Advancing marine soundscape ecology with low-cost recorders and machine learning

Submitted by Ben Williams to the University of Exeter as a thesis for the degree of Masters by Research in Biological Sciences, April 2021.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signed:

Acknowledgments

Firstly, I would like thank my supervisor Steve Simpson for offering me the opportunity to complete an MbyRes with the South West Aquatics Group, the expert supervision he provided and the confidence Steve gave me every step of way. I'm also very grateful to Lucille Chapuis for her co-supervision which has been integral to this thesis as well as Tim Gordon and Andy Radford for their close support that was essential throughout. I would also like to thank Isla Davidson and Emma Weschke for taking me under their wing as a research assistant and giving the opportunity to contribute to the kind of work I used to dream of being involved with whilst on the Great Barrier Reef. I would like to extend further gratitude to Harry Harding and Sophie Nedelec for taking the time to share their expertise on numerous occasions. I also wish to thank fellow MbyRes colleagues Isla Hely and Max Robinson for the frequent discussions and support we were able to offer each other and wish them the very best with their own projects. I would like to thank the University of Exeter Education Incubator, Mars Sustainable Solutions and Lizard Island Research Station for supporting work related to this thesis. I would like to extend additional gratitude to Tim Gordon and all other authors who contributed to the 2018 investigation performed in Spermonde, whose hard work was integral to providing the data I needed for my second chapter. The majority of this work was completed during the global pandemic and I would like to thank my supervisors, fellow group members and the University of Exeter whom I could not fault in any way for the support I was offered throughout this period. An especially large thank you should also go to all of those who got stuck with me during two national lockdowns!

Table of Contents

List of Figures and Tables

significant correlation ($p<0.05$). Blank cells indicate indices for which values from the corresponding frequency band were not calculated (see methods).

Figure 3.5: Violin plots of the three indices with the most significant differences between healthy and degraded habitat. (A) Medium band H (Mann-Whitney U; $U = 1.98$, $p<0.001$), (B) Full band ACI ($U = 1.78$, $p<0.001$), (C) Medium band th ($U = 1.63$, $p<0.001$).

Figure 3.6: Habitat classifications predicted by the machine learning model. Each cell indicates a single one minute recording from the 152 that were available across healthy and degraded habitats. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time for every recording. Values within cells represent the proportion of these 1000 iterations in which the recording was predicted as originating from a healthy site, with the remaining being predicated as degraded, also represented by the colour scale. Recordings taken on the left of the partition were taken during the day and recordings to the right were taken during crepuscular or night time periods. Although frequent gaps were present in the sampling regime, the order with which cells are presented within their respective blocks conserves the overall order with which they were sampled across time.

Figure 3.7: Plot from the principal component analysis of PC1 and PC2 for the Healthy and Degraded site recording samples. Samples from recordings of Restored sites are overlaid on this to help determine whether these conform with either of the two existing classes or whether the properties of their soundscape are distinct. Ellipses indicate the zone within which a new sample can be assigned to a class using the two principle components presented in this figure alone. Overlapping areas indicate ambiguous results which cannot be differentiated but nonetheless fit one of the existing classes.

Figure 3.8: Habitat classification predictions by the machine learning model for the restored site recording samples. Each cell indicates a single one-minute recording from the 110 that were taken from restored sites. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time for every recording. Values within cells represent the proportion of these 1000 iterations in which the recording was predicted as originating from a healthy site, with the remaining being predicated as degraded, also represented by the colour scale. Recordings taken on the left of the partition were taken

during the day and recordings to the right were taken during crepuscular or night time periods. Although frequent gaps were present in the sampling regime, the order with which cells are presented within their respective blocks conserves the overall order with which they were sampled across time.

Supplementary 3.2: Manuscript from Gordon *et al*., (in review) submitted to the Journal of Applied Ecology.

Supplementary 3.3: Results from a Spearmans collinearity test between all 12 indices in each of three frequency bands: low (0.05 – 0.8 kHz), medium (2 – 7 kHz), full (0.05 – 20 kHz). Darker cells represent a stronger correlation. Blank cells indicate values = -0.01 – 0.01.

Supplementary 3.4: Scatterplots between each of the eight indices selected for inclusion as features in the final model. Values from healthy and degraded sites alongside values from restored sites are included to enable divergence between these two groups to be observed if present.

Supplementary 3.5: Results from the Mann-Whitney U test between the index scores of healthy (n = 81) and degraded (n = 71) site recordings. Indices not calculated for specific frequency bands are left blank.

Chapter One:

Soundscape ecology and passive acoustic monitoring in the marine environment

Authors: Ben Williams[a], Stephen D. Simpson[a]

Affiliations: [a]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PS, United Kingdom

Author contributions: B.W. was primarily responsible for conducting and writing the review with comments from S.D.S.

1.1 Soundscape ecology

In any given habitat, there is typically an abundance of sensory information available. Sound is a ubiquitous component of this, the cumulative sum of which can be defined as the environmental 'soundscape' (Canada*, et al.,* 1978). Typically, sounds within a habitat can be divided into three major components: biophony, geophony and anthrophony (Pijanowski *et al.*, 2011; Duarte *et al.*, 2021). Biophony consists of all sounds emanating from a biological source (Krause, 1987). This may be through deliberate sound production, such as vocalisations (Lieberman, 1968; Tricas and Boyle, 2009; Potamitis, 2015) or use of objects in the environment to create sounds (Deakos, 2002). Alternatively, it may be unintentional, such as disturbance during motion (Larsson, 2012) or feeding (Krause, 2008; Pijanowski *et al.*, 2011). The other natural component of the soundscape of an environment is the geophony. This is comprised of sounds of a non-biological origin, including processes such as weather, the splashing of waves or rocks crumbling (Kull, 2006; Erbe *et al.*, 2015). The final component, anthrophony, encompasses all sounds of a human origin caused by activities including transportation and construction (Joo *et al*., 2011).

Soundscapes therefore contain a broad range of information about the processes occurring within a given habitat. Detecting, interpreting and contributing to this information can be essential to the survival of many organisms (Pijanowski *et al.*, 2011). Soundscape ecologists can also use this information to understand more about the ecology of a habitat and the life within. Until recently, this field of study has focussed on the terrestrial environment

Soundscape ecology in the marine environment

A common misconception is that the ocean soundscape is a void and empty domain (Cousteau and Dumas, 1953). However, the soundscape of the marine environment is in fact rich and varied. Sound is the dominant sensory process in the activities of many marine organisms. An excellent example of this is whale song, which enables the largest animals in the ocean to communicate across hundreds of kilometres, further than any terrestrial animal (Clark, 1990). Tiny organisms also rely on sound: fish, crustacean, mollusc and coral larvae have all been shown to use sound as a sensory cue to orientate towards suitable habitat before it is near enough to be visually or chemically detected (Simpson *et al.*, 2005; Lillis *et al.*, 2015; Lillis *et al.*, 2016; Gordon *et al.*, 2018). The ubiquitous sound of snapping shrimp is unmistakable, with the ceaseless crackling of their snaps revealing the presence of these highly cryptic invertebrates across almost all marine habitats (Versluis *et al.*, 2000). There are countless other organisms across a broad range of taxa which contribute further to the underwater biophony for a number of reasons (Tricas and Boyle, 2009; Gedamke and Robinson, 2010; Lobel *et al*., 2010; Coquereau *et al.*, 2016). These reasons may be passive, such as the distinctive scraping of parrotfish whilst feeding (Tricas and Boyle, 2014). Active sounds are also produced, including courtship displays of signallers which are in turn acted upon by receivers in mate selection (Lobel *et al*., 2010). Alternatively, these may be used to mediate agonistic interactions within and between species (Tricas and Boyle, 2014). Acoustic signals can also be intercepted by predators to detect prey, resulting in many species reducing their sound output in the presence of predators (Holt and Johnston, 2011). In addition to biophony, geophonic noise is also a significant component. This is primarily driven by waves and rain (Cazau *et al.*, 2017; Putland *et al.*, 2017) as well as being punctuated with rare occurrences such as thawing ice in glacial regions (Geyer *et al.*, 2016) and underwater earthquakes (Dziak *et al.*, 2015).

These natural sources vary between biological communities and habitats (Joseph and Margolina, 2014; Haver *et al.*, 2018), making the soundscape of the marine environment diverse and changeable over space and time (Matsinos *et al.*, 2008; Rodriguez *et al.*, 2014; Putland *et al.*,2017). Common trends in this are seen over diel, lunar and seasonal periods, which dictate the natural processes within habitats and the behavioural responses of organisms to this (Staaterman *et al.*, 2014; Buscaino *et al.*, 2016; Insley *et al.*, 2017; Kaplan *et al.*, 2018). Anthrophony, emanating from sources such as shipping, construction and seismic surveys, has also attracted growing interest (Williams *et al.*, 2015; Duarte *et al.*, 2021). This pollutant has doubled each decade for 60 years (Veirs *et al.*, 2018) and has increasingly been shown to negatively impact marine ecosystems and the life within (Wright *et al.*, 2007; de Soto, 2016).

Why studying marine soundscapes can be useful

Soundscape ecology in the terrestrial environment has led to many key discoveries, yet the advantages of studying sound in the underwater environment are perhaps even greater. Traditional surveying techniques typically rely on visual counts and identification of organisms. However, in almost every marine habitat, visibility is frequently too low for visual surveys to be conducted, and many are mostly inaccessible using this technique (Mooney *et al.*, 2020). Conversely, sound can travel much further and faster in water, at 4.3 times the speed of sound in air, allowing noises to be detected over great distances (Rogers and Cox, 1988). Sound in air can rarely be detected further than a few kilometres from its source, whereas, at the extrmeme end, some sounds in the ocean can be detected thousands of kilometres from the source (Lurton, 2002).

Capitalising on the information held within the diverse array of sounds present in the ocean, and the ability to detect this over scales often not possible with visual survey techniques, is therefore of interest to ecologists. Studying soundscapes can indicate key information about habitats and the behaviour of organisms present within these (Cotter, 2008; Rossi *et al.*, 2016; Gordon *et al.*, 2018). More focused components of bioacoustics and the marine soundscape, such as marine mammal communication, have been studied for some time (Cotter, 2008). However, the practice of studying the broader acoustic environment in the marine setting is an emerging field, only growing amongst ecologists in recent years (Lindseth and Lobel, 2018; Mooney *et al.*, 2020). Soundscapes have now been studied in a range of marine habitats, from tropical and temperate shallow reef habitats to the polar regions and the deep sea (Staaterman *et al.*, 2013; Harris *et al.*, 2016; Haver *et al.*, 2018; Lin and Tsao, 2018).

A broad range of biodiversity and functional attributes have been measured from marine soundscapes (Elise *et al.*, 2019; Mooney *et al.*, 2020). These attributes can be important indicators of a system's ability to provide ecosystem services and resist stressors (Ferrigno *et al.*, 2016; Hughes *et al.*, 2017). For example, soundscape properties such as amplitude, stochasticity and temporal trends have been found to have a relationship with the abundance and diversity of fish communities, essential for productive fisheries (Kaplan *et al.*, 2015; Elise *et al.*, 2019b). Such properties can also be an indicator of habitat quality (Piercy *et al.*, 2014; Bertucci *et al.*, 2016) which can be a key determinant of services such as coastal protection, income from tourism and provisions from biodiversity (Hicks *et al*., 2013; Pascal *et al.*, 2016; Otrachshenko and Bosello, 2017). Similar research has revealed differences across degraded and healthy systems allowing those in need of protection or restoration to be identified (Butler *et al.*, 2016; Gordon *et al.*, 2018). The changed soundscape on degraded systems has in turn been found to reduce recruitment from a variety of taxa, potentially spiralling sites into a negative feedback loop (Rossi *et al.*, 2017; Gordon *et al.*, 2018; Gordon *et al.*, 2019).

Although it is agreed that useful insights can be revealed from the study of the marine acoustic environment (Haver *et al.*, 2018; Lindseth and Lobel, 2018), the full potential of this endeavour is not certain (Buxton *et al.*, 2018) and contradictory findings on the relationship between properties of the soundscape and the wider ecosystem exist. For example, sound pressure level and spectral based acoustic metrics have been found to hold differing relationships with attributes such as fish diversity and benthic cover (Nedelec *et al.*, 2015; Freeman and Freeman, 2016; Buxton *et al.*, 2018; Elise *et al.*, 2019). The sound of snapping shrimp has also been found to both positively (Gordon *et al.*, 2018) and negatively (Nedelec *et al.*, 2015) correlate with habitat health, with a further study reporting no relationship (Kaplan *et al.*, 2015).

The validity of some of these findings has also been criticised due to short recording times (Mooney *et al.*, 2020). Misperceptions also exist about the attributes that some metrics may be quantifying, with many authors assuming fish vocalisations are driving results when in fact other factors may be the predominant determinants (Bohnenstiehl *et al.*, 2018; Bolgan *et al.*, 2018). Opposite trends in the values of modern computational metrics (discussed in Section 1.2) have also been observed, with different studies reporting the same metrics to be both affected (Bohnenstiehl *et al.*, 2018) and unaffected (Harris *et al*., 2016) by changes in spectral resolution of the recordings under analysis.

Further work is needed to elucidate which properties relate to particular aspects of ecology and how these may change spatially and temporally (Nedelec *et al.*, 2015; Elise *et al.*, 2019). The use of soundscape ecology has revolutionised the ability of terrestrial practitioners to perform rapid landscape assessments and gather important information on biological communities (Farina and Pieretti, 2012; Agnieszka, 2017). The same developments in the marine environment could set the path for equally valuable advances in this field, if not greater due to the advantages it offers over traditional underwater survey techniques.

1.2 Methods used to study marine soundscapes

Collecting marine soundscape recordings

Soundscape recordings are typically made by hydrophones, the underwater equivalent of a microphone (Lillis and Mooney, 2018). These devices consist of a piezoelectric transducer which converts acoustic pressure into an electrical current (Lau *et al.*, 2002). These devices are deployed at sites of interest for the collection of soundscape data in a practice known as passive acoustic monitoring (PAM) (Sousa-Lima *et al.*, 2013). Many considerations are needed during this stage such as the length of deployment and frequency range to sample from, determined by the sampling rate of a device. Oversampling can create very large datasets which require large amounts of digital storage to match and can lead to unworkable computing times during download and processing. Instead, investigations typically use duty cycles and frequency filters for long deployments where shorter recordings of a high quality are made periodically by these devices across the frequency band of interest to collect datasets which are more practical to archive and analyse (Farina, 2013)

Traditional methods to analyse recordings

Once soundscape recordings have been collected the next step is to analyse the information within these. There are a number of possible approaches, the choice of which is dependent on the questions under investigation. Auditory inspection is the most obvious of these, where investigators listen to playbacks of their recordings and note points of interest such as the frequency of occurrence of certain acoustic events (Putland *et al*., 2017; Archer *et al.*, 2018). However, this approach is slow and can limit investigations to signals within the range of human hearing sensitivity and difficulty can be had when interpreting multiple simultaneous signals. A more sophisticated and thorough approach is to convert the recording into an image, called a spectrogram, which displays the full spectrum of the recorded sounds and their intensity across time (Fig.1.1) (Archer *et al.*, 2018; Carriço *et al.*, 2020). Visual inspection of the spectrogram can then be used in complement to, or independently of, auditory inspections to more

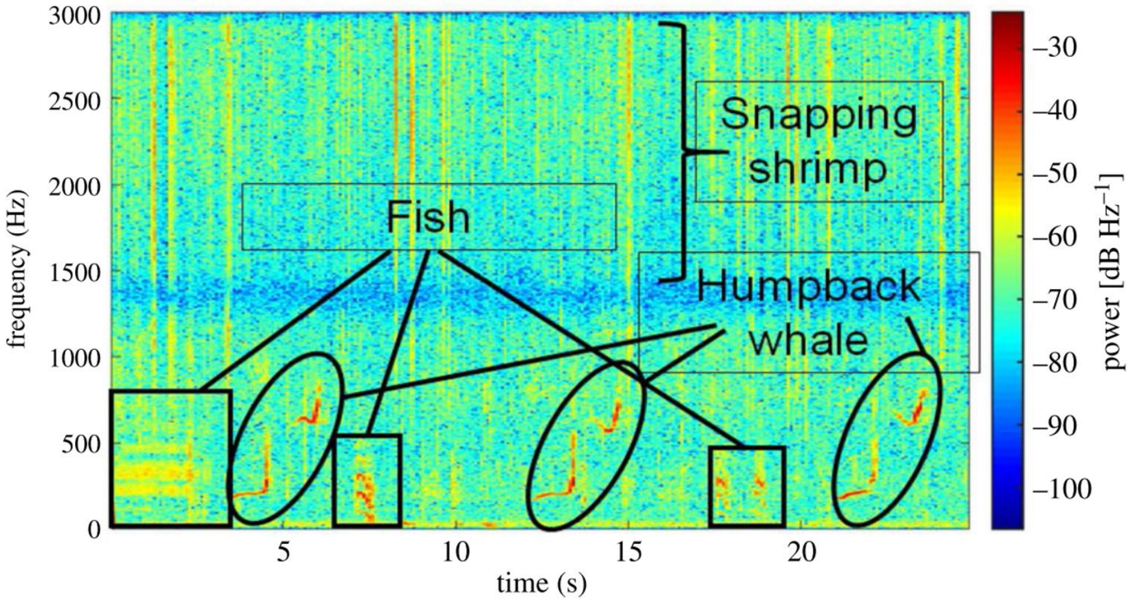


Figure 1.1 Spectrogram of a 30 second coral reef soundscape from Mooney *et al.* (2020). The plot shows frequency against time, relative intensity is indicated by colour, with darker colouration indicating low intensity sounds and lighter colouration indicating high intensity sounds. Sounds from biological sources are highlighted.

rapidly reveal conspicuous acoustic signatures or changes over time (Archer *et al*., 2018; McWilliam *et al*., 2018). Other plots such as power spectral density (PSD), which reveal the predominant frequencies present in a broadband signal, can be used to supplement this (Staaterman *et al*., 2014; McWilliam *et al*., 2018).

Assessing marine soundscapes using computer generated indices

Auditory and spectrogram approaches can be effective in some scenarios. However, PAM investigations can amass large quantities of data from long term recordings and human led inspection of these recordings can be limited by available time and repeatability with human investigators. Computer generated indices are an emerging solution that enables rapid assessments of soundscapes (Sueur *et al.*, 2008; Depraetere *et al.*, 2012; Buxton *et al.*, 2018). These approaches typically use algorithms to assess components of the spectrogram through searches for patterns or randomness across time, frequency and amplitude (Sueur, 2018a). They then quantify these findings and output values that correspond to these properties (Sueur *et al.*, 2014; Buxton *et al.*, 2018). Known as eco-acoustic indices, such approaches were primarily developed in the terrestrial environment with success in many areas (Depraetere *et al.*, 2012; Sueur *et al.*, 2014).

Recently, soundscape ecologists began applying these to recordings of marine habitats. This has revealed some interesting relationships between a number of indices and the ecological processes in certain habitats (Buxton *et al.*, 2018; Lindseth and Lobel, 2018; Mooney *et al.*, 2020). Over the last five years the use of acoustic indices has become standard practice in marine soundscape ecology investigations attempting to advance this field (Pieretti and Danovaro, 2020). Researchers now mostly use a combination of these indices alongside other surveying approaches to investigate their utility (Buxton *et al.*, 2018; Lindseth and Lobel, 2018). However, of the dozens of eco-acoustic indices in circulation amongst bio-acousticians and soundscape ecologists, only a select few have been investigated on marine soundscapes.

Summarising relevant indices

The following section outlines the nine indices that have previously been found to have a specific relationship with a component of marine ecology (Pieretti and Danovaro, 2020), and introduces three others that have been trialled for the first time in the marine setting as part of this thesis.

Amplitude index and sound pressure level

The most commonly used acoustic indices in the marine environment relate to amplitude. This is effectively a quantification of how loud a recording is. The amplitude index (M) calculates the median of the amplitude envelope of a recording (Fig. 1.2) (Sueur, 2018a), which is usually performed over short durations (e.g., seconds to minutes). However, this index can be subject to variability depending on the equipment and conditions in which recordings were taken. Sound pressure level (SPL) is similar to M but is measured in decibels (dB) relative to the intensity of a reference pressure to provide an exact intensity level

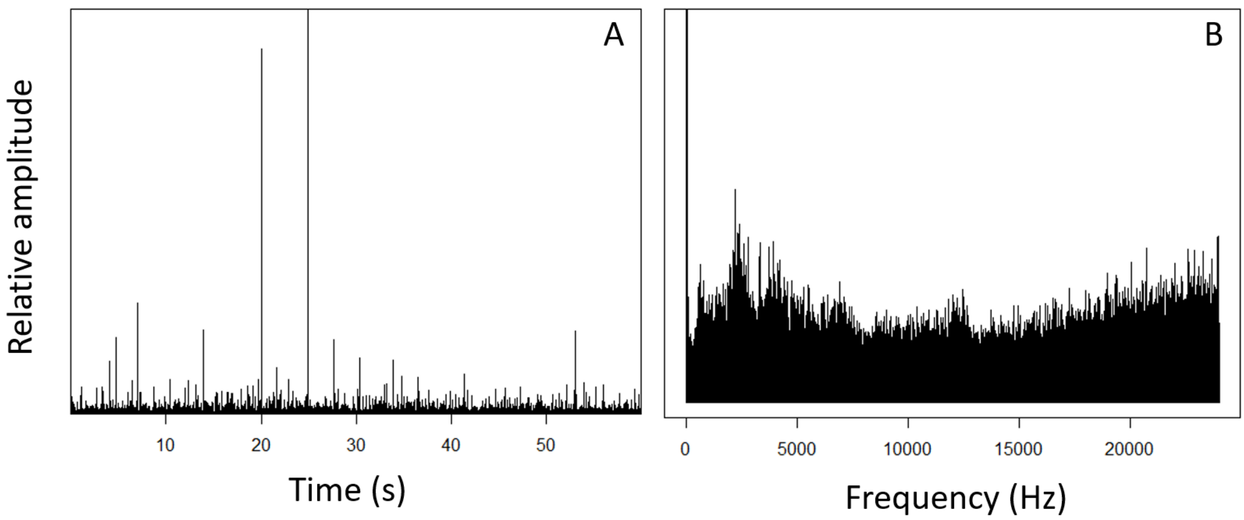


Figure 1.2. (A) Amplitude envelope of a one minute recording taken on a coral reef during this Masters study. The variability of this is assessed by the temporal entropy index (th). Amplitude is expressed relative to itself and as such no units are given. (B) A fast Fourier transform (FFT) of the same recording, this displays the relative amplitude for each frequency band. The variability of amplitude across frequency bands is assessed by the spectral entropy index (sh).

(Merchant *et al.*, 2015). This requires the calibration of instrumentation against reference sounds to be performed. The route mean square sound pressure level ($_{rms}$SPL) is commonly used to provide a working average of intensity across the duration of a recording.

Acoustic complexity index

The acoustic complexity index (ACI) is primarily designed to enable the discrimination of acoustically complex biotic sounds from background noise. ACI operates on the principle that biotic noises, especially vocalisations, exhibit an intrinsically higher variability of intensity across certain frequency bands (Pieretti *et al*., 2011). ACI operates by calculating the difference in intensity between each frequency bin in a temporal window and the intensity in the same frequency bin in adjacent windows. This is performed across the full length of the recording and the sum of all the differences between adjacent frequency bins is taken. ACI accounts for distance from the receiver by dividing this result by the total sum of intensities in the recording. High values of the ACI index indicates a greater abundance of complex signals that exhibit variable frequencies and intensities, such as frequent and varied vocalisations of birds, for which it was first developed (Bradfer-Lawrence *et al.*, 2019).

Acoustic diversity and evenness

The acoustic diversity index (ADI) and acoustic evenness index (AEI) were designed together to provide landscape ecologists with a way to quantify the level of variability in the intensities of different frequency bands present within a soundscape (Villanueva-Rivera *et al.*, 2011). ADI and AEI operate by dividing the spectrum of a recording into different frequency bins (e.g.,10), and then calculating the proportion of the total sound in the recording that is in each frequency bin. Each index then applies a diversity index to these results: ADI uses the Shannon evenness index (Spellerberg and Fedor, 2003) and AEI uses the Gini index (Hurlbert, 1971; Pekin *et al.*, 2012). These indices produce values from 0–1. Recordings in which the total sound is spread evenly across frequency bands report a high ADI value and a low AEI value, and *vice versa* when this distribution is uneven (Bradfer-Lawrence *et al.*, 2019).

Temporal entropy and spectral entropy

Temporal entropy (th) is used to measure the randomness in the total amplitude of a soundscape over time. The index takes the size of the amplitude envelope of a recording at each time point for which it is available (Fig.1.2A). It then executes the Shannon evenness on these values to provide the temporal entropy index (Sueur *et al.*, 2008). This index produces values from 0–1, with low values indicating a recording where the amplitude envelope has a high random variability. This value can be subtracted from 1 to yield temporal variability (TV), as used in Elise *et al.* (2019a). Spectral entropy (sh) operates on a similar principle but instead uses squared mean amplitude values of the full frequency spectrum. This is obtained from a fast Fourier transformation (FFT) of the recording (Fig. 1.2B) which when plotted appears similar to the amplitude envelope (Fig. 1.2A). It then applies the Shannon index to this transformation to yield spectral entropy. This is again reported between 0–1, with low values indicating a uniform distribution of amplitude across frequencies and high values indicating a random distribution (Sueur *et al.*, 2008).

Acoustic entropy

Acoustic entropy (H), also known as total entropy, measures randomness in the spectral and temporal domain of a soundscape. It is the product of temporal entropy and spectral entropy (Fig. 1.2), producing values between 0–1. Higher values of H indicate a recording with a greater evenness across all frequencies over time, often dominated by silence. Lower values indicate the dominance of a single frequency band across time, for example the continuous presence of chorusing from one species or from a source of chronic anthropogenic noise (Bradfer-Lawrence *et al.*, 2019).

Acoustic richness

The Acoustic Richness (AR) index was developed as an improvement upon the acoustic entropy index. The initial index was found to be less reliable in habitats with a reduced diversity and frequency of sounds and therefore AR was created (Depraetere *et al.*, 2012). AR ranks each recording by its median amplitude then multiplies each rank by the corresponding recordings temporal entropy. It then divides this value by the square of the number of samples. AR is different from other indices in that it must be applied to a group of recordings due to the ranking component. A value between 0–1 is produced, with higher values indicating a richer soundscape due to a greater diversity of vocalisations and other complex sounds (Depraetere *et al.*, 2012).

Bioacoustic index

The bioacoustics index (BI) is designed to quantify the total level of sound across the full spectrum. This uses a simple approach that begins by dividing the recording into a number of frequency bins determined by the user (e.g., 10). It then identifies the minimum intensity present at any point across the whole recording and calculates how much greater the mean intensity of each of the frequency bins is above this minimum. It then sums these to provide a BI value. This can be visualised by plotting a curve (Fig. 1.3.), with BI being the total area

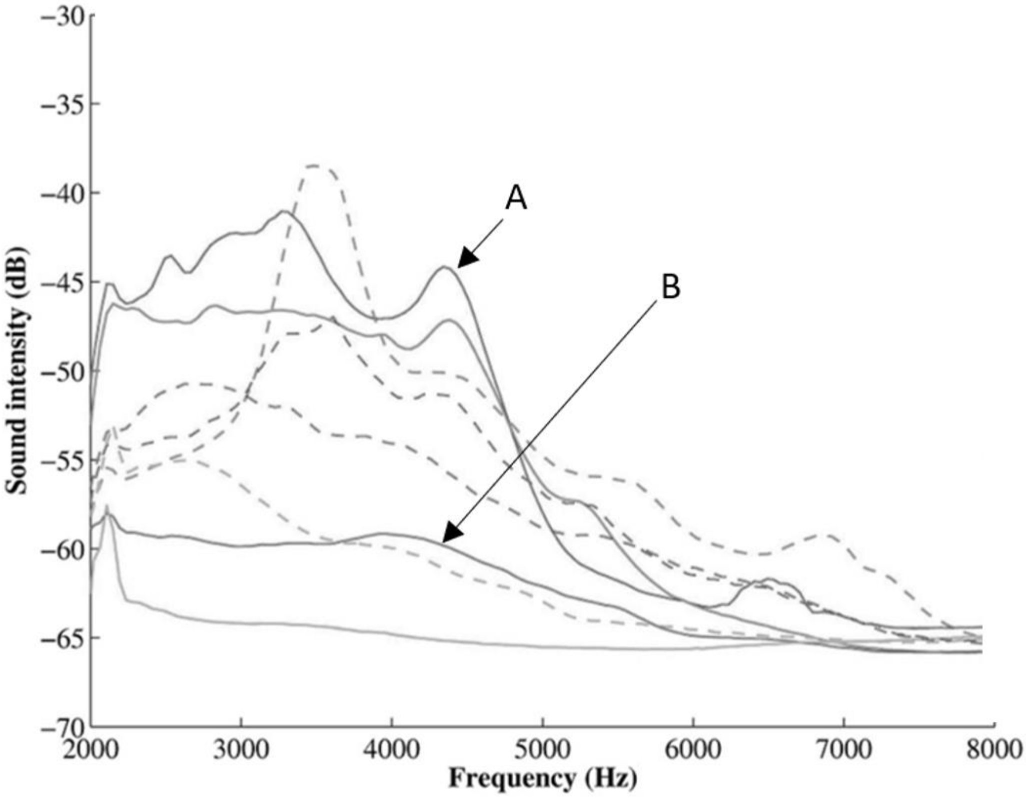


Figure 1.3. This plot, adapted from Boalman *et al.* (2007), shows the bioacoustics index (BI) curve for eight different recordings. These curves indicate sound intensity across the 0–8000 Hz frequency spectrum. The larger the area under this curve, the higher the BI value. For example, the total area under curve A is greater than the area under curve B and would therefore report a higher BI.

under the curve. Higher values of this index indicate a higher intensity level across a broader range of frequency bands, such as continuous high intensity broadband sounds produced by snapping shrimp or cicadas (Bradfer-Lawrence *et al.*, 2019).

Snap rate

Snap rate is an index unique to the marine environment. This index is used to calculate the level of snapping shrimp activity which can be an indicator of the wider ecosystem (Gordon *et al.*, 2018). Snapping shrimp are cryptic invertebrates ubiquitous in tropical and sub-tropical environments. These organisms perform rapid closure of their claws to produce cavitation bubbles which collapse to produce the characterise snapping sound, to both communicate and to stun prey or defend against aggressors (Versluis *et al.*, 2000). A standard approach used by snap rate calculators begins by squaring the amplitude value of all signals in a recording, eliminating negative values, and then taking the median amplitude value from this set. The approach then searches for instances where (i) the amplitude is exceeded by a set threshold (e.g., four standard deviations above the median), (ii) fall below this threshold within a short time frame (e.g., 0.125 ms), and, (iii) do not occur within a short window of time immediately after another snap to discount echoes (e.g., 1 ms) (Gordon *et al.*, 2018). This is usually converted into snaps per minute with higher values indicating a greater number of snaps.

Current consensus on the use of eco-acoustic indices

It is worth noting the validity of each index to determine useful ecological information from the marine environment is not yet fully established within the soundscape ecology community. Most of the indices presented here that have been used in multiple studies have been the subject of some of the contradictory findings discussed in Section 1.1. Some studies support the use of specific indices to reveal characteristics of biological communities, whilst others challenge these findings (Buxton *et al.*, 2018; Mooney *et al.*, 2020). Any index where this discrepancy has not yet been revealed has likely only been used in one or two studies and as such lacks sufficient validation as of yet. This trend is frequently seen in the field of terrestrial soundscape ecology where their use is more established (Buxton *et al.*, 2018). Additionally, any index that has not yet shown promise in the marine environment may simply have not been applied in an effective manner. Further experimentation with these may reveal a relationship with certain aspects of ecology. It is therefore to be expected that a significant period of trial and error will be required in the marine environment before computational indices can become a more exact science used to deliver accurate ecological and biological findings. Many practitioners also suggest that the use of eco-acoustic indices should never be treated as a replacement for less efficient traditional surveying techniques outright. Instead these indices could be used alongside traditional techniques in a complementary approach that builds a more comprehensive surveying effort (Mooney *et al.*, 2020).

1.3 Advances needed in marine soundscape ecology

Applications of soundscape ecology in the marine environment are still limited (Pieretti *et al.*, 2017; Lindseth and Lobel, 2018). Continued investigation to address the full potential of this field and better understand the relationship between soundscapes and the environment is needed (Buxton *et al.*, 2018; Lindseth and Lobel, 2018). Sampling from a variety of habitats around the globe over a range of spatial and temporal scales is required to better understand which properties of a soundscape may be of interest and to help identify commonly observable trends (Pieretti and Danovaro, 2020). Initially PAM will be needed alongside traditional surveying methods. This will enable novel PAM techniques to be compared to and validated against existing methods.

As a deeper understanding is developed, continued expansion of this field can progress. New users who can access recorders and the expertise on how to conduct a soundscape ecology investigation using these will be needed. Rapid analysis of recordings using computational approaches will be key to increasing the pace at which this growing volume of data can be interpreted. These advances will help PAM contribute to growing efforts to monitor the oceans to help better understand and protect marine habitats. PAM may be used in isolation or alongside existing monitoring techniques in complementary approaches.

This thesis presents two investigations aimed to advance the methods used by marine soundscape ecologists to collect and analyse PAM data. The first of these addresses the current inaccessibility of marine capable recorders; currently one of most limiting factors in PAM, holding back the widespread use of this practice in the marine environment. The second investigation demonstrates the potential of applying the modern analytical approach of machine learning to marine PAM data. This was performed using recordings from the world's largest active coral reef restoration programme and reveals an exciting new way to perform habitat assessments.

Chapter Two:

Monitoring marine ecosystems: the potential of low-cost action cameras for underwater eco-acoustic recordings.

Authors: Ben Williams[a], Lucille Chapuis[a], Timothy A. C. Gordon[a,b], Stephen D. Simpson[a]

Affiliations: [a]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PS, United Kingdom; [b]MARS Sustainable Solutions

Author contributions: B.W. was primarily responsible for data collection in the field, analysis and write up. L.C. contributed to analysis. L.C., T.A.C.G. and S.D.S. provided further comments on analysis and the written manuscript. A full manuscript submitted for publication is included in the supplementary material to this thesis (Supp. 2.1).

Abstract

1. Underwater passive acoustic monitoring (PAM) is an increasingly popular approach to monitor the health of aquatic environments through the analysis of soundscapes. Standard practices use hydrophones to record ambient sounds. They must either be cabled to surface recording devices or use autonomous instrumentation which comes at a premium cost. However, low-cost consumer-grade action cameras offer an accessible alternative also capable of autonomous underwater acoustic recordings.

2. The performance of two models of GoPro underwater action cameras when used as PAM recorders was evaluated. These were tested against a research-grade hydrophone in field conditions on shallow-water tropical coral reefs.

3. Simultaneous recordings of loudspeaker playbacks of known acoustic signals using all three instruments were taken first. Repeated deployments on different coral reef sites in which all three instruments were placed side by to side to record the same natural reef soundscape simultaneously were then undertaken. Eight of the most common eco-acoustic indices used in marine soundscape ecology from these GoPro recordings were calculated. These were used to assess the reliability and accuracy of results from the GoPros compared to the hydrophone.

4. Although not calibrated, GoPros appeared to provide recordings from which select eco-acoustic indices could be calculated reliably, including temporal variability, the acoustic complexity index and acoustic richness. Results from a GoPro can be compared against others of the same model but should not be used interchangeably with a hydrophone or those from another model. We outline the best settings that can be used to collect such soundscape data with GoPros.

5. Underwater action cameras are very popular with marine scientists and potential citizen scientists around the world. Their recordings represent a valuable tool for the global expansion of PAM techniques.

2.1 Introduction

Soundscape recordings are typically made by pressure-sensitive hydrophones, the underwater equivalent of a microphone (Lillis and Mooney, 2018). These devices consist of a piezoelectric transducer which convert pressure fluctuations into an electrical current (Lau *et al.*, 2002). Compared to microphones, the receivers on these devices have a reduced sensitivity matched to the lower acoustic impedance of water (Medwin, 2005). This allows louder sounds to be recorded without surpassing the receiver’s dynamic range which leads to distortion in recordings and reduces the intensity of quieter sounds. Hydrophones have been essential in advancing the field of marine soundscape ecology. However, these devices come with other drawbacks.

In every application of soundscape ecology, compromises are made between factors such as the number of sites at which recordings are taken, and the length of time these are taken for (Hill *et al.*, 2018; Elise *et al.*, 2019). Access to recording equipment often dictates this trade-off (Merchant *et al.*, 2015; Hill *et al.*, 2018). Terrestrial practitioners benefit from access to an extensive choice of equipment that is often able to overcome this limit to the point where other constraints become more limiting. High specification instruments are available to purchase for a few hundred pounds; e.g., Wildlife Acoustics Song Meter, £399 (Wildlife Acoustics, US) (Darras *et al.*, 2018; Beason *et al*., 2019). These devices house all the necessary components in a compact, self-contained and durable device, capable of long term deployments. New disruptive technologies that deliver many of these features in a small and much more cost effective device are also becoming available; e.g., the AudioMoth recorder, £47 (Open Acoustic Devices, UK) (Hill *et al.*, 2018).

However, for marine practitioners, access to equipment continues to be a limiting factor. This can result in investigations struggling to achieve the sampling depth they require or attract criticism when attempting to draw conclusions from only short windows of recording (Elise *et al.*, 2019; Mooney *et al.*, 2020). Often it is clear only one or two hydrophones could be accessed for even high impact investigations (Nedelec *et al.*, 2015; Bertucci *et al.*, 2016; Elise *et al.*, 2019), whereas dozens may be used in terrestrial investigations (Hill *et al.*, 2018; Sethi *et al.*, 2020).

This likely a result of the high expense associated with research-grade hydrophones. Low specification hydrophone sets ups typically come at costs similar to or above that of the highest specification terrestrial devices. These low cost options also require complex set ups which include cables to power sources and recorders above the surface, necessitating boat attendance or long cables attached along the sea floor to land nearby (Elise *et al.*, 2019). If an autonomous set up using these models is needed then floatation of dry components in a buoyant housing at the surface is required. This can be vulnerable to flooding, overheating and strong currents, as well as potentially introducing unwanted noise from surface splashing against the housing (Sousa-Lima *et al.*, 2013). Advances in the last decade have led to the development of self-contained research-grade hydrophones that remove these drawbacks and can be left unattended for extended periods of recording (Sousa-Lima *et al.*, 2013). However, these come at a premium cost that is typically an order of magnitude higher than the terrestrial equivalent; e.g., SoundTrap STD300, £2,230 (Ocean Instruments, New Zealand).

The high cost and/or technical expertise required to deploy hydrophones is therefore an obstruction to many marine practitioners. This prevents established marine soundscape ecologists from collecting recordings from multiple sites simultaneously on the scale that terrestrial practitioners are able to. It is also likely an obstruction to marine practitioners hoping to incorporate passive acoustic monitoring alongside existing work. The availability of low-cost equipment that requires minimal specialist knowledge for assembly and deployment would greatly increase the accessibility of PAM in the marine environment. This would allow existing soundscape ecologists to further expand the research in this field and enable new users to consider the use of PAM alongside existing work. There is therefore a demand to identify or develop low-cost equipment capable of collecting useful acoustic data in underwater environments.

GoPros: Low-cost audio-visual recorders commonly used in marine research

We identified GoPros (GoPro™, California, US) as a potential solution to some of the equipment challenges faced by marine soundscape ecologists. These devices are compact consumer-grade recorders primarily designed to collect high-definition video footage and images of action-adventure sports. They can be deployed in underwater housings and left to record for up to two hours before the battery dies. The widespread use of GoPros amongst marine practitioners makes it apparent that these devices are suitable for work in this setting. This includes the collection of images and video for surveying (Morgan *et al.*, 2017; Villon *et al.*, 2018), observing ecosystem functions (Rasher *et al*., 2017; Ford *et al.*, 2018; Lefcheck *et al.*, 2019), 3D photogrammetry (Raoult *et al.*, 2016; Young *et al.*, 2017) and in combination with ROV's and drones (Casella *et al.*, 2017; Corriero *et al.*, 2019). Additionally, audio can be collected by these devices. This can either be extracted from video footage to obtain an *MP3* file, or, from the GoPro 5 onwards, a 'raw' audio setting can be enabled which records an uncompressed, unmodified *wave* file, typically preferable to soundscape ecologists.

Whilst the use of microphones in place of hydrophones is unconventional, countless hours of video footage using GoPros have been collected by investigators for various reasons, all of which are accompanied by audio. These audio recordings contain the characteristic sounds expected of these habitats such as the background crackle of snapping shrimp, clearly distinguishable fish vocalisations and influences from the weather. This opens the possibility that ecologically relevant information could potentially be obtained from this audio for use in PAM investigations.

However, to our knowledge, no soundscape investigation in the marine environment has considered the use of terrestrial recorders, such as GoPros, to collect acoustic recordings. This makes the use of these devices an alternative approach to the current standard and introduces two key differences: (i) the use of microphones instead of hydrophones, (ii) the use of low cost consumer-grade devices. It is important to note also that no terrestrial soundscape ecology investigation has considered the use of GoPros for data collection. However, other purpose made consumer grade devices able to fulfil this demand are already available, precluding the use of GoPros (Hill *et al.*, 2018), but these do not yet offer underwater housing.

Limitations of consumer grade recorders

Although they offer improved accessibility, consumer grade devices are not without limitations. Notably, devices such as GoPros, do not come with microphones calibrated by the manufacturer. This is typically the case with hydrophones and calibration using alternative methods is not a simple or cost-effective process (Robinson *et al.*, 2014). Calibration is necessary to ascertain a true level of intensity relative to a standardised reference pressure. This allows comparisons of volume/intensity to be made across different devices (Robinson *et al.*, 2014).

As part of the calibration process, a detailed frequency response curve is produced which is essential when reporting the intensity of different sounds across the frequency spectrum. Often this frequency response curve is flat for high specification devices, which means the sensitivity is consistent across all relevant frequencies. This allows exact intensity to be calculated using the calibration data, and can applied with ease across the full spectrum, removing the need for complex adjustments at difference frequencies (Lurton, 2002; Robinson *et al.*, 2014). A flat response is not easily obtained and therefore is unlikely to be observed in consumer-grade devices. This in turn cannot be accounted for as they are uncalibrated, likely resulting in the intensity of sounds at different frequencies being misrepresented.

The dynamic range of microphones/hydrophones is also an important consideration (Robinson *et al.*, 2014). This is essentially the lowest and highest levels of sound intensity a device can record without failure of detection or distortion respectively. Details of this for the internal microphones are not provided by GoPro. System self-noise can also affect recordings, this is where the internal mechanisms of a device produce low level noise that is detected and recorded by the device (Robinson *et al.,* 2014). Again, no details on this are readily available for GoPros, meaning it could confound recordings. Finally, soundscape ecologists generally consider omnidirectional recordings which record audio from all directions equally (Robinson *et al.*, 2014). No deliberate attempt is made to provide this using GoPros and it is instead simulated using three internal microphones placed on the left, right and bottom of each device.

However, these limitations do not rule out the utility of lower specification devices such as GoPros. Although calibrated soundscape information is an important component to study some ecological characteristics (Lindseth and Lobel, 2018), the majority of indices used by soundscape ecologists do not rely on calibration (Sueur, 2018a). If deviations from a flat frequency response are consistent across recordings, an index can still be calculated accurately, providing a basis for relative comparisons. The lack of omnidirectionality does not always have to be a limitation. If a habitat being recorded is large enough then a (semi-)directional recording may still be adequate if the device is orientated towards the area of interest.

Aims of this investigation

This study was designed to test the utility of GoPros in marine soundscape ecology using modern PAM techniques. The first component of this used spectrogram and power spectral density (PSD) plots generated from recordings of two artificial signals produced from a loud-speaker and a reef soundscape recording. This was performed to explore any overt differences in GoPro and hydrophone recordings such as amplitude and frequency responses.

A more complex investigation was then designed to address the utility of eco-acoustic indices testing against the following hypotheses:

- H_1: Eco-acoustic indices calculated from GoPro camera audio can be used to produce reliable and accurate results in place of those calculated from research-grade hydrophone recordings.
- H_0: Eco-acoustic indices calculated from GoPro camera audio cannot be used to produce reliable and accurate results in place of those calculated from research-grade hydrophone recordings.

Reliability was defined as how consistent the relationship between the results calculated from the GoPros was to those from the hydrophone. Accuracy was defined as how closely the results calculated from GoPro recordings were to those from the hydrophone recordings, and is dependent first on a good reliability score.

We used a method comparison approach to perform this investigation. Such approaches are frequently used to compare alternative means of collecting measurements or data using new technologies or methodologies (Magari, 2002). In accordance with this, a SoundTrap STD300 hydrophone (Ocean Instruments, NZ), a commonly used model in published marine soundscape ecology investigations (Roland *et al.*, 2017; Bohnenstiehl *et al.*, 2018; Lillis *et al.*, 2018), was treated as the gold standard. Recordings on the hydrophone and several GoPro devices were taken simultaneously on different patches of a tropical reef.

2.2 Methods

Study site

All recordings were taken in November and December 2019, in the shallow-water coral reef environment, directly south-west of Lizard Island Research Station, Great Barrier Reef, Australia (14°40.8′S, 145°26.4′E; Fig. 2.1). Lizard Island is a mid-shelf island in the northern Great Barrier Reef, situated 27 km offshore and

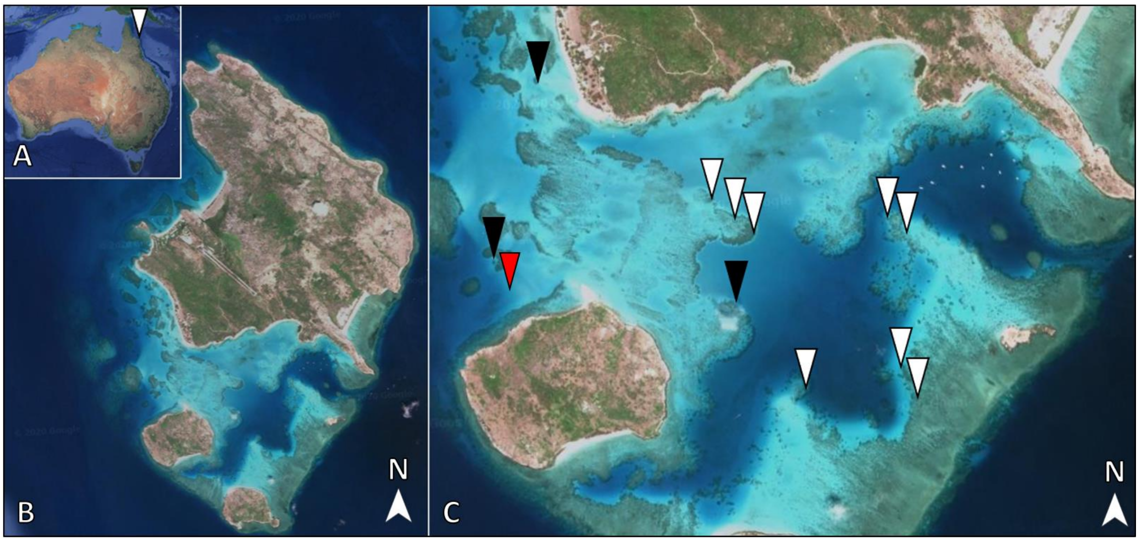


Figure 2.1. (A) Lizard Island (14°40.8'S, 145°26.4'E) relative to mainland Australia. (B) Aerial view of Lizard Island. (C) Approximate locations at which GoPros were placed alongside the hydrophone to collect recordings, white arrows indicate both GoPro 5 & 7 recordings were taken, black arrows indicate just GoPro 7 recordings, red arrow indicates location of the playback experiment (Maps: Google Earth, Maxar Technologies).

17 km from the outer Greater Barrier Reef. To sample the reef soundscape, the recording devices were deployed at 11 randomly selected reef flat sites by snorkelers at depths from 2 to 8 m (Fig. 2.1C). A playback experiment was also performed on a sand-flat site 50 m from the nearest reef at 4 m depth (Fig. 2.1C).

Collection of recordings

A SoundTrap (300 STD; Ocean Instruments, Auckland, NZ) was used as the industry standard hydrophone for this study. This self-contained device possesses an omnidirectional receiver with an inbuilt digital recorder (288 kHz maximal sampling rate, 16-bit resolution, 0.02–60 kHz ± 3dB frequency range, 34 dB re 1 μPa self-noise above 2 kHz, maximum gain before clipping 186 dB re 1 μPa, calibrated by manufacturer). For each deployment this was set to record continuously on the low gain setting at a sampling rate of 48 kHz.

Three GoPro HERO5 Black and two GoPro HERO7 Black (GoPro™, California, USA) were used, hereinafter referred to as GoPro 5 and GoPro 7 respectively. All cameras were enclosed in the manufacturer's 'Super Suit' underwater housing. Each GoPro contains three internal microphones with a sampling rate of 48 kHz. The cameras were set to record videos at the lowest resolution and frame rate per second possible for both cameras (720p, 60fps); note this does not affect audio quality. They were also set to record a 'raw' audio file in the .wav format using the 'Protune' settings accessible through the user interface.

Both the hydrophone and GoPro were suspended on vertical ropes at matching heights of 0.5 m above the seabed, less than 0.3 m apart, using small sub-surface buoys, dive weights, rope and cable ties (Fig. 2.2). The devices were left to record for 120 minutes, which was the maximum limit of the GoPro battery life. All recordings that resulted in battery or camera failure prior to 30 minutes were discarded. Those including excessive motorboat noise, in which five one minute segments could not be collected without boat disturbance, were discarded. Recordings were only taken when the wind was between 0 and 2 on the Beaufort scale, the sea state was calm or smooth on the Douglas scale, and there was no rain.

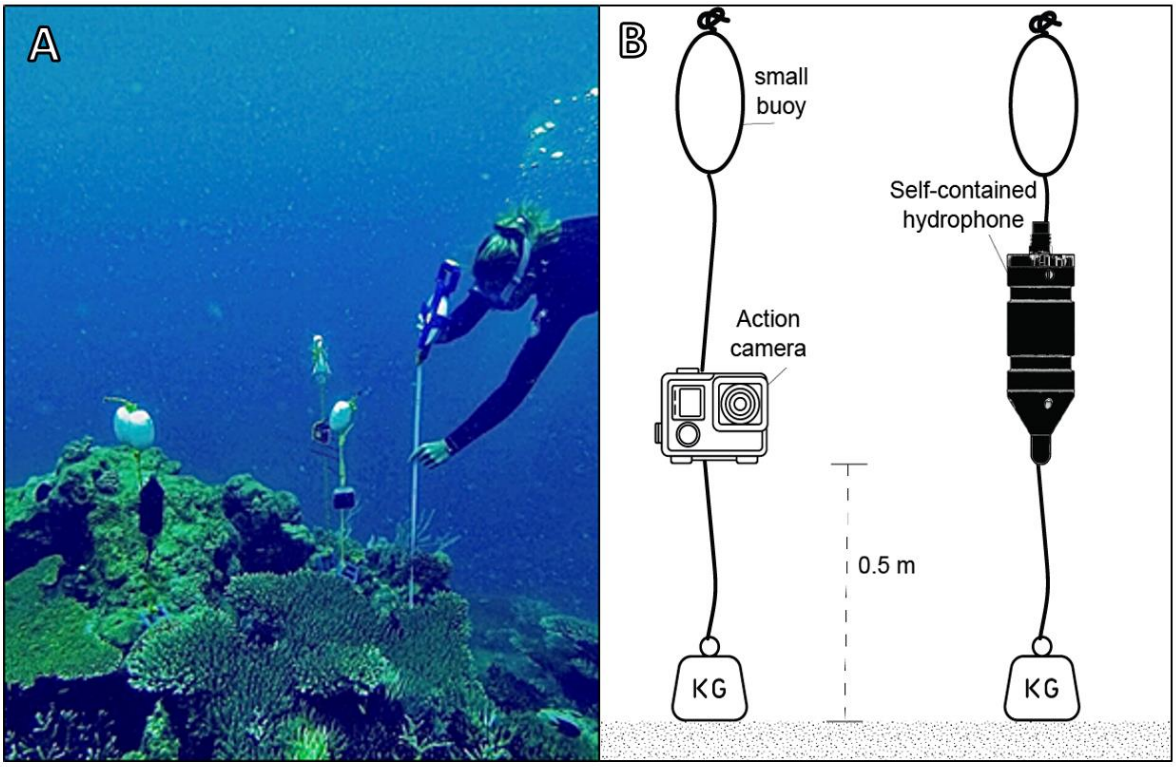


Figure 2.2. (A) GoPro 5 & 7 being deployed next to the hydrophone. (B) Schematic of how the devices were suspended. Devices were repositioned until all were within 0.3 m and suspended at equal heights.

Pre-processing recordings

Audio was extracted from the GoPro video footage using *WavePad* (v.9.6.3) and uploaded to *Audacity* (v2.3.1) alongside the raw GoPro audio recordings and hydrophone recordings. An approximate temporal alignment of each GoPro track against the hydrophone was then performed manually using a short section of speech present at the start of each recording. This section was then cropped from each track and used to perform a precise alignment using the *Signal Processing Toolbox* in *MATLAB* (v9.7.0) with the *AlignWave* plugin (Chen, 2020). This plugin uses waveform cross-correlation to output the difference between audio samples down to 1 ms. The difference was then removed from tracks as appropriate to produce an accurate alignment. Aligned tracks are referred to as one track from here on, with all subsequent changes made in parallel.

Additional recordings for Spectrogram and PSD inspection

Additional recordings were taken and aligned for spectrogram and power spectral density (PSD) comparisons. For this, a playback of two controlled acoustic signals were recorded. The acoustic signals were generated in *Raven Pro* (v1.6.1) and put into an audio track using *Audacity*, these were:

- Nine simultaneous pure tones of a one second duration. Beginning at 1 kHz up to 17 kHz at 2 kHz intervals.
- A sine sweep that increased linearly from 0 Hz to 20 kHz over a 10 second duration.

Each device was set to record whilst exposed to a playback of this track using an underwater loudspeaker (University Sound UW-30; max output 156 dB re 1 μPa at 1m, frequency response 0.1–10 kHz; Lubell Labs) powered by an amplifier (M033N, 18 W, frequency response 0.04–20 kHz; Kemo Electronic GmbH) and connected to an MP3 player (SanDisk Clip Jam), and a battery (12V 12Ah sealed lead acid). Here, it is important to note the loud-speaker may not produce sine sweeps or pure tones with consistent amplitudes across the full spectrum, especially below 100 Hz and above 10 kHz as this is the limit the manufacturer placed on its flat frequency response. However, the speaker could be relied on to output any imperfection consistently, allowing each device to record the same signal even if these were not true perfect representations of the input track.

These recordings were taken with the loudspeaker placed at 4 m depth on a sandflat >50 m from the nearest reef in the Lizard Island lagoon. A dive weight was placed as a marker 2 m from the speaker at this same depth. One by one, each device was placed on top of the marker and left for several cycles of the track. Each device was suspended and set to record in the same manor described previously. To prevent disturbance, recordings were taken whilst the sea-state was calm on the Douglas scale, the wind was 0 on the Beaufort scale and all snorkelers, divers and boats vacated the area during each cycle of recordings. One of the 30 s soundscape recordings, taken for the eco-acoustic index component, recorded by all three devices simultaneously was also selected for comparison. Recordings were processed in *Audacity* where an individual sine sweep and pure tone was isolated from each recording of the playback. These were processed in a purpose written *MATLAB* script used to produce spectrogram and PSD plots for each device.

Sampling recordings for index comparison

An acclimation period of 15 minutes was removed from the beginning of all aligned tracks to account for any disturbance to the experimental site during deployment. The next 15-minute period was used as the window of data to analyse as it encompassed the minimum length of time each GoPro was able to record for.

Sub-sampling of tracks to produce short windows of each recording is standard practice in PAM investigations. This reduces bottlenecks at the analysis stage that can be introduced by computationally intense approaches (Pieretti *et al.*, 2015; Elise *et al.*, 2019). However, no standardised regime is agreed (Elise *et al.*, 2019). Sub-samples typically range between five seconds to five minutes (Radford *et al*., 2014; Nedelec *et al.*, 2015) and may be taken once every few minutes of recording to once each hour (Staaterman *et al.*, 2013; Lombardi *et al*., 2016). A subsampling regime of 5 x 30 second non-overlapping segments chosen at random from each 15-minute window was therefore selected.

A further consideration is the presence of engine noise from motorboats that can dramatically affect the output of bioacoustics indices (Elise *et al.*, 2019). Recordings were therefore screened for boat noise through visually and acoustically inspecting spectrograms. Four full recording days were removed due to chronic boat disturbance. Boat noise was only observed for 3/95 sub-samples across the other recording days. These were removed and an alternative sample was randomly chosen from the corresponding 15 minute window from which each sub-sample was taken. In total, 40 sub-samples were generated for the GoPro5s and 55 for the GoPro7s from the eight and eleven successful recording blocks, with matched hydrophone sub-samples for each.

Eco-acoustic indices

Eight eco-acoustic indices (Table 2.1) were calculated in *R* using the packages *Seewave* (v2.1.6) and *Soundecology* (v.1.3.3). Snap rate was the only exception which was calculated using a modified *MATLAB* script from Gordon *et al.* (2018). Two frequency bands were used for every index. This included a low band (0.1–1.5 kHz) intended to be inclusive of most fish sounds (Lammers and Munger, 2016) with interference from low frequency ship noise and geophonic disturbance removed (Curtis *et al*., 1999). Above this a high band was used (1.5–20 kHz), dominated by broadband snapping shrimp (Versluis *et al.*, 2000). The upper limit of this high band was determined by the Nyquist frequency limit of the GoPros; that is, half the maximal sampling rate of a device for which recordings must be limited to adequately sample all sounds (Sueur, 2018b). The only exception to these bands was the index snap rate, for which the same high band was used and a broadband (0.1–20 kHz) range as snap rate is not typically measured in the low band alone (Bohnenstiehl *et al.*, 2018). These bands were created using *Seewave* and *Soundecology's* inbuilt frequency filters for indices where this is available. If not, sixth order Butterworth filters with a 40 dB roll off were implemented to tracks in *MATLAB*. Most indices were calculated using default windows, envelopes, and other settings from *Seewave* and *Soundecology* (Table 2.1). The exceptions were ADI and AEI's frequency bins. The default for

Table 2.1. The eco-acoustic indices calculated for each recording with a summary description of the mechanistic principle behind each and the software used. The settings and filter used for each index in their respective software is also included.

Index	**Mechanism**	**Software**	**Settings**	**Filter**	**Citation**
Acoustic Complexity Index (ACI)	Measures variability in intensity of frequencies across time.	*Seewave* in *R*	Window size = 512; type = Hamming; overlap = 0.	*Seewave* 'ffilter': Window = 1024; type = Hanning; overlap = 75.	(Pieretti *et al.*, 2011)
Acoustic Diversity Index (ADI)	Measures diversity across frequency bands.	*Soundecology* in *R*	Maximum frequency = (i) 1.5 kHz, (ii) 20 kHz; frequency bins = (i) 0.15 kHz, (ii) 2 kHz; threshold = -50 dB.	Filtered in function.	(Villanueva-Rivera *et al.*, 2011)
Acoustic Entropy (H)	Measures randomness across temporal and spectral domains.	*Seewave* in *R*	Window size = 512; envelope = Hilbert.	*Seewave* 'ffilter': Window = 1024; type = Hanning; overlap = 75.	(Sueur *et al.*, 2008)
Acoustic Evenness Index (AEI)	Measures evenness across frequency bands.	*Soundecology* in *R*	Maximum frequency = (i) 1500 Hz, (ii) 20 kHz; frequency bins = (i) 150 Hz, (ii) 2 kHz; threshold = -50 dB.	Filtered in function.	(Villanueva-Rivera *et al.*, 2011)
Acoustic Richness (AR)	Ranks recordings based on amplitude multiplied by randomness across the temporal domain.	*Seewave* in *R*	Envelope = Hilbert.	Seewave 'ffilter': Window = 1024; type = Hanning; overlap = 75.	(Depraetere *et al.*, 2012)
Bioacoustic Index (BI)	Measures cumulative intensity across frequency bands.	*Soundecology* in *R*	Minimum frequency = (i) 0.1 kHz, (ii) 1.5 kHz; Maximum frequency = (i) 1.5 kHz, (ii) 20 kHz; window size = 512.	Filtered in function.	(Boelman *et al.*, 2007)
Temporal Variability (TV)	Measures randomness across the temporal domain.	*Seewave* in *R*	No settings required.	*Seewave* 'ffilter': Window = 1024; type = Hanning; overlap = 75.	(Sueur *et al.*, 2008)
Snap Rate	Measures rate of snapping shrimp snaps.	*MATLAB*	Custom script.	6th order Butterworth filter; 40 dB roll off.	(Gordon *et al.*, 2018)

each of these uses ten frequency bins, therefore, ten 150 Hz and ten 1850 Hz bins were used for the low and high bands respectively.

Spectrogram and PSD plot comparisons

Qualitative inspection of spectrogram and PSD plots from the controlled playback experiment was used to compare recording properties across both models (5 and 7) and audio types (raw and video) against the hydrophone. Variance between each device of the same model and audio types was also investigated through manual listening and visual inspection of spectrograms.

Statistical analysis

In accordance with method comparison approaches, results generated from hydrophone recordings were treated as an ‘industry standard’ against which assessments of the reliability and accuracy of results generated from GoPro recordings could be made (Magari, 2002; Carstensen, 2011). Individual comparisons between the hydrophone and GoPro recordings were made for each eco-acoustic index using a single paired measure method comparison approach (Abu-Arafeh *et al*, 2016) for the GoPro 5’s (n = 40) and GoPro 7’s (n = 55). Non-parametric tests were selected, since little consistency between the distributions of each dataset was observed. All analysis was performed in *R* using the *MethComp* (v1.30.0) package.

Two tests were selected that are complementary to one another (Magari, 2002). The first was a Spearman’s rank-order correlation test, used to quantify the reliability of GoPros against the hydrophone by testing the strength of the linear relationship between these. The second test was a Passing-Bablok regression used to determine whether each GoPro-index combination under inspection was equivalent and can therefore be used interchangeably with results from a hydrophone. This is done by quantifying the proportional and constant bias (Bilic-Zulle, 2011). This test is specifically designed for method comparison studies and is non-sensitive to distribution errors or outliers. The test assumes a continuous distribution across datasets and a strong correlation between datasets which was determined in the first test using the Spearman's approach. Not all indices compared in this investigation met this second assumption. However, only those with a strong correlation have the potential to provide interchangeable results and as such only those should be considered in detail for this component. The Passing-Bablok regression analysis produces values for the intercept and slope which represent the constant and proportional bias respectively. Presence of a constant bias indicates a consistent deviation between results from the two devices in one direction. Presence of proportional bias indicates datasets do not agree equally across the range of measurements. For example, differences in results at the lower end of measurements may be small whilst differences present at the higher end are large (Ludbrook, 1997). The test also produces confidence interval bounds. If the confidence interval bounds calculated for the slope contain 1, and, the confidence interval bounds calculated for the intercept contain 0, then the two methods can be said to be in agreement and used interchangeably (Bilic-Zulle, 2011).

In combination, these two tests inform on the reliability and accuracy of GoPro recordings for calculating the eco-acoustic indices in question and whether they are equivalent to the hydrophones results. Without the Spearman test, the consistency of GoPro results against the hydrophone cannot be determined, and without the Passing-Bablok test, it would not be possible to determine whether a highly reliable GoPro-index could be used interchangeably or whether a bias is present.

2.3 Results

Spectrogram and power spectral density inspection

Spectrogram plots revealed a high level of consistency was observed between GoPro's of the same model (Supp 2.2). However, they did reveal notable differences between devices from different models (Fig. 2.3). Firstly, this included a difference in the gain applied, as indicated by the power/frequency colour scale. The GoPro video audio appeared to have the largest gain applied, followed by the hydrophone, with the lowest gain applied to the raw GoPro audio. A skewed frequency response is also revealed by the sine sweep in the video audio for both GoPro devices, and the raw audio for the GoPro 5 to a lesser degree. Here, lower frequencies in the sign sweep appear to be at a greater intensity whereas the hydrophone remains consistent. This skewed frequency response is also shown in the reef soundscape spectrograms, in which the hydrophone displays consistent broadband noise from a low frequency up to 20 kHz (also shown in the PSD plot (Fig. 2.4)). However, the video audio for both GoPros shows a greatly reduced intensity at frequencies above 15 kHz, and the GoPro 7 raw audio exhibits a slightly reduced intensity for signals above 10 kHz. The presence of more discrete discrepancies in the amplitude of some narrower frequency bands are also present in all the GoPro recordings whereas the hydrophone was uniform. The sine sweep also revealed an interesting artefact from the video audio files. As the sweep passed through the 8–10 kHz band, a second sweep appeared following the opposite gradient. This second sweep begins at approximately 15 kHz until it reaches 11 kHz where it stops.

The PSD plots (Fig. 2.4) of the pure tones revealed 18 clear peaks for the hydrophone, most likely a result of the nine pure tones and associated harmonics. The GoPro peaks are not as clearly defined, although it was evident that the raw recording from the GoPro 7 was the closest fit to the hydrophone with at least

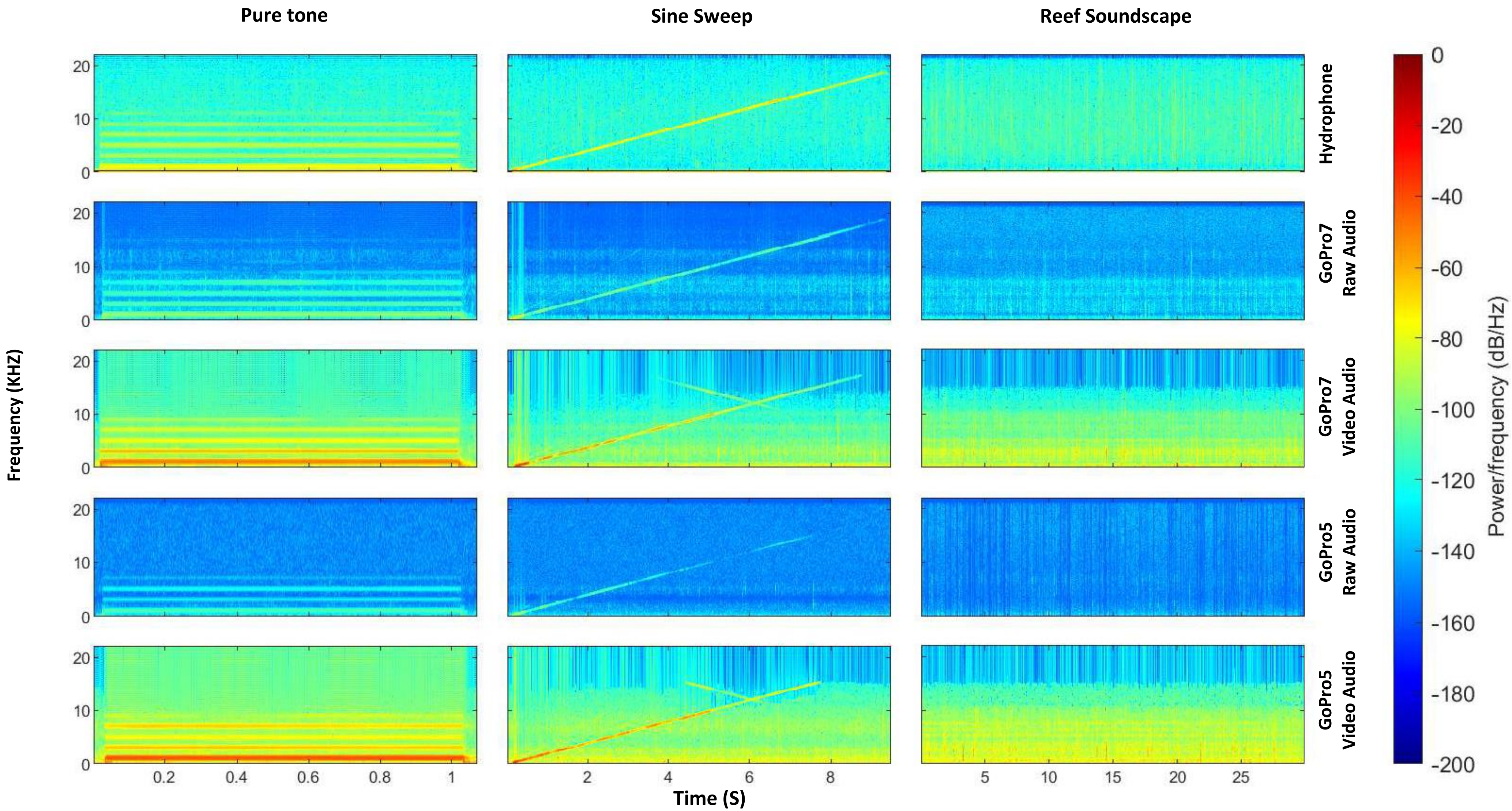


Figure 2.3. Spectrogram of a pure tone (left), a sine sweep (middle) and a coral reef soundscape (right) recorded by the hydrophone, GoPro 7 raw and video audio, and, GoPro 5 raw and video audio. A Hamming window was used with 75% overlap, pure tone window length = 128, sine sweep window length = 256, reef soundscape window length = 512.

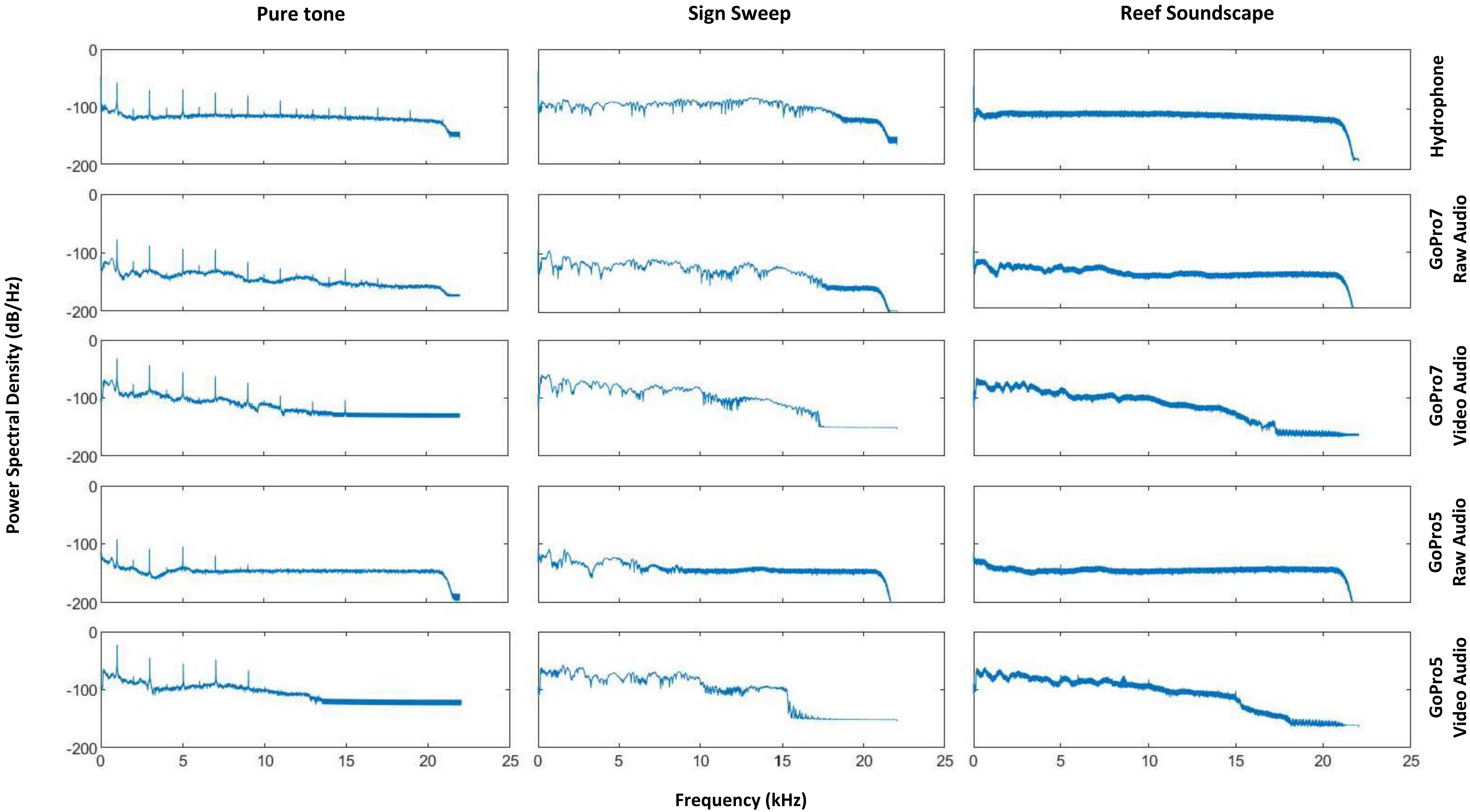


Figure 2.4. Welch's power spectral density (PSD) plots of a pure tone (left), a sine sweep (middle) and a coral reef soundscape (right) recorded by the hydrophone, GoPro 7 raw and video audio, and, GoPro 5 raw and video audio.

peaks (though many at a reduced amplitude), followed by the GoPro 7 video audio with 11. The GoPro 5 raw and video audio exhibited less peaks (around seven visible). The PSD of the sine sweep revealed a drop off in frequencies above 15 kHz in the video audio from both GoPros. Signals were attenuated above 15 kHz for the GoPro 5 and 16 kHz for the GoPro 7. PSD plots of the reef soundscape recorded by each device also showed uneven frequency responses for both GoPro devices. Conversely, the hydrophone showed a consistent intensity across the full spectrum up to 20 kHz. Relative to the hydrophone, the raw files taken by both GoPros showed a slightly increased intensity at the low end of the spectrum up to about 7 kHz, before exhibiting a flatter response up to 20 kHz. The videos for both GoPros showed an increased intensity at the lower end of the spectrum, relative to the hydrophone, which steadily decreased until 15 kHz where the intensity was reduced. At the lower end of the spectrum for both GoPro 7 audio types and the GoPro 5 video, many small inconsistencies in intensity between frequencies were also present, indicated by the 'zig-zagging' pattern along the spectrum.

Reef soundscape: reliability of acoustic indices generated by GoPros

Results from the correlation test between the GoPro and hydrophone recordings were used to estimate the reliability of the indices calculated from GoPro recordings (Fig. 2.5). Low band acoustic richness (AR) calculated from GoPro 7 raw audio recordings had the strongest correlation (*rho* = 0.98, $p<0.001$), with four others exhibiting highly significant correlations ($p<0.001$) above *rho* = 0.9. At the lower end, other GoPro-index correlations showed little to no correlation.

The GoPro 7 indices consistently reported stronger correlations with the hydrophone. The strongest correlation for 15/16 eco-acoustic indices was reported from the GoPro 7s. The exception was low band acoustic entropy (H) calculated from GoPro 5 raw audio recordings (*rho* = 0.75, $p<0.001$). Raw audio files had a stronger correlation with the hydrophone than video audio overall. Of the eight eco-acoustic indices across both

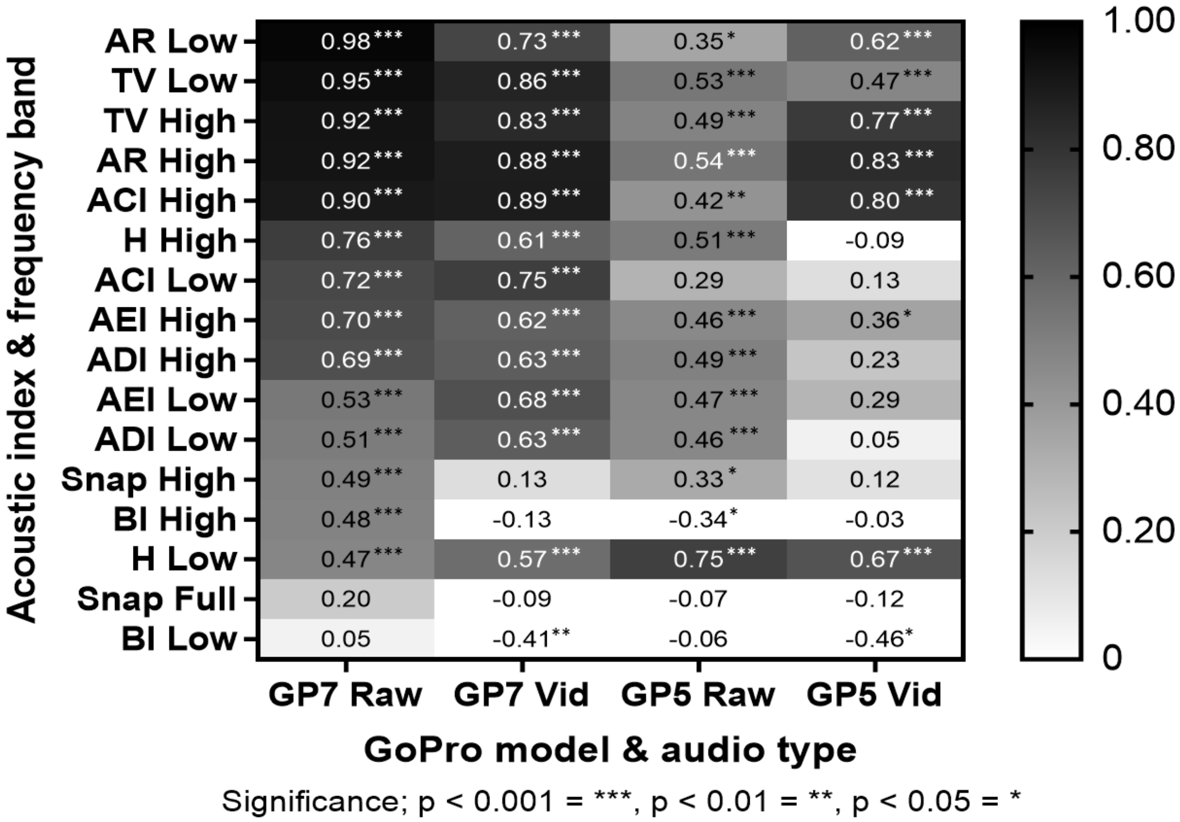


Figure 2.5. Spearman's rank-order correlation test scores between indices calculated from GoPro and Hydrophone recordings. Shading indicates strength of correlation, with no correlation (*rho* = 0) indicated by white and a perfect correlation (*rho* = 1 or -1) indicated by black. Rows are presented in ascending rank order of the GoPro 7 raw audio results.

audio types (raw and video), raw audio reported a stronger correlation in 12 and 11 cases for the GoPro 7 and 5 respectively. Results for any one index were often inconsistent across models and raw/video audio. However, some reported stronger correlations in general. Across both frequency bands AR, TV, ACI and H reported strong correlations for three out of four model/audio types (*rho*>0.5, *p*<0.001). At the lower end BI and snap rate showed little correlation in most instances. ADI and AEI exhibited a middling reliability (*rho* = 0.46–0.7, *p*<0.001) except when GoPro 5 video audio was used which showed little to no correlation (*rho*<0.4, *p*>0.05 or above).

Reef soundscape: accuracy of acoustic indices generated by GoPros

For results with a strong correlation, findings from the Passing-Bablok test (Table 2.2) could also be used to explore the relationship of indices calculated from the GoPros compared with the hydrophone. No GoPro-index combination fitted the criteria set by the Passing-Bablok regression needed to be considered indistinguishable from the hydrophone when confidence intervals were set to 90% and above. When confidence intervals were set to 80%, one measure, low band acoustic complexity index (ACI) calculated from GoPro7 raw audio recordings, reported confidence intervals for the slope (0.96 to 1.01) and intercept (-1.29 to 8.05) that encompassed the respective values of 1 and 0 needed to satisfy the criteria to be considered as an interchangeable measure.

High band ACI measured with GoPro 7 raw audio demonstrates a combination where index values have a strong correlation (*rho* = 0.9, $p<0.001$) but a proportional bias remains present (Fig. 2.6B). The values for this index are therefore reliable and can be used to calculate consistent results, but these results cannot be used interchangeably with or compared to the hydrophone without correcting for the proportional bias. Conversely, low band ADI calculated from GoPro 5 video audio (Fig. 2.6C) which suggests the constant bias could be removed to provide what appears to be an accurate measure. However, no correlation is observed (*rho* = 0.05, $p = 0.75$), indicating the index has a low reliability. This example demonstrates that the apparent absence of bias can be misleading and why indices with no correlation should not be considered accurate or reliable (Bilic-Zulle, 2011).

Moreover, high reliability cannot be confirmed without considering the accuracy, as demonstrated in Fig. 2.6D: a strong correlation is reported for high band ADI taken from GoPro 7 video audio (*rho* = 0.63, $p<0.001$). However, a very high proportional bias was also revealed (slope = 684.04). In this instance, the results for the hydrophone were highly similar for every recording whereas the GoPro reported a much greater variability. Therefore, any value for the GoPro would have fallen clos

Table 2.2. Slope and intercept values used to quantify proportional and constant bias calculated from the Passing-Bablok regression between each GoPro audio format and the hydrophone. A slope highly divergent from 1 indicates a large proportional bias, whereas an intercept highly divergent from 0 indicated a large constant bias. Rows are in rank order of the strength of correlation for the GoPro 7 raw recordings (Fig. 2.5). Low band ACI from GoPro 7 Raw recordings (in bold) was the only index to pass the test as equivalent to the hydrophone. Results indicated by inconclusive could not be computed by the test.

Index & Frequency Band	GoPro 7 Raw		GoPro 7 Video		GoPro 5 Raw		GoPro 5 Video	
	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept
AR Low	0.90	0.01	0.88	0.01	0.54	0.04	0.05	0.78
TV Low	0.93	0.00	0.91	0.00	0.63	0.01	0.01	0.27
TV High	1.18	0.00	1.27	0.00	0.93	0.03	0.77	0.01
AR High	0.93	0.00	0.97	0.00	0.97	-0.04	0.84	0.00
ACI High	0.50	82.67	0.83	31.52	0.72	49.86	0.47	97.77
H High	1.12	-0.14	3.69	-2.61	1.08	-0.09	4.42	-3.26
ACI Low	**0.99**	**3.49**	0.81	37.30	0.68	61.86	0.38	114.30
AEI High	15.70	-0.06	12.97	0.45	0.33	0.02	8.84	0.36
ADI High	162.41	-371.64	684.04	-1573.26	0.47	1.21	924.00	-2125.55
AEI Low	3.02	-0.10	2.74	-0.02	0.02	0.00	0.08	0.07
ADI Low	3.30	-5.27	7.66	-15.33	0.00	2.30	0.95	0.08
Snap High	2.83	-68.50	4.29	23.45	3.33	-150.43	6.12	-203.93
BI High	0.75	3.74	9.91	-58.29	1.17	-11.54	1.24	114.35
H Low	0.70	0.11	1.87	-0.46	0.25	0.36	0.51	0.24
Snap Full	3.15	-207.75	8.23	-869.19	3.64	-232.32	9.30	-906.50
BI Low	2.32	-0.75	Inconclusive	Inconclusive	0.13	1.42	0.50	3.53

to the regression line as the slope is vertical, resulting in a high *rho* value falsely indicating high reliability.

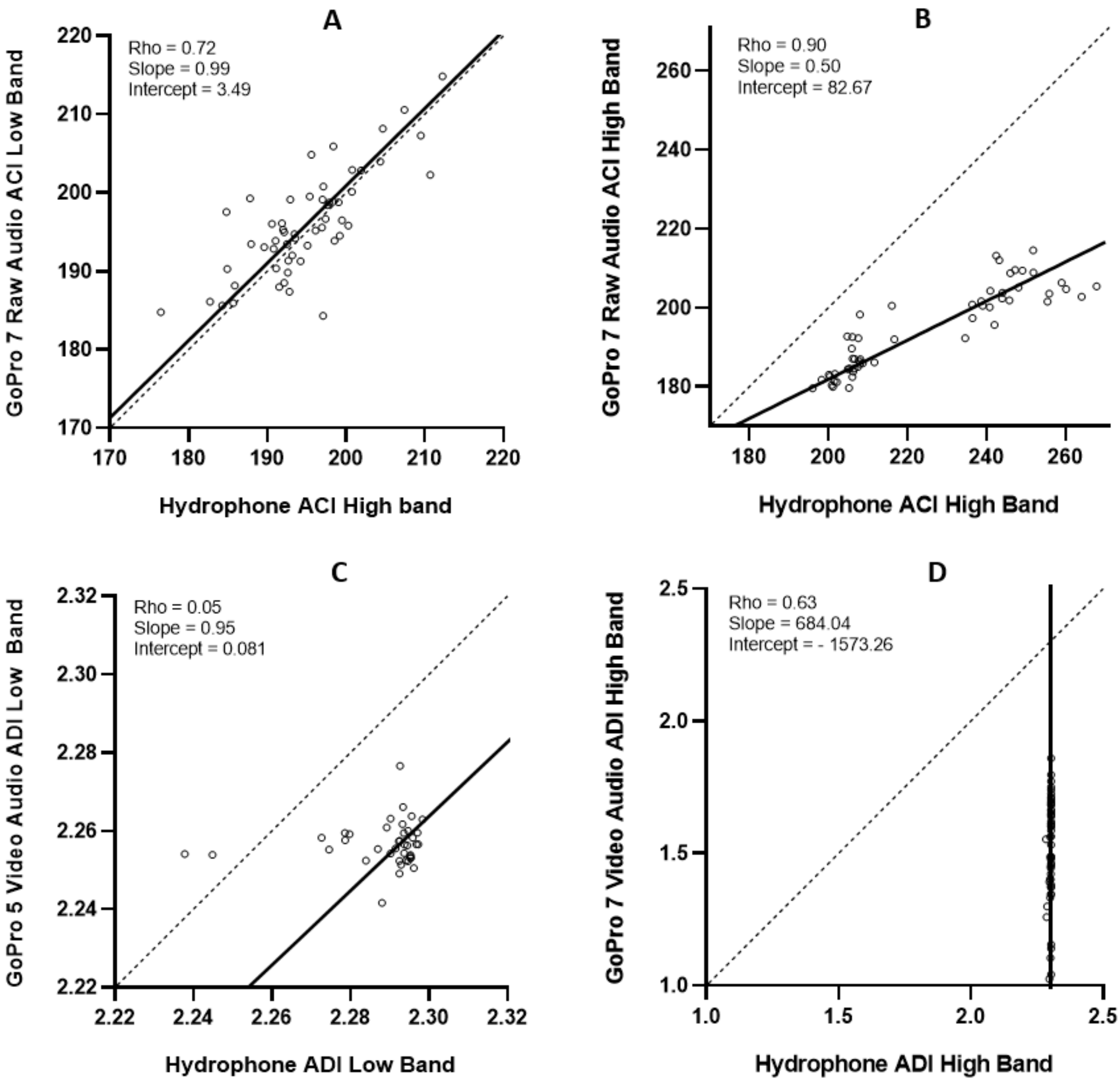


Figure 2.6. Passing-Bablok regression plots of selected GoPro-index combinations against the hydrophone results which demonstrate the range of possible relationships between the datasets. Solid line indicates the regression line, dashed line indicates the identity line (x = y). Correlation coefficients (*rho*) from the Spearman correlation test (Fig. 2.5) as well as slope and intercept values from the Passing-Bablok regression (Table 2.2) are included.

2.4 Discussion

Summary

The aim of this investigation was to determine whether eco-acoustic indices calculated from GoPro recordings taken in the marine environment are both accurate and reliable in relation to a research grade hydrophone. We trialled two models of GoPro (HERO5 Black and HERO7 Black). A correlation test supported the hypothesis that these devices can be used as reliable tools to collect recordings for many eco-acoustic indices in an underwater setting. However, neither GoPro model passed the Passing-Bablok regression analysis with 90% confidence. Therefore, neither offers results which can be used interchangeably with the hydrophone.

The most reliable GoPro from the two models tested appears to be the GoPro 7 and the most suitable audio format for both models tested appears to be the raw audio format. This reliability was assessed by the correlation test. Of the 64 GoPro-index combinations investigated, 17 reported a correlation equal to or above *rho* = 0.7 and 15 others were above *rho* = 0.5. This demonstrates that reliable comparisons can be made between these indices when recordings are collected by the same model of GoPro.

The accuracy was assessed by the regression analysis. The results from this show that values for acoustic indices determined from GoPro recordings predominantly are not suitable to be used interchangeably with the Hydrophone. Only one GoPro-index combination passed this test: low band ACI taken from GoPro7 raw audio recordings. However, for many GoPro-index combinations with a strong correlation to the hydrophone the differential bias (indicated by slope) was approaching 0. This shows in most cases that, although different, the relationship was not highly divergent.

The spectrogram and PSD plots offer an explanation for some of the statistical findings from the correlation and regression analysis. The hydrophone appears to capture more peaks for the pure tones and shows a flatter frequency response for the sinusoidal sweep than GoPros. These differences also appear more prominent for GoPro video audio which may explain each model's greater reliability when raw audio was used. Differences in overall amplitude are also clear between each audio format with the largest difference between raw files and the hydrophone. However, as each index assesses the spectrogram relative to itself the overall amplitude of a recording should not change the result. This is supported by these findings indicating that lack of calibration is not a major drawback for devices being used to calculate the indices trialled.

Qualitative assessments of the variance between each device of the same model through manual listening and visual inspection of spectrograms revealed that all major recording properties (e.g. frequency response, frequency filtering) were conserved between individuals from the same class of model (5 or 7) and audio type (raw or video) (Supp 2.2). Model and audio type were therefore the key drivers behind differences. As differences within groups were negligible, this indicates that comparisons of audio recordings within the same group can be performed, but not comparisons across different groups.

For practitioners who may wish to utilise GoPro recordings, this means that some eco-acoustic indices calculated from recordings taken by the same model of GoPro can be used in comparisons against one another. This could be between treatments, timepoints, sites or other such variables. For example, an investigation may be looking to determine whether there is a difference in an eco-acoustic index between two different reef sites. For indices which reported a strong correlation, the GoPro and Hydrophone should reveal a comparable conclusion between the two sites. This may not be reported with matching index values, but an equivalent level of statistical difference, or lack thereof, will be observed between the two sites.

Possible Ecological Applications

A key question for practitioners interested in using GoPros to collect soundscape recordings is what characteristics and attributes can be monitored or investigated using GoPro recordings. To address this question, it is important to note there is still much to be learnt about using soundscapes to study marine habitats (Bohnenstiehl *et al.*, 2018; Haver *et al.*, 2018; Chary *et al.*, 2020). However, initial findings and established trends using the eco-acoustic indices trialled in this investigation are emerging (Lindseth and Lobel, 2018). Using the knowledge currently available on the relationship between these indices and the marine environment there are a number of ecological attributes that could potentially be explored with GoPros. Indices with at least one instance of a strong correlation (*rho*>0.7) across models and audio types are discussed with this in mind.

One attribute often considered a direct provider of ecosystem services is the abundance and diversity of fish species present in a habitat (Holmlund and Hammer, 1999; Hicks *et al*., 2013). The highest performing measures reported here by GoPros have all been found to directly indicate trends in fish diversity and/or abundance in the past, these were: ACI (Bertucci *et al.*, 2016; Harris *et al*., 2016), TV (Elise *et al.*, 2019), high band H and AR (Harris *et al*., 2016). The use of ACI and TV has also been shown to correlate with the abundance of fish within certain trophic levels (Elise *et al.*, 2019). H can also be used to detect harmonic chorusing of certain species (Bohnenstiehl *et al.*, 2018), which in turn can be an indicator of higher richness and abundance of other cryptic species (Staaterman *et al.*, 2014).

Habitat type and quality is also an important consideration for many marine ecologists (Magris *et al.*, 2016; Terrado *et al.*, 2016; Chary *et al.*, 2020). Some of the indices that have the potential to be reliably determined using GoPro recordings such as ACI and TV have been found to change depending on the type of habitat present (Ceraulo *et al.*, 2018; Elise *et al.*, 2019). ACI has been shown to correlate with ecosystem state (Elise *et al.*, 2019) and both ACI and AR have been used previously to differentiate between healthy and degraded habitats (Butler *et al*., 2016; Gordon *et al.*, 2018). Differences in ACI between fished and unfished zones has also been shown (Davies *et al.*, 2020). Differences in the soundscape between habitats has not been well explored and further study may improve insights into this (Parsons *et al.*, 2016).

Many ecological functions in marine habitats are dictated by temporal trends (Radford *et al.*, 2008; Park *et al.*, 2019; Shlesinger and Loya, 2019; Smale *et al.*, 2019) and investigations have found this is reflected in the soundscape of marine habitats. The indices, H and ACI have previously been observed to show diel periodicity (Staaterman *et al.*, 2014; Kaplan *et al.*, 2015; Bertucci *et al.*, 2016; Rice *et al.,* 2017), as well as lunar and seasonal trends for ACI (Staaterman *et al.*, 2014; Pieretti *et al.*, 2017) and all showed strong potential to be calculated reliably using GoPros. Trends for other indices over these longer temporal periods have not yet been studied, and warrant further investigation.

Indices and ecological uses not yet supported

The indices that GoPro recordings reliably produce could be used to investigate a variety of ecological functions. However, it is important to highlight indices that either could not be well interpreted from the results of this study, performed poorly or are not yet possible with GoPros.

The next consideration after indices with the strongest correlations with the hydrophone are those that reported a weak positive relationship (*rho* = 0.5–0.7). The acoustic diversity index (ADI) and acoustic evenness index (AEI) occupy this group. Both these indices assess the variation across different frequency bins within the spectrogram (Villanueva-Rivera *et al.*, 2011). Studying the spectrogram and PSD plots (Figs. 2.3 & 2.4) provides potential explanations as to why these indices did not perform so well. Here, the plots revealed inconsistencies in intensity recorded across frequency bins of the GoPros. This would affect outputs from ADI, and AEI as they are designed to directly measure differences across frequency bins. Shortcomings of these indices are further revealed by the regression analysis when considering the proportional bias reported by each (Table 2.2). Across both models of GoPro, this was typically either close to zero or very high. This indicates that values output by either the GoPros or the hydrophone were highly variable, whereas the distribution of the results from the other were much narrower. For example. The slope for high band ADI across all GoPros was high (Table 2.2), indicating that the hydrophone output highly similar results across all sites undermining the regression test (Fig. 2.6D). This makes it difficult to interpret the correlation with confidence.

Both AEI and ADI have been applied in the marine environment on a limited number of occasions and their utility needs further investigation (McWilliam and Hawkins, 2013; Rice *et al.,* 2017; Roca and Van Opzeeland, 2019). However, initial findings have found ADI to be a promising measure to detect or quantify boat noise and harmonic fish chorusing (Rice *et al.,* 2017; Siddagangaiah *et al.*, 2019). There is just one instance of the use of AEIs in the marine environment, where it was found to correlate with species richness of marine mammals on shelf habitats in the Southern Ocean (Roca and Van Opzeeland, 2019).

Further down the strength of correlations is the bioacoustics index (BI) and snap rate. Both performed best in the high band from GoPro 7 raw recordings (*rho* =

0.48, p<0.001; *rho* =0.49, p<0.001 respectively). However, the low band for each exhibited no significant correlation (*rho* = 0.05, p = 0.72; *rho* = 0.2, p = 0.14 respectively). BI has only been used once in the marine environment, but showed promise by outperforming other eco-acoustic indices, where it reported strong correlations with planktivore biomass and laminar, foliose and helmet-shaped coral cover in a tropical reef system (Elise *et al.*, 2019). Snap rate has been broadly investigated in the marine environment and likely constitutes a useful measure if properly applied. Diel, lunar and seasonal trends have been observed for this index (Lillis *et al.*, 2016; Ricci *et al.*, 2016; Lillis and Mooney, 2018; Lyon *et al.*, 2019). It has also been implicated with tropical reefs in different states of health (Butler *et al*., 2016; Gordon *et al.*, 2018) and one instance has also found snap rate correlates with habitat complexity (Lyon *et al.*, 2019). However, the findings from all these studies demonstrated similar results with alternative eco-acoustic indices which the GoPros reported more reliably. This means GoPros do not have to be ruled out from performing these kind of habitat assessments as other indices may be suitable alternatives.

A key limitation of GoPros is their inability to report calibrated reference pressure values such as sound pressure level (SPL). Certain frequency bands of SPL have been found to correlate with an extensive list of ecosystem functions in marine habitats including: habitat complexity, coral cover, fish abundance, fish diversity benthic diversity, benthic invertebrate density, predator density, algal density, porites cover, coral growth forms, the presence of dead coral, encrusting coral cover and other habitat traits (Kennedy *et al.*, 2010; Kaplan *et al.*, 2015; Parmentier *et al.*, 2015; Bertucci *et al.*, 2016; Freeman and Freeman, 2016; Elise *et al.*, 2019). These are all either through a direct correlation, or through the strength of the metrics diel trend. This list is more extensive than the combined list of ecosystem functions that the other eco-acoustic indices trialled on GoPro recordings in this study have been shown to be indicators of previously. However, SPL is the most widely used acoustic metric in the literature on marine soundscapes currently (Lindseth and Lobel, 2018). As such, a reporting bias due to the metrics frequent use may be responsible for the more comprehensive list of ecosystem functions found to be indicated by SPL. In support of this, one study looked for correlations between the acoustic indices SPL, ACI, TV, SE, BI and H across five frequency bands and six ecosystem functions (Elise *et al.*, 2019). Here, SPL had the strongest correlation with just one out of the six ecosystem functions, suggesting that other eco-acoustic indices have perhaps been overlooked previously and that uncalibrated devices such as GoPros could potentially be used to perform many of these assessments.

A limitation of GoPros for acoustics monitoring is their greatly reduced battery life compared to many commercially available hydrophones. Throughout this study, GoPros rarely exceeded two hours of continuous recording before the battery died. In comparison, the hydrophone used in this study is capable of continuous recording for two weeks and can achieve significantly longer periods with an external battery extension or duty cycling. This limits the suitability of GoPros for long term passive acoustic monitoring without battery changes. Memory capacity is an additional limitation. GoPros accept commercially available SD cards which, though costly, are currently available with a capacity up to 1 TB. However, GoPros are inherently less efficient with storage as currently they can only be set to record audio whilst video is also being recorded. Ten minutes of raw audio recorded as a .wav file from the models used in this study typically took up 450–540 MB, which must be collected alongside the same length of video which was 3.2–4 GB when set at the lowest quality. The ability to record audio in isolation, duty cycle a device and storage of files in a compressed format would reduce the storage required for acoustic data collected by these devices and likely improve the battery life.

Additionally, GoPros only record frequencies up to 22 kHz when the raw setting is used and 16 kHz with audio extracted from video. This is likely because the upper limit of the human threshold for hearing is around 20 kHz, and in practice is usually more like 15 kHz (Ashihara, 2007). Although not considered in detail in this investigation, many marine mammal studies rely on detecting frequencies above these values (Matthews *et al.*, 1999) which would rule out GoPros. Some soundscape investigations sample higher than 20 kHz (Elise *et al.*, 2019), although this is not ubiquitous and a strong case of support for sampling above this limit has not yet been given (Lindseth and Lobel, 2018).

Benefits of using GoPros for soundscape investigations

GoPros offer some useful advantages over hydrophones, with the key benefit being their reduced cost. The self-contained fully submersible SoundTrap STD300 hydrophone used in this study can be purchased from the supplier for £2,230. In comparison, for £215 the GoPro HERO 7 Black with a 32 GB SD card and battery included, alongside Super Suit housing that enables underwater deployment, can be purchased from GoPro directly, with other retailers listing these for lower prices elsewhere. This makes the hydrophone over ten times more costly than the GoPro, and this is before international shipping and import tax for the hydrophone is included, which may be avoidable for GoPros in many countries.

The inexpensive nature of GoPros relative to hydrophones would allow an investigation to increase the number of sites from which soundscape recordings are taken by an order of magnitude with the same equipment budget. This could be used to greatly increase the spatial scale of the sampling effort in a soundscape-based investigation. Soundscapes have been found to differ over small spatial scales, habitats and depths (Radford *et al.*, 2014; Elise *et al.*, 2019; Ceraulo *et al.*, 2020; Davies *et al.*, 2020). This often requires investigators to sacrifice a more comprehensive sampling regime over a broader extent of the reef system they are hoping to study due to equipment limitations. Increasing the spatial range sampled could improve the confidence in conclusions regarding the value of different metrics for characterising a habitat's soundscape.

Other aspects worth highlighting are GoPros ability to be fully submersed with no components kept above the surface whilst recording; this has only been achieved with research hydrophones in the last decade (Sousa-Lima *et al.*, 2013). They also possess multifunctionality, being able to record video simultaneously with audio. This presents the opportunity to collect complementary visual and acoustic data. Their frequent occurrence in published literature also demonstrates that these devices are commonplace amongst research groups studying the marine environment. The potential for GoPros to be used for acoustic data collection opens the possibility for practitioners already in possession of GoPros to begin incorporating acoustic analysis into their work. Other studies have also demonstrated the potential of consumer-grade cameras such as GoPros for use in citizen science projects (Letessier *et al.*, 2015; Raoult *et al.*, 2016; Florisson *et al.*, 2018). In July 2018, GoPro announced that they had sold over 30 million GoPro HERO cameras around the world (Hillary K Grigonis, 2018). This opens up a significant potential for citizen science to contribute to the study of marine soundscapes.

Example Case Study

A recent marine soundscape study provides a suitable case study in which the use of this technology could be utilised (Elise *et al.*, 2019). Here, an investigation was performed on a reef system around Europa Island in the South West Indian Ocean which studied the relationship between its soundscape and several key ecosystem functions. Acoustic recordings were taken with a RESON TC 4014-5 hydrophone (Teledyne Marine, US) for a period of two hours during daylight at nine different sites. The indices ACI, BI, H, TV and SPL were calculated from these recordings across several frequency bins up to 50 kHz. Two GoPro cameras were also used to collect a 90 minute video recording in stereo at each site to identify fish and estimate abundance and biomass of those observed. These recordings were coupled with community assemblage surveys performed by divers.

This paper revealed some potentially exciting findings but was criticised in other published work as it only took two hour recording blocks from each site (Mooney *et al.*, 2020). The results presented in this chapter show that ACI, H and TV can all be reliably taken from GoPro recordings. Therefore, an alternative approach could have used consumer-grade recorders (e.g., the GoPros they were already using), and expanded the sampling effort at these sites. Acoustic data taken exclusively from GoPro recordings alone would limit some elements of the study. However, if taken in addition to hydrophone recordings these could have multiplied the replication of some measures. Additionally, this could instead be performed alongside the hydrophone with no need for access to any extra equipment beyond batteries and data storage.

Using GoPro 7's set to record raw audio as an example, the following costs and benefits would be provided:

Disadvantages

- GoPros do not record beyond 22 kHz (Fig. 6 & 7), this study recorded up to 50kHz.

- SPL could not be calculated.

- Results from the present study indicate BI may not be as reliable using a GoPro compared to a hydrophone (*rho* = 0.48).

Advantages

- Acoustic sampling effort could be increased. Recordings could be taken at each site every day, as opposed to just once, replicates at each site could be taken simultaneously, or, additional sites could be sampled.

- Video sampling effort could be increased. Additional video footage would be collected by each device, increasing the sampling depth of the fish assessments.

- Temporal trends could be considered due to the increased sampling capability. Changes over the lunar cycle could be studied if recordings were taken over multiple days. Alternatively, diel trends could be investigated if recordings were taken at different times of day at each site.

- Use of GoPros would remove the need for boat attendance or floatation of dry components during recording.

There are clearly costs and benefits of using both approaches and the use of consumer-grade recorders may not be appropriate to all the goals set in the investigation used as an example. However, this demonstrates that consumer grade recorders present researchers with additional options when designing their investigations that may provide important benefits.

Recommendations for use

A walkthrough on how to set up, deploy and retrieve data is detailed in the supplementary material (Supp. 2.3). For prospective users, this study reveals some important considerations for use. Firstly, raw audio recordings available with the GoPros appear to provide the most reliable file format to work with. An added benefit of this is the reduced file size of raw audio files which therefore download significantly faster and require less data storage.

Additionally, several models of GoPro are available alongside many other consumer grade recorders that may be useful. This study only focused on two models, the GoPro HERO5 Black and GoPro HERO7 Black deployed inside Super Suit underwater housing. Notable differences between the reliability of eco-acoustic indices between the two models was observed, with the GoPro 7 demonstrating the greater reliability for most indices. The difference between these two models highlights that other models cannot be assumed to provide a level of reliability similar to those presented here. Validation of the reliability of an index on alternative devices in a similar approach to this study would be required for each index before it is used in an ecoacoustics study.

Some considerations should be made by prospective users when selecting reliable eco-acoustic indices to calculate from GoPro recordings. Some correlations reported here are very strong, however, no finite limit of suitability is attainable. It is instead up to prospective users to apply their own discretion when selecting these. If using results from this investigation as a starting point when selecting an index or suite of indices, users should first check the strength of correlation, used to indicate reliability, for the index in question (Fig. 2.5). If this is satisfactory, results from the regression analysis (Table 2.2) should also be consulted to determine whether the proportional bias is reasonable and does not indicate that the GoPro or hydrophone failed to collect adequate results (e.g., Fig. 2.6D). For less reliable measures, a hydrophone could still be used to complement the GoPros if available and could even be used to reveal more about the reliability of these indices in different settings.

Limitations of this study

Some important limitations of the approach used in this study should be acknowledged. Firstly, only three GoPro 5s and two GoPro 7s were used and all GoPros from each model were treated as one unit. Due to logistical limitations these recorders could not all be deployed simultaneously every recording day. Although auditory and visual inspections of spectrograms from each recorder model and audio type (Supp 2.2) revealed minimal differences within these, some small variance may still be present. This may therefore have reduced correlation values for certain indices due to the introduction of an additional variable, which could have led to some indices being dismissed as they were tested across multiple devices whereas each individual device may have had a consistent relationship with the hydrophone. Due to equipment failure, different sample sizes were collected for the GoPro 7s (n = 55) and GoPro 5s (n = 40), this may have affected results from the statistical analysis.

Another limitation is the trialling of this in just one habitat, a coral reef ecosystem. The same reliability using GoPro recordings cannot be guaranteed for markedly different soundscapes present on alternative habitats, such as those dominated by anthropogenic (e.g., boat traffic) or geophonic (e.g., strong tidal flow) noise. Additionally, only one (recently calibrated) hydrophone was used for comparisons. Use of a single hydrophone is commonplace in most soundscape investigations (Nedelec *et al.*, 2015; Gordon *et al.*, 2018; Elise *et al.*, 2019). However, when validating a new approach, method comparison approaches often test with multiple devices to account for error in these (Carstensen, 2011; Taffé, 2018). To the authors knowledge there has been no comparison in the variance between indices taken from different hydrophone models and this could be an important consideration in itself.

Future study

As the first study to consider the use of a consumer-grade recorders for marine soundscape investigations, not every potential aspect has been explored. There is therefore room for further investigations that may help better elucidate the full extent of their potential in this field.

The next to step to determine the utility of consumer-grade recorders for soundscape investigations goes beyond the proof of concept presented here. This study took recordings at haphazardly selected locations and no habitat assessment was performed. Instead, this could be validated in a real-world experimental setting that shows whether indices calculated from GoPro recordings come to the same ecological conclusion as those from a hydrophone recording. This could be an assessment of the relationship between an index and an ecological parameter such as species richness, habitat types or ecosystem health.

To the authors knowledge, no coral reef soundscape investigation has considered as many eco-acoustic indices as trialled here in a single study to date. However, there are still many indices that were not explored, opening up other avenues of future study. Only alpha indices that output a single value independent of other recordings were considered as they could be more rapidly compared using a consistent methodology. However, beta indices were not tested in this study. These indices perform between group comparisons to assess how acoustically similar or dissimilar two or more groups are relative to one another (Sueur *et al.*, 2014). Beta indices applied to the marine field in the past include spectral dissimilarity (Lillis *et al*., 2014; Lindseth and Lobel, 2018) and acoustic dissimilarity (Bertucci *et al.*, 2016; Lindseth and Lobel, 2018).

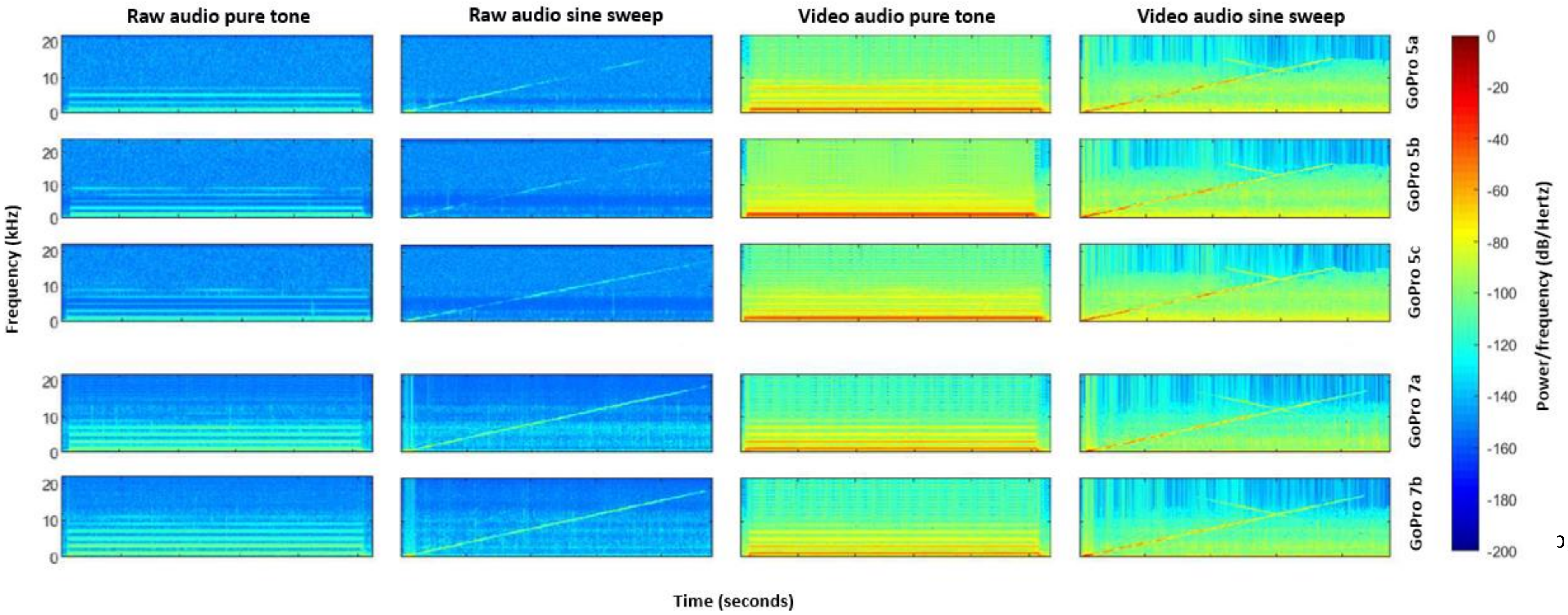
Additionally, only two frequency bands were trialled for each index in this investigation. Whilst an attempt was made to select the most suitable frequency bands, there is still little consensus over the most appropriate bands for each index (Nedelec *et al.*, 2015; Elise *et al.*, 2019). Further investigations may find alternative frequency bands offer improved or decreased reliability for eco-acoustic indices calculated from GoPro recordings. Likewise, only one combination of window size, envelopes, thresholds and other settings were trialled for each index when many are available and worth considering (Sueur, 2018a). The reliability of indices may improve or decline if alternative settings were to be trialled. Appropriate frequency bands and index settings may well change on a case by case basis depending on the habitat, time of day, community composition and other variables. This could limit the use of an index if an inappropriate band or setting is the only reliable choice, or, alternatively changes could improve the performance of indices, including those reported as less reliable in this investigation.

The combination of acoustic indices with the video capability of GoPros is another obvious route for further investigation (Reeves-Ozanich *et al.*, 2019). A multimedia approach applied to marine investigations utilising these methods could lead to the development of audio-visual indices not yet considered. For example, variables such as species richness and abundance, or visual attributes such as turbidity or lighting taken from video footage could be coupled with acoustic indices from the audio. Used together, these may provide indications of variables such as community resilience or presence of cryptic species.

Other consumer-grade recorders that overcome cost and availability barriers otherwise in place are also worth considering. Lower cost uncalibrated hydrophones are becoming available, such as the LSTN2 (£210) (Seiche, Holsworthy, UK) capable of on-the-go recording by connecting with smartphones. Cheap, autonomous, self-contained devices that can be left on unattended deployments for longer periods than GoPros are also likely to be available in the near future. The AudioMoth recorder (£47) (Open Acoustics, UK) has recently filled this niche in terrestrial PAM and adaptations for the marine environment using waterproof casing could meet this demand in the marine context.

Finally, passive acoustic monitoring of marine habitats is still an emerging field (Parsons *et al.*, 2016; Buxton *et al.*, 2018; Lindseth and Lobel, 2018). The need for a more complete understanding of the relationship between marine habitats, their soundscapes and the methods that can be used to study these has been recognised (Pieretti *et al.*, 2017; Buxton *et al.*, 2018; Obura *et al.*, 2019). During the early stages of developing marine soundscape ecology as a field it is understandably important that only equipment of a high specification is appropriate to fully explore and understand the acoustic domain. However, as key properties of soundscapes become recognised, expansion to a wider community of practitioners is necessary for the field to fulfil its potential. If these key components can be adequately assessed by lower specification consumer-grade recorders then these devices may be able to meet this demand. In turn, this expansion may contribute advances to the field of marine soundscape ecology in a mutually beneficial relationship.

2.5 Supplementary



Supplementary 2.3. A Guide to collecting soundscape recordings with GoPro cameras

Set up:

1. Switch the GoPro on and navigate to the ‘Protunes’ setting and turn this on.
2. In ‘Protunes’, select ‘RAW audio’ and turn this on and select, ‘low’. This will set the GoPro to output an additional .wav file with no compression, gain or other adjustments to the audio.
3. Whilst still in ‘Protunes’, select ‘Mics’ and change this to ‘Stereo’. This will prevent the wind reduction mechanism from switching any mics off whilst recording.
4. Return to the home screen and set video quality and FPS to the lowest available setting to save battery and memory if desirable.

The GoPro can be suspended in the same manor used for hydrophones. A simple approach can be used:

1. Place the GoPro in ‘Super Suit’ or other appropriate underwater housing.
2. Secure the device with two rotations of rope and cable ties.
3. Tie the rope off to a dive weight or other holdfast at the bottom and a small buoy for floatation at the top.
4. Ensure no moving parts are likely to produce rubbing or knocking noises that may disturb recordings.

Deployment:

1. Place on the desired location.
2. Press the record button once placed if aiming to conserve battery or storage.
3. GoPros will typically run for two hours on these settings if the battery being used hasn’t aged. Approximately 3–4.5 GB of data storage will be required for every ten minutes of recording.

Retrieving data:

1. Once retrieved, the GoPro can be connected to a personal computer to download the recordings via any file manager.
2. Unless video is required, it will be quicker and require considerably less storage to copy over the raw .wav files and delete the MP4 video and other files. Do not use the GoPro App, this will delete the raw .wav files.
3. The GoPro will have split longer recordings into multiple segments. These can be combined using Audacity or another appropriate software.
4. Alternatively, SD cards can be stored for this to be completed at a later date.

Chapter Three:

The sounds of recovery: machine learning reveals coral restoration enhances reef soundscapes

Authors: Ben Williams[a], Timothy A. C. Gordon[a,b], Lucille Chapuis[a], Andrew N. Radford[c], Stephen D. Simpson[a]

Affiliations: [a]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PS, United Kingdom; [b]MARS Sustainable Solutions, Makassar, Indonesia; [c]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom

Author contributions: B.W. was primarily responsible for analysis, modelling design and writing the manuscript. T.A.C.G. was responsible for field collection of data. T.A.C.G., L.C., A.N.R. and S.D.S. provided comments on analysis, modelling and the written manuscript. A full manuscript submitted for publication is included in the supplementary material to this thesis and will be submitted imminently (Supp. 3.1).

Abstract

Widespread degradation of tropical coral reefs around the world has resulted in them becoming amongst the most threatened habitats globally. This has led to an increased demand for conservation and restoration of these habitats. Adequate monitoring of restored sites is essential to assess their success and identify further areas in need of attention. This investigation builds on previous research that used labour intensive manual listening to explore how PAM can be used to assess the progress of actively restored sites at one of the world's largest tropical reef restoration projects, in South Sulawesi, Indonesia. The new work presented here applies modern computational approaches to recordings from the same sites to determine whether these could be used to more rapidly assess restoration using PAM data. A set of 12 eco-acoustic indices were calculated for

up to three frequency bands; a low (50–800 Hz), medium (2–7 kHz) and full band (0.05–20 kHz), for a total of 33 index-frequency band combindations. Fifteen of these 33 combinations reported a significant difference between healthy and degraded habitats. However, high variability in the distribution of results was observed, offering a limited ability for any one index to discriminate between these two habitats without extensive sampling. This investigation therefore attempted to construct a machine learning model which could better discriminate between these two habitat classes using an optimised set of combined eco-acoustic indices. This used a supervised approach (regularised discriminant analysis) that was trained on labelled one minute recordings from both habitats and then tested blind. The pooled misclassification rate of 1000 cross-validated iterations of the model was 8.27% (± 0.84), demonstrating the first ever successful implementation of PAM and machine learning to determine tropical reef health from acoustic recordings. 1000 repeats of the model were then executed on a set of artificially restored reef recordings from three sites. This reported that a recently restored site (<12 months) that still exhibited a reduced coral cover (25.6% ± 2.6) received a majority classification of its recordings as degraded (27/33), whereas two sites restored >24 months previously that now exhibit an increased coral cover (A: 79.1% ± 3.9; B: 66.5% ± 3.8) received a majority classification of their recordings as healthy (A: 33/39; B: 37/38). Future work should validate this method by investigating trends observed when this tool is applied to additional restored sites. If this method continues to report promising results, this approach could offer a valuable tool that allows marine practitioners to assess habitats rapidly using short snapshot recordings, or effectively monitor habitat recovery over time, with a reduced reliance on frequent labour intensive in-water surveys.

3.1 Introduction

The end goal of advancing marine ecosystem monitoring using soundscape ecology is to develop methods that provide useful real world applications. This includes efforts in conservation and restoration where novel methods of monitoring can be used to assess the effectiveness of these projects or to evaluate further sites that may be in need of attention. Improvements to monitoring can help conservation and restoration keep up with the growing demand for their implementation. This chapter will focus on the use of soundscape ecology to monitor the effectiveness of an active restoration approach used on tropical reef habitats.

Demand for restoration on tropical reefs

Tropical reefs cover less than 0.1% of the ocean surface, yet provide habitat for over 25% of described marine species (Plaisance *et al.*, 2011). However, tropical reefs are also among the most threatened ecosystems globally (Pratchett *et al*., 2014). Pressures such as overfishing, coastal development, resource extraction and climate change are some of the primary causes of this loss (Bridge *et al.*, 2013; Hughes *et al.*, 2017). These pressures have left the remaining tropical reefs in a vulnerable state (Hughes *et al.*, 2017). On many reef habitats, local stressors continue to rise, alongside the threat of climate change across reefs globally (Hoegh-Guldberg *et al.*, 2017; Hughes *et al.*, 2017). The number of communities reliant upon tropical reefs is also overrepresented, with over 275 million people dependent on the ecosystem services these habitats provide (Gattuso *et al*., 2014). Additionally, these communities are mostly in regions set to see population booms in the coming decades (Sale *et al.*, 2014). Tropical reefs are therefore high priority habitats requiring attention to protect the biodiversity present in these ecosystems and the hundreds of millions of people reliant upon the services they provide.

Restoration is one approach that can help support tropical reefs into the future. Restoration can be defined as the "*process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed*" (Society for Ecological Restoration International, 2004). The use of restoration approaches on tropical reefs is on the increase: a 2020 study identified 1,600 reported restoration projects, up from <400 prior to the year 2000 (Duarte *et al.*, 2020). Restoration approaches can be passive, such as the implementation of marine protected areas and reduced exploitation quotas, or they can be active, involving direct human interventions and management such as the cultivation of keystone organisms or the introduction of physical substrates (Becker *et al.*, 2018).

Monitoring restoration projects is important to assess the successes or shortcomings of these at achieving habitat and conservation goals. A recent report on the state of tropical reef restoration projects outlined a lack of clear objectives and sufficient monitoring as key issues inhibiting the success of these projects (Boström-Einarsson *et al.*, 2020). The report cautions that these issues can result in poorly implemented projects as the knowledge gained from previous endeavours is not built upon. On the other hand, improved monitoring and assessment of restoration projects can help quantify their effectiveness and inform decision makers and stakeholders on which approach is most appropriate for the habitats they are trying to restore and the communities that depend on them (Seaman, 2007; Baldera *et al*., 2018).

The potential of passive acoustic monitoring to assess reef restoration

This investigation was performed using existing soundscape recordings and coral cover survey data from one of the world's largest coral reef restoration projects, based in South Sulawasi, Indonesia (Williams *et al.*, 2019). Here, practitioners have used a novel restoration approach centred on the use of ‘Reef Stars’. These structures use steel frames 54 cm in diameter which are coated in course beach sand. Multiple live coral fragments from the surrounding environment are attached using cable ties and adjacent stars are bound together. This provides a stable substrate and improved flow of nutrients over the surface of coral fragments due to the 28 cm elevation given by the structure. Two years after deployment, coral cover on restored sites increased from less than 10% to over 60%, and supported a diverse community of 42 different coral species. As of 2015, 11,000 of these had been placed to produce an area of artificially restored reef over 7,000 m^2, with this number having continued to increase. The presence of further degraded habitats and naturally healthy habitats in the local area offered useful reference sites against which to compare ecological data from the restored sites.

Previous work used the recording and survey data to successfully demonstrate the application of PAM as a tool with which to measure the progress of restored sites (Gordon *et al*., in review (Supp. 3.2)). The authors used phonic richness—a novel manual listening approach developed for the purposes of their investigation—to assess fish-produced sounds; an important contributor to tropical reef soundscapes (McWilliam *et al.*, 2018; Carriço *et al.*, 2020). This new measure scored the diversity of fish sounds within a recording in the lower frequency range (<800 Hz) and was performed on multiple one minute recording samples from each site. Results from this analysis revealed that restored sites exhibit an enhanced diversity of fish sounds compared to degraded sites, converging with those observed on naturally healthy sites. These results supported the hypothesis that PAM can be used to differentiate healthy and degraded habitats at this location and that the progress of restored sites can therefore be assessed using phonic richness of fish-produced sounds as a reference.

Limitations of passive acoustic monitoring to assess reef restoration

While previous investigation has shown that PAM has the potential to be a useful tool to monitor tropical reef restoration, the method developed in Gordon *et al*. (in review) relies on manual listening which is a slow and labour intensive process. The ideal PAM tool should also provide a means to rapidly collect and analyse soundscape data. The researchers in this instance therefore also trialled the use of two acoustic indices which were calculated from the recordings used for phonic richness assessments. These were sound pressure level (SPL) and the acoustic complexity index (ACI), in a low (50–800 Hz) and high (2–7 kHz) frequency band for each. Tests using these indices found no meaningful correlation with the novel phonic richness measure, indicating that SPL and ACI do not appear to be dictated by the diversity of fish calls.

Comparisons between these indices and habitat type showed more promise, with significant differences between healthy and degraded habitats reported for ACI during daylight hours. However, the strength of this difference was still low and the distribution of index values included a significant overlap between values from either habitat type. This means the value of a random selected sample still has a high likelihood of originating from either habitat type. Extensive sampling is therefore required to build a robust dataset from each habitat which can be tested against one another. This undermines the goal of PAM which is to provide rapid assessments that can be performed with snapshot recordings. An ability to assess many more sites in this rapid fashion would facilitate a much broader spatial scale of assessment. This would help reduce the time needed to survey sites of interest and overcome equipment access limitations which can limit available hydrophones to a small number of sites when deployed to record for extended periods.

Aims of this investigation

The previous investigation provided a proof of concept that PAM can be used to discriminate between healthy and degraded habitats with a high degree of accuracy using manual inspection (Gordon *et al*., in review). The new hypothesis that, if manual analysis can be used to discriminate between these, then computational analyses may also be able to perform this task to a similar or improved degree of accuracy, was therefore created. This investigation set out to test this hypothesis and determine whether it may be possible to develop a computationally driven tool which can perform rapid assessments of habitat class in the context of the restoration project presented. Such a tool would need to be able to classify sites into the two classes of surveyed habitat, healthy or degraded, using short snapshot recordings from these sites. This would allow restored sites to also be compared to these reference habitats, enabling the success of restored sites to be assessed.

Exploring the use of individual eco-acoustic indices

This investigation first set out to explore further whether any index not already tested may exhibit a strong relationship with phonic richness using a greater suite of indices than those presented in Gordon *et al*. (in review). A strong correlation between any of these and phonic richness could offer a useful replacement for the more time intensive phonic richness approach. These same indices were also compared between habitat types to determine whether any could be used to accurately assign sites into degraded or healthy classes.

Numerous published studies performed on reef habitats have employed the use eco-acoustic indices in their investigation (Pieretti and Danovaro, 2020). However, thus far the highest number of these utilised in any such investigation on tropical reefs has been six (Elise *et al.*, 2019), with the majority of studies using three or less indices (Kaplan *et al.*, 2015; Nedelec *et al.*, 2015; Freeman and Freeman, 2016; Elise *et al.*, 2019; Bertucci *et al.*, 2020; Carriço *et al.*, 2020). This investigation therefore explored their use in further depth, utilising 12 of the most practical indices to perform a comprehensive assessment of the utility of each of these in the context presented. This includes the first reported use on reef habitats of three indices, the normalised-differences soundscape index (NDSI), number of peaks (NP) and the amplitude index (M).

Exploring the use of compound eco-acoustic indices

Recent advances in terrestrial soundscape ecology have seen soundscape ecologists advocate for the use of multiple indices in unison to generate multivariate compound indices (Eldridge *et al.*, 2018; Bradfer-Lawrence *et al.*, 2019). These compound indices more comprehensively capture the ecosystem functioning of a habitat than any one individual index can. The most advanced assessments of these use the modern analytical approach of machine learning. This can deliver an increased resolution of insight into these large multivariate datasets through identifying hidden patterns that traditional statistics are unable to reveal (Kendrick *et al.*, 2016; Eldridge *et al.*, 2018; Sethi *et al.*, 2020). A common technique is the use of supervised machine learning which requires the input of training data in the form multiple indices. An algorithm is then employed which attempts to build a model that can perform the regression or classification problem at hand. A supervised machine learning model was therefore constructed to determine whether this can deliver an improved ability to classify short recordings compared to individual indices used on their own.

Machine learning on acoustic recordings has seen some use in the marine environment, primarily for investigations using call identification to monitor marine mammal populations, due to their distinctive vocalisations (Bittle and Duncan, 2013; Roca and van Opzeeland, 2019). A similar machine learning approach was also trialled to compare diversity, richness and total number of manually identified sounds between different marine sites in the Everglades national park (Buxton *et al.*, 2018). However, this revealed no strong relationship between the models findings and these variables ($R^2 \geq 0.40$, MSE≥195).

This investigation presents the first use of machine learning as a PAM tool to monitor coral reef habitat and to test its utility to monitor marine restoration. The advantages of this approach compared to manual inspection and single index methods are discussed, and the future implications of such methods highlighted.

3.2 Methods

Data collection

Recordings and percentage coral cover values were obtained from previous work conducted by Gordon *et al.* (in review) (Supp. 3.2). Here, we collected data from seven sites around the islands Badi (Fig. 3.1B) and Bontosua (Fig. 3.1C) in the Spermonde Archipelago (South Sulawesi, Central Indonesia; 4°56.9′S, 119°18.1′E; Fig. 3.1A). These sites were representative of four distinct habitat types present within the system, these sites were: Healthy A & B, Degraded A & B, Mature Restored A & B, and Newly Restored (one site only), for a total of seven sites. The two healthy sites exhibited naturally high coral cover (A: 91.2% ± 2.0; B: 93.1% ± 2.6; mean ± SE) whereas the degraded sites exhibited low coral cover (A: 2.1% ± 0.9; B: 17.6% ± 4.6) as a result of coral mining and persistent destructive dynamite fishing at these sites (Williams *et al.*, 2019). The remaining three sites were comprised of previously degraded sites which have been restored using the novel restoration methodology developed in the region (Williams *et al.*, 2019). The two mature restored sites were established >24

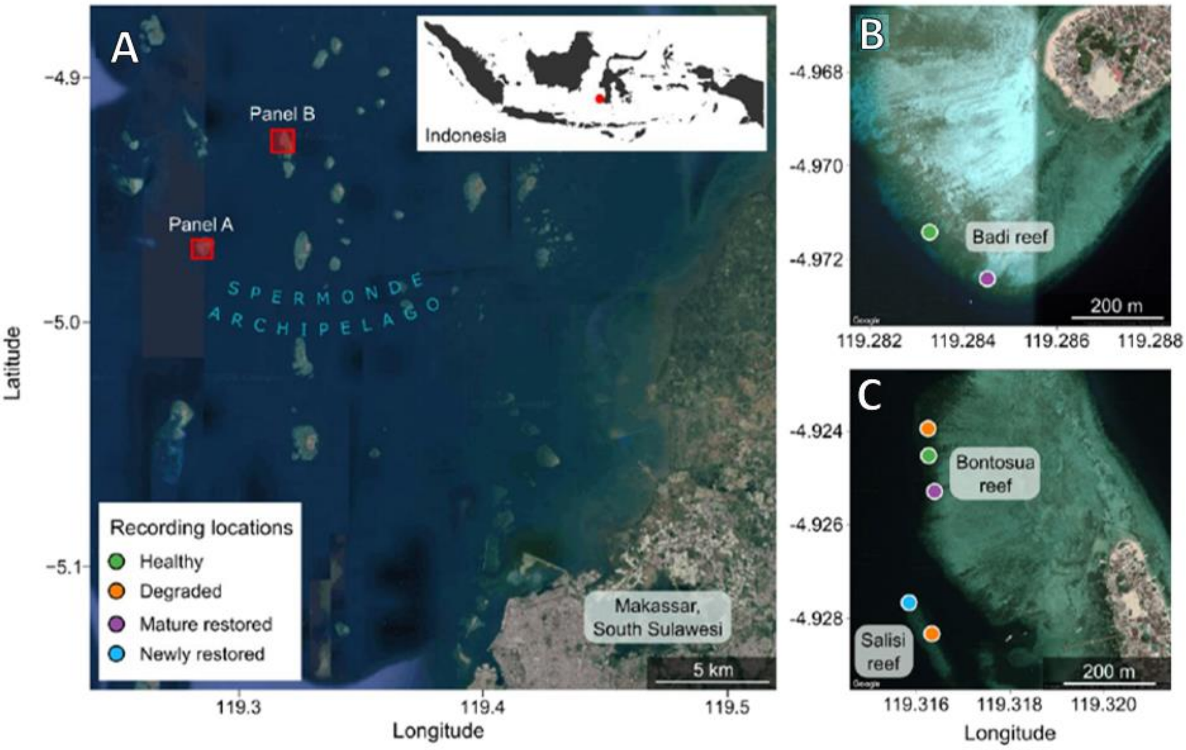


Figure 3.1. Location and habitat class of the seven reef sites, present within the broader Spermonde Archipelago (3.1A), from which soundscape recordings were collected. Fringing reefs from two nearby islands: Badi (3.1B) and Bontosua (3.1C) were used. This figure is adapted from Gordon *et al.,* (in review).

months previously to the collection of data for this study and exhibited an increased coral cover (A: 79.1% ± 3.9; B: 66.5% ± 3.8) over the newly restored site (25.6% ± 2.6) established <12 months previously. Further details are in Gordon *et al.* (in review).

Two-hundred and sixty-two one-minute soundscape recordings were produced across the seven sites using SoundTrap hydrophones (SoundTrap 300 STD, Ocean Instruments, Auckland, NZ) by Gordon *et al.* (in review). These hydrophones were suspended 0.5 m above the seabed and set to record at a sampling rate of 48 kHz. The recordings were collected using a regime which sampled sites three days across the 2018 new moon (September 10th) and five days across the following full moon (August 26th) during daylight (09:00–15:00)

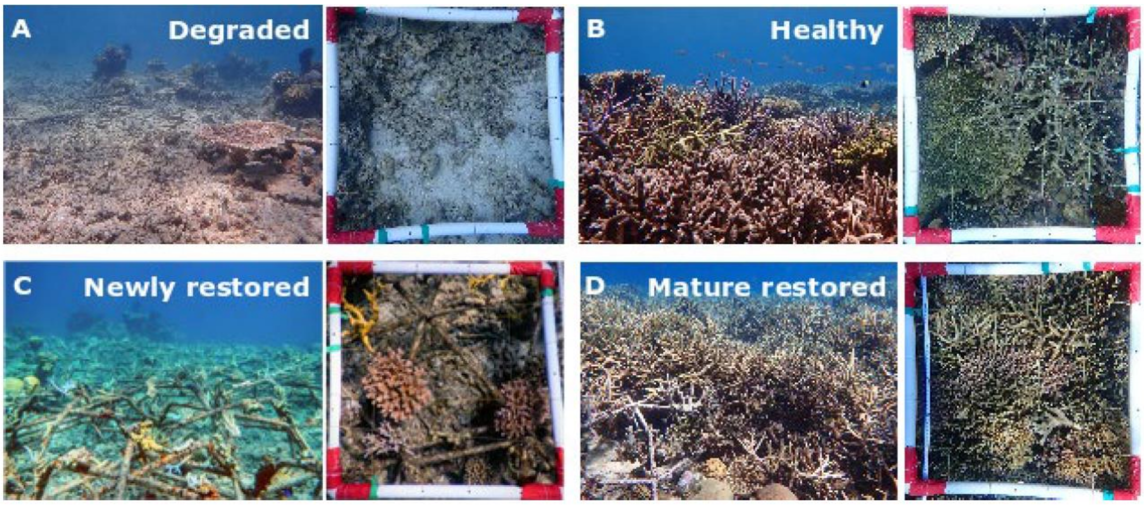


Figure 3.2. Representative habitat and coral cover images from the four habitat classes at which soundscape recordings were taken. (A) Degraded, (B) healthy, (C) newly restored and (D) mature restored. This figure originates from Gordon *et al.* (in review).

and night time periods (half an hour either side of: sunrise, sunset and midnight). The 262 samples were produced by sub-sampling five non-overlapping one-minute segments from each of these hour long periods at random. Only samples which were recorded under calm conditions (wind speed <20 km h^{-1}) and which were clear of motorboat noise were included in the sample set. As a limited number of hydrophones were available these were rotated between sites, for which an approximately even spread between each period was attempted, further details should be obtained from Gordon *et al.* (in review).

Processing recordings

Each one-minute recording from the 262 strong sample set was band-pass filtered using a short-term Fourier transform filter into three frequency bands: a low band (5–800 Hz), a medium band (2–7 kHz) and a full band (0.05–20 kHz).

The first two bands were selected in accordance with Gordon *et al.* (in review) which utilised a low band to encompass the known fish vocalisations present within the recording and a high band dominated by invertebrate sound. The additional full band produced for this investigation was selected to encompass the full spectrum of potentially relevant frequencies and is typically used in coral reef soundscape investigations (Kaplan *et al.*, 2015; Lyon, 2018). Frequencies <0.05 kHz were excluded from low and full band recordings to reduce the impact of low frequency shipping noise and geophonic noise from waves (Curtis *et al*., 1999). A new audio file for every recording in each frequency band was written to produce tracks filtered using a uniform method for subsequent analysis. All processing was performed in *R* (v3.4.2. R Development Core Team, 2020), audio files were read and written using the *tuneR* (v.1.3.3) package and the filter was implemented using *Seewave* (v2.1.6).

Calculating eco-acoustic indices

Twelve eco-acoustic indices were obtained (Table 3.1.). Every index was calculated for all three frequency bands with two exceptions. The first was snap rate which was only calculated for the middle and full bands as this index is designed to detect the sound of snapping shrimp cavitation bubbles which do not preside at lower frequencies (Bohnenstiehl *et al*., 2016). The second was the normalised difference soundscape index (NDSI). In the terrestrial context this index is typically used to quantify discrepancies in amplitude between an anthropogenic noise band up to 1 kHz and a biophonic noise band at selected higher frequencies (Kasten *et al.*, 2012). For the first time, this index was implemented in the marine environment to instead quantify differences in the 1 kHz band where fish noise dominates, and the 2–5 kHz band where snapping shrimp sound is reportedly at its highest intensity (Au and Banks, 1998). This was

Table 3.1. The 12 eco-acoustic indices calculated from recordings. A summary description of the mechanistic principle, the software used and the respective settings employed is detailed for each.

Index	**Mechanism**	**Software**	**Settings**	**Citation**
Acoustic Complexity Index (ACI)	Measures varibility in intensity of frequencies across time.	*Seewave* in *R*	Window size = 512; type = Hamming; overlap = 0.	(Pieretti 2011)
Acoustic Entropy (H)	Measures randomness across temporal and spectral domains.	*Seewave* in R	Window size = 512; envelope = Hilbert.	(Sueur 2008)
Acoustic Eveness Index (AEI)	Measures diversity across frequency bands.	*Soundecology* in *R*	Max freq = audio tracks maximum; freq step = max freq/10; threshold = -50 dB.	(Villanueva-Rivera 2011)
Amplitude Index (M)	Measures median of amplitude envelope.	*Seewave* in *R*	Envelope = Hilbert.	(Sueur 2008)
Acoustic Richness (AR)	Ranks recordings based on amplitude multiplied by randomness across the temporal domain.	*Seewave* in *R*	Envelope = Hilbert.	(Depraetere 2012)
Bioacoustic Index (BI)	Measures cumalitive intensity across frequency bands.	*Soundecology* in *R*	Min and max freq matched to track as approproate; window size = 512.	(Boelman 2007)
Normalised mean difference index (NDSI)	Measures ampltitude difference between two selected frequency bands.	*Seewave* in *R*	Min and max freq matched to track as approproate; window size = 512.	(Kasten 2012)
Number of peaks	Number of major frequency peaks on obtained from a mean spectrum	*Seewave* in *R*	Window size = 512; type = Hanning; overlap = 0.	(Sueur 2008)
Spectral entropy (sh)	Measures randomness across the frequency domain.	*Seewave* in *R*	No settings required.	(Sueur 2008)
Temporal Entropy (th)	Measures randomness across the temporal domain.	*Seewave* in *R*	No settings required.	(Sueur 2008)
Snap Rate	Measures rate of snapping shrimp snaps.	*MATLAB*	Custom script.	(Gordon 2018)
Sound Pressure Level (SPL)	Calibrated measure of root mean squared sound pressure level	*paPAM* in *MATLAB*	Window length = 1024; type = Hamming; Overlap = 50%	(Nedelec 2016)

therefore implemented on the full band recordings alone, to capture both the fish and shrimp bands. Therefore, for each of the original 262 samples, a feature set of 33 index values were created across the three frequency bands. All indices were calculated using the *R* package *Seewave* where possible and *Soundecology* (v.1.3.3) for remaining indices.

Comparison of indices to phonic richness and habitat class

Phonic richness scores were obtained from Gordon *et al*. (in review) for each recording. This novel metric quantified the diversity of fish sounds present within recordings and demonstrated an ability to discriminate between healthy and degraded habitat types using the same recordings from which indices were calculated for this study. A relationship between the full set of index results from all 262 recordings and phonic richness was tested for one by one for each of the 33 indices. This was performed using a Pearson's correlation test between each respective index and phonic richness. Note that Gordon *et al*. (in review) only used phonic richness scores from 100 recordings in their investigation to maintain a balanced design, however, phonic richness assessments had been completed for 262 recordings.

The results for each index were also plotted against each other to compare the degree of overlap between the distribution of results from healthy and degraded habitats. If little to no overlap was observed between the two classes for any index, then the respective index would provide a very promising measure with which to differentiate between healthy and degraded habitats. Additionally, the difference between the results of each of the 33 indices calculated from healthy habitats (n = 81) and those calculated from degraded habitats (n = 71) was also tested for using a Mann-Whitney U test.

Applying machine learning to soundscape data

The novel approach applied here attempted to develop a supervised machine learning model which could be used to accurately assign recordings to either healthy or degraded habitat classes. This was performed using a selection of indices in combination, referred to as a ‘compound index’ (Eldridge *et al.*, 2018). Three key steps to develop the models were undertaken: (i) a suitable machine learning algorithm for the task was selected, (ii) feature selection was performed to select the optimum combination of indices with which to construct the compound index, (iii) the algorithm was trained and tested using the compound index results from sub-samples of the full dataset. The finalised model was implemented on the full dataset of healthy and degraded recordings to assess the models ‘skill’; its ability to correctly classify new data when presented with it. It was also implemented on recordings from the three restored sites to provide a prospective assessment of the progress that restored sites had made towards converging with the soundscape of the healthy sites.

Selecting a machine learning approach

The first step involved selecting an algorithm able to deliver a predictive classification of samples into one of two habitat states using the continuous features available. This could be conducted using supervised learning, as the correct classification of each sample was known prior to its input and could therefore be used by the model to self-train on the data. Predictive discriminant analysis fulfils these requirements, of which there are three commonly implemented options: linear (LDA), quadratic (QDA) and regularised (RDA). An important assumption of the first two options is the absence of collinearity between features meaning there should be little to no correlation between each of the eco-acoustic indices input. A Pearson’s correlation test was therefore performed between each index, as in Gordon *et al*. (in review). The results (Supp. 3.3) indicated many of the features selected exhibit a high level of collinearity between one another. RDA was therefore selected as this test is robust to collinearity in features (Friedman, 1989).

RDA makes less assumptions than LDA and QDA. RDA's three assumptions are: (i) an approximately equal distribution of samples across the classes specified by the user; (ii) the minimum number of samples within a class is greater than the number of features; (iii) features are continuous and normally distributed within each separate class. For healthy and degraded sites 81 and 71 samples were available which was deemed satisfactory to pass the first assumption, and surpassed the maximum number of features available (n = 33), satisfying the second assumption. A Shapiro-Wilks test was implemented on all features from the datasets of both habitat classes to determine whether normal distributions were present. Only eight features from degraded habitats and nine from healthy habitats satisfied this assumption (p>0.05). However, histogram analysis revealed the remaining features exhibited sub-Gaussian distributions, with long tails on one or both sides that did not deviate considerably from a normal distribution. Discriminant analysis has been shown to be robust to violations of these four assumptions, including the normality of feature sets, allowing predictive models to be assembled without significantly detrimental effects to their accuracy (Lachenbruch and Goldstein, 1979). Deviations from these assumptions over and above what can be tolerated will result in a model with a high misclassification rate. If present, this will be revealed when the performance of the model is tested.

Feature selection

The next step was the selection of appropriate features from the 33 available to be used as a compound index with which to construct the most accurate model.

It is desirable to select an optimised combination of features to prevent overtraining, where the model becomes over fitted to the training data and no longer responds accurately to new data, as well as reducing computation time (Kuhn and Johnson, 2013). Two tests were performed to gather suggested feature combinations. The first was recursive feature elimination (RFE). This operates by selecting subsets of features and adding or removing a small number of other features progressively over multiple iterations until an optimised combination is found (Kuhn and Johnson, 2019). The second approach used was a multivariate adaptive regression spline (MAR) which constructs models using the specified algorithm (RDA in this instance) with different combinations of features. It then progressively adds the remaining features and scores the associated increase or decrease in parameters, such as the predictive error in the model, to determine the importance of a feature (Kuhn and Johnson, 2019).

One-hundred iterations of each approach were performed. The list of suggested features from the RFE included eight index/frequency band combinations; these were: ACI full, H middle, NDSI full, BI middle, th full, H full, ACI middle and th middle (details in Table 3.1). This was highly congruent with rankings obtained from the relative importance scores using the MAR (Fig. 3.3). From here, further trial and error was used by executing the full model (outlined below) to select a final feature set with the lowest misclassification rate. This was performed by sequential removal and addition of highly correlated features and features with a middle of the range relative importance score and above, using re-iterations of the model to test each new combination. This led to the discarding of th in both the full and middle bands and introduction of low band ACI and middle band AR into the final set (ACI low, ACI mid, AR middle, BI middle, ACI full, H full, NDSI full). Feature selection was performed using the *R* packages *mlbench* (v2.1.1) and *Caret* (v.6.0-86).

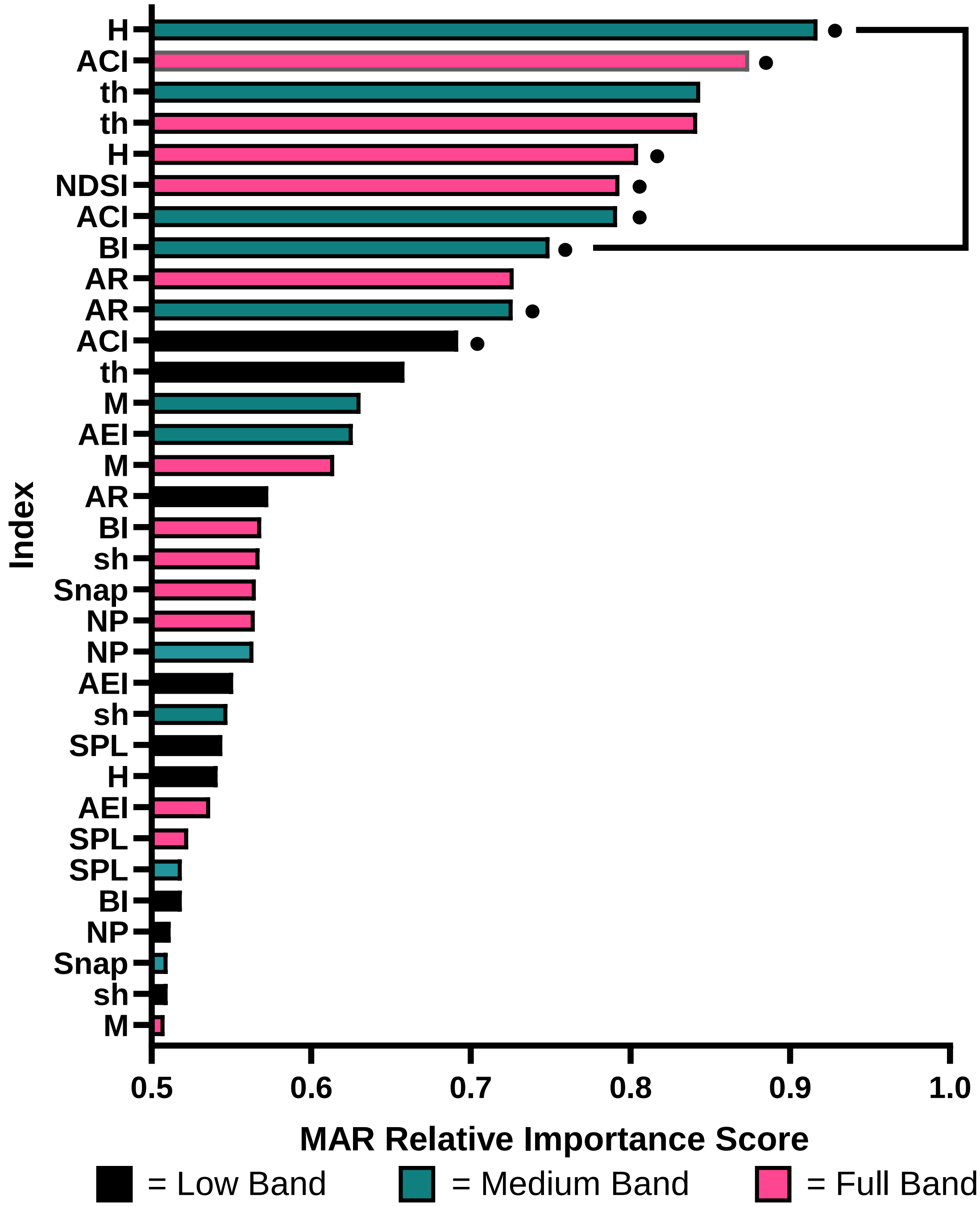


Figure 3.3. Relative importance rankings of indices obtained from the MAR analysis. The eight recommendations obtained from the RFE analysis are indicated by the black line. The top eight indices of the MAR analysis were congruent with the RFE's eight recommendations, though the order was not conserved. Black dots to the right of bars indicate features which were selected for the final model after further trial and error.

Constructing the final model

Using the healthy and degraded datasets, an RDA could then be executed. The model's prediction for each sample was reported enabling a classification error to be produced. This error is the proportion of samples that were assigned to the incorrect class; habitat type in this case, and is an indicator of the model's skill.

However, it is important to note that models constructed using the full dataset available typically overestimate their own accuracy. It is therefore essential to perform cross-validation of the model if a more representative estimate of its accuracy is required. Cross-validation involves splitting the data into two groups. The first is a 'training set' in which the model is provided with samples and informed of the correct classification for each, enabling it to construct its predictors which will be used to classify new data. The second is a 'test set', upon which the model is executed whilst blind to the true class of each sample. This yields a prediction of the class for each sample within the test set, allowing the accuracy of the model to be obtained when presented with new data that was not used in its construction (Stone, 1974). There are several varieties of cross-validation. In this instance, K-fold cross validation using 10 folds was identified as a suitable and conservative technique for estimating error (Hastie *et al*., 2009). This split the data into 10 groups, treating nine of the ten as the training set and then testing the model on the remaining fold which acted as the test set. This process was then repeated for all combinations of the initial 10 folds and the accuracy reported.

Importantly, like many machine learning approaches, RDA employs random processes in its establishment of a model (Wu *et al.*, 1996). Combined with the random division points used to sample the data in cross-validation, this leads to a classification error which is not fixed and varies with each re-iteration of the model. To better report on the accuracy it is recommended to construct multiple iterations of the cross-validated model (Kuhn and Johnson, 2013). Results from these can then be used to produce pooled values for the mean and standard error that better reflects the level of accuracy, and potential variability of this. This can inform the user asto the likely accuracy that can be expected, and the associated level of variance, when selecting one final instance of the model to implement on further datasets; in this case the compound index results for the restored reefs. One-thousand repeats of the cross-validated model construction were therefore performed to provide a suitable level of depth for accuracy to be assessed (Rao *et al*., 2008). The RDA model was constructed using the *R* packages *MASS* (v.7.3-53) and *KlaR* (v.0.6-15).

The next step was then to repeat this process now using the full dataset to construct 1000 models which could be used to classify the 110 audio samples from the three restored sites. However, the suitability of the restored data for entry into the model also had to be confirmed. Should the restored sites exhibit soundscape properties highly distinct from both healthy and degraded sites, the model would be forced to attempt to fit them into a classification that is inappropriate. The presence of divergence from both classes was therefore explored using cluster analysis. This employed a principal component analysis (PCA) conducted on the feature set of the eight selected indices and pairs plot which was also performed in *R* between every combination of two indices against one another (Supp. 3.4).

3.3 Results

Comparing indices to phonic richness

Results from the Pearson's correlation test between each index and phonic richness revealed no strong relationship between phonic richness and any of the 33 indices trialled (Fig. 3.4). The strongest finding was a negative correlation with the acoustic entropy index (H) in the full band (Pearson correlation; rho = -0.43; *p*<0.001), with all others reporting weaker correlations.

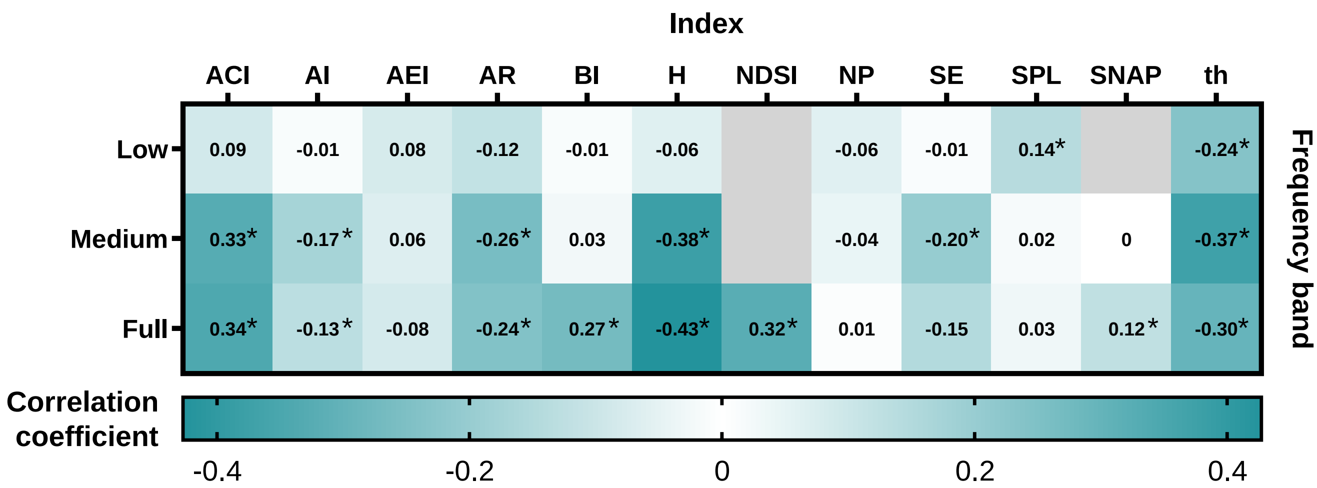


Figure 3.4. This heat map shows results from the Pearson correlation test between eco-acoustic index and phonic richness scores for the full set of 262 recordings in the three frequency bands employed. Strength of correlation is indicated by the colour bar. Cells marked with an asterisk indicate those with a significant correlation (*p*<0.05). Blank cells indicate indices for which values from the corresponding frequency band were not calculated (see methods).

Comparing indices between healthy and degraded sites

Results from the Mann-Whitney U test between healthy and degraded habitat index scores revealed there was a significant difference between 15 of the 33 indices between these two habitats (Supp. 3.5). The strongest significant difference was observed for medium band H (Mann-Whitney U; U = 1.98; *p*<0.001). Violin plots of the three most significantly different index results between the healthy and degraded sites show a large overlap between the distributions of the values between both classes (Fig. 3.5). This revealed that no index was able to differentiate between the two classes as a randomly selected value from one class has a high chance of also originating from the other.

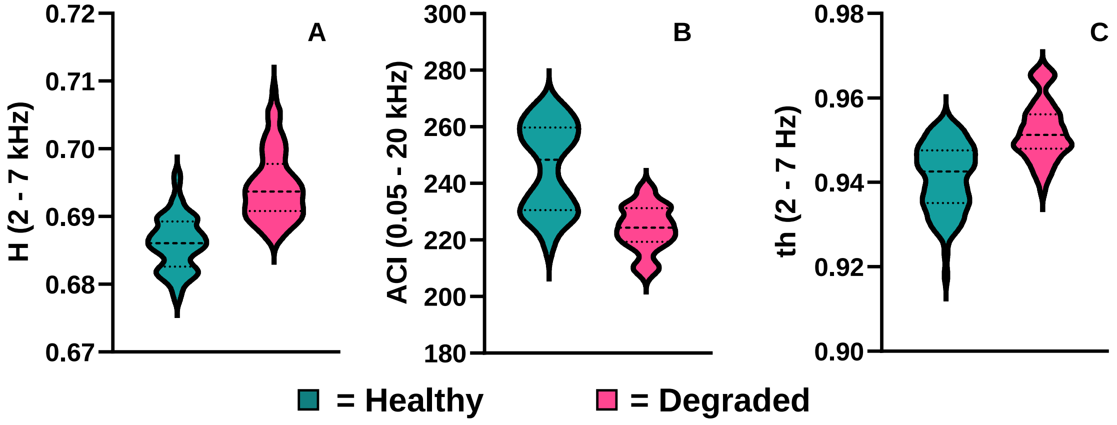


Figure 3.5. Violin plots of the three indices with the strongest significant differences between healthy and degraded habitat sites. (A) Medium band H (Mann-Whitney U; U = 1.98, p<0.001), (B) Full band ACI (U = 1.78, p<0.001), (C) Medium band th (U = 1.63, p<0.001).

Machine learning

From the 1000 repeated constructions of the cross-validated model using the 152 recordings taken across healthy and degraded sites, the pooled mean misclassification rate was 8.27% (± 0.84, SE). Results from the confusion matrix of these repeats showed that, of the 81 recording samples taken from the two healthy sites, 72.96 (± 0.11) of these were correctly classified as healthy, with 8.04 (± 0.11) misclassified as degraded. Of the 71 recordings taken from the two degraded sites, 67.22 (± 0.09) of these were classified as degraded, with 3.74 (± 0.09) misclassified as healthy. Individual results for each recording sample were also reported (Fig. 3.6).

Cluster analysis using the principal component analysis (Fig. 3.7) and pairs plot (Supp. 3.4) were then used to examine whether the 110 samples taken from recordings of the three restored sites were suitable for the model. Results from

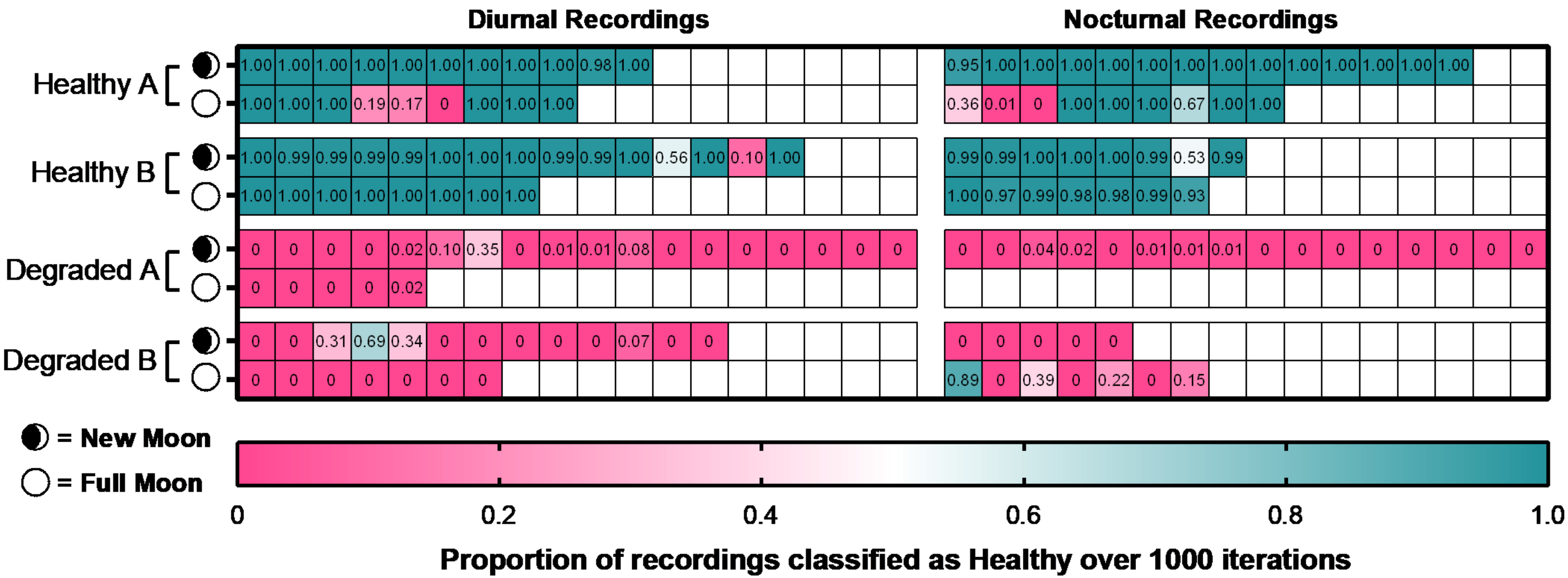


Figure 3.6. Habitat classifications predicted by the machine learning model. Each cell indicates a single one minute recording from the 152 that were available across healthy and degraded habitats. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time for every recording. Values within cells represent the proportion of these 1000 iterations in which the recording was predicted as originating from a healthy site, with the remaining being predicated as degraded, also represented by the colour code. Recordings taken on the left of the partition were taken during the day and recordings to the right were taken during crepuscular or night time periods. Although frequent gaps were present in the sampling regime, the order with which cells are presented within their respective blocks conserves the overall order with which they were sampled across time.

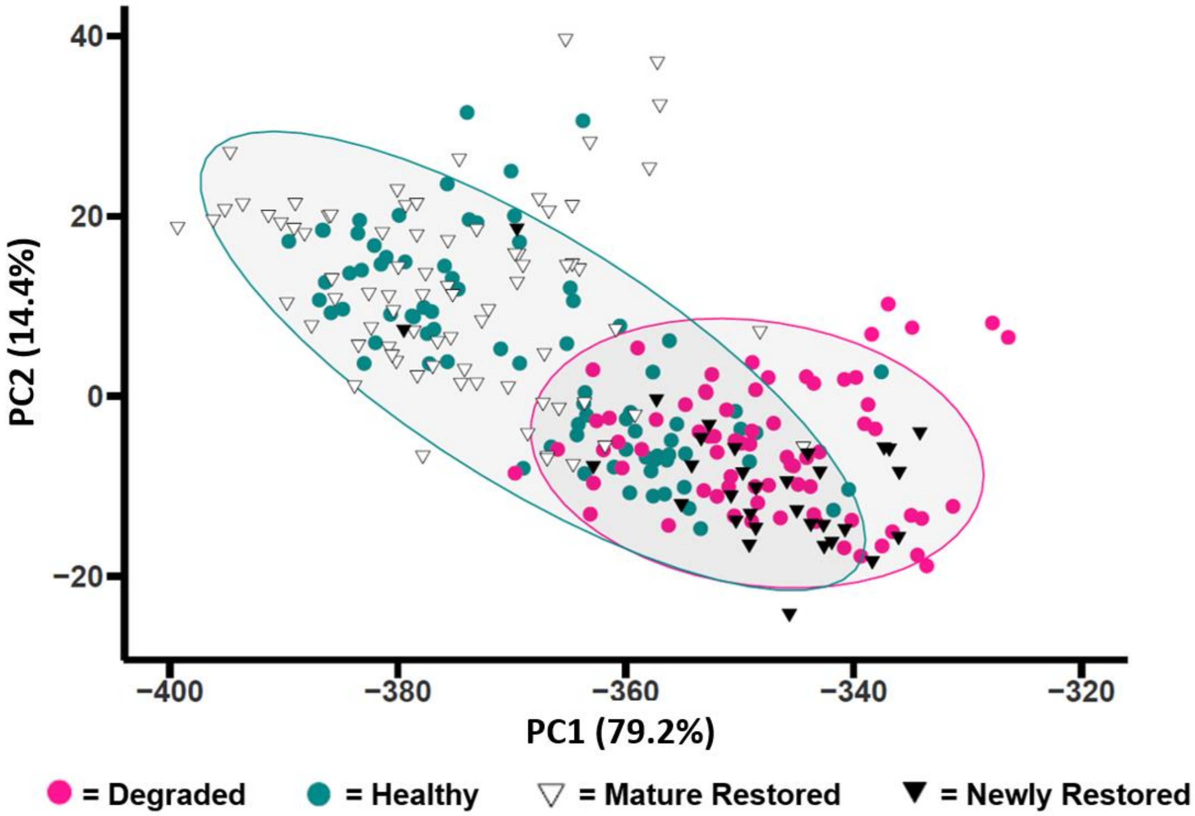


Figure 3.7. Plot from the principal component analysis of PC1 and PC2 for the Healthy and Degraded site recording samples. Samples from recordings of Restored sites are overlaid on this to help determine whether these conform with either of the two existing classes or whether the properties of their soundscape are distinct. Ellipses indicate the zone within which a new sample can be assigned to a class using the two principle components presented in this figure alone. Overlapping areas indicate ambiguous results which cannot be differentiated but nonetheless fit one of the existing classes.

the plots showed a strong overlap with both the healthy and degraded habitat classes. For the Mature Restored and Newly Restored sites 70/81 and 70/71 samples respectively fell within one or both of the predictive ellipses for the two existing classes. This indicates that the soundscapes of the restored sites did not diverge from the soundscape present on the other two habitat types when using the properties investigated here. This supports the inputting of restored samples into the model as this is likely to generate an estimation of classification with a similar level of accuracy observed for the original two sites from which it was constructed. Additionally, the PCA showed that 61/81 samples from the Mature Restored sites fell within the ellipse that could be used to predict healthy sites, whereas 24/27 samples of recordings from the Newly Restored site fell within the ellipse that can be used to predict degraded sites. However, it is important to note that there was a large region of overlap between the healthy and degraded class, with most of the ellipse of the degraded classes encompassed by that of the healthy class as well.

Execution of the model on the restored site samples was therefore performed in the same manner (Fig. 3.8). A key observation from the results of this is the majority classification of samples from mature restored sites as healthy, and samples from the newly restored site as degraded. A more decisive classification of Mature Restored site B was reported over Mature Restored site A, with 37/38 and 33/39 samples reporting a majority classification of healthy respectively. The six samples which reported a majority classification as degraded on Mature Restored site A occurred consecutively on the new moon at night. On the Newly Restored site, 27/33 samples reported a majority classification as degraded, all of these were during the full moon (though only four new moon samples were available) and five of these were at night.

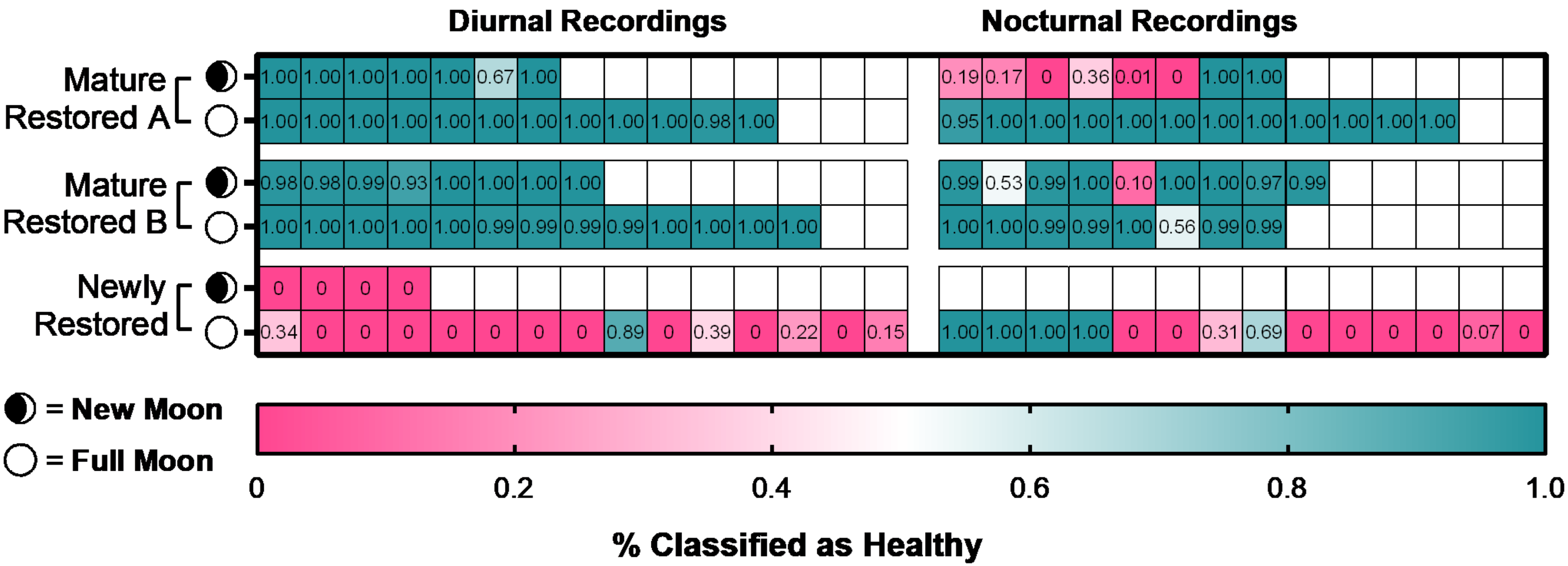


Figure 3.8. Habitat classification predictions by the machine learning model for the restored site recording samples. Each cell indicates a single one-minute recording from the 110 that were taken from restored sites. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time for every recording. Values within cells represent the proportion of these 1000 iterations in which the recording was predicted as originating from a healthy site, with the remaining being predicated as degraded, also represented by the colour scale. Recordings taken on the left of the partition were taken during the day and recordings to the right were taken during crepuscular or night time periods. Although frequent gaps were present in the sampling regime, the order with which cells are presented within their respective blocks conserves the overall order with which they were sampled across time.

3.4 Discussion

This investigation set out to determine whether computational analysis of soundscape data collected using passive acoustic monitoring (PAM) from tropical coral reefs could be used to perform rapid classification of these habitats into two eco-sates: healthy or degraded. Eco-states were classified using coral cover as a measure of this and the soundscapes of two sites from both of these classes were sampled over the new and full moon periods to build a dataset with which to test this hypothesis. Findings from this investigation support previous work which suggests no single eco-acoustic index can be used to discriminate between these two eco-states. However, use of a supervised machine learning approach that considered an optimised set of multiple indices in unison demonstrated an ability to accurately predict habitat class from randomly drawn acoustic samples with a misclassification rate of just 8.27% ($\pm$ 0.84). To the author's knowledge, this is the first successful demonstration of combining PAM with machine learning to monitor coral reef habitats, highlighting the potential of this approach.

Furthermore, this model was executed on soundscape recordings taken from nearby coral reef habitats that had been actively restored (Williams *et al.*, 2019). This was used to demonstrate the ability of this approach to perform a rapid assessment of these restored sites using one-minute soundscape recordings as an indicator. This has strong implications for marine restoration practitioners interested in using PAM to monitor the progress of restored sites against reference habitats. More generally, it also demonstrates the potential for further applications of machine learning using PAM data from coral reef habitats. Such applications could be used to investigate other aspects of reef ecology that may be represented in the soundscape. These kind of advances are needed to help marine practitioners capitalise on the ability of PAM to collect data over broad temporal resolutions using easily reproducible low effort surveying methods.

Relationship between eco-acoustic indices and manual listening

Previous work by Gordon *et al.* (in review) used phonic richness, a novel manual listening technique, to quantify the diversity of fish noise within the recording samples used for this investigation. Their results showed that phonic richness significantly outperformed two existing eco-acoustic indices (ACI and SPL) in discriminating between healthy and degraded habitats. This supports findings from previous investigations which have shown results from manual assessments of fish noise diversity can relate to other factors such as diversity of fishes or their distribution within distinct habitats (McWilliam *et al.*, 2017; Desiderà *et al.*, 2019). However, the scoring of phonic richness is labour intensive due to the demanding process of manual listening and annotation, and is restricted by the requirement of a single human assessor to produce consistent results. The development of computational approaches that can rapidly produce comparable results to this method is therefore highly desirable.

The present investigation therefore expanded on work from Gordon *et al*. (in review) by running a more comprehensive suite of 12 acoustic indices on the same recordings across three alternative frequency bands. Whilst 15 of the 33 indices showed a significant correlation (Fig. 3.4), these were all weak, with the strongest being the acoustic entropy index (H) in the full band (Pearson correlation; rho = -0.43; $p<0.001$). This suggests that it may be an unrealistic assumption that the diversity of fish noise present in a recording is a direct driver of results generated by any of the acoustic indices currently available. This has been an assumption of numerous investigations to date which act on the hypothesis that certain indices are quantifying the diversity of fish noise (Bertucci *et al.*, 2016; Buxton *et al.*, 2018).

However, it is important to note that the phonic richness score assessed the diversity of fish noise by noting the number of unique fish sounds present within a recording, but did not quantify the frequency of occurrence of each of these within a recording. Previous research has shown that frequency of occurrence of fish sounds can have a significant impact on the values of eco-acoustic indices taken from naturally occurring coral reef soundscapes (Staaterman *et al.*, 2017). Other investigation have also tested the robustness of certain indices to simulated recordings with alternative levels of fish vocalisations. They report similar findings that show indices can be sensitive to both diversity and frequency of occurrence of fish sounds making it difficult to discern between these variables (Bohnenstiehl *et al.*, 2018; Bolgan *et al.*, 2018). A combined diversity and abundance metric for fish noise may reveal a stronger relationship between indices and these assessments.

Additionally, the 50–800 Hz band consistently reported the weakest correlation between phonic richness and the suite of indices trialled (Fig. 3.4). This band fully encompasses the range of audible fish sounds within the recordings used to generate phonic richness scores (Gordon *et al.*, in review). It is therefore interesting to note that indices in the higher frequency band and full band reported slightly stronger correlations. This further supports the hypothesis that indices are not directly quantifying the diversity of fish sound. Instead a combination of diversity ad abundance of fish produced sounds, and/or other elements of the soundscape that may be captured in these alternative bands. However, this does open the possibility that these two elements are still related, meaning that a recording with a high diversity of fish sound may also be likely to exhibit some other unknown traits to which indices are more sensitive. Future studies that further explore this hypothesis using simulated tracks may help reveal more about these observations.

Using eco-acoustic indices to compare healthy and degraded sites

Having failed to show a strong correlation with phonic richness, a suite of eco-acoustics indices calculated for each recording sample was also compared directly between healthy and degraded habitats. This was done to determine whether any of the indices could reliably be used to discriminate between these two habitat states. Of the 33 index and frequency band combinations trialled, 15 reported a significant difference between the two habitats classes (Supp. 3.5).

The present investigation trialled a greater number of indices than previous tropical reef soundscape studies. However, it was clear from these results that no single index demonstrated the ability to discriminate between healthy and degraded sites with the desired power. The high test statistic values from the Mann-Whitney U tests (Supp. 3.5) highlighted that the separability between both classes was still low; for any given value from one habitat there was still a high chance it would overlap with values reported for the other habitat. Violin plots of the three most significant results help to visualise this by revealing the high degree of overlap between the results of each habitat class (Fig. 3.8). This shows that a single recording could not reliably be used to discriminate between the two habitat classes.

This is congruent with findings from previous research. Studies directly reporting on the difference between index values of healthy and degraded tropical reef habitats revealed significant differences between the two habitat classes using extensively replicated samples (Butler *et al.*, 2016; Gordon *et al.*, 2018). However, once again their results show high within-group variability that precludes the use of any single index to report a reliable habitat classification on individual samples alone. Additional investigations have explored the relationship between indices and other aspects of reef ecology that can be strong indicators of overall reef health such as: fish diversity, structural complexity, algal cover and more (Nedelec *et al.*, 2015; Freeman and Freeman, 2016; Buxton *et al.*, 2018; Elise *et al.*, 2019). Once again these studies consistently reported that indices do not constitute reliable indicators of these other aspects of reef ecology without extensive replication, and often they still fail to provide any discriminatory power when previous investigations did report this (Kaplan *et al.*, 2015; Harris *et al*., 2016).

Evaluating the features selected for the machine learning model

Although no single index demonstrated an ability to accurately classify individual recordings as degraded or healthy, many revealed an ability to perform this when considered as part of a larger dataset.

The first stage of this process was the feature selection step which reduced the full set of 33 indices to a smaller optimised set that best captured the soundscape properties of interest whilst reducing the risk of overtraining. Of the eight indices selected for the final feature set (Fig. 3.3), the Acoustic Complexity Index (ACI) was the only index to be included from all three frequency bands. This index has become one of the predominant indices with which to compare reef soundscapes (Harris *et al*., 2016; Bolgan *et al.*, 2018) and its selection across all three bands further supports its value for assessing reef recordings. The individual index with the highest importance ranking from the multivariate adaptive regression spline (MAR), and also the strongest significant difference between healthy and degraded sites as revealed by the Mann-Whitney U tests, was medium band H. This index has seen less use than ACI in reef soundscape ecology investigations thus far, but, has been found to reveal useful trends in some studies (Staaterman *et al.*, 2014; Harris *et al*., 2016). Investigations into the have shown this index is driven by snapping shrimp and fish chorusing (Staaterman *et al.*, 2014; Kaplan *et al.*, 2015; Siddagangaiah *et al.*, 2019). Next in the MAR importance rankings was th in the medium and full bands. However, this index was highly correlated with H (Supp. 3.3) and was therefore excluded from the final feature set. Whilst the findings here support its use in discriminating between healthy and degraded recordings, the results suggest future users conducting a similar investigation may wish to consider H in its place if necessary. Both medium band BI and AR were also included in the final feature set. BI has seen little use on reef habitats, however, a recent investigation found it outperformed other indices when being used to indicate planktivorous fish abundance and laminar, foliose and helmet shaped coral abundance (Elise *et al.*, 2019). AR has also seen limited use but has previously been shown to significantly differ between pre-bleaching and bleached coral reef sites as well as being an indicator of fish diversity in certain settings (Harris *et al*., 2016; Gordon *et al.*, 2018).

The last index included in the final feature set was the normalised mean difference index (NDSI), which was used on a tropical reef soundscape for the first time in this investigation. This index was used to quantify differences between the 0.05–1 kHz band, which includes most sounds produced by fish (Gordon *et al*. in review), and the 2–5 kHz band where snapping shrimp are reportedly at their most intense (Coquereau *et al.*, 2016). Several indices did not feature, most notably of which is sound pressure level (SPL). This index is the most utilised index for assessing the soundscapes of reef habitats (Pie*retti et al*., 2017; Lindseth and Lobel, 2018). However, when a broad suit of indices are utilised it does not necessarily come out as the strongest performer for testing ecosystem functions (Elise *et al.*, 2019). A reporting bias due to its frequent use may be responsible for its prevalence as an ecological indicator. Further research into its performance against other indices may be useful in determining whether this should be considered a default index of choice for studies on reef soundscapes in the way it has been in the past (Nedelec *et al.*, 2015; Bertucci *et al.*, 2020).

It is important to note that findings during the feature selection stage of this investigation are specific to the data and questions under consideration here. These should be applied with caution in other studies. However, with that in mind, these findings may offer a useful starting place for future work performing similar investigations. Comparing and contrasting optimised feature sets of alternative models addressing similar questions may help elucidate indices and frequency bands that consistently yield useful findings over other combinations. Further research into this aspect of reef PAM and machine learning will be necessary to properly explore these comparisons.

Future research may also find success using alternative indices for capturing acoustic features, in addition to the indices trialled here. A recent study used machine learning to estimate the species richness of marine mammal communities in two different Southern Ocean habitats. They used a similar index approach to the one applied here, and also a second approach which split recordings into 256 evenly spaced frequency bands up to 2.5 kHz, from which they input amplitude values for each bin into a random forest algorithm. They found both approaches delivered similar results, suggesting this additional method may be worth trialling on reef habitats as well. Another approach worth considering was demonstrated in a study which used recordings of forest habitats. Here, researchers utilised a 128 feature set and cross-convolutional neural network (CNN) developed by Googles DeepMind team that was created using sound samples from AudioSet, a database of 70 million labelled audio files taken from YouTube (Hershey *et al.*, 2017). This enabled the soundscape ecologists to apply this CNN using an unsupervised machine learning approach on 2750 hrs worth of recordings taken from several different forest habitats around the globe. This was able to assign samples to many key categories, enabling the team to observe divisions between samples taken from different habitats around the globe as well as habitat quality and temporal cycles within individual forests (Sethi *et al.*, 2020).

Assessing model skill

When assessing the model’s overall accuracy to discriminate between the two eco-states it is important to note that limited spatial replication was used in this investigation. Only two healthy and two degraded sites were explored, which does not constitute an extensively representative sample set of similar habitats. However, with this in mind, the effectiveness of the final model showed an impressive ability to discriminate between the two classes of habitat that separate these sites. Across the 1000 iterations of the cross-validated model, the habitat class for each of 152 recording samples taken across both habitats were predicted with just an 8.27% ($\pm$ 0.84) misclassification rate.

To explore how the models skill could have been improved further it is worth considering the sources of the observed error rate. The presence of this error could be due to several factors in isolation or in combination. As highlighted in the methods, the sub-Gaussian distribution of several of the feature sets technically violates an assumption of the RDA (Wu *et al.*, 1996). This effect was likely due to the inclusion of samples from alternative times of day and multiple sites. Diel trends are frequently observed in reef soundscapes and this is reflected in the output of eco-acoustic indices (Kaplan *et al.*, 2015; Bertucci *et al.*, 2020; Carriço *et al.*, 2020). Assessing the discrepancy between the nocturnal and diurnal periods has even been proposed as a useful indicator of certain ecosystem functions (Kaplan *et al.*, 2015). Additionally, reef soundscapes are known to differ over small spatial scales (Putland *et al*., 2017). Considering samples were taken from spatially separated sites to provide replicates, it is to be expected that differences across the same habitat class will have occurred. Both of these factors may have skewed the distributions of the feature sets. Furthermore, the dataset used to train the model itself was likely imperfect and will have contained natural outliers through ecological randomness that cannot be resolved at the sampling resolution employed. A recording regime will never truly be representative of the soundscape of the habitat for which it is employed unless every minute of every day is recorded at each site for the duration of the study, and after termination of this it is no longer necessarily representative into the future. Even if such a strict regime was employed, ecological randomness that cannot be explained will likely still impact results. For many investigations hoping to study broader temporal periods this level of sampling is impractical and as such compromises are made which include the loss of data that could fill gaps needed to explain observed trends, increasing the chance that randomness impacts results further.

Application of the model on restored sites

The misclassification rate reported by the 1000 cross-validation iterations was considered low enough for the model to be proficient at performing preliminary classifications of samples taken from the restored sites. However, it was first important to confirm whether the samples from the restored sites could be fitted to the samples on which the model was trained. If the properties of this new test data were highly divergent then it may not be appropriate to apply the model to them. For example, if a model was trained to classify recordings from healthy and degraded coral reef sites, and was then used to classify recordings of the open ocean, or even a terrestrial habitat, the model would of course be inappropriate as there would be no meaningful similarity between the properties of these soundscapes.

Results from the principal component analysis (Fig. 3.7) and pairs plot (Supp. 3.4) revealed a high degree of overlap between the training data and the restored sites. This indicates that the new test data, in the form of the restored samples, appeared appropriate for entry into the model. As discussed, conclusions from this assessment may need to be considered with caution given the low spatial replication used to train the model. However, some interesting observations were made when studying the results taken from 1000 iterations of the model on the restored site recording samples.

One of the key observations from these results was the apparent disparity between the Mature Restored sites and the Newly Restored site. Of the recording samples from the two Mature Restored sites, 70/77 were given a majority classification of healthy, whereas only 6/33 samples from the Newly Restored sites received this (Fig. 3.7). The Mature Restored sites off Badi and Bontosua islands were established >24 months prior to recordings and a higher percentage coral cover was observed on these (A: 79.1% ± 3.9; B: 66.5% ± 3.8) compared to <12 month old Newly Restored site off Bontosua (25.6% ± 2.6). Once again, the spatial replication performed on restored sites was low, with just two and one Mature and Newly restored sites respectively. This limits the confidence that can be given to conclusions associated with this data. However, of the data available, this supports the hypothesis that the stage of recovery these sites have reached can indeed be indicated by their soundscape and that this can be detected using a machine learning approach. This therefore offers an exciting new way for marine conservationists and restoration practitioners to monitor the progress of their efforts on similar tropical reef habitats, and perhaps elsewhere. Future research to further validate these findings and develop tools to assess this is therefore essential.

If this hypothesis is correct, it would suggest there exists a point, between the stages of recovery observed on the Mature and Newly Restored sites, at which the soundscape begins to converge more closely with healthy sites. It is likely that this is not a sudden switch, and the convergence happens gradually. This would be observable when inspecting the classification predictions, which may also reveal the stage at which samples report this switch is dependent on when recordings are taken, such as alternative lunar or diel periods. Whilst such a trend is observed in the data reported here, more samples would be needed to verify this.

Further development of machine learning approaches

This investigation demonstrates that a model based on marine soundscape properties can be constructed using machine learning and used to make reliable predications of which eco-state (healthy or degraded) a recording most closely aligns with. This offers exciting prospects to expand the utility of PAM on reef habitats. In theory, if time was taken to collect training data from appropriate sites, a model could then be constructed and used to perform a much more rapid assessment of additional sites within the same reef network using significantly reduced sampling regimes. This could overcome the limitation of equipment availability and greatly increase the number of sites that could be tested, allowing informative indications of their health, or other functions, to be obtained. For example, an investigation such as this, could first collect a robust dataset from which to construct a model to quantify an ecological attribute of interest. Short term deployments just minutes in length could then be performed on a larger rotation of sites, capturing snapshots of each from which a classification could be obtained. An increased confidence in these conclusions could be delivered through repeats on these across broader temporal scales, perhaps through performing repetitive cycles a few minutes in length over many sites during one day. Further validation of this method against traditional survey techniques would certainly be needed. But, if continued success is seen, this would allow mapping of ecosystem functions which can be predicted using soundscape properties to be performed over entire reef networks more rapidly than traditional survey techniques.

Future investigations employing the approach used here could develop this further in a number of ways. Firstly, this study employed a binary classification of reef health. In reality reefs exist on a broad gradient of eco-states that are not as simplified as this classification (Downs *et al.*, 2005; Smith *et al.*, 2008). Recording samples used in this investigation therefore could only be assigned to one of two classes. This overlooks a broad level of variability possible within each class and makes it difficult to handle sites which lie somewhere in between this as neither class is fully appropriate. One work around in this situation may be to sample multiple recordings from sites across different times of day and lunar phases to observe how many of these samples were assigned to each class, allowing this to be placed on a discrete scale dependant on these results. However, an improved approach would be to have applied the same level of sampling depth completed on each site in this investigation across many more sites on this sliding gradient of eco-states. An alternative machine learning algorithm such as random forests, neural networks or logistic regression could then be trained on these data to produce models which can make predictions on a continuous scale.

Further quantification of reef health beyond coral cover alone could also have been employed. Coral cover can be a strong indicator of overall reef health (Smith *et al.*, 2016; Dietzel *et al.*, 2020). However, other attributes should be considered to properly determine the eco-state of a site to appraise whether these were truly comparable across the sites sampled here. For example, fish abundance can be an important consideration in assessing reef health. A previous investigation into fish abundance at these sites found that the introduction of Reef Stars significantly increased the number of fish present within one month, and can double fish abundance after four years (Seraphim *et al.*, 2018). Research into the relationship of this attribute and the soundscape of natural and restored reefs may reveal new ways to rapidly estimate fish abundance at these sites.

Additionally, other variables universally applicable to reef sites, such as depth or reef zone (e.g., fore, back, flat), could also have been included in model development. Existing research supports the premise that these factors can have an impact on soundscape (McWilliam *et al.*, 2018; Elise *et al.*, 2019). The model presented here therefore cannot be applied to reefs with qualities that significantly diverge from the sites used to collect training data without reducing the level of confidence in its predictions. However, if additional sites with these variations were sufficiently sampled to provide training data from these sites, a model could potentially be constructed which accounts for this. Although, if described using categorical data, an alternative algorithm able to accommodate this would be needed.

Additionally, the model in this investigation was trained on 81 healthy and 71 degraded recording samples across the new and full moon periods. Across the four sites this may be an adequate level of sampling to minimise the risk of overtraining whilst delivering sufficient predictive skill in this instance. However, to increase the confidence in the applicability of this method over neighbouring reef systems, increased spatial replication using a similar sampling depth to this on each new site could have been applied.

It may also be desirable to sample across a broader temporal scale. However, this depends on the intended output of such an investigation. Samples could have been taken across the full month allowing the lunar phase to be included as a feature. But, this would have required further sampling across the full month to collect a dataset with which to train the model. Considering that this technique offers the ability to obtain classifications using rapidly obtained snapshot samples from other sites of interest, it may not be necessary to go to such efforts when sampling if the desired outcome is achievable with less. However, if additional ecological attributes were under investigation using more complex models, these narrow periods may not necessarily contain detectable acoustic fingerprints which could be used to quantify such attributes. For example, if eco-state were to be classified over a gradient, or fish diversity was assessed on a continuous scale, information on these attributes may only be contained in soundscapes over broader temporal periods, if at all. The discriminant analysis approach used here can be used to predict multiple categories which may be sufficient for some applications. However, once again other machine learning algorithms can offer more precision in these assessments through prediction of continuous values. Further research into the application of these will reveal more about the utility of such approaches and contribute important steps towards unlocking the full potential of PAM and machine learning on reef habitats.

Conclusion

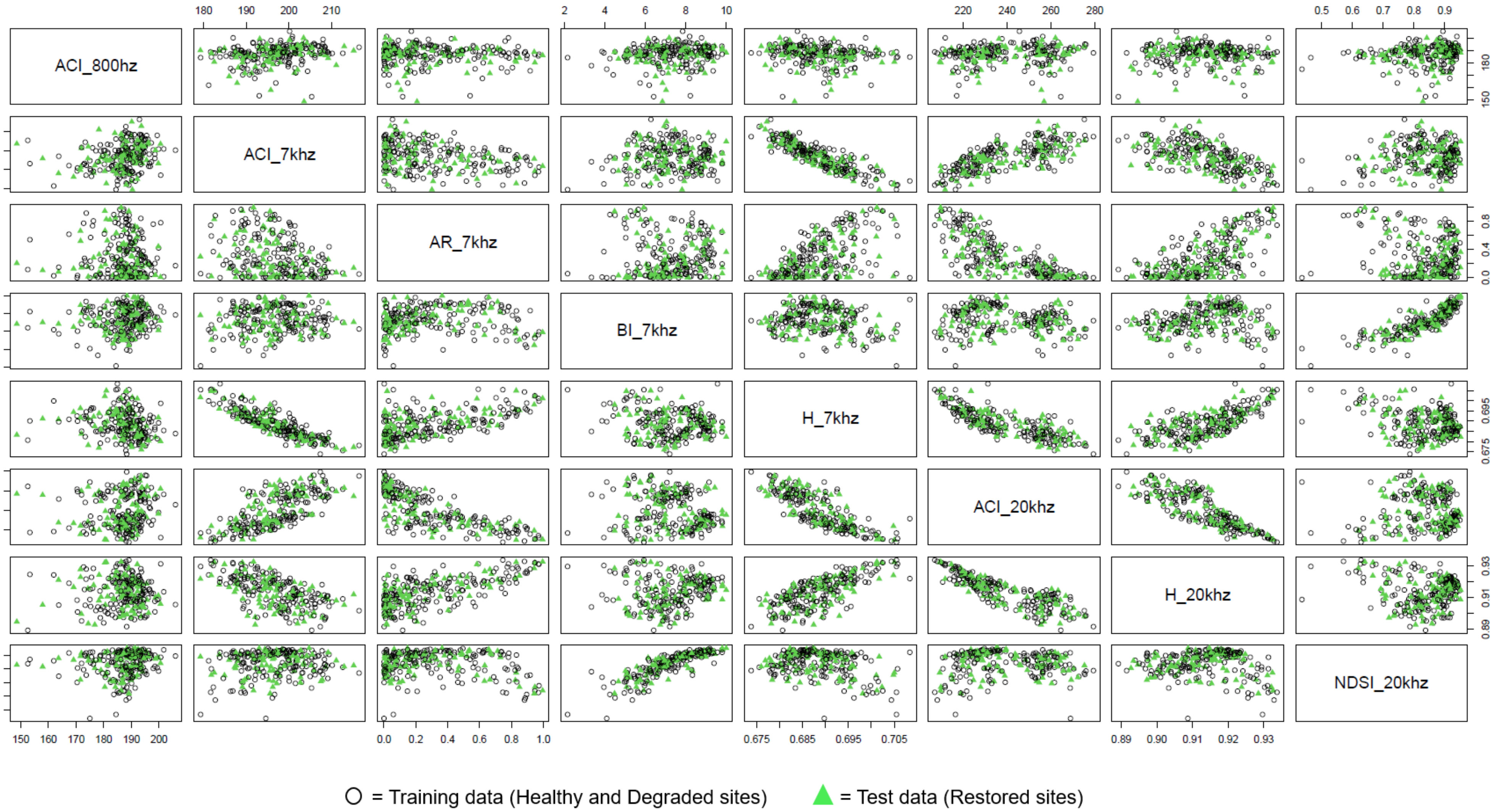
Modern investigations attempting to advance the field of PAM in the marine environment recognise that computationally obtained metrics from acoustic recordings may hold the potential to rapidly assess habitats. To date, existing studies have primarily assessed the utility of individual eco-acoustic indices to predict ecosystem attributes such as diversity or benthic cover. This investigation outlines a novel approach which combines multiple indices to deliver an improvement upon the predictive capacity of any single index, offering a potentially significant development in the field, with future investigations likely to benefit from adopting a similar approach. This investigation also revealed valuable insights to marine practitioners attempting to restore large areas of reef in the Spermonde Archipelago. The findings here reveal that acoustic recordings can be used to give additional evidence, alongside coral cover, to track the progress of restored sites, and that mature sites (>24 months old) appear to be converging with naturally healthy sites.

3.5 Supplementary information

Supplementary 3.3. Results from a Spearmans collinearity test between all 12 indices in each of three frequency bands: low (0.05 – 0.8 kHz), medium (2 – 7 kHz), full (0.05 – 20 kHz). Darker cells represent a stronger correlation. Blank cells indicate values = -0.01 – 0.01.

	ACI Low	ACI Medium	ACI Full	AEI Low	AEI Medium	AEI Full	AI Low	AI Medium	AI Full	AR Low	AR Medium	AR Full	BI Low	BI Medium	BI Full	H Low	H Medium	H Full	NP Low	NP Medium	NP Full	SE Low	SE Medium	SE Full	TE Low	TE Medium	TE Full	SPL Low	SPL Medium	SPL Full	SNAP Medium	SNAP Full
ACI Medium	0.35																															
ACI Full	0.24	0.65																														
AEI Low	-0.07	0.02	0.02																													
AEI Medium	-0.16	-0.23	0.14	-0.07																												
AEI Full	-0.04	-0.10	0.27	0.43	0.60																											
AI Low	-0.13	-0.16	-0.28	-0.11	0.13	-0.21																										
AI Medium	0.04	0.09	-0.46	0.14	-0.46	-0.34	0.03																									
AI Full	0.09	0.06	-0.50		-0.42	-0.43	0.05	0.89																								
AR Low	-0.21	-0.29	-0.43	-0.20	0.08	-0.32	0.92	0.07	0.13																							
AR Medium	-0.05	-0.15	-0.68	0.12	-0.50	-0.39	0.16	0.87	0.83	0.24																						
AR Full	-0.02	-0.22	-0.73		-0.35	-0.42	0.21	0.77	0.88	0.32	0.89																					
BI Low	-0.13	0.06	0.09	0.16	-0.07	0.09	-0.13	-0.11	-0.15	-0.19	-0.10	-0.13																				
BI Medium	0.10	0.17	0.10	-0.21		-0.20	0.29	-0.29	-0.16	0.28	-0.24	-0.14	0.02																			
BI Full	0.10	0.14	0.18	-0.06	-0.22		-0.24	-0.13	-0.24	-0.26	-0.16	-0.38	0.37																			
H Low	0.14	0.13	-0.02	-0.48	-0.17	-0.23	-0.22	0.02		-0.18		-0.07	-0.22	0.03	0.28																	
H Medium	-0.27	-0.83	-0.75	0.08		-0.02	0.08	0.31	0.27	0.21	0.49	0.48	-0.08	-0.43	-0.18	-0.13																
H Full	-0.16	-0.54	-0.80	0.01	0.04	-0.23	0.29	0.43	0.53	0.42	0.61	0.79	-0.17	-0.09	-0.56	-0.17	0.67															
NP Low	-0.23	0.07	0.20	0.07	0.03	0.16	-0.39	-0.13	-0.14	-0.42	-0.19	-0.21	0.01	-0.22	-0.02	0.09	-0.08	-0.18														
NP Medium	0.07	-0.13	0.05	0.02	0.15	0.20	-0.10	-0.02	-0.09	-0.14	-0.05	-0.07	-0.11	-0.27	0.03	-0.01	0.17	-0.01	-0.07													
NP Full	0.25	0.11	0.12	0.03	0.08	0.09	-0.13	0.18	0.19	-0.12	0.14	0.16	-0.25	-0.27	-0.19	0.08	-0.01	0.15	-0.02	0.18												
SE Low	0.25	0.25	0.13	-0.36	-0.20	-0.10	-0.41	-0.02	-0.05	-0.42	-0.08	-0.17	-0.12		0.35	0.95	-0.22	-0.31	0.16	0.03	0.08											
SE Medium	-0.08	-0.31	0.36	0.02	0.54	0.51	-0.22	-0.53	-0.58	-0.26	-0.59	-0.55	-0.04	-0.32	0.05	-0.09	0.22	-0.22	0.17	0.32	0.11	-0.04										
SE Full	0.04	-0.18	-0.12	0.03	0.36	0.06	0.10	0.03	0.14	0.13	0.06	0.31	-0.23	-0.08	-0.70	-0.24	0.21	0.65		0.04	0.42	-0.30	0.20									
TE Low	-0.38	-0.40	-0.48	-0.29	0.10	-0.38	0.63	0.11	0.18	0.80	0.26	0.35	-0.28	0.10	-0.28	-0.01	0.34	0.50	-0.23	-0.11	-0.02	-0.32	-0.15	0.24								
TE Medium	-0.19	-0.54	-0.91	0.06	-0.36	-0.36	0.22	0.63	0.63	0.36	0.84	0.81	-0.05	-0.18	-0.19	-0.05	0.76	0.76	-0.18	-0.06	-0.08	-0.17	-0.47	0.06	0.41							
TE Full	-0.20	-0.58	-0.91		-0.07	-0.29	0.31	0.50	0.57	0.45	0.70	0.83	-0.12	-0.08	-0.44	-0.12	0.72	0.97	-0.21	-0.03	0.03	-0.27	-0.33	0.44	0.51	0.87						
SPL Low	-0.08	0.10	-0.02	0.01	-0.06	-0.16	0.39	0.06	-0.01	0.27	0.08	-0.05	-0.18		0.07		-0.13	-0.16	-0.03	-0.06	-0.10	-0.04	-0.24	-0.30	0.15	0.05	-0.09					
SPL Medium	-0.02	0.49	-0.14	0.08	-0.51	-0.39	0.08	0.79	0.70	0.05	0.65	0.50	-0.06	-0.10	-0.02	0.09	-0.16	0.07	-0.07	-0.10	0.05	0.08	-0.68	-0.22	0.01	0.31	0.15	0.30				
SPL Full	0.05	0.43	-0.15	-0.09	-0.47	-0.45	0.06	0.66	0.75	0.06	0.55	0.54	-0.05	-0.06	0.01	0.14	-0.15	0.06	-0.06	-0.09	0.01	0.13	-0.60	-0.23	0.01	0.27	0.15	0.25	0.85			
NAP Medium	0.13	0.43	-0.18	0.11	-0.44	-0.32	-0.02	0.86	0.78	-0.05	0.68	0.58	-0.07	-0.16	-0.05	0.12	-0.06	0.19	-0.08	-0.09	0.19	0.12	-0.57	-0.02	-0.03	0.33	0.23	0.10	0.85	0.74		
SNAP Full	0.23	0.46	-0.12	-0.09	-0.41	-0.39	-0.02	0.71	0.80	-0.01	0.55	0.56	-0.11	-0.03	-0.03	0.13	-0.18	0.10	-0.07	-0.13	0.15	0.13	-0.60	-0.10	-0.04	0.24	0.15	0.05	0.77	0.84	0.84	
NDSI Full	0.09	0.22	0.22	-0.21	0.05	-0.11	0.25	-0.42	-0.32	0.21	-0.40	-0.33	0.10	0.87	0.12	0.04	-0.51	-0.30	-0.21	-0.22	-0.31	0.04	-0.24	-0.24	-0.01	-0.30	-0.27	0.11	-0.16	-0.10	-0.25	-0.12

Supplementary 3.4. Scatterplots between each of the eight indices selected for inclusion as features in the final model. Values from healthy and degraded sites alongside values from restored sites are included to enable divergence between these two groups to be observed if present.



Supplementary 3.5. Results from the Mann-Whitney U test between the index scores of healthy (n = 81) and degraded (n = 71) site recordings. Indices not calculated for specific frequency bands are left blank.

	AI		**AEI**		**ACI**		**AR**		**BI**		**H**	
	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value
Full	0.54253	0.015255	0.173825	0.436942	1.777053	1.91E-15	1.076803	1.47E-06	0.326178	0.144644	1.446746	9.80E-11
Medium	0.620978	0.005485	0.597857	0.007502	1.383988	6.04E-10	1.073501	1.58E-06	1.183326	1.21E-07	1.980192	8.32E-19
Low	0.037985	0.865108	0.242776	0.2776	0.911648	4.56E-05	0.346823	0.120893	0.090009	0.687293	0.195707	0.381449

	NDSI		**NP**		**SE**		**Snap Rate**		**SPL**		**TE**	
	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value	u-score	p-value
Full	1.39142	4.89E-10	0.307273	0.169389	0.320398	0.151897	0.109001	0.625926	0.309259	0.16665	1.622635	3.97E-13
Medium			0.303865	0.17417	0.225435	0.313371	0.090009	0.687293	0.047069	0.833277	1.631718	2.94E-13
Low			0.056349	0.801039	0.047069	0.833279			0.209745	0.34824	0.752275	0.000767

Chapter Four:

Discussion

Authors: Ben Williams[a], Stephen D. Simpson[a]

Affiliations: [a]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PS, United Kingdom

Author contributions: Written by B.W. with comments from S.D.S.

4.1 Advances needed in PAM

Passive acoustic monitoring (PAM) and soundscape ecology are receiving a growing level of attention as a means by which to study the marine environment. The attraction to these approaches primarily stems from the low effort required to collect large amounts of data. Recorders can be left to autonomously monitor habitats for long periods of time, offering a significant advantage over traditional in-water surveying methods which can be logistically challenging and labour intensive. PAM can collect data across broader temporal scales with ease, allowing this approach to better capture information on the state of marine habitats across the natural cycles present. It also reduces the requirement of the advanced training needed to conduct these surveys and eliminates complications that arise when comparing underwater visual census data collected using different techniques and survey teams.

The end goal of developing this field is to provide a valuable tool which is of use to the understanding and proper management of our marine environment. For example it could be used to assess habitat health or the progress of restoration, as demonstrated in Chapter Three. It could also be used to quantify the success of conservation approaches such as MPAs or other management practices (Bertucci *et al.*, 2016). Additoinaly, PAM has been found to be an indicator of the scale of disturbance events, such as bleaching and cyclones (Gordon *et al.*, 2019). It also offers a tool with which fisheries could be assessed (Lindseth and Lobel, 2018; Elise *et al.*, 2019).

However, two key limitations have been preventing PAM from achieving widespread uptake. These are: (i) the requirement of expensive high specification equipment, and (ii) the lack of rapid analytical approaches that can reliably be used to ascertain habitat attributes from recordings. This thesis contributes significant advancements towards overcoming both these obstacles.

4.2 Low cost recorders

Existing marine soundscape investigations have relied on high specification hydrophone recorders. This is likely in part due to the lack of alternatives and has subsequently become an established norm that has not been re-considered. However, lower specification recorders are frequently used with success in the terrestrial environment (Whytock and Christie, 2017; Sethi *et al.*, 2020). Chapter Two of this thesis therefore set out to identify whether alternative options do exist. We used consumer grade GoPro action cameras that can be deployed underwater. Soundscape recordings captured by these devices contain many identifiable biophonic, geophonic and anthropogenic sounds characteristic of the marine environment.

This investigation formally demonstrated that these devices are indeed capable of collecting recordings that exhibit many of the same properties as those collected by research grade hydrophones. It also showed that some computational indices, which are gaining significant traction due to the rapidity with which they can be calculated, can also be extracted from these recordings. Future study that is able to use these devices to generate ecological data of relevance to a broader investigation would help to further test the extent of their utility.

GoPros have several unique advantages, including their ability to simultaneously record video data, and their widespread possession by marine practitioners and potential citizen scientists. However, other acoustic recorders offer advantages specific to acoustics. One of these is the AudioMoth recorder recently developed for terrestrial habitats which we have begun to test in the marine setting (Hill *et al.*, 2018). This open source device with a case, batteries and SD card is 30 times cheaper than existing hydrophones and three times cheaper than GoPros. Importantly, it is also able to record for much greater durations than GoPros (up to nine days at 20 kHz sampling frequency) due to its improved battery and memory capacity (Hill *et al.*, 2019). These smart recorders also offer a duty cycling capability enabling the length of period over which they can sample to be greatly expanded. Further still, these devices can be programmed to automatically generate metrics from recordings in real time, including eco-acoustic indices. Importantly, this removes the two most computationally limiting steps in analysis currently: (i) the need to upload hundreds to thousands of hours of recordings to a computer, (ii) the need to calculate all indices in one stage. In the future, smart recorders such as these are likely to become the dominant technology amongst those entering the field of PAM to monitor their habitats. Further investigations that continue to explore these emerging technologies will help other researchers, marine practitioners and citizen scientists to begin capitalising on the benefits offered by PAM.

4.3 PAM data analysis with machine learning

The use of machine learning to analyse soundscape data is and emerging practice in terrestrial investigations (Buxton *et al.*, 2018; Eldridge *et al.*, 2018; Metcalf *et al.*, 2020). Some uses of machine learning have been made in the marine environment, primarily in attempts to produce identifiers for vocalisations of specific organisms (Noda *et al*., 2016; Lin *et al*., 2018; Roca and Van

Opzeeland, 2019), but no existing study has used this approach to ascertain other habitat related metrics from recordings. However, a few dozen published investigations have attempted to produce these outcomes using traditional methods that compare individual eco-acoustic metrics or use labour intensive manual listening to compare habitats (see reviews: Lindseth and Lobel, 2018; Mooney *et al.*, 2020, Pieretti and Danovaro, 2020). As demonstrated in Chapter Three, novel machine learning algorithms offer previously unobtainable levels of accuracy using rapidly obtained computational indices and intelligent algorithms.

Machine learning algorithms offer a new approach for analysis that future studies may wish to consider. This also offers the opportunity to revisit many existing datasets from previous studies that may yet reveal further insights from raw recording and index data. Using both new and existing datasets, the capacity of models to predict other habitat attributes should be further explored. This may include classifiers able to assign habitats to specified categories, as demonstrated in this thesis, or, regression based models which can make predictions along a gradient to better predict specific ecological attributes.

4.4 Combining advances from this thesis

Together, contributions from this thesis to the field of PAM in the marine environment have the potential to help overcome two of the biggest obstacles to the fields expansion. Low cost recorders will help expand the scale of data collection possible, and machine learning provides improved analytical techniques that can maximise ecological inferences that can be made.

However, low cost recorders make data collection easier, whereas advanced analytical processes such as machine learning can make this stage more complex. This puts the field is at risk of a data deluge. Marine practitioners and scientists not firmly rooted in soundscape ecology and machine learning may struggle to extract the full potential from the data they are collecting. Standardisation is therefore required.

As these datasets build, unsupervised models offer an opportunity to simplify the analytical process. These use algorithms such as convolutional neural networks that generate an extensive set of common features that can be used to discriminate between recordings. For example, Google's DeepMind team recently published work outlining a new approach which used a cross-convolutional nueral network (CNN) on 5.24 million hours of recordings from YouTube. This CNN generated a 128 strong feature set that best captures the variation across 30'871 specific video labels used (e.g speech, lawn mower, etc) (Hershey *et al.*, 2017). This feature set has subsequently been found to also have relevance in an ecological context. This was revealed by an investigation which calculated ran the full feature set on 2750 hours of terrestrial soundscape recordings (Sethi *et al.*, 2020). This enabled the investigators to find strong relationships between the feature set and accompanying ecological data such as temporal periods and habitat quality using random forest machine learning algorithms. They were also able to identify anamolous events such as logging or shotgun blasts used in poaching. This approach could just as easily be applied to existing marine soundscape datasets to test its utility here. Universal feature sets such as these could help standardise methods used in this field. Software or centralised systems which enable practitioners to upload their acoustic data and receive feature set values for recordings could then significantly advance the ease with which this data can be obtained. Further software or workflows which enables users to apply the corresponding labels to this data (e.g habitat quality, species diversity, etc) and then implement the most appropriate machine learning algorithms would further support soundscape ecologists with this kind of advanced analysis.

Furtherstill, smart recorders could be programmed to calculate these kind of acoustic features on the go, without keeping the computationally costly raw recordings, streamlining the process further. This numeric feature data of feature could be easily uploaded to the Web, unlike the terabytes of acoustic recordings that soundscape ecologists currently have to work with. This would allow online tools to perform the analysis once given the appropriate inputs and large collaborative datasets to be assembled with ease.

With continued progress, PAM in the marine environment may be able to realise this potential in the near future. These advances would help PAM become a valuable tool that could significantly improve the scale of ecological monitoring that can occur in important marine habitats, allowing feedback on management practices to be gathered.

Bibliography

Abu-Arafeh, A., Jordan, H., & Drummond, G. (2016). Reporting of method comparison studies: a review of advice, an assessment of current practice, and specific suggestions for future reports. *British Journal of Anaesthesia*, *117*, 569–575.

Agnieszka, O. (2017). Scientific ideas included in the concepts of bioacoustics, acoustic ecology, ecoacoustics, soundscape ecology, and vibroacoustics. *Archives of Acoustics*, *42*, 415–421.

Archer, S. K., Halliday, W. D., Riera, A., Mouy, X., Pine, M. K., Chu, J. W. F., … Juanes, F. (2018). First description of a glass sponge reef soundscape reveals fish calls and elevated sound pressure levels. *Marine Ecology Progress Series*, *595*, 245–252.

Ashihara, K. (2007). Hearing thresholds for pure tones above 16kHz. *The Journal of the Acoustical Society of America*, *122*, EL52–EL57.

Au, W. W. L., & Banks, K. (1998). The acoustics of the snapping shrimp *Synalpheus parneomeris* in Kaneohe Bay . *The Journal of the Acoustical Society of America*, *10*, 41–47.

Baldera, A., Hanson, D. A., & Kraft, B. (2018). Selecting indicators to monitor outcomes across projects and multiple restoration programs in the Gulf of Mexico. *Ecological Indicators*, *89*, 559–571.

Beason, R. D., Riesch, R., & Koricheva, J. (2019). AURITA: an affordable, autonomous recording device for acoustic monitoring of audible and ultrasonic frequencies. *Bioacoustics*, *28*, 381–396.

Becker, A., Taylor, M. D., Folpp, H., & Lowry, M. B. (2018). Managing the development of artificial reef systems: The need for quantitative goals. *Fish and Fisheries*, *19*, 740–752.

Bertucci, F., Guerra, A. S., Sturny, V., Blin, E., Sang, G. T., & Lecchini, D. (2020). A preliminary acoustic evaluation of three sites in the lagoon of Bora Bora, French Polynesia. *Environmental Biology of Fishes*, *103*, 1–12.

Bertucci, F., Parmentier, E., Lecellier, G., Hawkins, A. D., & Lecchini, D. (2016). Acoustic indices provide information on the status of coral reefs: an example from Moorea Island in the South Pacific. *Scientific Reports*, *6*, 33326.

Bilic-Zulle, L. (2011). Lessons in biostatistics Comparison of methods : Passing and Bablok regression. *Biochemia Medica*, *21*, 49–52.

Bittle, M., & Duncan, A. (2013). A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring. In *Proceedings of Acoustics* (Vol. 2013).

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*, 2137–2144.

Bohnenstiehl, Delwayne R., Lillis, A., & Eggleston, D. B. (2016). The curious acoustic behavior of estuarine snapping shrimp: Temporal patterns of snapping shrimp sound in sub-tidal oyster reef habitat. *PLoS ONE*, *11*, 1–21.

Bohnenstiehl, DelWayne R., Lyon, R. P., Caretti, O. N., Ricci, S. W., & Eggleston, D. B. (2018). Investigating the utility of ecoacoustic metrics in marine soundscapes. *Journal of*

Ecoacoustics, *2*, R1156L.

Bolgan, M., Amorim, M. C. P., Fonseca, P. J., Di Iorio, L., & Parmentier, E. (2018). Acoustic complexity of vocal fish communities: A field and controlled validation. *Scientific Reports*, *8*, 1–11.

Boström-Einarsson, L., Babcock, R. C., Bayraktarov, E., Ceccarelli, D., Cook, N., Ferse, S. C. A., … McLeod, I. M. (2020). Coral restoration – A systematic review of current methods, successes, failures and future directions. *PLoS ONE*, *15*, 1–24.

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S. G., & Dent, D. H. (2019). Guidelines for the use of acoustic indices in environmental research. *Methods in Ecology and Evolution*, *10*, 1796–1807.

Bridge, T. C. L., Hughes, T. P., Guinotte, J. M., & Bongaerts, P. (2013). Call to protect all coral reefs. *Nature Climate Change*, *3*, 528-530.

Buscaino, G., Ceraulo, M., Pieretti, N., Corrias, V., Farina, A., Filiciotto, F., … Giuseppe, A. (2016). Temporal patterns in the soundscape of the shallow waters of a Mediterranean marine protected area. *Scientific Reports*, *6*, 34230.

Butler, J., Stanley, J. A., & Butler, M. J. (2016). Underwater soundscapes in near-shore tropical habitats and the effects of environmental degradation and habitat restoration. *Journal of Experimental Marine Biology and Ecology*, *479*, 89–96.

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

Canada, A. R. C. of, Truax, B., & Project, W. S. (1978). *The world soundscape project's handbook for acoustic ecology*. ARC Publications: Aesthetic Research Centre: World Soundscape Project.

Carriço, R., Silva, M. A., Vieira, M., Afonso, P., Menezes, G. M., Fonseca, P. J., & Amorim, M. C. P. (2020). The Use of Soundscapes to Monitor Fish Communities: Meaningful Graphical Representations Differ with Acoustic Environment. *Acoustics*, *2*, 382–398.

Carstensen, B. (2011). *Comparing Clinical Measurement Methods: A Practical Guide - Bendix Carstensen*. John Wiley & Sons.

Casella, E., Collin, A., Harris, D., Ferse, S., Bejarano, S., Parravicini, V., … Rovere, A. (2017). Mapping coral reefs using consumer-grade drones and structure from motion photogrammetry techniques. *Coral Reefs*, *36*, 269–275.

Cazau, D., Bonnel, J., Jouma'a, J., Le Bras, Y., & Guinet, C. (2017). Measuring the marine soundscape of the Indian Ocean with southern elephant seals used as acoustic gliders of opportunity. *Journal of Atmospheric and Oceanic Technology*, *34*, 207–223.

Ceraulo, M, Papale, E., Caruso, F., Filiciotto, F., Grammauta, R., Parisi, I., … Buscaino, G. (2018). Acoustic comparison of a patchy Mediterranean shallow water seascape: *Posidonia oceanica* meadow and sandy bottom habitats. *Ecological Indicators*, *85*, 1030–1043.

Ceraulo, Maria, Sal Moyano, M. P., Bazterrica, M. C., Hidalgo, F. J., Papale, E., Grammauta, R., … Buscaino, G. (2020). Spatial and temporal variability of the soundscape in a Southwestern Atlantic coastal lagoon. *Hydrobiologia*, *847*, 1–23.

Chary, K. L., Sreekanth, G. B., Deshmukh, M. K., & Sharma, N. (2020). Marine soundscape and fish chorus in an archipelago ecosystem comprising bio-diverse tropical islands off Goa

Coast, India. *Aquatic Ecology*, *54*, 1–19.

Chen, W. (2020). Align Wave. MATLAB Central File Exchange.

Clark, C. W. (1990). Acoustic behavior of mysticete whales. In *Sensory abilities of cetaceans* (pp. 571–583). Springer.

Coquereau, L., Grall, J., Chauvaud, · Laurent, Gervaise, C., Clavier, J., Aurélie Jolivet, ·, & Iorio, L. Di. (2016). Sound production and associated behaviours of benthic invertebrates from a coastal habitat in the north-east Atlantic. *Marine Biology*, *3*, 127.

Corriero, G., Pierri, C., Mercurio, M., Marzano, C. N., Tarantini, S. O., Gravina, M. F., … Valenzano, E. (2019). A Mediterranean mesophotic coral reef built by non-symbiotic scleractinians. *Scientific Reports*, *9*, 1–17.

Cotter, A. J. R. (2008). The “soundscape” of the sea, underwater navigation, and why we should be listening more. *Advances in Fisheries Science*, *50*, 451–471.

Cousteau, J.-Y., & Dumas, F. (1953). The silent world. *New York and London*.

Curtis, K. R., Howe, B. M., & Mercer, J. A. (1999). Low-frequency ambient sound in the North Pacific: Long time series observations. *The Journal of the Acoustical Society of America*, *106*, 3189.

Darras, K., Kolbrek, B., Knorr, A., Meyer, V., & Zippert, M. (2018). Assembling cheap, high-performance microphones for recording terrestrial wildlife: the Sonitor system. *F1000Research*, *7*.

Davies, B. F. R., Attrill, M. J., Holmes, L., Rees, A., Witt, M. J., & Sheehan, E. V. (2020). Acoustic Complexity Index to assess benthic biodiversity of a partially protected area in the southwest of the UK. *Ecological Indicators*, *111*, 106019.

de Soto, N. A. (2016). Peer-reviewed studies on the effects of anthropogenic noise on marine invertebrates: from scallop larvae to giant squid. *The Effects of Noise on Aquatic Life II* (pp. 17–26). Springer.

Deakos, M. H. (2002). Humpback whale (*Megaptera novaeangliae*) communication: The context and potential functions of pec-slapping behavior on the Hawaiian wintering grounds. Thesis, University of Hawaii at Manoa.

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

Desiderà, E., Guidetti, P., Panzalis, P., Navone, A., Valentini-Poirrier, C. A., Boissery, P., … Iorio, L. Di. (2019). Acoustic fish communities: Sound diversity of rocky habitats reflects fish species diversity. *Marine Ecology Progress Series*, *608*, 183–197.

Dietzel, A., Bode, M., Connolly, S. R., & Hughes, T. P. (2020). Long-term shifts in the colony size structure of coral populations along the Great Barrier Reef. *Proceedings of the Royal Society*, *287*, 20201432.

Dornelas, M., Madin, E. M. P., Bunce, M., DiBattista, J. D., Johnson, M., Madin, J. S., … Bates, A. E. (2019). Towards a macroscope: leveraging technology to transform the breadth, scale and resolution of macroecological data. *Global Ecology and Biogeography*, *28*, 1937–1948.

Downs, C. A., Woodley, C. M., Richmond, R. H., Lanning, L. L., & Owen, R. (2005). Shifting the paradigm of coral-reef ‘health’ assessment. *Marine Pollution Bulletin*, *51*, 486–494.

Duarte, C M, Agusti, S., Barbier, E., L, B. G., Castilla, J. C., Gattuso, J., … Worm, B. (2020). Rebuilding marine life. *Nature*, *580*, 39–51.

Duarte, Carlos M., Chapuis, L., Collin, S. P., Costa, D. P., Devassy, R. P., Eguiluz, V. M., … Juanes, F. (2021). The soundscape of the Anthropocene ocean. *Science*, *371,* 6529.

Dziak, R. P., Bohnenstiehl, D. R., Stafford, K. M., Matsumoto, H., Park, M., Lee, W. S., … Mellinger, D. K. (2015). Sources and levels of ambient ocean sound near the Antarctic Peninsula. *PLoS One*, *10*, e0123425.

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., & Peck, M. (2018). Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. *Ecological Indicators*, *95*, 939–952.

Elise, S., Bailly, A., Urbina-Barreto, I., Mou-Tham, G., Chiroleu, F., Vigliola, L., … Bruggemann, J. H. (2019). An optimised passive acoustic sampling scheme to discriminate among coral reefs’ ecological states. *Ecological Indicators*, *107*, 105627.

Elise, S., Urbina-Barreto, I., Pinel, R., Mahamadaly, V., Bureau, S., Penin, L., … Bruggemann, J. H. (2019). Assessing key ecosystem functions through soundscapes: A new perspective from coral reefs. *Ecological Indicators*, *107*, 105623.

Erbe, C., Verma, A., McCauley, R., Gavrilov, A., & Parnum, I. (2015). The marine soundscape of the Perth Canyon. *Progress in Oceanography*, *137*, 38–51.

Farina, A. (2013). *Soundscape ecology: principles, patterns, methods and applications*. Springer.

Farina, A., & Pieretti, N. (2012). The soundscape ecology: A new frontier of landscape research and its application to islands and coastal systems. *Journal of Marine and Island Cultures*, *1*, 21–26.

Ferrigno, F., Bianchi, C. N., Lasagna, R., Morri, C., Russo, G. F., & Sandulli, R. (2016). Corals in high diversity reefs resist human impact. *Ecological Indicators*, *70*, 106–113.

Florisson, J. H., Tweedley, J. R., Walker, T. H. E., & Chaplin, J. A. (2018). Reef vision: A citizen science program for monitoring the fish faunas of artificial reefs. *Fisheries Research*, *206*, 296–308.

Ford, A. K., Eich, A., McAndrews, R. S., Mangubhai, S., Nugues, M. M., Bejarano, S., … Ferse, S. C. A. (2018). Evaluation of coral reef management effectiveness using conventional versus resilience-based metrics. *Ecological Indicators*, *85*, 308–317.

Freeman, L. A., & Freeman, S. E. S. (2016). Rapidly obtained ecosystem indicators from coral reef soundscapes. *Marine Ecology Progress Series*, *561*, 69–82.

Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American Statistical Association*, *84*, 165–175.

Gattuso, J. P., Hoegh-Guldberg, O., & Pörtner, H. O. (2014). Cross-chapter box on coral reefs. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change* (pp. 97–100). Cambridge University Press.

Gedamke, J., & Robinson, S. M. (2010). Acoustic survey for marine mammal occurrence and distribution off East Antarctica (30-80 E) in January-February 2006. *Deep Sea Research Part II: Topical Studies in Oceanography*, *57*, 968–981.

Geyer, F., Sagen, H., Hope, G., Babiker, M., & Worcester, P. F. (2016). Identification and quantification of soundscape components in the Marginal Ice Zone. *The Journal of the Acoustical Society of America*, *139*, 1873–1885.

Gordon, T. A. C., Harding, H. R., Wong, K. E., Merchant, N. D., Meekan, M. G., McCormick, M. I., … Simpson, S. D. (2018). Habitat degradation negatively affects auditory settlement behavior of coral reef fishes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 5193–5198.

Gordon, T. A. C., Radford, A. N., Davidson, I. K., Barnes, K., McCloskey, K., Nedelec, S. L., … Simpson, S. D. (2019). Acoustic enrichment can enhance fish community development on degraded coral reef habitat. *Nature Communications*, *10*, 1–7.

Harris, S. A., Shears, N. T., & Radford, C. A. (2016). Ecoacoustic indices as proxies for biodiversity on temperate reefs. *Methods in Ecology and Evolution*, *7*, 713–724.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Haver, S. M., Gedamke, J., Hatch, L. T., Dziak, R. P., Van Parijs, S., McKenna, M. F., … Hanson, B. (2018). Monitoring long-term soundscape trends in US waters: The NOAA/NPS ocean noise reference station network. *Marine Policy*, *90*, 6–13.

Hershey, S., Chaudhuri, S., Ellis, D. P. W., Gemmeke, J. F., Jansen, A., Moore, R. C., … Wilson, K. (2017). CNN architectures for large-scale audio classification. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (pp. 131–135). Institute of Electrical and Electronics Engineers Inc.

Hicks, C. C., Graham, N. A. J., & Cinner, J. E. (2013). Synergies and tradeoffs in how managers, scientists, and fishers value coral reef ecosystem services. *Global Environmental Change*, *23*, 1444–1453.

Hill, A. P., Prince, P., Piña Covarrubias, E., Doncaster, C. P., Snaddon, J. L., & Rogers, A. (2018). AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, *9*, 1199–1211.

Hill, A. P., Prince, P., Snaddon, J. L., Doncaster, C. P., & Rogers, A. (2019). AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. *HardwareX*, *6*, e00073.

Hillary K Grigonis. (2018). How Popular is GoPro? Company Has Sold More Than 30 Million Cameras | Digital Trends. Retrieved 7 November 2020, from https://www.digitaltrends.com/photography/gopro-sells-30-million-heros/

Hoegh-Guldberg, O., Poloczanska, E. S., Skirving, W., & Dove, S. (2017). Coral reef ecosystems under climate change and ocean acidification. *Frontiers in Marine Science*, *4*, 158.

Holmlund, C. M., & Hammer, M. (1999). Ecosystem services generated by fish populations. *Ecological Economics*, *29*, 253–268.

Holt, D. E., & Johnston, C. E. (2011). Can you hear the dinner bell? Response of cyprinid fishes to environmental acoustic cues. *Animal Behaviour*, *82*, 529–534.

Hughes, T.P., Barnes, M. L., Bellwood, D. R., Cinner, J. E., Cumming, G. S., Jackson, J. B. C., … Scheffer, M. (2017). Coral reefs in the Anthropocene. *Nature*, *546*, 82–90.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, *52*, 577–586.

Insley, S. J., Halliday, W. D., & de Jong, T. (2017). Seasonal patterns in ocean ambient noise near Sachs Harbour, Northwest Territories. *Arctic*, 239–248.

Balensiefer, M., Rossi, R., Ardinghi, N., Cenni, M., & Ugolini, M. (2004). SER international primer on ecological restoration. *Society for Ecological Restoration International Science & Policy Working Group*, *2*.

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

Joseph, J. E., & Margolina, T. (2014). Quantifying the ocean soundscape at a very busy southern California location. *The Journal of the Acoustical Society of America*, *135*, 2368.

Kahle, D., & Wickham, H. (2013). ggmap: Spatial visualization with ggplot2. *The R Journal*, *5*.

Kaplan, M. B., Lammers, M. O., Zang, E., & Aran Mooney, T. (2018). Acoustic and biological trends on coral reefs off Maui, Hawaii. *Coral Reefs*, *37*, 121–133.

Kaplan, M. B., Mooney, T. A., Partan, J., & Solow, A. R. (2015). Coral reef species assemblages are associated with ambient soundscapes. *Marine Ecology Progress Series*, *533*, 93–107.

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.

Kendrick, P., Waddington, D., Lopez, L., & Young, R. (2016). Assessing the robustness of soundscape complexity indices. *International Congress on Sound and Vibration, 23*, 1-8.

Kennedy, E. V, Holderied, M. W., Mair, J. M., Guzman, H. M., & Simpson, S. D. (2010). Spatial patterns in reef-generated noise relate to habitats and communities: evidence from a Panamanian case study. *Journal of Experimental Marine Biology and Ecology*, *395*, 85–92.

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

Krause, B. (2008). Anatomy of the soundscape: evolving perspectives. *Journal of the Audio Engineering Society*, *56*, 73–80.

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, 1247.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.

Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press.

Kull, R. C. (2006). Natural and urban soundscapes: the need for a multi-disciplinary approach. *Acta Acustica United with Acustica*, *92*, 898–902.

Lachenbruch, P. A., & Goldstein, M. (1979). Discriminant analysis. *Biometrics*, *45*, 69–85.

Lammers, M. O., & Munger, L. M. (2016). From Shrimp to Whales: Biological Applications of Passive Acoustic Monitoring on a Remote Pacific Coral Reef (pp. 61–81). Springer, New York, NY.

Larsson, M. (2012). Incidental sounds of locomotion in animal cognition. *Animal Cognition*, *15*, 1–13.

Lau, S. T., Kwok, K. W., Chan, H. L. W., & Choy, C. L. (2002). Piezoelectric composite hydrophone array. *Sensors and Actuators A: Physical*, *96*, 14–20.

Lefcheck, J. S., Innes-Gold, A. A., Brandl, S. J., Steneck, R. S., Torres, R. E., & Rasher, D. B. (2019). Tropical fish diversity enhances coral reef functioning across multiple scales. *Science Advances*, *5*, eaav6420.

Letessier, T. B., Juhel, J.-B., Vigliola, L., & Meeuwig, J. J. (2015). Low-cost small action cameras in stereo generates accurate underwater measurements of fish. *Journal of Experimental Marine Biology and Ecology*, *466*, 120–126.

Lieberman, P. (1968). Primate Vocalizations and Human Linguistic Ability. *The Journal of the Acoustical Society of America*, *44*, 1574.

Lillis, A., Apprill, A., Suca, J. J., Becker, C., Llopiz, J. K., & Mooney, T. A. (2018). Soundscapes influence the settlement of the common Caribbean coral *Porites astreoides* irrespective of light conditions. *Royal Society Open Science*, *5*, 181358.

Lillis, A., Bohnenstiehl, D. R., & Eggleston, D. B. (2015). Soundscape manipulation enhances larval recruitment of a reef-building mollusk. *PeerJ*, *3*, e999.

Lillis, A., Bohnenstiehl, D. W., Peters, J. W., & Eggleston, D. (2016). Variation in habitat soundscape characteristics influences settlement of a reef-building coral. *PeerJ*, *4*, 1–16.

Lillis, A., Eggleston, D. B., & Bohnenstiehl, D. R. (2014). Estuarine soundscapes: distinct acoustic characteristics of oyster reefs compared to soft-bottom habitats. *Marine Ecology Progress Series*, *505*, 1–17.

Lillis, A., & Mooney, T. A. (2018). Snapping shrimp sound production patterns on Caribbean coral reefs: relationships with celestial cycles and environmental variables. *Coral Reefs*, *37*, 597–607.

Lin, T.H., Tsao, Y., & Akamatsu, T. (2018). Comparison of passive acoustic soniferous fish monitoring with supervised and unsupervised approaches. *The Journal of the Acoustical Society of America*, *143*, EL278–EL284.

Lindseth, A. V., & Lobel, P. S. (2018). Underwater soundscape monitoring and fish bioacoustics: A review. *Fishes*, *3*, 36.

Lobel, P. S., Kaatz, I. M., & Rice, A. N. (2010). Acoustical behavior of coral reef fishes. Reproduction and sexuality in marine fishes: evolutionary patterns and innovations. C. KS Berkley. CA, University of California Press.

Lombardi, A. R., Hay, A. E., & Barclay, D. R. (2016). Soundscape characterization in a dynamic acoustic environment: Grand Passage, Nova Scotia, a planned in-stream tidal energy site. *Citation: Proceedings of Meetings on Acoustics, 27*, 5001.

Ludbrook, J. (1997). Special article comparing methods of measurement. *Clinical and Experimental Pharmacology and Physiology*, *24.*

Lurton, X. (2002). *An introduction to underwater acoustics: principles and applications*. Springer Science & Business Media.

Lyon, R.P., Eggleston, D. B., Bohnenstiehl, D. R., Layman, C. A., Ricci, S. W., & Allgeier, J. E. (2019). Fish community structure, habitat complexity, and soundscape characteristics of patch reefs in a tropical, back-reef system. *Marine Ecology Progress Series*, *609*, 33–48.

Magari, R. T. (2002). Statistics for laboratory method comparison studies. *BioPharm*, *15*, 28–32.

Magris, R. A., Treml, E. A., Pressey, R. L., & Weeks, R. (2016). Integrating multiple species connectivity and habitat quality into conservation planning for coral reefs. *Ecography*, *39*, 649–664.

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*, *23*, 945–959.

Matthews, J. N., Rendell, L. E., Gordon, J. C. D., & Macdonald, D. W. (1999). A review of frequency and time parameters of cetacean tonal calls. *Bioacoustics*, *10*, 47-71.

McWilliam, J. N., & Hawkins, A. D. (2013). A comparison of inshore marine soundscapes. *Journal of Experimental Marine Biology and Ecology*, *446*, 166–176.

McWilliam, J. N., McCauley, R. D., Erbe, C., & Parsons, M. J. G. (2017). Patterns of biophonic periodicity on coral reefs in the Great Barrier Reef. *Scientific Reports*, *7*, 1–13.

McWilliam, J. N., McCauley, R. D., Erbe, C., & Parsons, M. J. G. G. (2018). Soundscape diversity in the Great Barrier Reef: Lizard Island, a case study. *Bioacoustics*, *27*, 295–311.

Medwin, H. (2005). *Sounds in the sea: From ocean acoustics to acoustical oceanography*. Cambridge University Press.

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*, 257–265.

Metcalf, O. C., Barlow, J., Devenish, C., Marsden, S., Berenguer, E., & Lees, A. C. (2020). Acoustic indices perform better when applied at ecologically meaningful time and frequency scales. *Methods in Ecology and Evolution*, *12*, 421-431.

Mooney, T. A., Di Iorio, L., Lammers, M., Lin, T.-H., Nedelec, S. L., Parsons, M., … Stanley, J. (2020). Listening forward: approaching marine biodiversity assessments using acoustic methods. *Royal Society Open Science*, *8*, 201287.

Morgan, K. M., Perry, C. T., Johnson, J. A., & Smithers, S. G. (2017). Nearshore turbid-zone corals exhibit high bleaching tolerance on the Great Barrier Reef following the 2016 ocean warming event. *Frontiers in Marine Science*, *4*, 224.

Nedelec, S. L., Simpson, S. D., Holderied, M., Radford, A. N., Lecellier, G., Radford, C., & Lecchini, D. (2015). Soundscapes and living communities in coral reefs: Temporal and spatial variation. *Marine Ecology Progress Series*, *524*, 125–135.

Noda, J. J., Travieso, C. M., & Sánchez-Rodríguez, D. (2016). Automatic taxonomic classification of fish based on their acoustic signals. *Applied Sciences*, *12*, 443.

Obura, D. O., Aeby, G., Amornthammarong, N., Appeltans, W., Bax, N., Bishop, J., … Wongbusarakum, S. (2019). Coral reef monitoring, reef assessment technologies, and ecosystem-based management. *Frontiers in Marine Science*, *6*, 443.

Otrachshenko, V., & Bosello, F. (2017). Fishing for answers? Impacts of marine ecosystem quality on coastal tourism demand. *Tourism Economics*, *23*, 963–980.

Park , J.Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019). Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Science*, *365*, 284–288.

Parmentier, E., Berten, L., Rigo, P., Aubrun, F., Nedelec, S. L., Simpson, S. D., & Lecchini, D. (2015). The influence of various reef sounds on coral-fish larvae behaviour. *Journal of Fish Biology*, *86*, 1507–1518.

Parsons, M., Erbe, C., McCauley, R., McWilliam, J., Marley, S., Gavrilov, A., & Parnum, I. (2016). Long-term monitoring of soundscapes and deciphering a usable index: Examples of fish choruses from Australia. *Proceedings of Meetings on Acoustics*, *27*, 010023.

Parsons, M. J. G., Salgado Kent, C. P., Recalde-Salas, A., & McCauley, R. D. (2016). Fish choruses off Port Hedland, Western Australia. *Bioacoustics*, *26*, 135–152.

Pascal, N., Allenbach, M., Brathwaite, A., Burke, L., Le Port, G., & Clua, E. (2016). Economic valuation of coral reef ecosystem service of coastal protection: A pragmatic approach. *Ecosystem Services*, *21*, 72–80.

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*, 1513–1522.

Piercy, J. J. B., Codling, E. A., Hill, A. J., Smith, D. J., & Simpson, S. D. (2014). Habitat quality affects sound production and likely distance of detection on coral reefs. *Marine Ecology Progress Series*, *516*, 35–47.

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

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

Pieretti, N., Martire, M., Farina, A., & Danovaro, R. (2017). Marine soundscape as an additional biodiversity monitoring tool: A case study from the Adriatic Sea (Mediterranean Sea). *Ecological Indicators*, *83*, 13–20.

Pieretti, Nadia, & Danovaro, R. (2020). Acoustic indexes for marine biodiversity trends and ecosystem health. *Philosophical Transactions of the Royal Society*, *375*, 20190447.

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*, 1213–1232.

Plaisance, L., Caley, M. J., Brainard, R. E., & Knowlton, N. (2011). The Diversity of Coral Reefs: What Are We Missing? *PLoS ONE*, *6*, e25026.

Potamitis, I. (2015). Unsupervised dictionary extraction of bird vocalisations and new tools on assessing and visualising bird activity. *Ecological Informatics*, *26*, 6–17.

Pratchett, M. S., Hoey, A. S., & Wilson, S. K. (2014). Reef degradation and the loss of critical ecosystem goods and services provided by coral reef fishes. *Environmental Sustainability*. *7*, 37-43.

Putland, R. L., Constantine, R., & Radford, C. A. (2017). Exploring spatial and temporal trends in the soundscape of an ecologically significant embayment. *Scientific Reports*, *7*, 1–12.

Radford, C. A., Stanley, J. A., & Jeffs, A. G. (2014). Adjacent coral reef habitats produce different underwater sound signatures. *Marine Ecology Progress Series*, *505*, 19–28.

Radford, C.A., Jeffs, A. G., Tindle, C. T., & Montgomery, J. C. (2008). Temporal patterns in ambient noise of biological origin from a shallow water temperate reef. *Oecologia*, *156*, 921–929.

Rao, R. B., Fung, G., & Rosales, R. (2008). On the dangers of cross-validation. An experimental evaluation. *Proceedings of the 2008 SIAM international conference on data mining* (pp. 588–596).

Raoult, V., David, P. A., Dupont, S. F., Mathewson, C. P., O'Neill, S. J., Powell, N. N., & Williamson, J. E. (2016). GoPros^{TM} as an underwater photogrammetry tool for citizen science. *PeerJ*, *4*, e1960.

Rasher, D. B., Hoey, A. S., & Hay, M. E. (2017). Cascading predator effects in a Fijian coral reef ecosystem. *Scientific Reports*, *7*, 1–10.

Reeves Ozanich, E., Gerstoft, P., Toole, C., Freeman, L., Freeman, S., & Johnson, A. (2019). Detecting shifts in coral reef soundscape with unsupervised learning. *The Journal of the Acoustical Society of America*, *146*, 2885.

Ricci, S. W., Eggleston, D. B., Bohnenstiehl, D. R., & Lillis, A. (2016). Temporal soundscape patterns and processes in an estuarine reserve. *Marine Ecology Progress Series*, *550*, 25–38.

Rice, A. N., Soldevilla, M. S., & Quinlan, J. A. (2017). Nocturnal patterns in fish chorusing off the coasts of Georgia and eastern Florida. *Bulletin of Marine Science*, *93*, 455–474.

Robinson, S P, Lepper, P. A., Hazelwood, R. A., & National Physics Laboratory. (2014). Good Practice Guide. *Underwater Noise Measurement*, *133*, 1–97.

Roca, I. T., & Van Opzeeland, I. (2019). Using acoustic metrics to characterize underwater acoustic biodiversity in the Southern Ocean. *Remote Sensing in Ecology and Conservation,* 6¸262-273.

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.

Rogers, P. H., & Cox, M. (1988). Underwater Sound as a Biological Stimulus. In *Sensory Biology of Aquatic Animals* (pp. 131–149). Springer New York.

Roland, A., Wage, K. E., & Parsons, E. C. M. (2017). The marine soundscape off the Isle of Mull in Scotland's Inner Hebrides. *The Journal of the Acoustical Society of America*, *142*, 2685.

Rossi, T., Connell, S. D., & Nagelkerken, I. (2017). The sounds of silence: Regime shifts impoverish marine soundscapes. *Landscape Ecology*, *32*, 239–248.

Rossi, T., Nagelkerken, I., Pistevos, J. C. A., & Connell, S. D. (2016). Lost at sea: ocean acidification undermines larval fish orientation via altered hearing and marine soundscape modification. *Biology Letters*, *12*, 20150937.

Sale, P. F., Agardy, T., Ainsworth, C. H., Feist, B. E., Bell, J. D., Christie, P., … Sheppard, C. R. C. (2014). Transforming management of tropical coastal seas to cope with challenges of the 21st century. *Marine Pollution Bulletin*, *85*, 8–23.

Seaman, W. (2007). Artificial habitats and the restoration of degraded marine ecosystems and fisheries. In *Biodiversity in Enclosed Seas and Artificial Marine Habitats* (pp. 143–155). Springer.

Seraphim, M., Alexander, M., Janetski, N., Snellgrove, D., Jompa, J., Rappe, R., … Sloman, K. (2018). Coral reef restoration: interactions between fish communities and artificial reefs. A poster for: *STEM for Britain*. Westminster.

Sethi, S. S., Jones, N. S., Fulcher, B. D., Picinali, L., Clink, D. J., Klinck, H., … Ewers, R. M. (2020). Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 17049–17055.

Shlesinger, T., & Loya, Y. (2019). Breakdown in spawning synchrony: A silent threat to coral persistence. *Science*, *365*, 1002–1007.

Siddagangaiah, S., Chen, C.-F., Hu, W.-C., & Pieretti, N. (2019). A Complexity-Entropy Based Approach for the Detection of Fish Choruses. *Entropy*, *21*, 977.

Simpson, Stephen D., Meekan, M., Montgomery, J., McCauley, R., & Jeffs, A. (2005). Homeward Sound. *Science*, *308*, 221.

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M., Harvey, B. P., Straub, S. C., … Donat, M. G. (2019). Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nature Climate Change*, *9*, 306–312.

Smith, J. E., Brainard, R., Carter, A., Grillo, S., Edwards, C., Harris, J., … Sandin, S. (2016). Re-evaluating the health of coral reef communities: baselines and evidence for human impacts across the central Pacific. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20151985.

Smith, T. B., Nemeth, R. S., Blondeau, J., Calnan, J. M., Kadison, E., & Herzlieb, S. (2008). Assessing coral reef health across onshore to offshore stress gradients in the US Virgin Islands. *Marine Pollution Bulletin*, *56*, 1983–1991.

Sousa-Lima, R. S., Norris, T. F., Oswald, J. N., & Fernandes, D. P. (2013). A Review and Inventory of Fixed Autonomous Recorders for Passive Acoustic Monitoring of Marine Mammals. *Aquatic Mammals*, *39*, 23-53.

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*, 177–179.

Staaterman, E., Rice, A. N., Mann, D. A., & Paris, C. B. (2013). Soundscapes from a Tropical Eastern Pacific reef and a Caribbean Sea reef. *Coral Reefs*, *32*, 553–557.

Staaterman, Erica, Paris, C. B., DeFerrari, H. A., Mann, D. A., Rice, A. N., & D'Alessandro, E. K. (2014). Celestial patterns in marine soundscapes. *Marine Ecology Progress Series*, *508*, 17–32.

Staaterman, Erica, Ogburn, M. B., Altieri, A. H., Brandl, S. J., Whippo, R., Seemann, J., … Duffy, J. E. (2017). Bioacoustic measurements complement visual biodiversity surveys: Preliminary evidence from four shallow marine habitats. *Marine Ecology Progress Series*, *575*, 207–215.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*, 111–133.

Sueur, Jérôme. (2018a). Indices for Ecoacoustics. In *Sound Analysis and Synthesis with R* (pp. 479–519). Springer, Cham.

Sueur, Jérôme. (2018b). What Is Sound? In *Sound Analysis and Synthesis with R* (pp. 7–36). Springer, Cham.

Sueur, Jerome, Aubin, T., & Simonis, C. (2008). Equipment review: Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics*, *18*, 213–226.

Sueur, Jérôme, & Farina, A. (2015). Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. *Biosemiotics*, *8*, 493–502.

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

Sueur, Jérôme, Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, *3*, 1–11. Taffé, P. (2018). Effective plots to assess bias and precision in method comparison studies. *Statistical Methods in Medical Research*, *27*, 1650–1660.

Team, R. C. (2019). R: a language and environment for statistical computing. R Foundation for Statistical Computing.

Terrado, M., Sabater, S., Chaplin-Kramer, B., Mandle, L., Ziv, G., & Acuña, V. (2016). Model development for the assessment of terrestrial and aquatic habitat quality in conservation planning. *Science of the Total Environment*, *540*, 63–70.

Tricas, T. C., & Boyle, K. (2009). Validated reef fish sound scans of passive acoustic monitors on Hawaiian coral reefs. *The Journal of the Acoustical Society of America*, *125*, 2589.

Tricas, T. C. T., & Boyle, K. S. K. (2014). Acoustic behaviors in Hawaiian coral reef fish communities. *Marine Ecology Progress Series*, *511*, 1–16.

Veirs, S., Veirs, V., Williams, R., Jasny, M., & Wood, J. (2018). A key to quieter seas: half of ship noise comes from 15% of the fleet. *PeerJ Preprints*, *6*, e26525v1.

Versluis, M., Schmitz, B., Von der Heydt, A., & Lohse, D. (2000). How snapping shrimp snap: Through cavitating bubbles. *Science*, *289*, 2114–2117.

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

Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., & Villéger, S. (2018). A deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecological Informatics*, *48*, 238–244.

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

Williams, R., Wright, A. J., Ashe, E., Blight, L. K., Bruintjes, R., Canessa, R., … Erbe, C. (2015). Impacts of anthropogenic noise on marine life: Publication patterns, new discoveries, and future directions in research and management. *Ocean & Coastal Management*, *115*, 17–24.

Williams, S. L., Sur, C., Janetski, N., Hollarsmith, J. A., Rapi, S., Barron, L., … Mars, F. (2019). Large-scale coral reef rehabilitation after blast fishing in Indonesia. *Restoration Ecology*, *27*, 447–456.

Wright, A. J., Soto, N. A., Baldwin, A. L., Bateson, M., Beale, C. M., Clark, C., … Godinho, A. (2007). Anthropogenic noise as a stressor in animals: a multidisciplinary perspective. *International Journal of Comparative Psychology*, *20*.

Wu, W., Mallet, Y., Walczak, B., Penninckx, W., Massart, D. L., Heuerding, S., & Erni, F. (1996). Comparison of regularized discriminant analysis, linear discriminant analysis and quadratic discriminant analysis, applied to NIR data. *Analytica Chimica Acta*, *329*, 257–265.

Young, G. C., Dey, S., Rogers, A. D., & Exton, D. (2017). Cost and time-effective method for multi-scale measures of rugosity, fractal dimension, and vector dispersion from coral reef 3D models. *PloS One*, *12*, e0175341.