40	1,105.63	1,023.93	1,019.30
8	1,192.40	1,192.40	1,112.03	1,052.14	1,040.70
9	1,192.40	1,192.40	1,097.48	1,023.84	1,034.57
10	1,192.40	1,192.40	1,112.92	**996.51**	1,031.56
11	1,192.40	1,192.40	1,112.61	1,035.74	1,026.51
12	1,192.40	1,192.40	1,087.09	1,011.76	1,034.42
13	1,192.40	1,192.40	1,098.34	1,003.85	1,037.21
14	1,192.40	1,192.40	1,113.04	1,025.92	1,055.41
15	1,192.40	1,192.40	1,096.56	1,047.31	1,037.45

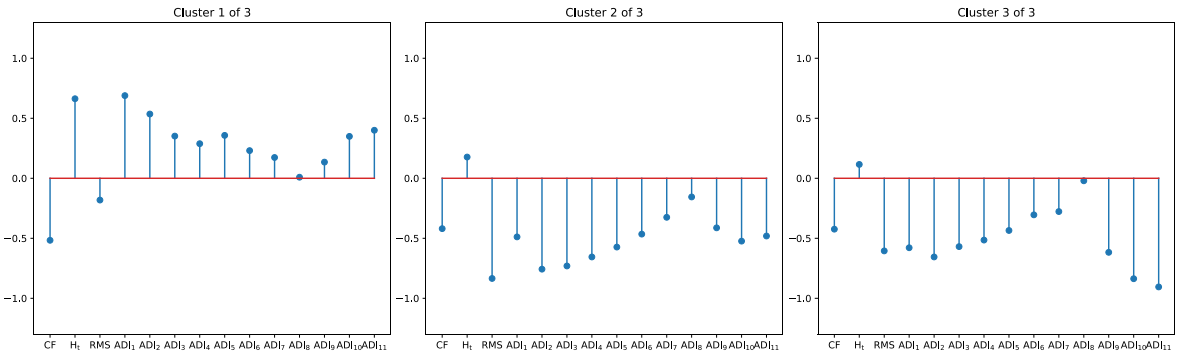

Figure 3.10 Mean vectors for the resulting clusters in the Humboldt application.

On the other hand, state one presented more randomicity expressed by its high levels of Ht and $ADIs$. Therefore, this state should represent more chaotic soundscapes.

Finally, notice that CF was approximately the same for the three states. This indicates that in general, the signal peaks did not separate as much from the effective value (RMS).

The transition probabilities indicate that state two would be the intermediate state for states one and three, given that its transition probability to either of these states was the same. Nonetheless, once the sequence reached state one, it was very difficult for it to pass to a different state. Notice that state one was the most

stable and isolated cluster but was also the most likely starting point, as happened to cluster six in Jaguas' case. Lastly, state three had a high probability to pass to state two, which makes sense given their similarity, however it was almost impossible for this state to pass to state one.

Table 3.13 Transition matrix (A) of the resulting model in the forest transformation case. The initial probability for each state (π) is represented by row 0. Each cell represents the transition probability of the row state (i) to the column state (j). The greener the cell color, the higher probability for that transition, and the red cells indicate low transition probability. Notice that the row's sum equals one.

Current State (i)	**Next State (j)**		
	1	**2**	**3**
0	0.5556	0.2593	0.1852
1	0.9268	0.0488	0.0244
2	0.0714	0.8572	0.0714
3	0.0000	0.2632	0.7368

3.2.6. Model Implementation Results

In this stage we implemented the Viterbi algorithm to find the final classification and we inspected the relation between the obtained clusters and the ecosystem transformations. In Figure 3.11 we used pie charts to estimate the transformation composition of each of the resulting states for the training dataset. We found that each cluster could be associated to an ecosystem transformation (see Table 3.14), but the average proportion for the assigned transformations was 56.57% for all states. In other words, we should not expect that a given cluster would produce always the same transformation, although most of the emissions would correspond to the assigned label.

Table 3.14 Cluster's label according to their training data distribution in the Humboldt case.

Label	Low Transformation	Medium Transformation	High Transformation
Cluster number	1	2	3

Given this assignation and looking back at Figure 3.10, we might also conclude that the high and medium transformations presented similar soundscapes and the low transformation soundscape tended to be very different. However, low transformation had an important presence in the medium and high clusters, which made it the most changing soundscape of the three.

Next, with the ecosystem type assignation to each cluster and the Viterbi algorithm results for both training and testing data, we computed the confusion matrices to assess the generalization capacity and classification accuracy. The compound matrix is displayed in Table 3.15.

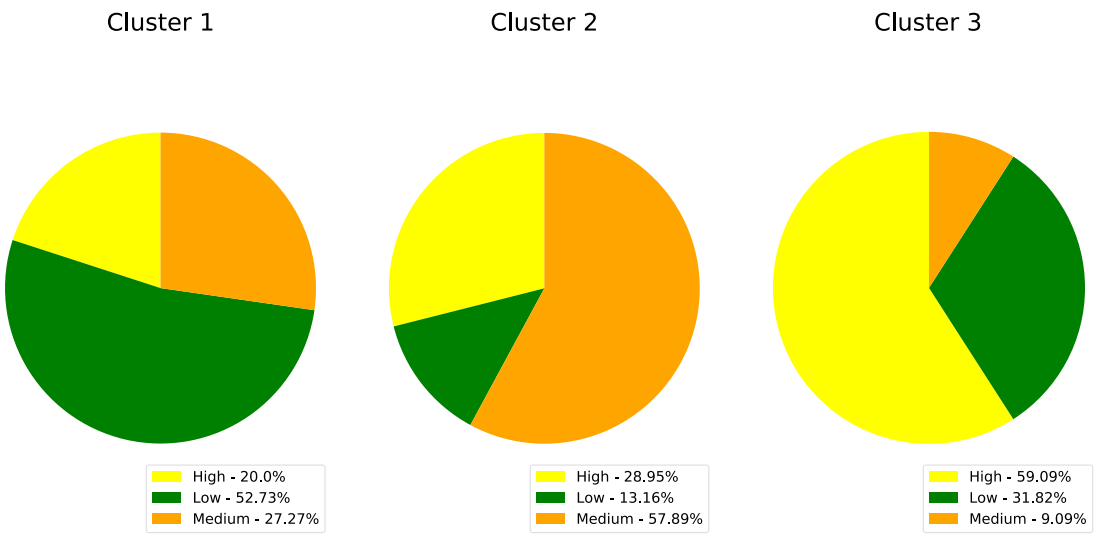

Figure 3.11 Training distribution of ecosystem transformations in the resulting clusters.

Table 3.15 Normalized confusion matrix for the training and testing Humboldt datasets. Notice that the sum of the rows equals one. Greener cells indicate higher values, redder cells indicate lower values.

True Transformation	Dataset	Predicted Label		
		3 (High)	1 (Low)	2(Medium)
High	Training	0.37	0.31	0.31
	Testing	0.11	0.44	0.44
Low	Training	0.17	0.71	0.12
	Testing	0	0.56	0.44
Medium	Training	0.05	0.38	0.56
	Testing	0.1	0.2	0.7

The matrix showed that the model was not sufficiently generalized because the dataset was too small. On the other hand, the classification results indicated that the model identified 68% of the total low transformation samples, 58.8% of the medium transformation samples and only 31.8% of the high transformation samples.

A temporal analysis was not presented because the locations were not continuously monitored, and the sequences were very short (the average sequence lasted 4.63 ± 3.62 days)

In conclusion, this methodology was not the best for this application. Although the original Humboldt's study analyzed image information for approximately 30 years (Rodríguez-Buritica, 2017), the acoustical analysis period was much shorter and irregular, presenting considerable gaps. GMMHMM is a sequential model, which requires long sequences to be properly trained and the dataset was both small and too heterogeneous to have long sequences. Consider also that in theory the methodology was applied to the same ecosystem type: tropical dry forests but with different ages. From this point of view, it seemed reasonable that the frequency complexity (related to the species diversity) did not vary between transformations. Additionally, the transformation

classification was applied to large areas, whereas this methodology is rather for local problems given the scope of the recorders.

Nonetheless, we might conclude that in this application, a high general entropy (temporal and biophonic) would represent older forests, whereas younger forests would present mostly monotonous chorus. Besides, both studies indicated that the better-preserved habitats would have high transition probability from the assigned state and itself, and that their initial probability would be high as well.

Chapter 4. Conclusions

4.1. Summary

In this work, a new methodology for assessing ecosystem change through ecoacoustics analysis was proposed. Firstly, the elimination of noisy recordings was executed using an automatic threshold on the PSD values in the 600 – 1200 Hz band. Secondly, the most common ecoacoustics indices and other acoustical features for soundscape analysis were analyzed using the Pearson correlation coefficient and the calculation of ANOVA F-value. The first was used to detect similar features and remove the duplicates to avoid redundancy, the second led to find the fittest features for the application and comprehend the main acoustical differences between the ecosystem types. After selecting the soundscape features, the biodiversity features were computed per band as a modification of the ADI. These values allow to recognize the frequency bands with more information or entropy, due to either random calls or geophony when generalized across the spectrum. Lastly, the model was adjusted using hidden Markov models with Gaussian mixture emissions. This model identified the main acoustical patterns that produced the observations. Each of the resulting states could be associated to a specific ecosystem type, but this was not the main purpose, but to recognize stationary and transitory acoustical patterns, which could be common to all types of habitat. The GMMHMM was also used to analyze the data as a time series and recognize the probable transitions between states on a daily basis.

The methodology was developed for the Jaguas study in the first place. The purpose of this study was to identify the acoustical signature of three ecosystem habitats by distributing two recorders per ecosystem type and registering the acoustical activity for over a year. The model identified six clusters: two were mostly assigned to forest samples, three had mostly stubble data and one presented a pattern almost unique for pasture sites. The classification led to a correct recognition of the 81% for forest, 51.2% for pasture and 96.6% for stubble. We could also recognize that the forest soundscapes presented more uniformity in their sounds, more occupied frequency bands (related to the number of species) and more biophonic intensity. On the other hand, the stubble soundscape had more entropy in their sounds, which was related to geophony but less biophonic activity. Finally, the pasture sites had the most versatile soundscapes with both geophony and biophony alternating during the year.

A second case in the Colombian Caribbean was used to apply the methodology. In this case, the recordings comprised a longer time span, but there were few continuous daily records. Additionally, the ecosystem types corresponded to forest transformations, which could be best interpreted as forest longevity. The resultant model did not present an accurate generalization capacity, given the heterogeneity and small size of the dataset. However, it showed that there was not a striking acoustical difference between high and medium transformed forests, whose age is less than 30 years old; but showed that older or low transformed forests presented a different acoustical signature to the rest, with more entropy across the spectrum. The model identified three underlying states, which we associated to the transformation levels and resulted in the recognition of 68% of the low transformed forest samples, 58.8% for the medium transformation and 31.8% for the high transformation.

In conclusion, this methodology is more appropriate for long term studies with uninterrupted sampling, preferably several times per day. It also requires the correct classification of the monitored habitats, for obtaining clusters that, associating to the studied ecosystem types, would be easier to interpret.

4.2. Contributions of this work

The proposed methodology made the following contributions to the ecoacoustics field:

4.2.1. Automatic removal of noisy recordings due to geophony

Based on the algorithm proposed by Bedoya et al. for rain detection (Bedoya et al., 2017), we proposed an automatic threshold for identifying the recordings with very high levels of geophony that would introduce noise to the analysis. We noticed that this threshold did not only discard recordings in which heavy rain was present, but it also detected intense wind, which could also mask most of the acoustical activity. Notice that our interest was not to discard all recordings with geophonic presence, because these sounds are an important part of the soundscape and may explain some characteristics of biophonic behavior (such as the acoustic adaptation hypothesis (Sueur & Farina, 2015)). Therefore, this threshold is suggested for soundscape applications and not for specific species study, in which biophony needs to be isolated and all geophony should be removed.

Although noise filtering is desired for most ecoacoustic studies, we only found Towsey's report addressing this problem in this field (Towsey, 2013). However, as himself put it, this method was aimed at removing constant sounds during a recording, that could correspond to any of the soundscape components, probably leading to an important loss of information. Since our method focuses in a frequency band typically occupied by geophony, we consider that it gives more information to the user about the type of noise that it is rejecting.

4.2.2. Analysis of ecoacoustic indices similarity

Most of the studies involving ecoacoustic indices are directed towards the identification of certain ecological qualities of the monitored habitats. These ecological qualities could include species richness (e.g. (Harris, Shears, & Radford, 2016)), ecosystem type (e.g. (Eldridge, Casey, Moscoso, & Peck, 2016)), landscape configuration (e.g. (Fuller et al., 2015)) etc. However, we did not find an actual study that compared the indices measurements to identify those who provided the same information. In this work, we computed the Pearson correlation coefficient for every pair of acoustical features to find the most representative features of all. For instance, we found that ADI is highly correlated to AEI, H and H_s, therefore using the four indices for a study would be redundant. Although these results, specially Jaguas', should show a general tendency; smaller scaled studies could present considerable differences in the features similarities given the bias of limited sampling.

4.2.3. Automatic recognition of daily soundscape patterns

Automatic recognition algorithms have been used mainly on the biophonic level for species identification (Bedoya et al., 2014; Eldridge et al., 2018; Tucker et al., 2014), but rarely used for characterization of terrestrial soundscape patterns and connection to habitat structure (Bormpoudakis et al., 2013; Ulloa, Aubin, Llusia, Bouveyron, & Sueur, 2018).

When the state of the art of this research was established, only Bormpoudakis et al. work had used clustering algorithms for grouping and linking soundscape patterns to ecosystem type (Bormpoudakis et al., 2013). However, this study was made on a small dataset, for a very specific time of the day and year and avoided other elements of soundscape as geophony and technophony. Therefore, the scope of its findings was very reduced.

Towards the end of our research, we discovered Ulloa et al. work, which identified regions of interest (ROIs) in the recordings, grouped them by unsupervised algorithms and characterized two types of ecosystems by using heat maps showing the distribution of the ROIs clusters (Ulloa et al., 2018). This study was also tested on a small dataset, for diurnal recordings and focused mainly on biophonic calls, given the procedure for the detection of the ROIs.

Although our methodology highlights the importance of biophonic sounds by extracting a higher number of features for characterizing this component (the $ADIs$), it does not discard the remaining soundscape components (geophony and technophony) from the recordings, because they are a relevant part of the ecosystem acoustical signature and model the communication behavior of communities as well. Besides, our analysis unit, which is the daily average of the recordings, encompasses all the acoustical information in the day, detailing the totality of the acoustical processes in the monitored site.

4.2.4. Temporal relation of soundscape patterns

Many studies have addressed the temporal variation of soundscapes (Bernie Krause et al., 2011; Matsinos et al., 2008; Mullet et al., 2016) but none of them has proposed an inner model that explains the relation between soundscape patterns and their stationary or transitional probabilities.

This feature allows our method to identify the temporal stability of soundscape patterns, indicating the permanence of singing communities; and the natural sequence of acoustical states, finding the seasonality of the ecosystem in the acoustical level. Furthermore, the transition matrix could help us detect forthcoming ecosystem changes and take early conservation decisions based on that information.

4.3. Publications

During the development of this master's project, the author published the following preliminary studies:

- *Duque-Montoya, D. C., Isaza, C. V., & Cano Rojas, E. (2017). Comparación de vocalizaciones de Espadarana prosoblepon en Dos Tipos de Ecosistemas usando Agrupamiento Jerárquico. Revista Iberica de Sistemas e Tecnologias de Informacao, 24, 13–21. http://doi.org/10.17013/risti.24.13–21*

This work consisted on the comparison of the vocalizations of two *Espadarana prosoblepon* populations located in different sites, a disturbed and a conserved forest. The results showed that the degradation of this species' habitat affected its calls structure.

- *Duque-Montoya, D. C., & Isaza, C. (2018). Automatic Ecosystem Identification Using Psychoacoustical Features. In Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence - PRAI 2018 (pp. 1–4). New York, New York, USA: ACM Press. http://doi.org/10.1145/3243250.3243251*

In this work, psychoacoustical features and two types of machine learning algorithms were tested for forest and stubble discrimination. The results indicated that neural networks and two kernel functions of support vector machines were able to recognize the difference between forest and stubble by their psychoacoustical description.

- *Daza, J. M., Isaza, C., Bedoya, C., Cano, E., Duque-Montoya, D. C., Gómez, W. E., & Sánchez-Giraldo, C. (2018). Automating biological monitoring on the Northern Andes of South America: combining biology and machine learning for conservation. In J. Gaikad, B. König-Ries, & F. Recknagel (Eds.), Proceedings of the 10th International Conference on Ecological Informatics (p. 17). Jena. Retrieved from https://icei2018.uni-jena.de/*
- *Duque-Montoya, D. C., Isaza, C., & Daza, J. (2018). Assessing Ecosystem Change using Soundscape Analysis. In J. Gaikad, B. König-Reis, & F. Recknagel (Eds.), Proceedings of the 10th International Conference on Ecological Informatics (pp. 229–230). Jena. Retrieved from https://icei2018.uni-jena.de/*
- *Sánchez-Giraldo, C., Cano, E., Gómez, W. E., Duque-Montoya, D. C., Isaza, C., Bedoya, C., & Daza, J. M. (2018). Understanding the Relationship between Soundscape and Landscape Features in a Tropical Andean Environment. In J. Gaikad, B. König-Ries, & F. Recknagel (Eds.), Proceedings of the 10th International Conference on Ecological Informatics (p. 18). Jena. Retrieved from https://icei2018.uni-jena.de/*

The previous three works were presented during the 10th international conference on ecological informatics (ICEI 2018) in Jena, Germany. The first presentation showed the general results from the application of passive acoustic monitoring and machine learning techniques on the conservation of different sites in the northern Andes of south America. The second presentation highlighted the results of our research group in soundscape analysis for ecosystem change assessing. And the third presentation showed a study on the relationship between ecoacoustic indices and landscape features in the tropical Andean environment.

4.4. Future Work

Extensive work is still to be done in this field. Firstly, ecoacoustics highlights the importance of three soundscape components on the acoustical communities: biophony, geophony and technophony. However, little research has been done on the latter two components, as the studies keep focusing in the communities' communication behavior. Additionally, the acoustical properties of each of these components is still unclear. Traditional methods have assumed each element to be limited on a certain frequency band, but we know that in practice,

that is not true, and the three elements constantly overlap on the spectrum. This overlapping is precisely one of the explanations for changes in the species communication behavior, whereby a key reason for studying all soundscape elements is being practically ignored.

Secondly, further exploration of the soundscape and ecological processes is required. The studies have focused on correlating soundscape and biodiversity or landscape configuration but are not going deeper to describe soundscape as a manifestation of the underlying ecosystem processes. For instance, an interesting study would relate soundscape characteristics to species migration, as an evidence to the acoustic adaptation hypothesis (Sueur & Farina, 2015) and its importance to ecosystem constitution.

Nonetheless, the limited scope of acoustical recording is still an obstacle to such studies. This observation leads to another necessity in soundscape studies, which is the clear definition of the intrinsic limitations of acoustical research. The reliability of soundscape findings can be severely affected by the actual spatial scope of the study, which is limited by the physical configuration of the monitored site and the nature of the acoustical signals of interest.

Lastly, the relevance of ecoacoustic indices is still uncertain. The proposal of many similar indices indicates that biodiversity abstraction from acoustical complexity is not clear yet. Furthermore, recent studies questions the validity of many acoustic indices, either for their theoretical basis (Sandoval, Barrantes, & Wilson, 2018) or for their actual applicability (Eldridge et al., 2018). For future applications a decision must be made either to continue to find a universal descriptor of ecoacoustic complexity, or to define application-specific features that maximize soundscape description.

Personally, I think that the current state of the art demands for more studies of the last type, for which machine learning algorithms present as valuable tools. Nonetheless, a better understanding of specific soundscape patterns would lead us inevitably to associate similar findings and slowly construct a formal theory from where the definitive ecoacoustical features would emerge.

References

Agranat, I. (2009). Automatically identifying animal species from their vocalizations. *Fifth International Conference on Bio-Acoustics, …*, 1–22. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Automatically+identifying+animal+species+from+their+vocalizations#0

Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ*, *1*, 19. http://doi.org/10.7717/peerj.103

Baker, M. C., & Logue, D. M. (2003). Population differentiation in a complex bird sound: A comparison of three bioacoustical analysis procedures. *Ethology*, *109*(3), 223–242. http://doi.org/10.1046/j.1439-0310.2003.00866.x

Bedoya, C., Isaza, C., Daza, J. M., & López, J. D. (2014). Automatic recognition of anuran species based on syllable identification. *Ecological Informatics*, *24*, 200–209. http://doi.org/10.1016/j.ecoinf.2014.08.009

Bedoya, C., Isaza, C., Daza, J. M., & López, J. D. (2017). Automatic identification of rainfall in acoustic recordings. *Ecological Indicators*, *75*, 95–100. http://doi.org/10.1016/j.ecolind.2016.12.018

Bilmes, J. A. (1998). A gentle tutorial of the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *ReCALL*, *4*(510), 126. http://doi.org/10.1080/0042098032000136147

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. *Pattern Recognition* (Vol. 4). http://doi.org/10.1117/1.2819119

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., … Kirschel, A. N. G. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus. *Journal of Applied Ecology*, *48*(3), 758–767. http://doi.org/10.1111/j.1365-2664.2011.01993.x

Blumstein, D. T., & Munos, O. (2005). Individual, age and sex-specific information is contained in yellow-bellied marmot alarm calls. *Animal Behaviour*, *69*(2), 353–361. http://doi.org/10.1016/j.anbehav.2004.10.001

Boelman, N. T., Asner, G. P., Hart, P. J., & Martin, R. E. (2007). Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. *Ecological Applications*, *17*(8), 2137–2144. http://doi.org/10.1890/07-0004.1

Bormpoudakis, D., Sueur, J., & Pantis, J. D. (2013). Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications. *Landscape Ecology*, *28*(3), 495–506. http://doi.org/10.1007/s10980-013-9849-1

Collier, T. C., Blumstein, D. T., Girod, L., & Taylor, C. E. (2010). Is alarm calling risky? Marmots avoid calling from risky places. *Ethology*, *116*(12), 1171–1178. http://doi.org/10.1111/j.1439-0310.2010.01830.x

De Coensel, B. (2007). *Introducing the temporal aspect in environmental soundscape research*. Ghent University. Retrieved from http://users.ugent.be/~bdcoense/content/data/pdf/phd/DecoenselPHD07.pdf

De Coensel, B., Botteldooren, D., Debacq, K., Nilsson, M. E., & Berglund, B. (2007). Soundscape classifying ants. In *Internoise*. http://doi.org/10.1260/135101007781447993

De Coensel, B., Botteldooren, D., Debacq, K., Nilsson, M. E., & Berglund, B. (2008). Clustering outdoor

soundscapes using fuzzy ants. *2008 IEEE Congress on Evolutionary Computation, CEC 2008*, 1556–1562. http://doi.org/10.1109/CEC.2008.4630999

De Solla, S. R., Fernie, K. J., Barrett, G. C., & Bishop, C. A. (2006). Population trends and calling phenology of anuran populations surveyed in Ontario estimated using acoustic surveys. *Biodiversity and Conservation*, *15*(11), 3481–3497. http://doi.org/10.1007/s10531-004-6905-9

Delaney, D. K., Grubb, T. G., Beier, P., Pater, L. L., & Hildegard Reiser, M. (1999). Effects of helicopter noise on Mexican spotted owls. *Journal of Wildlife Management*, *63*(1), 60–76. http://doi.org/10.2307/3802487

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*(1), 46–54. http://doi.org/10.1016/j.ecolind.2011.05.006

Eldridge, A., Casey, M., Moscoso, P., & Peck, M. (2015). A New Direction for Soundscape Ecology ? Toward the Extraction and Evaluation of Ecologically-Meaningful Soundscape Objects using Sparse Coding Methods. *Peer J Computer Science*, *2020*, 1–17. http://doi.org/10.7717/peerj.2108

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., & Eyre-walker, Y. (2018). Sounding out ecoacoustic metrics : Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. *Ecological Indicators*, *94*(March), 1–14. http://doi.org/10.1016/j.ecolind.2018.06.012

Erbe, C., King, A. R., Yedlin, M., & Farmer, D. M. (1999). Computer models for masked hearing experiments with beluga whales (Delphinapterus leucas). *Journal of the Acoustical Society of America*, *105*(5), 2967–2978. http://doi.org/10.1121/1.426945

Everest, F. A., & Shaw, N. A. (2001). *Master Handbook of Acoustics, Fourth Edition*. *The Journal of the Acoustical Society of America* (Vol. 110). http://doi.org/10.1121/1.1398048

Farina, A., Pieretti, N., Salutari, P., Tognari, E., & Lombardi, A. (2016). The Application of the Acoustic Complexity Indices (ACI) to Ecoacoustic Event Detection and Identification (EEDI) Modeling. *Biosemiotics*, *9*(2), 227–246. http://doi.org/10.1007/s12304-016-9266-3

Fletcher, N. (2007). Animal Bioacoustics. In T. D. Rossing (Ed.), *Springer Handbook of Acoustics* (1st ed., pp. 473–491). New York: Springer New York. http://doi.org/10.1007/978-0-387-30425-0_19

Forney G.D., J. (1973). The viterbi algorithm. *Proceedings of the IEEE*. http://doi.org/10.1109/PROC.1973.9030

Fuller, S., Axel, A. C., Tucker, D., & Gage, S. H. (2015). Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? *Ecological Indicators*, *58*(February 2016), 207–215. http://doi.org/10.1016/j.ecolind.2015.05.057

Gage, S. H., & Axel, A. C. (2014). Visualization of temporal change in soundscape power of a Michigan lake habitat over a 4-year period. *Ecological Informatics*, *21*, 100–109. http://doi.org/10.1016/j.ecoinf.2013.11.004

Gage, S. H., Wimmer, J., Tarrant, T., & Grace, P. R. (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. http://doi.org/10.1016/j.ecoinf.2017.01.002

Gaitán Forero, D. (2017). *Índice para estimar cambios en ecosistemas a partir del análisis del paisaje acústico*. Universidad de Antioquia.

Gasc, A., Sueur, J., Pavoine, S., Pellens, R., & Grandcolas, P. (2013). Biodiversity Sampling Using a Global Acoustic Approach: Contrasting Sites with Microendemics in New Caledonia. *PLoS ONE*, *8*(5), e65311. http://doi.org/10.1371/journal.pone.0065311

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, *405*(6783), 220–7.

http://doi.org/10.1038/35012228

Harris, S. A., Shears, N. T., Radford, C. A., & Reynolds, J. (2016). Ecoacoustic indices as proxies for biodiversity on temperate reefs. *Methods in Ecology and Evolution*, *7*(6), 713–724. http://doi.org/10.1111/2041-210X.12527

Hmmlearn developers. (2016). Hmmlearn. Retrieved from https://hmmlearn.readthedocs.io/en/latest/

Hobson, K. a., Rempel, R. S., Greenwood, H., Turnbull, B., & Van Wilgenburg, S. L. (2002). Acoustic surveys of birds using electronic recordings: New potential from an omnidirectional microphone system. *Wildlife Society Bulletin*, *30*(3), 709–720. http://doi.org/10.2307/3784223

Job, J. R., Myers, K., Naghshineh, K., & Gill, S. A. (2016). Uncovering Spatial Variation in Acoustic Environments Using Sound Mapping. *PLoS ONE*, *11*(7), 19. http://doi.org/10.1371/journal.pone.0159883

Jones, E., Oliphant, E., & Peterson, P. (2001). Scipy: Open Source Scientific Tools for Python. Retrieved from https://scipy.org

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*(3–4), 259–276. http://doi.org/10.1016/j.landurbplan.2011.08.001

Jorge, F. C., Machado, C. G., da Cunha Nogueira, S. S., & Nogueira-Filho, S. L. G. (2018). The effectiveness of acoustic indices for forest monitoring in Atlantic rainforest fragments. *Ecological Indicators*, *91*, 71–76. http://doi.org/10.1016/j.ecolind.2018.04.001

Kasten, E. P., Gage, S. H., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, *12*, 50–67. http://doi.org/10.1016/j.ecoinf.2012.08.001

Kirschel, A. N. G., Earl, D. A., Yao, Y., Escobar, I. A., Vilches, E., Vallejo, E. E., & Taylor, C. E. (2009). Using Songs To Identify Individual Mexican Antthrush Formicarius Moniliger: Comparison of Four Classification Methods. *Bioacoustics-the International Journal of Animal Sound and Its Recording*, *19*(1–2), 1–20. http://doi.org/10.1080/09524622.2009.9753612

Kogan, J. A., & Margoliash, D. (1998). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. *J Acoust Soc Am*, *103*(4), 2185–2196. http://doi.org/10.1121/1.421364

Krause, B. (1987). Bioacoustics, habitat ambience in ecological balance. *Whole Earth Review*, *57*, 14–18.

Krause, B., Gage, S. H., & Joo, W. (2011). Measuring and interpreting the temporal variability in the soundscape at four places in Sequoia National Park. *Landscape Ecology*, *26*(9), 1247–1256. http://doi.org/10.1007/s10980-011-9639-6

Krause, B. L., & Farina, A. (2016). Using ecoacoustic methods to survey the impacts of climate change on biodiversity. *Biological Conservation*, *195*, 245–254. http://doi.org/10.1016/j.biocon.2016.01.013

Langbauer, W. R. (2000). Elephant communication. *Zoo Biology*, *19*(5), 425–445. http://doi.org/10.1002/1098-2361(2000)19:5<425::AID-ZOO11>3.0.CO;2-A

Lengagne, T., Aubin, T., Lauga, J., & Jouventin, P. (1999). How do king penguins (Aptenodytes patagonicus) apply the mathematical theory of information to communicate in windy conditions? *Proceedings of the Royal Society B: Biological Sciences*, *266*(1429), 1623–1628. http://doi.org/10.1098/rspb.1999.0824

Lengagne, T., & Slater, P. J. B. (2002). The effects of rain on acoustic communication: Tawny owls have good reason for calling less in wet weather. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1505), 2121–2125. http://doi.org/10.1098/rspb.2002.2115

Leong, K. M., Ortolani, A., Burks K. D, Mellen, J. D., & Savage, A. (2003). Quantifying acoustic and temporal characteristics of vocalizations for a group of captive African elephants Loxodonta africana. *Bioacoustics*, *13*(3), 213–231. http://doi.org/10.1080/09524622.2003.9753499

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., … Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: Assessing change in ecological communities through time. *Trends in Ecology and Evolution*, *25*(10), 574–582. http://doi.org/10.1016/j.tree.2010.06.016

Marcellini, D. L. (1974). Acoustic Behavior of the Gekkonid Lizard, Hemidactylus frenatus. *Herpetologia*, *30*(1), 44–52. Retrieved from http://www.jstor.org/stable/3891412

Mathworks, I. (2013). MATLAB. Natick. Retrieved from https://la.mathworks.com/products/matlab.html?s_tid=hp_products_matlab

Matsinos, Y. G., Mazaris, A. D., Papadimitriou, K. D., Mniestris, A., Hatzigiannidis, G., Maioglou, D., & Pantis, J. D. (2008). Spatio-temporal variability in human and natural sounds in a rural landscape. *Landscape Ecology*. http://doi.org/DOI 10.1007/s10980-008-9250-7

Mcwilliam, J. N., & Hawkins, A. D. (2013). A comparison of inshore marine soundscapes. *Journal of Experimental Marine Biology and Ecology*, *446*, 166–176. http://doi.org/10.1016/j.jembe.2013.05.012

Merchant, N. D., Fristrup, K. M., Johnson, M. P., Tyack, P. L., Witt, M. J., Blondel, P., & Parks, S. E. (2015). Measuring acoustic habitats. *Methods in Ecology and Evolution*, *6*(3), 257–265. http://doi.org/10.1111/2041-210X.12330

Mitrović, D., Zeppelzauer, M., & Breiteneder, C. (2010). Features for Content-Based Audio Retrieval. *Advances in Computers*, *78*(10), 71–150. http://doi.org/10.1016/S0065-2458(10)78003-7

Mullet, T. C., Gage, S. H., Morton, J. M., & Huettmann, F. (2016). Temporal and spatial variation of a winter soundscape in south-central Alaska. *Landscape Ecology*, *31*(5), 1117–1137. http://doi.org/10.1007/s10980-015-0323-0

Pekin, B. K., Jung, J., Villanueva-Rivera, L. J., Pijanowski, B. C., & Ahumada, J. A. (2012). Modeling acoustic diversity using soundscape recordings and LIDAR-derived metrics of vertical forest structure in a neotropical rainforest. *Landscape Ecology*, *27*(10), 1513–1522. http://doi.org/10.1007/s10980-012-9806-4

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*(3), 868–873. http://doi.org/10.1016/j.ecolind.2010.11.005

Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., & Krause, B. L. (2011). What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, *26*(9), 1213–1232. http://doi.org/10.1007/s10980-011-9600-8

Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., … Pieretti, N. (2011). Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*, *61*(3), 203–216. http://doi.org/10.1525/bio.2011.61.3.6

Qi, J., Gage, S. H., Joo, W., Napoletano, B., & Biswas, S. (2007). Soundscape characteristics of an environment: a new ecological indicator of ecosystem health. In *Wetland and Water Resource Modeling and Assessment: A Watershed Perspective* (Vol. 20071553, pp. 201–214). http://doi.org/10.1201/9781420064155

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. http://doi.org/10.1109/5.18626

Rasmussen, C. E., & Williams, C. K. I. (2004). *Gaussian processes for machine learning. International journal of neural systems*. http://doi.org/10.1142/S0129065704001899

Razali, M. H. M., Zakaria, A., Shakaff, A. Y. M., Saad, R. V. F. S. A., Kamarudin, L. M., & Abdullah, N. S. H. (2015). Visualisation of acoustic entropy index for rainforest health monitoring system. *2015 IEEE SENSORS - Proceedings*, 5–8. http://doi.org/10.1109/ICSENS.2015.7370531

Rodríguez-Buritica, S. (2017). *GEF-Tropical Dry Forest*. Bogota.

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. http://doi.org/10.1016/j.ecoinf.2013.12.006

Rychtáriková, M., & Vermeir, G. (2013). Soundscape categorization on the basis of objective acoustical parameters. *Applied Acoustics*, *74*(2), 240–247. http://doi.org/10.1016/j.apacoust.2011.01.004

Sandoval, L., Barrantes, G., & Wilson, D. R. (2018). Conceptual and statistical problems with the use of the Shannon-Weiner entropy index in bioacoustic analyses. *Bioacoustics*, pp. 1–15. http://doi.org/10.1080/09524622.2018.1443286

Sankupellay, M., Dema, T., Tarar, S., Towsey, M., Truskinger, A., Brereton, M., & Roe, P. (2016). Visual Analytics of Eco-Acoustic Recordings: The Use of Acoustic Indices to Visualise 24-Hour Recordings. In *2016 Big Data Visual Analytics, BDVA 2016*. http://doi.org/10.1109/BDVA.2016.7787051

Sankupellay, M., Towsey, M., Truskinger, A., & Roe, P. (2015). Visual Fingerprints of the Acoustic Environment: The Use of Acoustic Indices to Characterise Natural Habitats. In *2015 Big Data Visual Analytics, BDVA 2015*. http://doi.org/10.1109/BDVA.2015.7314306

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is It Really Robust? *Methodology*. http://doi.org/10.1027/1614-2241/a000016

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(July 1928), 379–423. http://doi.org/10.1145/584091.584093

Slabbekoorn, H. (2004). Singing in the wild: The ecology of birdsong. In *Nature's Music: The Science of Birdsong* (pp. 178–205). http://doi.org/10.1016/B978-012473070-0/50009-8

Slabbekoorn, H., & Bouton, N. (2008). Soundscape orientation: a new field in need of sound investigation. *Animal Behaviour*, *76*(4). http://doi.org/10.1016/j.anbehav.2008.06.010

Smith, J. O. (2007). *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications*. W3K Publishing. Retrieved from https://ccrma.stanford.edu/~jos/mdft/

Spellerberg, I. F., & Fedor, P. J. (2003). A tribute to Claude-Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the "Shannon-Wiener" Index. *Global Ecology and Biogeography*, *12*(3), 177–179. http://doi.org/10.1046/j.1466-822X.2003.00015.x

Sueur, J., & Farina, A. (2015). Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. *Biosemiotics*, *8*(3), 493–502. http://doi.org/10.1007/s12304-015-9248-x

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*(4), 772–781.

Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid Acoustic Survey for Biodiversity Appraisal. *PLoS ONE*, *3*(12), e4065. http://doi.org/10.1371/journal.pone.0004065

Tchernichovski, O., Nottebohm, F., Ho, C., Pesaran, B., & Mitra, P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour*, *59*(6), 1167–1176. http://doi.org/10.1006/anbe.1999.1416

Tobias, J. A., Aben, J., Brumfield, R. T., Derryberry, E. P., Halfwerk, W., Slabbekoorn, H., & Seddon, N. (2010). Song divergence by sensory drive in amazonian birds. *Evolution*, *64*(10), 2820–2839. http://doi.org/10.1111/j.1558-5646.2010.01067.x

Torija, A. J., Ruiz, D. P., & Ramos-Ridao, a F. (2013). Application of a methodology for categorizing and differentiating urban soundscapes using acoustical descriptors and semantic-differential attributes. *The Journal of the Acoustical Society of America*, *134*(1), 791–802. http://doi.org/10.1121/1.4807804

Towsey, M. (2013). *Noise removal from waveforms and spectrograms derived from natural recordings of the environment.* Retrieved from http://eprints.qut.edu.au/61399/

Towsey, M., Wimmer, J., Williamson, I., & Roe, P. (2014). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*, *21*, 110–119. http://doi.org/10.1016/j.ecoinf.2013.11.007

Truax, B. (1998). Models and Strategies for Acoustic Design. *Hey Listen!*, 8–16. Retrieved from http://www.academia.edu/1355504/Models_and_strategies_for_acoustic_design

Tucker, D., Gage, S. H., Williamson, I., & Fuller, S. (2014). Linking ecological condition and the soundscape in fragmented Australian forests. *Landscape Ecology*, *29*(4), 745–758. http://doi.org/10.1007/s10980-014-0015-1

Ulloa, J. S., Aubin, T., Llusia, D., Bouveyron, C., & Sueur, J. (2018). Estimating animal acoustic diversity in tropical environments using unsupervised multiresolution analysis. *Ecological Indicators*, *90*(November 2017), 346–355. http://doi.org/10.1016/j.ecolind.2018.03.026

Urazghildiiev, I. R., & Clark, C. W. (2006). Acoustic detection of North Atlantic right whale contact calls using the generalized likelihood ratio test. *Journal of the Acoustical Society of America*, *120*(4), 1956–1963. http://doi.org/10.1121/1.2257385

Villanueva-Rivera, L. J., Pijanowski, B. C., Doucette, J., & Pekin, B. (2011). A primer of acoustic analysis for landscape ecologists. *Landscape Ecology*, *26*(9), 1233–1246. http://doi.org/10.1007/s10980-011-9636-9

Wildlife Acoustics, 2011. Song Meter User Manual.

Aval

Thesis

Methodology for ecosystem change assessing using ecoacoustics analysis

Advisor

Claudia Victoria Isaza Narváez