Original Articles

Enhancing automated analysis of marine soundscapes using ecoacoustic indices and machine learning

Ben Williams [a,b,*], Timothy A.C. Lamont [a,c], Lucille Chapuis [a], Harry R. Harding [d], Eleanor B. May [a], Mochyudho E. Prasetya [e], Marie J. Seraphim [f], Jamaluddin Jompa [g], David J. Smith [h,i], Noel Janetski [e], Andrew N. Radford [d], Stephen D. Simpson [a,d]

[a] College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4PS, UK
[b] Centre for Biodiversity and Environment Research, University College London, UK
[c] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[d] School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK
[e] MARS Sustainable Solutions, Makassar, Indonesia
[f] School of Health and Life Sciences, University of the West of Scotland, PA1 2BE, UK
[g] Graduate School, Hasanuddin University, 90245 Makassar, Indonesia
[h] Coral Reef Research Unit, School of Life Sciences, University of Essex, Colchester, Essex CO3 4SQ, UK
[i] Mars Incorporated, 4 Kingdom Street, Paddington, London W2 6BD UK

ARTICLE INFO

ABSTRACT

Historically, ecological monitoring of marine habitats has primarily relied on labour-intensive, non-automated survey methods. The field of passive acoustic monitoring (PAM) has demonstrated the potential of this practice to automate surveying in marine habitats. This has primarily been through the use of 'ecoacoustic indices' to quantify attributes from natural soundscapes. However, investigations using individual indices have had mixed success. Using PAM recordings collected at one of the world's largest coral reef restoration programmes, we instead apply a machine-learning approach across a suite of ecoacoustic indices to improve predictive power of ecosystem health. Healthy and degraded reef sites were identified through live coral cover surveys, with 90–95% and 0–20% cover respectively. A library of one-minute recordings were extracted from each. Twelve ecoacoustic indices were calculated for each recording, in up to three different frequency bandwidths (low: 0.05–0.8 kHz, medium: 2–7 kHz and broad: 0.05–20 kHz). Twelve of these 33 index-frequency combinations differed significantly between healthy and degraded habitats. However, the best performing single index could only correctly classify 47% of recordings, requiring extensive sampling from each site to be useful. We therefore trained a regularised discriminant analysis machine-learning algorithm to discriminate between healthy and degraded sites using an optimised combination of ecoacoustic indices. This multi-index approach discriminated between these two habitat classes with improved accuracy compared to any single index in isolation. The pooled classification rate of 1000 cross-validated iterations of the model had a 91.7% 0.8, mean SE) success rate at correctly classifying individual recordings. The model was subsequently used to classify recordings from two actively restored sites, established >24 months prior to recordings, with coral cover values of 79.1% (±3.9) and 66.5% (±3.8). Of these recordings, 37/38 and 33/39 received a classification as healthy respectively. The model was also used to classify recordings from a newly restored site established <12 months prior with a coral cover of 25.6% (±2.6), from which 27/33 recordings were classified as degraded. This investigation highlights the value of combining PAM recordings with machine-learning analysis for ecological monitoring and demonstrates the potential of PAM to monitor reef recovery over time, reducing the reliance on labour-intensive in-water surveys by experts. As access to PAM recorders continues to rapidly advance, effective automated analysis will be needed to keep pace with these expanding acoustic datasets.

1. Introduction

Ecological monitoring of marine habitats is key to understanding these ecosystems and successfully measuring the outcomes of conservation and restoration programmes happening in our oceans. This kind of ecological monitoring often relies on visual census surveys. However, these come with limitations which include the requirement of expert data collectors, logistical complexities, are often expensive, are typically poor at monitoring cryptic organisms within the ecological community, and, only collect a snapshot in time of the target site, rather than continous long-term data (Mooney et al., 2020; Munger et al., 2022). Moreover, conservation and restoration programmes are typically limited by time and resources, making it a challenge to sufficiently report on their progress using these methods (Boström-Einarsson et al., 2020; Rilov et al., 2020).

Automated passive acoustic monitoring (PAM) of whole soundscapes has the potential to address many of these limitations (Lindseth and Lobel, 2018; Mooney et al., 2020, Lamont et al., 2021). Low-cost acoustic recording technology capable of recording continuously for several days, or longer with duty cycling, is becoming available (Chapuis et al., 2021; Lamont et al., 2022). These devices can be deployed rapidly by non-expert data collectors and left to record autonomously. These can collect data on cryptic organisms that disproportionately rely on acoustic communication compared to non-cryptic species (Lamont et al., 2022). A growing number of studies have found relationships between the soundscapes of marine habitats and traditional ecological metrics such as benthic cover, fish communities and overall habitat quality using automated approaches (Nedelec et al., 2015; Butler et al., 2016; Freeman and Freeman, 2016; Harris et al., 2016; Gordon et al., 2018; Elise et al., 2019). As well as being useful for the tracking of habitat characteristics, soundscapes also constitute important components of ecosystem functioning, especially for orientation and recruitment of a variety of organisms (Simpson et al., 2005; Lecchini et al., 2018; Gordon et al., 2019). Surveying reefs using acoustics can therefore provide new understandings that may explain both ecological and behavioural processes.

Given recent innovations in autonomous hydrophone technology, our ability to capture large databases of long-term soundscape recordings is expanding (Sousa-Lima et al., 2013; Lamont et al., 2022). However, analytical approaches must keep pace with this if the potential of these data is to be maximised. Early analysis was based on aural assessments or visual inspection of spectrograms to score characteristics such as the frequency of occurrence and diversity of acoustic events like fish vocalisations (Putland et al., 2017; Archer et al., 2018; McWilliam et al., 2017, 2018; Bertucci et al., 2020; Lamont et al., 2021). However, these approaches can be slow and labour intensive, introducing a severe limit on the speed at which PAM data can be analysed.

Computationally generated ecoacoustic indices are becoming a popular approach to overcome this limitation (Gibb et al., 2019). Ecoacoustic indices have primarily been developed for terrestrial habitats (Sueur et al., 2014) where they are used to quantify soundscape attributes including variability across time and/or frequency bands (Stowell and Sueur, 2020). These indices can be automatically derived from long-term acoustic recordings, enabling extended recordings to be analysed. Several indices have been tested recently in the marine environment, revealing relationships between these and attributes of the ecological community, habitat quality and ecological functioning of marine habitats (Harris et al., 2016; Lindseth and Lobel, 2018; Gordon et al., 2018; Elise et al., 2019b; Mooney et al., 2020).

These studies have primarily used individual ecoacoustic indices to test for differences between sites (e.g., low- and high-quality habitats, reefs pre- and post-bleaching on reefs), or relationships with other ecological metrics (e.g., biodiversity, abundance etc) so far. However, these indices do not perform consistently across all marine investigations (Kaplan et al., 2015; Bertucci et al., 2016; Dimoff et al., 2021). Results from any individual index can also be biased by

individual components of the soundscape, such as a high density of snapping shrimps or repetitive fish chorusing, limiting their utility to characterise the wider community (Staaterman et al., 2013; Bolgan et al., 2018; Dimoff et al., 2021).

Some terrestrial soundscape ecology investigations have attempted to overcome similar performance issues through combining several ecoacoustic indices to generate multivariate representations of acoustic recordings known as 'compound indices' to generate a more holistic representation of the soundscape (Eldridge et al., 2018). These compound indices can then be input into machine learning algorithms which are able to identify relationships between this 'feature set' of indices and the task asked of the algorithm through finding patterns and interactions between the indices. (Eldridge et al., 2018; Bradfer-Lawrence et al., 2019; Gibb et al., 2019; Sethi et al., 2020).

Such tasks include supervised approaches such as classification problems which group soundscape recordings into categories specified by the researchers (e.g logged or unlogged forest) or regression tasks which can place recordings along a gradient (e.g avian diversity) (Sethi et al,. 2020, 2021). Alternatively, unsupervised approaches can be used such as clustering which can be used to identify key groups from recordings, or, anomaly detection can be used to identify recordings which significantly deviate from the majority (e.g containing noise pollution or a rare animal chorus) (Sethi et al., 2020).

In this study, we use whole soundscape recordings to test whether a machine learning algorithm could be trained using a compound index to accurately classify recordings from two different ecostates of a marine habitat. We used coral reef soundscapes due to the diversity of sounds present on these ecosystems and known relationships between their soundscape and ecological attributes (Kaplan et al., 2015; McWilliam et al., 2017; Nedelec et al., 2015). We then trialled this model in a real world application using recordings from neighbouring reefs which had been actively restored by one of the world's largest reef restoration programmes, located in Sulawesi, Indonesia. Coral restoration programmes typically fail to collect adequate monitoring data (Bayraktarov et al., 2019; Boström-Einarsson et al., 2020). Through demonstrating the utility of this approach, we intend to highlight a more efficient and cost-effective means of field data collection for the monitoring of restoration and other marine conservation projects alike.

2. Methods

2.1. Study sites

Recordings were made at sites in the Spermonde Archipelago (South Sulawesi, Central Indonesia; 4°56.9′S, 119°18.1′E; Fig. 1C) around Pulau Badi (Fig. 1A) and Pulau Bontosua (Fig. 1B). In 2006, the Mars assisted reef restoration system (MARRS) (buildingcoral.com) was developed to provide a novel methodology to re-establish coral cover at sites that have been degraded to rubble fields by historical coral mining and dynamite fishing (Williams et al., 2019). Coral fragments are attached to 'reef stars', interlinked metal frames, which provide a stable substrate. This prevents the smothering of new coral recruits through the turn over of rubble by waves and tides which typically slows the recovery process to a decadal timescale or longer (Fox et al., 2003). Between 2013 and 2017 this practice increased coral cover from approximately 10% to 60% across 7,000 m^2 of reef (Williams et al., 2019). Recordings were made at seven sites which encompassed four distinct types of reef habitat: healthy (Badi & Bontosua), degraded (Bontosua & Salisi), mature restored (Badi & Bontosua) and newly restored (Salisi). Coral cover was assessed to provide a measure of reef health (methodology in Supp. 1). The two healthy sites exhibited naturally high coral cover (Badi: 91.2% ± 2.0; Bontosua: 93.1% ± 2.6; mean ± SE) whereas the degraded sites exhibited low coral cover (Salisi: 2.1% ± 0.9; Bontosua: 17.6% ± 4.6) (Fig. 2). The two mature restored sites were established >24 months previously and exhibited increased coral cover (Badi: 79.1% ± 3.9; Bontosua: 66.5% ± 3.8) compared to
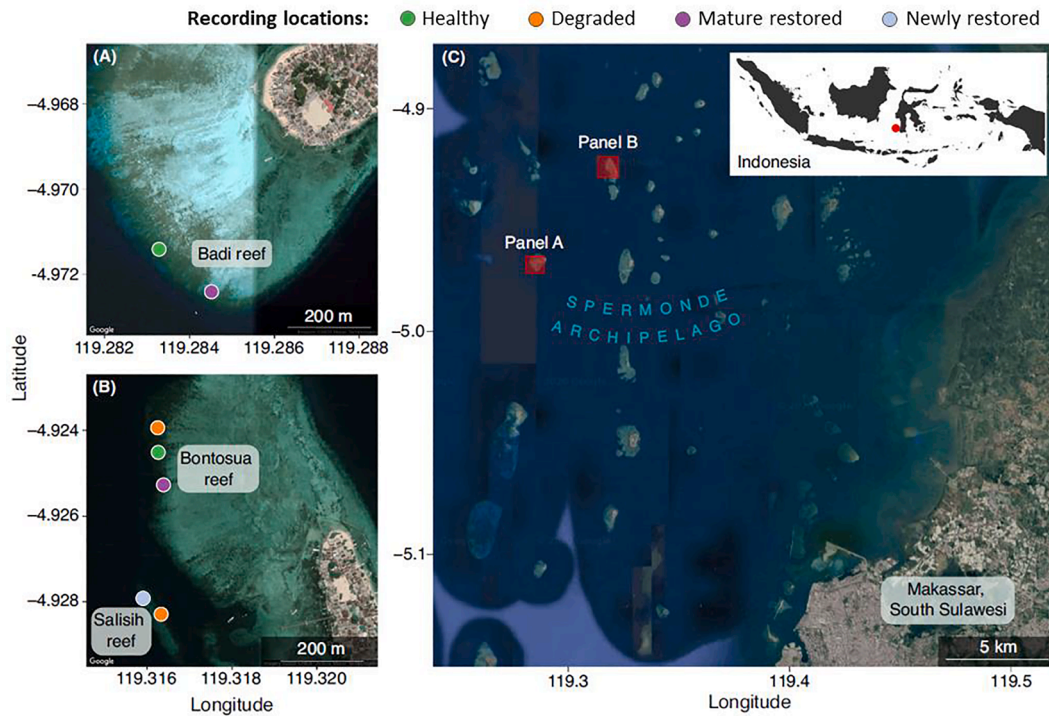
Fig. 1. Location and habitat class of the seven reef sites, present within the broader Spermonde Archipelago, Indonesia (A) where soundscape recordings were collected. Fringing reefs from two nearby islands: Bontosua (B) and Badi (C) were used. Modified from Lamont et al. (2021).
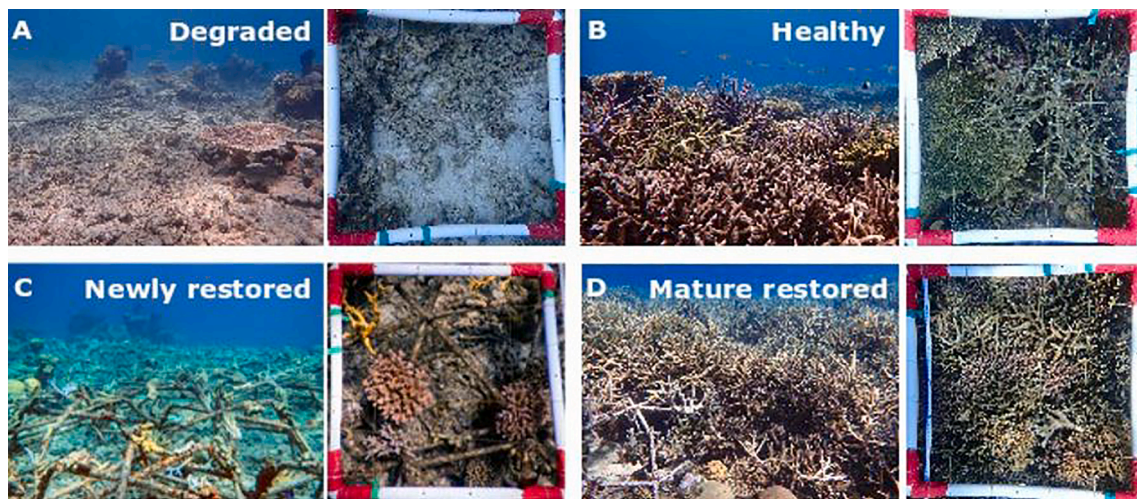

Fig. 2. Representative habitat and coral cover images from the four habitat classes at which soundscape recordings were made. (A) Degraded, (B) healthy, (C) newly restored and (D) mature restored.

the newly restored site (25.6% $\pm$ 2.6), which was established <12 months previously.

2.2. Acoustic recordings

The recordings were made across the seven sites during August–September 2018 as part of the MARRS monitoring programme. We used a regime which sampled one hour blocks from sites five days either side of the full moon (26th August 2018) and three days either side of the following new moon (10th September 2018) during daylight (09:00–15:00), twilight (half an hour either side of sunrise and sunset) and night time (half an hour either side of midnight) periods. Three SoundTraps (SoundTrap 300STD, 48 kHz sampling rate, Ocean

Instruments, Auckland, NZ) were used, with one suspended 0.5 m above the seabed for each deployment at a site. In each new round of deployments, SoundTraps were assigned randomly to recording sites within a counterbalanced blocking design, in order to control for potential instrument error.

We sub-sampled five non-overlapping one-minute segments from each of the hour-long periods at random. Only samples recorded during calm conditions (wind speed < 20 km h^{-1}) were used. These samples were also screened for motorboat noise and any recordings with this disturbance were removed, resulting in 262 recordings in the final sample set. This sample set was previously used in Lamont et al. (2021) to compare fish sound diversity between sites.

2.3. Processing recordings

Each recording was filtered using a short-term Fourier transform band-pass filter into three frequency bands: low-frequency (0.05–0.8 kHz), medium-frequency (2–7 kHz) and a broadband (0.05–20 kHz). The low-frequency band covered all the known fish vocalisations within the dataset (Lamont et al., 2021), while the medium-frequency band comprised invertebrate (primarily snapping shrimp) sound (Elise et al., 2019a). The broad-frequency band was used to encompass the full spectrum of potentially relevant frequencies, as previously used in coral reef soundscape investigations (Kaplan et al., 2015; Lyon, 2018). Frequencies below 0.05 kHz were excluded from the low- and broad-frequency band recordings to remove geophonic noise and self-noise from the recording system (Curtis et al., 1999). All processing was performed in *R* (v3.4.2. R Development Core Team, 2020): audio files were read and written using *tuneR* (v.1.3.3) (Ligges et al., 2018) and the filters were implemented using *Seewave* (v2.1.6) (Sueur et al., 2008).

2.4. Ecoacoustic indices

Twelve ecoacoustic indices were chosen from a range of previous soundscape studies (Table 1). Each index was calculated for all three frequency bands, with two exceptions: Snap rate was only calculated for the middle- and broad-frequency bands, because snapping shrimp cavitation bubbles are not audible at lower frequencies (Bohnenstiehl et al., 2016), and the normalised difference soundscape index (NDSI) was only calculated for the broad-band recordings. NDSI 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). We instead used this index to quantify differences in the 1 kHz band where fish noise dominates, and, a higher 2–7 kHz frequency band where snapping shrimp sound dominates (Au and Banks, 1998). Thus, we established a feature set of 33 index values across 12 indices and three frequency bands for each of the 262 one-minute recordings. All indices were calculated using the *R* package *Seewave* (Sueur et al., 2008) where possible, all remaining indices were calculated in *Soundecology R* package (v.1.3.3) (Villanueva-Rivera et al., 2018) other than SPL which was calculated in paPAM (Nedelec et al., 2016) and snap rate which calculated using a custom MATLAB script (Gordon et al., 2018).

2.5. Selection of indices to differentiate healthy and degraded habitats

All 33 indices were examined for separation between recordings from healthy (n = 81) and degraded (n = 71) habitats using Mann-Whitney U tests. Since the purpose was to explore the potential of each candidate index for model development, we did not control for spatial and temporal pseudoreplication within the library of recordings, nor control for cross-correlation between indices with similar trends. Violin plots were used to visualise the degree of overlap between the distributions for indices with indicative differences. Where minimal overlap was observed between the two habitats, the index could be considered likely to provide a promising measure with which to differentiate between healthy and degraded habitats.

2.6. Machine-learning approaches to develop a compound index

Following analysis of individual indices, we developed a supervised machine-learning model to assign recordings to either healthy or degraded habitat classes. A regularised discriminant analysis (RDA) algorithm was selected to account for the high level of collinearity reported between indices (Supp. 1). An optimised set of indices was selected in a 'feature selection' stage, using recursive feature elimination (RFE) and a multivariate adaptive regression spline (MAR) (Kuhn and Johnson, 2019) (Supp. 1). The RFE highlighted increases in model accuracy with the multi-index approach as additional indices were added sequentially (Supp. 1, Fig. S1). Predictive accuracy was greatest with eight indices, followed by a gradual decline as the addition of further indices introduced noise and/or caused model overtraining. The list of suggested features from RFE included the following index/frequency band combinations: broadband ACI, H, NDSI and H_t; and medium-frequency band ACI, BI, H and H_t. This was highly congruent with rankings obtained from the relative importance scores using the MAR (Fig. 3).

Following RFE, further manual feature selection was conducted by systematic removal and addition of indices whilst executing the full model, to select a final feature set with the lowest misclassification rate. This led to discarding H_t in both the broad-range and middle-frequency bands, and the introduction of low-frequency band ACI and middle-frequency band AR. Thus, the final set was: low-frequency band ACI, medium-frequency band ACI, AR and BI, broadband ACI, H and NDSI. Feature selection was performed using the *R* packages *mlbench* (v2.1.1) (Leisch and Dimitriadou, 2010) and *Caret* (v.6.0–86) (Kuhn, 2020).

2.7. Constructing the final model

An RDA model was constructed using the healthy and degraded datasets, using the *R* packages *MASS* (v.7.3-53) (Venables and Ripley,

Table 1

Twelve ecoacoustic indices calculated from recordings with summary description of the mechanistic principle, software used and respective settings employed.

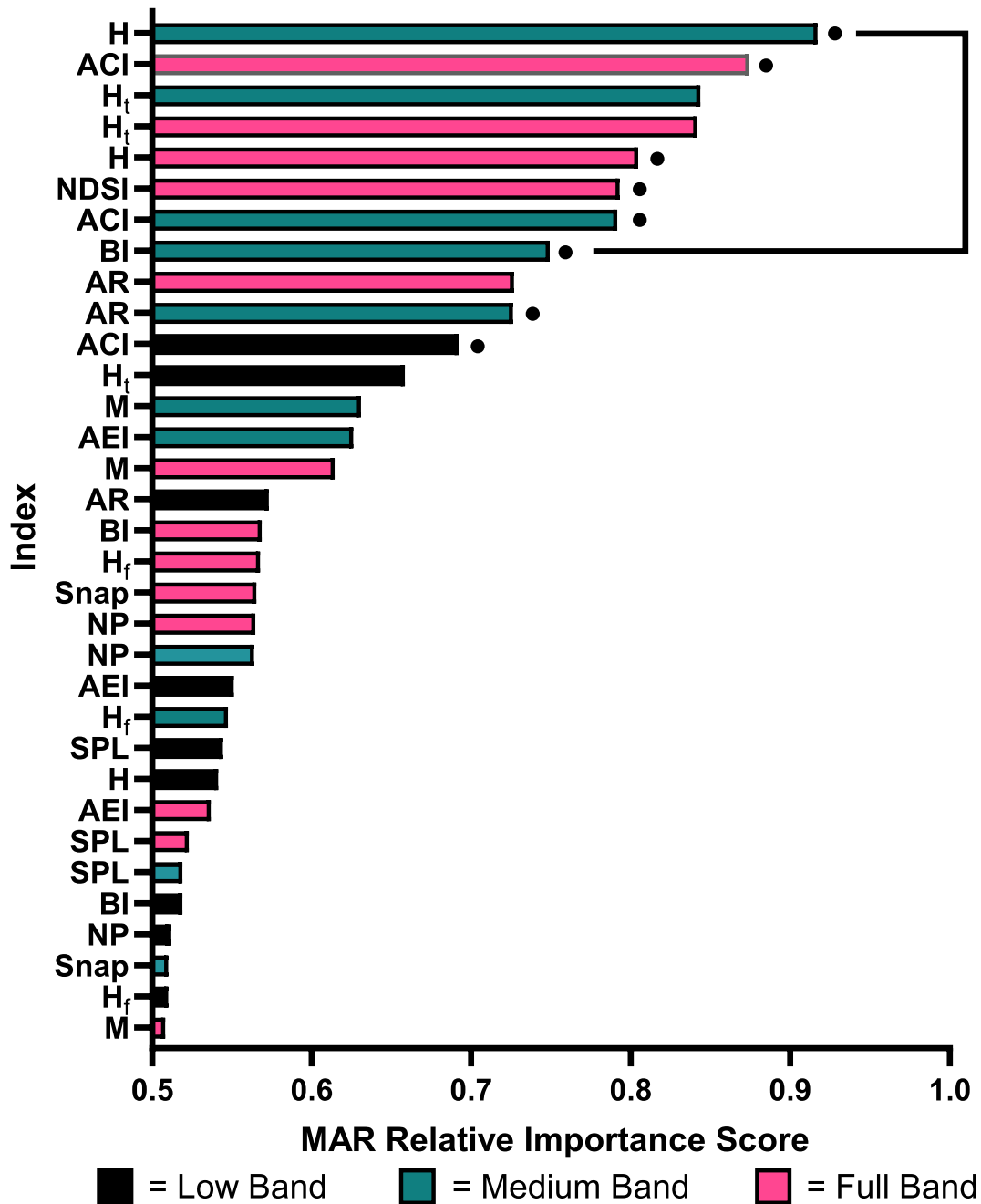| Index	Mechanism	Software	Settings	Origin
Acoustic complexity index (ACI)	Measures variability in intensity of frequencies across time	*Seewave* in *R*	Window size = 512; type = Hamming; overlap = 0	(Pieretti, 2011)
Acoustic entropy (H)	Measures randomness across temporal and spectral domains	*Seewave* in R	Window size = 512; envelope = Hilbert	(Sueur, 2008)
Acoustic eveness index (AEI)	Measures diversity across frequency bands	*Soundecology* in *R*	Max freq = upper bound of band in use; 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 cumulative intensity across frequency bands	*Soundecology* in *R*	Min and max frequency matched to band in use; window size = 512	(Boelman, 2007)
Normalised mean difference index (NDSI)	Measures amplitude difference between two selected frequency bands	*Seewave* in *R*	Min and max frequency matched to band in use; window size = 512	(Kasten, 2012)
Number of peaks	Number of major frequency peaks obtained from a mean spectrum	*Seewave* in *R*	Window size = 512; type = Hanning; overlap = 0	(Sueur, 2008)
Spectral entropy (Ht)	Measures randomness across the frequency domain	*Seewave* in *R*	No settings required	(Sueur, 2008)
Temporal entropy (Hf)	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 from Gordon et al., (2018)	Widely used
Sound pressure level (SPL)	Calibrated measure of root mean squared sound pressure level	*paPAM* in *MATLAB*	Window length = 1024; type = Hamming; Overlap = 50%	Widely used

Fig. 3. Relative importance rankings of indices obtained from the multivariate adaptive regression (MAR) analysis used for feature selection. The eight recommendations obtained from the recursive feature elimination (RFE) analysis are indicated by the black lines. The top eight indices of the MAR analysis were congruent with the eight recommendations from RFE, though the order was not conserved. Black dots to the right of bars indicate features which were selected for the final model after further manual feature selection.

200) and *KlaR* (v.0.6-15) (Weihs et al., 2005). Model accuracy was assessed using k-fold cross validation (10 folds), whereby the dataset was partitioned into 'training' and 'test' sets to prevent overestimation of model accuracy when presented with new data (Supp. 1). To account for variation in the RDA, 1000 repeats of the cross-validated model construction were performed to assess accuracy (Rao et al. 2008).

A principal component analysis (PCA) and a pairs plot comparing each combination of the eight selected indices were generated for all recordings. These were used to test whether soundscape properties from restored reefs diverged from the healthy and degraded classes, which would lead to potentially inappropriate classifications using the RDA trained on the healthy and degraded recordings. Both tests were conducted in R using inbuilt functions (Supp. 1, Fig. S4).

3. Results

3.1. Comparing indices between healthy and degraded sites

Exploratory Mann-Whitney U tests revealed significant differences between healthy and degraded habitat index scores for 12 of the 33 indices (Fig. 4). Bonferroni corrections were also used to reduce the likelihood of false positives in the search for a significant difference; the original alpha value of 0.05 was therefore divided by 33 to provide a new value of 0.00152. Using this more conservative approach no longer reveals AI in the broad and medium-frequency bands, as well as AEI in the medium-frequency band, to be significantly different.

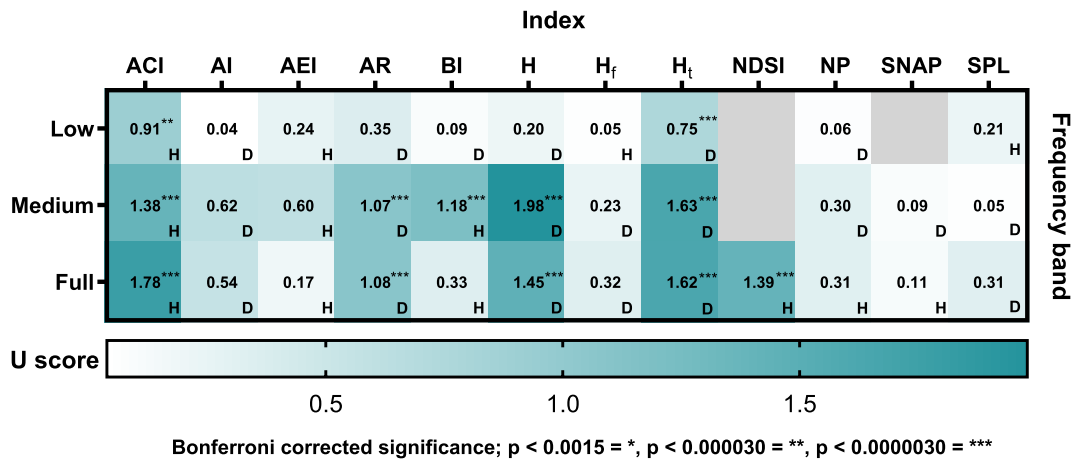Violin plots of the three most significantly different index results

Fig. 4. Heat map displaying results from Mann-Whitney U tests between the ecoacoustic index scores calculated from recordings of healthy (n = 81) and degraded (n = 71) sites in low-, medium- and broad-frequency bands. The habitat class with the higher mean is indicated by the letter in the bottom right corner of each cell (H = Healthy; D = Degraded). Blank cells indicate indices for which values from the corresponding frequency band were not calculated (see Methods).

between the healthy and degraded sites revealed large zones of overlap between values for these indices between habitat classes (Fig. 5). The strongest significant difference was reported for H in the 2–7 kHz band, here 71 of the 152 (47 %) of recordings reported results that did fall in range of both habitat types.

3.2. Comparing indices to phonic richness

We searched for a correlation between each ecoacoustic index and the diversity of fish sounds present in each recording using the 'phonic richness' method which counts the number of unique sound types audible in each recording (Lamont et al., 2021) (Supp. 1). This revealed no strong relationships (Pearson correlations) between phonic richness and any of the 33 indices trialled (Supp. 1, Fig. S3). The strongest relationship was a weak negative correlation with the acoustic entropy index (H) for the broad-frequency band (Pearson correlation; rho = -0.43; $p < 0.001$), with all other indices reporting weaker correlations than this.

3.3. Regularised discriminant analysis

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, mean SE). Across these model constructions, of the 81 recording samples taken from the two healthy sites, 73.0 (0.1) of these were correctly classified as healthy,

with 8.0 (0.1) misclassified as degraded. Of the 71 recordings taken from the two degraded sites, 67.2 (0.1) of these were classified as degraded, with 3.7 (0.1) misclassified as healthy (see Fig. 6 for individual results for each recording sample).

Cluster analysis using the principal component analysis (PCA; Fig. 7) and a pairs plot (Supp. 1, Fig. S4) were used to examine whether the 110 samples taken from recordings of the three restored sites overlapped with recordings from the control sites. If they deviated then it would be likely that the model would provide inappropriate classifications to these sites. 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 at the other two habitat types when using the properties investigated here. This supports the inputting of restored samples into the model as the data were not divergent from the original training data. 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 could 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 when using only the two dimensions shown by the PCA, with most of the ellipse of the degraded classes encompassed by that of the healthy class.

Analysis of the restored site samples revealed that the majority of samples from mature restored sites were classified as healthy, but
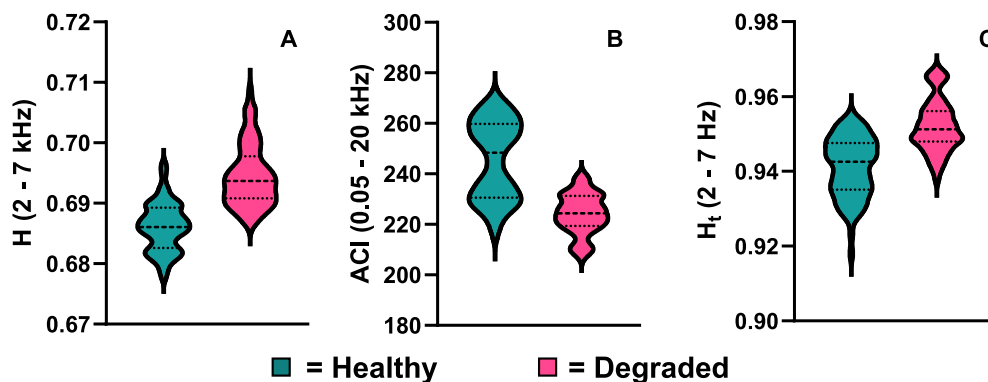

Fig. 5. Violin plots of the three indices with the most significant differences between healthy (n = 81) and degraded (n = 71) habitat. (A) Medium-frequency band Entropy Index (H) (Mann-Whitney U; U = 1.98, p < 0.001), (B) Broad-frequency band Acoustic Complexity index (ACI) (U = 1.78, p < 0.001), (C) Medium-frequency band Temporal Entropy (H_t) (U = 1.63, p < 0.001).
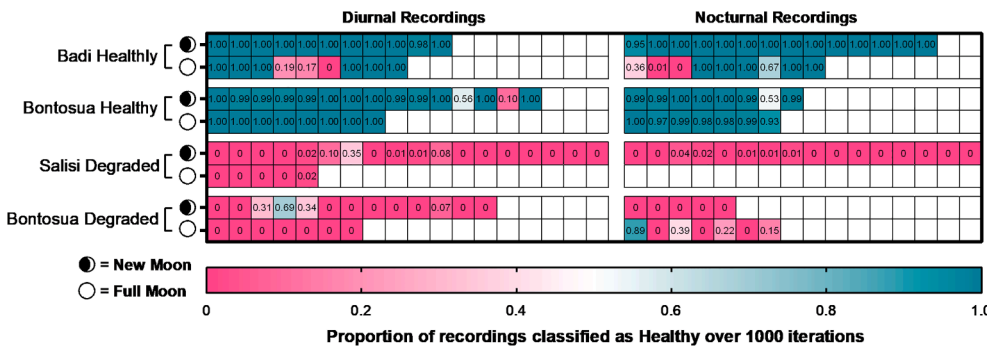
Fig. 6. Machine learning classification of acoustic samples from the healthy and degraded sites. Each cell indicates a single one-minute recording from the 152 taken in healthy and degraded habitats. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time 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 predicted as degraded (green shading: >0.5; pink shading: < 0.5). Recordings on the left of the partition were made during the day and recordings to the right were made during crepuscular or nighttime 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. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
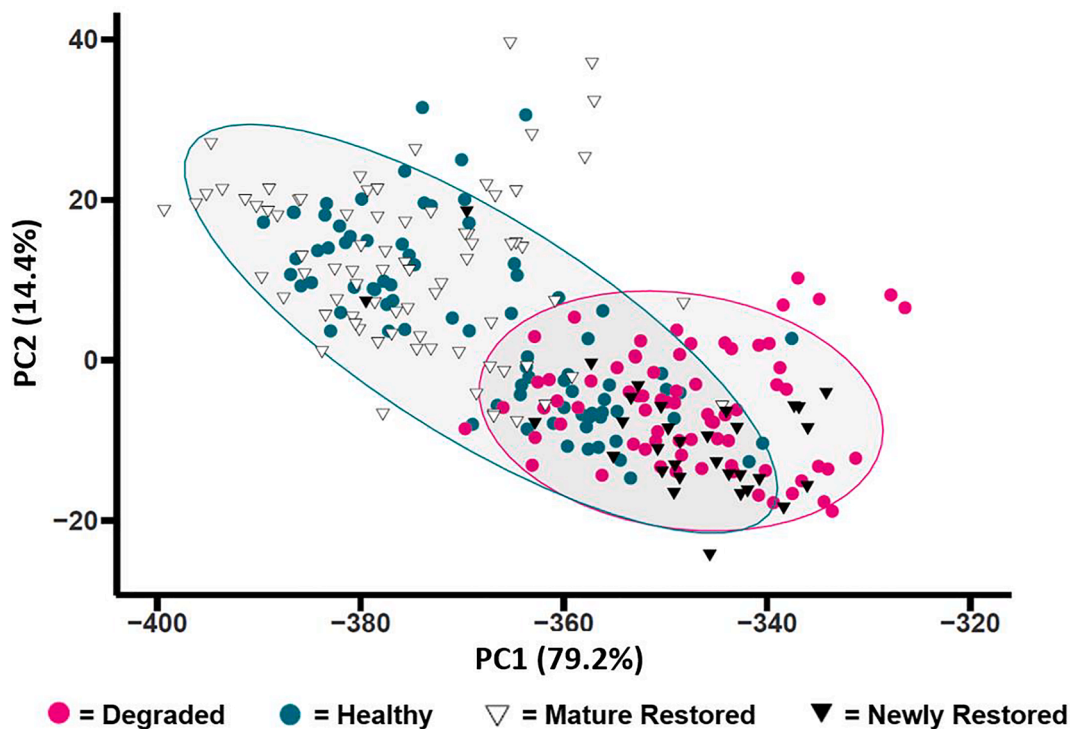

Fig. 7. Plot from the principal component analysis of PC1 and PC2 scores for the Healthy and Degraded site recording samples. Samples from recordings of Restored sites are overlaid on this to help determine whether these conform with either of the two existing classes or whether the properties of their soundscape are distinct. Ellipses indicate the zone within which a new sample can be assigned to a class using the two principal components presented in this figure. Overlapping areas indicate ambiguous recordings which cannot be differentiated by PCA.

samples from the newly restored site were mainly classified as degraded (Fig. 8). The Bontosua mature restored site was classified more clearly than the Badi mature restored site, with 37/38 and 33/39 samples classified as healthy respectively. The six samples classified as degraded from the Badi mature restored site occurred consecutively on the new moon at night. At the newly restored site, 27/33 samples were classified as degraded, and all of these were during the full moon (though only four new moon samples were available) and five of these were at night.

The model trained on the 2018 recordings was also tested on a smaller number of recordings taken at the same sites 10 months later (June/July 2019). Here, the model provided similar predictions for six of the seven sites; while one site (Healthy Bontosua) exhibited a change in prediction between 2018 and 2019, changing from primarily being classified as healthy to degraded (Table 2, full results in Fig. S5).

4. Discussion

This study compared the value of individual ecoacoustic indices and a machine-learning model trained on a compound index to discriminate between coral reef ecostates. While no single ecoacoustic index could reliably discriminate between healthy and degraded reefs, a supervised machine-learning approach more accurately predicted habitat class from randomly drawn acoustic samples. This highlights the potential of combining PAM with machine learning for monitoring the health of marine ecosystems.

Twelve individual ecoacoustic indices were calculated in up to three frequency bandwidths, totalling 33 values, of which 12 indices were significantly different between healthy and degraded reefs (Fig. 4). There were no strong correlations between any of these indices and
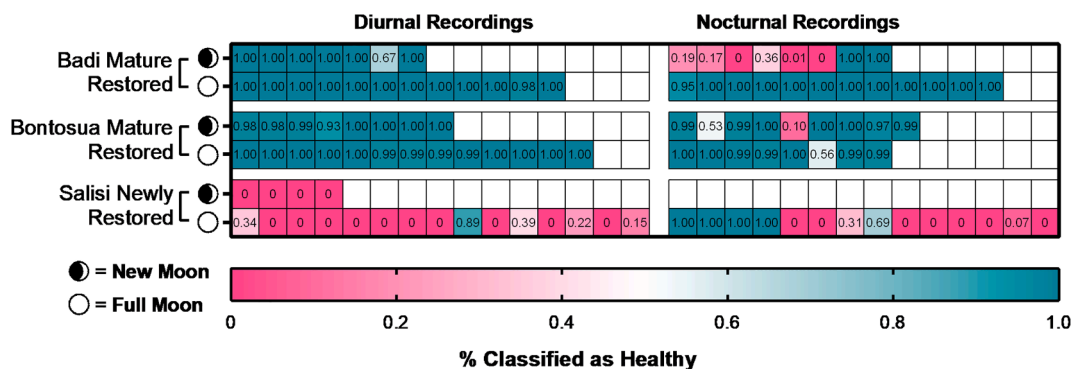
Fig. 8. Machine-learning classification of acoustic samples from the restored sites. Each cell indicates a single one-minute recording from the 110 taken from restored sites. The model was executed 1000 times on the dataset, generating a new habitat class prediction each time for every recording. Values 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 (green shading: >0.5; pink shading: <0.5). Recordings on the left of the partition were made during the day and recordings to the right were made during crepuscular or nighttime periods. Despite gaps in the sampling regime, the order within blocks conserves the overall order with which they were sampled across time. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Results from the application of the 2018 model when tested on recordings taken at the same sites in 2019.

	Badi Healthy	Bontosua Healthy	Salisi Degraded	Bontosua Degraded	Badi Mature Restored	Bontosua Mature Restored	Salisi Newly Restored
Recordings classified as Healthy	9/9	2/12	0/5	5/12	9/9	12/12	8/9
Proportion classified as Healthy	1.0	0.17	0	0.42	1.0	1.0	0.89

phonic richness (Supp. 1), indicating that fish sound diversity alone is not the dominant driver of these results; rather alternative aspects of the soundscape are responsible. A metric that combines abundance alongside diversity of fish vocalisations may reveal more about the role these play in driving index values. Other contributors may originate from alternative biotic or abiotic sources. Invertebrates are well documented contributors to reef soundscapes including snapping shrimp (Bohnenstiehl et al., 2016), urchin feeding activities (Radford et al., 2008) and the movement of hard shelled organisms (Freeman et al., 2014). The photosynthetic process of macro algae has also been reported as a sound producer (Freeman et al., 2018). Though understudied, abiotic attributes may also influence the soundscape as ecostates change (Duarte et al. 2021). Though recordings were taken in calm conditions, low level geophonic noise produced by waves and wind may propagate differently through a rubble field compared to a more structurally complex reef, or the regular movements of unconsolidated coral rubble due to hydrodynamic forcing might also contribute to the soundscape (Kenyon et al. 2020).

The distribution of values within indices showing significant differences between healthy and degraded reefs exhibited substantial overlap between the two habitat classes. The best performing individual index was H in the 2–7 kHz band. For this index, 71 of the 152 (47 %) of recordings had non-overlapping values, meaning they could not be correctly classified as healthy or degraded, all other recordings were ambiguous as the values fell within the range reported for both habitat types. This means that the ability to distinguish between habitat classes from a single recording using an individual index is low, as any given value from one class has a high chance of being reported from a recording in the other class. Violin plots of the three most significant results demonstrate this large overlap between the values of each class (Fig. 5). In isolation, these indices could discriminate between habitats if extensive sampling is achievable for all sites of interest to build up a dataset that can be tested for statistical significance, as demonstrated here. However, their potential to deliver reliable results from short 'snapshot' recordings is lower and their ability to deliver insights into

more complex tasks, beyond a coarse healthy-degraded comparison, may be limited.

In contrast to bivariate analyses, combining multiple indices with regularised discriminant analysis (RDA) gave a strong predictive ability to classify habitats based on single recordings. Recursive feature elimination (RFE) highlights the increase in accuracy attainable through constructing an optimised set of multiple indices (Supp. 1, Fig. S1) compared to using individual indices (Fig. 5). The misclassification rate of the final RDA model was 8.27% (0.84, mean SE) when applied to recordings from the same season. The model made accurate predictions despite being kept blind to diel and lunar period, which are both known to influence marine soundscapes (Staaterman et al., 2014); this highlights its robustness to temporal changes in soundscapes. The model also reliably delivered the same classification for recordings from six of the seven sites taken ten months later. The feature-selection stage of this approach is specific to the data and questions considered in this study. However, indices within the final feature set may offer a useful starting place for similar investigations elsewhere. To produce optimised models, investigations at new locations addressing new questions should carry out an independent feature selection process on their own data.

Following the successful classification of healthy and degraded habitats, our compound index based model was used to examine soundscape recordings taken from nearby coral reef habitats that had been restored (Williams et al., 2019). This tested the ability of this approach to perform a rapid assessment of these restored sites using one-minute soundscape recordings. The model was able to detect differences between the two mature restored sites and the newly restored site. Of the recording samples from the two mature restored sites, 33/39 and 37/38 were primarily classified as healthy, whereas 27/33 samples from the newly restored sites were classified as degraded (Fig. 8). The mature restored sites were more than twice as old as the newly restored site (restoration started >24 months prior to recordings on mature restored sites, compared to <12 months for the newly restored site), and had approximately three times more live coral cover (79.1% ± 3.9 and 66.5% ± 3.8 for the mature restored sites, 25.6% ± 2.6 for the newly

restored site; values all % live coral cover mean ± SE; full data in Supp. 1). Restoration progress was clearly detected in the soundscape, with better classification made possible by using a machine-learning-driven approach, suggesting PAM can be a useful tool for monitoring restoration against reference sites. More generally, this study highlights the potential for using machine-learning approaches to explore PAM data to provide greater analytical power in coral reef monitoring programmes.

Further improvements include considering sources of observed error in the model, which could be due to several factors working in isolation or in combination. The RDA approach operates best when input features have Gaussian distributions (Wu et al., 1996), but some features used in this study exhibited sub-Gaussian distributions. This is likely due to inclusion of samples from various times of day and from multiple sites. Diel trends are frequently detected in reef soundscapes across a range of ecoacoustic indices (Kaplan et al., 2015; Bertucci et al., 2020b; Carriço et al., 2020). Additionally, reef soundscapes can differ over small spatial scales (Putland et al., 2017). In this study, samples were taken from spatially separated sites to reduce pseudoreplication, thus differences within habitat classes are likely. Both these factors may have skewed the distributions of the feature sets. Furthermore, the dataset used to train the model will also contain natural outliers through ecological randomness that cannot be resolved at the sampling resolution employed. Longer periods of recording can be used to minimise impacts of this natural variation (Bradfer-Lawrence et al., 2019), though we show here that short periods of recording can still be used to identify accurate classifications between significantly different habitat states.

Six of the seven sites studied in 2018 retained similar classifications when resampled 10 months later in 2019. The outlier was Bontosua Healthy, for which 10/12 recordings were incorrectly classified as degraded. Recordings at this site were only collected during the day in 2019, and 9/12 of these were taken during the new moon period. The soundscape may therefore have been inadequately sampled. Alternatively, this could be an early indicator of a changing state of reef health at this site, not yet seen in the coral cover data, which was similar in both years (Supp. 1), or, this could demonstrate the specificity of this model to the time it was taken.

Future investigations could build on the present study by considering a more nuanced approach to classifying ecostate. For example, this study employed a binary classification of reef health but, in reality, marine habitats occur across gradients of ecostates (Downs et al., 2005; Smith et al., 2008). Sampling across these gradients, and using regression-based algorithms such as logistic regression, random forests or neural networks could support models that can predict on a continuous scale. Additionally, within the context of coral reefs, although live coral cover may be a strong indicator of overall reef health (Smith et al., 2016; Dietzel et al., 2020), other attributes of interest could be incorporated to better determine the ecostate of a site. For example, soundscape-based machine-learning models could be trained to predict metrics which are effort and training-intensive, such as fish or invertebrate abundance or diversity, and other habitat attributes could be explored such as structural complexity or ecosystem stability. Similar approaches could also be applied to other kinds of marine habitats where soundscape research has so far been limited (Pieretti and Danovaro, 2020). By drawing comparisons with a wider range of traditional metrics used in marine monitoring, the potential for machine-learning-based analyses of ecoacoustic recordings can be further developed.

5. Conclusion

Given the increasing availability of hydrophone technology (Chapuis et al., 2021; Lamont et al., 2022), acoustic datasets from the marine environment are set to rise.

Automated analyse is needed to efficiently process and analyse large acoustic datasets so that insights from the information held within can be maximised. However, this is so far underdeveloped. Our investigation presents an automated approach, through the use of a compound index

and machine-learning, that improves upon existing approaches used in the marine environment. We first demonstrate the use of this to classify coral reef habitats into healthy or degraded ecostates based on short-term recordings. We then demonstrate this in an applied setting, highlighting the utility of this approach when assessing areas of restored reef, revealing that restoration progress is detectable in the soundscape. This investigation provides the first evidence that compound indices and machine learning are able to outperform the use of single ecoacoustic indices on a tropical reef and that this approach should be considered for use in other marine and terrestrial habitats applications.

CRediT authorship contribution statement

Ben Williams: Conceptualization, data curation, formal analysis, visualization writing (original draft). **Timothy A.C. Lamont:** Investigation, data curation, writing (review & editing), funding acquisition. **Lucille Chapuis:** Investigation, writing (review & editing), funding acquisition. **Harry R. Harding:** Investigation, writing (review & editing). **Eleanor B. May:** Investigation. **Mochyudho E. Prasetya:** Investigation. **Marie J. Seraphim:** Investigation. **Noel Janetski:** project administration, resources. **Jamaluddin Jompa:** project administration, resources. **Dave Smith:** project administration, resources. **Andrew N. Radford:** writing (review & editing), supervision. **Stephen D. Simpson:** writing (review & editing), funding acquisition and supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2022.108986.

References

Archer, S.K., Halliday, W.D., Riera, A., Mouy, X., Pine, M.K., Chu, J.W.F., Dunham, A., Juanes, F., 2018. First description of a glass sponge reef soundscape reveals fish calls and elevated sound pressure levels. Mar. Ecol. Prog. Ser. 595, 245–252. https://doi.org/10.3354/meps12572.

Au, W.W.L., Banks, K., 1998. The acoustics of the snapping shrimp Synalpheus parneomeris in Kaneohe Bay. J. Acoust. Soc. Am. 103, 41–47. https://doi.org/10.1121/1.423234.

Bayraktarov, E., Stewart-Sinclair, P.J., Brisbane, S., Boström-Einarsson, L., Saunders, M. I., Lovelock, C.E., Possingham, H.P., Mumby, P.J., Wilson, K.A., 2019. Motivations, success, and cost of coral reef restoration. Restor. Ecol. 27, 981–991. https://doi.org/10.1111/rec.12977.

Bertucci, F., Guerra, A.S., Sturny, V., Blin, E., Sang, G.T., Lecchini, D., 2020a. A preliminary acoustic evaluation of three sites in the lagoon of Bora Bora, French Polynesia. Environ. Biol. Fishes 103 (8), 891–902. https://doi.org/10.1007/s10641-020-01000-8.

Bertucci, F., Maratrat, K., Berthe, C., Besson, M., Guerra, A.S., Raick, X., Lerouvreur, F., Lecchini, D., Parmentier, E., 2020b. Local sonic activity reveals potential partitioning in a coral reef fish community. Oecologia 193, 125–134. https://doi.org/10.1007/s00442-020-04647-3.

Bertucci, F., Parmentier, E., Lecellier, G., Hawkins, A.D., Lecchini, D., 2016a. Acoustic indices provide information on the status of coral reefs: an example from Moorea Island in the South Pacific. Sci. Rep. 6, 1–9. https://doi.org/10.1038/srep33326.

Bertucci, F., Parmentier, E., Lecellier, G., Hawkins, A.D., Lecchini, D., 2016b. Acoustic indices provide information on the status of coral reefs: an example from Moorea Island in the South Pacific. Sci. Rep. 6, 33326. https://doi.org/10.1038/srep33326.

Bohnenstiehl, D.R., Lillis, A., Eggleston, D.B., Radford, C.A., 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), e0143691. https://doi.org/10.1371/journal.pone.0143691.

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. Sci. Rep. 8, 1–11. https://doi.org/10.1038/s41598-018-28771-6.

Boström-Einarsson, L., Babcock, R.C., Bayraktarov, E., Ceccarelli, D., Cook, N., Ferse, S.C. A., Hancock, B., Harrison, P., Hein, M., Shaver, E., Smith, A., Suggett, D., Stewart-Sinclair, P.J., Vardi, T., McLeod, I.M., Bianchi, C.N., 2020. Coral restoration – a systematic review of current methods, successes, failures and future directions. PLoS One 15 (1), e0226631. https://doi.org/10.1371/journal.pone.0226631.

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S.G., Dent, D.H., Zamora-Gutierrez, V., 2019. Guidelines for the use of acoustic indices in environmental research. Methods Ecol. Evol. 10 (10), 1796–1807. https://doi.org/10.1111/2041-210X.13254.

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. J. Exp. Mar. Bio. Ecol. 479, 89–96. https://doi.org/10.1016/j.jembe.2016.03.006.

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. https://doi.org/10.3390/acoustics2020022.

Curtis, K.R., Howe, B.M., Mercer, J.A., 1999. Low-frequency ambient sound in the North Pacific: long time series observations. J. Acoust. Soc. Am. 106, 3189–3200. https://doi.org/10.1121/1.428173.

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. doi: 10.1098/rspb.2020.1432.

Dimoff, S.A., Halliday, W.D., Pine, M.K., Tietjen, K.L., Juanes, F., Baum, J.K., 2021. The utility of different acoustic indicators to describe biological sounds of a coral reef soundscape. Ecol. Indic. 124, 107435 https://doi.org/10.1016/j.ecolind.2021.107435.

Downs, C.A., Woodley, C.M., Richmond, R.H., Lanning, L.L., Owen, R., 2005. Shifting the paradigm of coral-reef "health" assessment. Mar. Pollut. Bull. 51, 486–494. https://doi.org/10.1016/j.marpolbul.2005.06.028.

Duarte, C.M., Chapuis, L., Collin, S.P., Costa, D.P., Devassy, R.P., Eguiluz, V.M., Erbe, C., Gordon, T.A.C., Halpern, B.S., Harding, H.R., Havlik, M.N., Meekan, M., Merchant, N.D., Miksis-Olds, J.L., Parsons, M., Predragovic, M., Radford, A.N., Radford, C.A., Simpson, S.D., Slabbekoorn, H., Staaterman, E., Van Opzeeland, I.C., Winderen, J., Zhang, X., Juanes, F., 2021. The soundscape of the Anthropocene ocean. Sci 371 (6529). https://doi.org/10.1126/science.aba4658.

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

Elise, S., Bailly, A., Urbina-Barreto, I., Mou-Tham, G., Chiroleu, F., Vigliola, L., Robbins, W.D., Bruggemann, J.H., 2019a. An optimised passive acoustic sampling scheme to discriminate among coral reefs' ecological states. Ecol. Indic. 107, 105627 https://doi.org/10.1016/j.ecolind.2019.105627.

Elise, S., Urbina-Barreto, I., Pinel, R., Mahamadaly, V., Bureau, S., Penin, L., Adjeroud, M., Kulbicki, M., Bruggemann, J.H., 2019b. Assessing key ecosystem functions through soundscapes: A new perspective from coral reefs. Ecol. Indic. 107, 105623 https://doi.org/10.1016/j.ecolind.2019.105623.

Fox, H.E., Pet, J.S., Dahuri, R., Caldwell, R.L., 2003. Recovery in rubble fields: long-term impacts of blast fishing. Mar. Pollut. Bull. 46, 1024–1031. https://doi.org/10.1016/S0025-326X(03)00246-7.

Freeman, S.E., Rohwer, F.L., D'Spain, G.L., Friedlander, A.M., Gregg, A.K., Sandin, S.A., Buckingham, M.J., 2014. The origins of ambient biological sound from coral reef ecosystems in the Line Islands archipelago. J. Acoust. Soc. Am. 135, 1775–1788. https://doi.org/10.1121/1.4865922.

Freeman, L.A., Freeman, S.E.S., 2016. Rapidly obtained ecosystem indicators from coral reef soundscapes. Mar. Ecol. Prog. Ser. 561, 69–82. https://doi.org/10.3354/meps11938.

Freeman, S.E., Freeman, L.A., Giorli, G., Haas, A.F., Radford, C.A., 2018. Photosynthesis by marine algae produces sound, contributing to the daytime soundscape on coral reefs. PloS one 13 (10), e0201766. https://doi.org/10.1371/journal.pone.0201766.

Gibb, R., Browning, E., Glover-Kapfer, P., Jones, K.E., Börger, L., 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods Ecol. Evol. 10 (2), 169–185. https://doi.org/10.1111/2041-210X.13101.

Gordon, T.A.C., Harding, H.R., Wong, K.E., Merchant, N.D., Meekan, M.G., McCormick, M.I., Radford, A.N., Simpson, S.D., 2018. Habitat degradation negatively affects auditory settlement behavior of coral reef fishes. Proc. Natl. Acad. Sci. 115, 5193–5198. https://doi.org/10.1073/pnas.1719291115.

Gordon, T.A.C., Radford, A.N., Davidson, I.K., Barnes, K., McCloskey, K., Nedelec, S.L., Meekan, M.G., McCormick, M.I., Simpson, S.D., 2019. Acoustic enrichment can enhance fish community development on degraded coral reef habitat. Nat. Commun. 10, 1–7. https://doi.org/10.1038/s41467-019-13186-2.

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

Kaplan, M., Mooney, T., Partan, J., Solow, A., 2015. Coral reef species assemblages are associated with ambient soundscapes. Mar. Ecol. Prog. Ser. 533, 93–107. https://doi.org/10.3354/meps11382.

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. Ecol. Inform. 12, 50–67. https://doi.org/10.1016/j.ecoinf.2012.08.001.

Kenyon, T.M., Doropoulos, C., Dove, S., Webb, G.E., Newman, S.P., Sim, C.W., Arzan, M., Mumby, P.J., 2020. The effects of rubble mobilisation on coral fragment survival, partial mortality and growth. J. Exp. Mar. Biol. Ecol 533, 151467. https://doi.org/10.1016/j.jembe.2020.151467.

Kuhn, M., Johnson, K., 2019. Feature Engineering and Selection: A Practical Approach for Predictive Models. CRC Press.

Lamont, T.A.C., Williams, B., Chapuis, L., Prasetya, M., Seraphim, M.J., Harding, H.R., May, E.B., Janetski, N., Jompa, J., Smith, D.J., Radford, A.N., Simpson, S.D., 2022. The sound of recovery: Coral reef restoration success is detectable in the soundscape. J. Appl. Ecol 59 (3), 742–756. https://doi.org/10.1111/1365-2664.14089.

Lamont, T.A., Chapuis, L., Williams, B., Dines, S., Gridley, T., Frainer, G., Fearey, J., Maulana, P.B., Prasetya, M.E., Jompa, J. and Smith, D.J., 2022. HydroMoth: Testing a prototype low-cost acoustic recorder for aquatic environments. Remote. Sens. Ecol. Conserv. doi: 10.1002/rse2.249.

Lecchini, D., Bertucci, F., Gache, C., Khalife, A., Besson, M., Roux, N., Berthe, C., Singh, S., Parmentier, E., Nugues, M.M., Brooker, R.M., Dixson, D.L., Hédouin, L., 2018. Boat noise prevents soundscape-based habitat selection by coral planulae. Sci. Rep. 8, 9283. https://doi.org/10.1038/s41598-018-27674-w.

Lindseth, A., Lobel, P., 2018. Underwater soundscape monitoring and fish bioacoustics: a review. Fishes 3 (3), 36. https://doi.org/10.3390/fishes3030036.

Lyon, R.P., 2018. Fish biodiversity, habitat complexity, and soundscape characteristics of patch reefs in a tropical, back-reef nursery 42.

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. Sci. Rep. 7, 1–13. https://doi.org/10.1038/s41598-017-15838-z.

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. https://doi.org/10.1080/09524622.2017.1344930.

Mooney, T.A., Di Iorio, L., Lammers, M., Lin, T.-H., Nedelec, S.L., Parsons, M., Radford, C., Urban, E., Stanley, J., 2020. Listening forward: approaching marine biodiversity assessments using acoustic methods. doi: 10.1098/rsos.201287.

Munger, J., Herrera, D.P., Haver, S.M., Waterhouse, L., McKenna, M.F., Dziak, R.P., Gedamke, J., Heppell, S.A., Haxel, J.H., 2022. Machine learning analysis reveals relationship between pomacentrid calls and environmental cues. Mar. Ecol. Prog. Ser. 681, 197–210. https://doi.org/10.3354/MEPS13912.

Nedelec, S.L., Campbell, J., Radford, A.N., Simpson, S.D., Merchant, N.D., Fisher, D., 2016. Particle motion: the missing link in underwater acoustic ecology. Methods Ecol. Evol. 7 (7), 836–842. https://doi.org/10.1111/2041-210X.12544.

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. Mar. Ecol. Prog. Ser. 524, 125–135. https://doi.org/10.3354/meps11175.

Pieretti, N., Danovaro, R., 2020. Acoustic indexes for marine biodiversity trends and ecosystem health. Acoustic indexes for marine biodiversity trends and ecosystem health. 375 (1814), 20190447. https://doi.org/10.1098/rstb.2019.0447.

Putland, R.L., Constantine, R., Radford, C.A., 2017. Exploring spatial and temporal trends in the soundscape of an ecologically significant embayment. Sci. Rep. 7, 1–12. https://doi.org/10.1038/s41598-017-06347-0.

Radford, C., Jeffs, A., Tindle, C., Montgomery, J.C., 2008. Resonating sea urchin skeletons create coastal choruses. Mar. Ecol. Prog. Ser. 362, 37–43. https://doi.org/10.3354/meps07444.

Rao, R.B., Fung, G., Rosales, R., 2008. On the dangers of cross-validation. An experimental evaluation, in: Proceedings of the 2008 SIAM International Conference on Data Mining. SIAM, pp. 588–596.

Rilov, G., Fraschetti, Simonetta, Gissi, Elena, Pipitone, Carlo, Badalamenti, Fabio, Tamburello, Laura, Menini, Elisabetta, Goriup, Paul, Mazaris, Antonios D., Garrabou, Joaquim, Benedetti-Cecchi, Lisandro, Danovaro, Roberto, Loiseau, Charles, Claudet, Joachim, Katsanevakis, Stelios, Rilov, C., Fraschetti, S., Gissi, E., Pipitone, C., Badalamenti, F., Tamburello, L., Menini, E, Goriup, P. Mazaris, A.D., Garrabou, J., Benedetti-Cecchi, L. Danovaro, R., Loiseau, C., Claudet, J., Katsanevakis, S., 2020. A fast-moving target: achieving marine conservation goals under shifting climate and policies. Ecol. Appl. 30, e02009. doi: 10.1002/EAP.2009.

Sethi, S.S., Ewers, R.M., Jones, N.S., Sleutel, J., Shabrani, A., Zulkifli, N., Picinali, L., n.d. Soundscapes predict species occurrence in tropical forests. https://doi.org/10.1101/2020.09.24.311381.

Sethi, S.S., Jones, N.S., Fulcher, B.D., Picinali, L., Clink, D.J., Klinck, H., Orme, C.D.L., Wrege, P.H., Ewers, R.M., 2020. Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. Proc. Natl. Acad. Sci. U. S. A. 117, 17049–17055. https://doi.org/10.1073/pnas.2004702117.

Simpson, S.D., Meekan, M., Montgomery, J., McCauley, R., Jeffs, A., 2005. Homeward Sound. Science (80-.). 308, 221. doi: 10.1126/science.1107406.

Smith, J.E., Brainard, R., Carter, A., Grillo, S., Edwards, C., Harris, J., Lewis, L., Obura, D., Rohwer, F., Sala, E., Vroom, P.S., Sandin, S., 2016. Re-evaluating the health of coral reef communities: baselines and evidence for human impacts across the central Pacific. Proc. R. Soc. B Biol. Sci. 283, 20151985. https://doi.org/10.1098/rspb.2015.1985.

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. Mar. Pollut. Bull. 56, 1983–1991. https://doi.org/10.1016/j.marpolbul.2008.08.015.

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. Aquat. Mamm. 39. https://doi.org/10.1109/RIOAcoustics.2013.6683984.

Staaterman, E., Paris, C.B., DeFerrari, H.A., Mann, D.A., Rice, A.N., D'Alessandro, E.K., 2014. Celestial patterns in marine soundscapes. Mar. Ecol. Prog. Ser. 508, 17–32. https://doi.org/10.3354/meps10911.

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. https://doi.org/10.1007/s00338-012-1007-8.

Stowell, D., Sueur, J., 2020. Ecoacoustics: acoustic sensing for biodiversity monitoring at scale. Remote Sens. Ecol. Conserv. 6, 217–219. https://doi.org/10.1002/rse2.174.

Sueur, J., Aubin, T., Simonis, C., 2008. Equipment review: Seewave, a free modular tool for sound analysis and synthesis. Bioacoustics 18, 213–226. https://doi.org/10.1080/09524622.2008.9753600.

Sueur, J., Farina, A., Gasc, A., Pieretti, N., Pavoine, S., 2014. Acoustic indices for biodiversity assessment and landscape investigation. Acta Acust. United Acust. 100, 772–781. https://doi.org/10.3813/AAA.918757.

Villanueva-Rivera, L.J., Pijanowski, B.C., Villanueva-Rivera, M.L.J., 2018. Package 'soundecology.'.

Williams, S.L., Sur, C., Janetski, N., Hollarsmith, J.A., Rapi, S., Barron, L., Heatwole, S.J., Yusuf, A.M., Yusuf, S., Jompa, J., Mars, F., 2019. Large-scale coral reef rehabilitation after blast fishing in Indonesia. Restor. Ecol. 27, 447–456. https://doi.org/10.1111/rec.12866.

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. Anal. Chim. Acta 329, 257–265. https://doi.org/10.1016/0003-2670(96)00142-0.