TECNIACÚSTICA
CUENCA 2023

ANALYSIS OF THE UNDERWATER SOUNDSCAPE USING ECOACOUSTIC INDICES: THEORETICAL COMPILATION OF ECOACOUSTIC PARAMETERS

Bautista Parra, Nerea[1]; Affatati, Alice[2]; Poveda Martínez, Pedro[1*]; Carbajo San Martín, Jesús[1*]; Ramis Soriano, Jaime[1,2*]*

[1] Departamento de Física, Ingeniería de Sistemas y Teoría de la Señal, Universidad de Alicante
[2] Geophysics Research Section l, National Institute of Oceanography and Applied Geophysics (OGS), Trieste, Italy

ABSTRACT

Ecoacoustics deals with the study of the properties, evolution, and function of sounds in the environment. Ecoacoustics analysis is becoming an important way to quantify the ecological aspects of animal communication in a given terrestrial or underwater landscape (also referred to as "soundscape"). This approach relies on the importance of the identification of the different acoustic signals that occur in the environment and the ability to analyse how they affect the soundscape and the individuals living in it. This soundscape is complex and requires procedures to transform the data collected into information that is useful for understanding the environment. The purpose of this work is to provide a theoretical compilation of ecoacoustic indices that facilitate the understanding and study of the landscape, to be applied later, in the experimental part, in an underwater environment and to discuss the results obtained.

Keywords— ecoacoustic indices, anthropophony, biophony, soundscape, underwater acoustics.

1. INTRODUCTION

Ecoacoustics, known as the study of environmental sound, is becoming an important tool for quantifying ecological aspects of the landscape and is useful for investigating critical ecological issues such as biodiversity loss and even climate change [1].

Therefore, ecoacoustics is an expanding field that has great potential for biodiversity monitoring and is also an interesting alternative to the visual study of the landscape. Visual analysis and observations of the landscape have been used in geography since its beginnings, as they are the most direct way of studying the physiognomy of the Earth. However, in recent years, researchers have implemented different approaches that recognize other sensory channels which compensate the importance of what is seen [2]. Talking about sound, the soundscape has biological ("biophony"), geological ("geophony"), and anthropogenic ("anthropophony") contributions [3].

Nowadays, ecoacoustic analysis is an integral part of research into the ecological aspects of the landscape. The convenience and low cost of acoustic recording facilitates the accumulation of enormous amounts of sound data, that are difficult to listen to in their entirety. Sound acquisition systems provide a time series of recordings that can be transformed into the correct metrics and subsequently analysed to examine the acoustic temporal patterns present in different environments [4].

Research in this area has led in recent years to the birth of methods that improve the understanding and relationship of signals collected during biological and ecological processes. In this way, parameters and indices which help in the characterization of the soundscape emerge. These, among others, are the ecoacoustic indices. These indices have a great potential to expand the ways of monitoring underwater ecosystems, solving the lack of algorithms that limit such acoustic monitoring and being at the same time a non-invasive method, unlike traditional approaches.

Passive Acoustic Monitoring (PAM) is the technique mainly used to record soundscapes and provides a way to study marine organisms in their natural environment [5], without introducing any interference or stress factor or altering the soundscape.

In this work, an initially theoretical compilation of those ecoacoustic indices with a greater presence in underwater environments [1] will be given to understand how they are

[*] ***Contact author**: nbp27@alu.ua.es; pedro.poveda@ua.es; jesus.carbajo@ua.es; jramis@ua.es*

used and how they are applied to a sound environment through acoustic recordings, which are studied in the second part of this work.

2. ECOACOUSTIC INDICES

Ecoacoustic indices are mathematical functions designed to evaluate certain aspects of biodiversity and the soundscape [6]. The main use of these indices is exploited in ecological studies, so they can also be described as indicators of sound ecology that assess the spectral and temporal variation of sound emissions in the landscape, either terrestrial or underwater.

Numerous indices have been developed over time to evaluate the environment and analyse facets of the diversity of animal and plant communities. It is convenient to know that some of these indices mainly differentiate three types of sound emitted according to its source or origin. Bernard Krause [3], musician and ecologist, differentiated these type of sounds as follows:

- Geophony: natural sounds of non-biological origin, such as the sea, water currents, wind, earthquakes.
- Biophony: sounds produced by living organisms in a given environment or habitat.
- Anthropophony: sounds generated by humans and the activities they perform in the environment. These sounds can range from more controlled sounds, such as music, theatre, or language, to more incoherent and chaotic sounds, such as noise.

The indices aim to quantify richness, evenness, regularity, divergence, or rarity in species abundance, among many other traits. To facilitate the classification of the indices developed so far, they may be differentiated into α and β indices [6].

3. α INDICES

The α indices are those related to the number of entities (also known as richness) and the relative abundance of each entity (also known as evenness). Therefore, these types of indices aim to represent different attributes of an environment or habitat. Examples of these are richness, complexity, evenness, or heterogeneity of an acoustic community or soundscape. There are three categories:

- Intensity: indices that analyse the sound sample from the amplitude or intensity between signals.
- Complexity: indices that estimate a level of heterogeneity with the signal spectrum (frequency) or amplitude envelope (time).
- Type of soundscape: indices that consider the relationship between biophony, geophony, and anthropophony [3] of a soundscape.

The α indices used in the experimental part are the most used in underwater acoustics, which are shown in Table 1.

Table 1. α indices used in the experimental part of the work.

Name	Principle	Ref.
Acoustic Complexity Index (ACI)	Spectrogram complexity	[7]
Normalised Difference Soundscape Index (NDSI)	Ratio of anthrophony to biophony	[8]
Acoustic Entropy Index (H)	Envelope and spectrum complexity	[9]
Temporal Entropy (H_t)	Envelope complexity	[9]
Spectral Entropy (H_f)	Spectrum complexity	[9]
Acoustic Richness (AR)	Envelope complexity and intensity	[10]
Median of amplitude envelope (M)	Median of amplitude envelope	[10]
Acoustic Diversity Index (ADI)	Spectrum complexity	[11]

These indices, as mentioned below, may be affected by factors such as background noise (transient or permanent), relative loudness, temporal and/or frequency overlap between sounds from different sources. However, these indices are the most used in research as they aim to give a single value to characterise a soundscape.

3.1. Acoustic Complexity Index (ACI)

The ACI [7] is the second most widely used index in underwater acoustics. It was first developed to produce a direct quantification of the complexity of the soundscape by measuring the intensity variations that occur throughout a sound recording, despite the presence of anthropogenic sound [12], in the different frequency bands. It provides greater value to signals with higher amplitude. These frequency bins are usually default to 1 kHz. ACI were originally used by researchers to analyse bird vocalisations.

The theoretical basis for this index is that biotic sounds are characterised by a high variability of intensity, whereas anthropogenic noise or geophonic sounds are usually of constant intensity. For example, in the terrestrial environment, both geophonic sound (wind, water…) and the sound of roads or air traffic are usually monotonous and of lower intensity than those produced by animals such as birds and amphibians., which is why the ACI assigns them lower values. The calculation of this index is made by adding the absolute difference between two adjacent intensity values from the intensity matrix obtained by means of a Short-Time Discrete Fourier Transform (STDFT).

$$\begin{array}{cc} & \begin{array}{ccccc} n_1 & \cdots & n_j & \cdots & n_J \end{array} \\ \begin{array}{c} \omega_1 \\ \vdots \\ \omega_k \\ \vdots \\ \omega_K \end{array} & \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1J} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kj} & \cdots & a_{kJ} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{K1} & \cdots & a_{Kj} & \cdots & a_{KJ} \end{pmatrix} \end{array} \quad (1)$$

where a_{kj} are the Fourier coefficients obtained from the STDFT calculation, K is the number of frequencies, and J is the number of Fourier windows calculated in the signal.

If we consider a single frequency interval, which is a single row k of an STDFT matrix, the ACI calculates the derivative of the coefficients scaled by the sum of the coefficients as follows:

$$ACI_j = \sum_{j=1}^{J-1} \left(\frac{|a_{j+1} - a_j|}{\sum_{j=1}^{J} a_j} \right) \quad (2)$$

The calculation is processed for each frequency interval k, and the total is summed so that we obtain for a single time interval i. Starting from Eq. 2, we obtain this in the form:

$$ACI_{kj} = \sum_{k=1}^{K} ACI_j \quad (3)$$

Finally, if all the intervals i are considered, the final ACI is obtained.

$$ACI_{kji} = \sum_{i=1}^{I} ACI_{kj} \quad (4)$$

Theoretically, ACI is reliable for determining biodiversity, which will be lower with lower values of ACI.

3.2. Normalised Difference Soundscape Index (NDSI)

The NDSI was developed to estimate the relative amount of anthropogenic and biological components in a soundscape [8] and it categorizes the soundscape into anthropophony and biophony (or in some cases geophony and biophony). It. It compares the level of anthropogenic disturbance of the soundscape by calculating the ratio of human-generated to animal-generated acoustic components. It has originally been used in environments collected in an acoustic library. It has also been used in other terrestrial studies with forest vegetation, looking for minimum human intervention [13]. The ratio is calculated as follows:

$$NDSI = \frac{b-a}{b+a} \quad (5)$$

where b is the level of biophony and a is the level of anthropophony.

The soundscape energy level is separated into frequency components by evaluating the frequency spectrum (power spectral density) with the Welch's method [14], which is discretised to a resolution of 1 kHz. Each component is the sum of the power based on 1 kHz frequency bins in each range.

The frequency ranges defined for each level consider mechanical signals (anthropophony) to occupy the frequency range from 1 to 2 kHz (sometimes also considered 0.2 to 1.5 kHz), while biophony occupies the bands from 2 to 8 kHz (sometimes also considered up to 11 kHz) [15, 16]. However, these values can be modified depending on the soundscape being explored and the sampling rate of the recording.

The NDSI can take values from -1 to 1, where the last one indicates a biological sound, free of anthropophony. Therefore, the more negative the value, the more anthropogenic disturbance the soundscape will present.

It may also be the case that geophonic sounds - such as wind or rain -occur in the same frequency range as anthropophony. These sounds can cover the spectrum at lower frequencies with more energy. In these study cases, the NDSI may reflect the relationship between geophony and biophony rather than between anthropophony and biophony. Even so, the calculation will be considered the same.

3.3. Acoustic Entropy Index (H)

Acoustic entropy (H), or simply entropy, determines the species richness in an acoustic habitat or space. Mathematically it is the product of spectral entropy (H_f or H_s) and temporal entropy (H_t) [9].

$$H = H_t \times H_f \quad (6)$$

Its results are on a scale from 0 to 1, where 0 indicates purer tones, and tends to 1 for random noise. The value increases as the number of frequency bands and amplitude modulations increases.

One of the main issues of entropy is that, if there is a lot of geophony or anthropophony, the index is less reliable and produces false high values. This seems to occur when background noise dominates the recordings, as is often the case in habitats where diversity is lower. On the other hand, the original study, where biological sound is high and background noise is very low, provides expected results [9].

Temporal Entropy (H_t) is calculated from the Hilbert amplitude envelope of the signal. H_t estimates the Shannon uniformity [17] of this envelope. If it is considered a signal $x(t)$ with length n, the oscillation amplitude envelope is obtained by the analytical signal $\xi(t)$ of $x(t)$. This signal is defined as:

$$\xi(t) = x(t) + i x_H(t) \quad (7)$$

where i^2 is proportional to -1, and $x_H(t)$ is the Hilbert transform of the signal $x(t)$.

The analytical signal can give access to both the instantaneous amplitude envelope and the instantaneous frequency, using the latter in the H_f. The probability mass function of the amplitude envelope *A(t)* is obtained as:

$$A(t) = \frac{|\xi(t)|}{\sum_{t=1}^{n} |\xi(t)|} \tag{8}$$

where:

$$\sum_{t=1}^{n} A(t) = 1 \tag{9}$$

This means that the amplitude envelope is scaled by its sum, so that the sum of the sample values is equal to 1.

Referring to signal theory, the entropy H of a random variable X with a probability mass function $p_X(x)$ is defined as:

$$H(X) = -\int_{-\infty}^{+\infty} p_X(x) \times log_2 p_X(x)\ dx \tag{10}$$

Considering this and applying Shannon uniformity [17], the temporal entropy is calculated as:

$$H_t = -\frac{\sum_{i=1}^{n} A(t) \log_2(A(t))}{\log_2(n)} \tag{11}$$

where $log_2(n)$ is the number of categories. Categories are usually species that are differentiated by their relative abundance in a community. Temporal entropy ranges from 0 to 1.

A noisy signal with a large number of amplitude modulations tends towards 1, while a calm signal will tend towards 0. However, it can also happen that a sustained sound with a nearly flat envelope shows a high temporal entropy.

Spectral entropy (H_f) is similar, but it works in the frequency domain. It is obtained by applying Shannon uniformity [17] to the mean frequency spectrum scaled by its integral, replacing species by frequency intervals. In this way, we start by applying a STFT to the signal to obtain the mean spectrum *s(f)*, which is transformed into a probability mass function, as it is done with H_t. It then becomes known as *S(f)* and has a size of N samples. Mathematically it is expressed as:

$$H_f = -\frac{\sum_{f=1}^{N} S(f) \log_2(S(f))}{\log_2(N)} \tag{12}$$

The index is also limited between 0 and 1, following the same reasoning as that taken for temporal entropy.

3.4. Acoustic Richness (AR)

The AR [10] is an index based on the temporal entropy index [9] and the median of the amplitude envelope, resulting in the index M. It reflects the amplitude of an acoustic recording. M is calculated as:

$$M = median\big(A(t)\big) \times 2^{1-deph} \tag{13}$$

where *A(t)* is the amplitude envelope and *depth* are the digitalization bits of the recorded signal.

The index can take values between 0 and 1. Louder recordings will give higher values, reflecting louder soundscapes and more presence of geophony (especially in the case of storms), while lower values will be produced by very quiet recordings, with almost no biophony or geophony. The AR takes the signal size of the M index, and relates it to H_t.

$$AR = \frac{rank(M) \times rank(H_t)}{n^2} \tag{14}$$

where M is the median amplitude envelope, H_t is the time entropy, and n is the number of files analysed.

The calculation is based on a set of files, as shown in the equation. First the M and H_t indices will be calculated for each file individually and then the result is sorted in ascending order. The position of each recording is used to calculate the AR index, which is highly dependent on the selected set of files.

The results of this index were successfully used in bird species identification and take the same values as the indices on which it depends. Higher values indicate a greater richness of the soundscape.

3.5. Acoustic Diversity Index (ADI)

The ADI [11] uses the Shannon uniformity index [17] applied to the spectral content, as does the spectral entropy H_f [9]. This index calculates the STDFT of the signal and divides it into short frequency bins (e.g., 1 kHz). From each division, the relative amplitude above a certain threshold (usually specified in dB) is selected and the Shannon index is then applied to the chosen intervals. Thus, the index determines how full the selected intervals are, which indicates the degree of occupancy of the different acoustic niches in the recordings. It is expressed as follows:

$$ADI = \sum_{i=1}^{S} p_i \ln(p_i) \tag{15}$$

where p_i is the fraction of sound in each frequency band i and S is the number of frequency bands.

By default, the bandwidth between 0 and 10 kHz is divided into 10 intervals, where the selected threshold is -50 dB to eliminate background noise [18]. However, these parameters can be modified according to the specifications of the soundscape.

This index increases with the uniformity of the frequency bands. A uniform signal (noisy or completely silent in all frequency bands) will give a high value, while a punctual tone (energy concentrated in a short period of time) will be closer to 0.

Therefore, higher values are usually produced by high levels of geophony or anthropophony covering the noise spectrogram, or otherwise quiet recordings with no variations in the frequency bands. On the other hand, low values are produced by the dominance of a narrow band of frequencies, such as for example the night-time sound of some insects.

4. β INDICES

The β indices focus on estimating the acoustic dissimilarity or disparity between communities of organisms or between two-time intervals in the same community [6]. In other words, these indices help determine the extent to which two or more acoustic communities (or soundscapes) are acoustically different. They can also be used to evaluate changes in the same soundscape or community between two different dates.

Despite their similarity to traditional measures for these types of communities, some of the studies using these types of indices have several problems related to frequency, time, and amplitude. Some of them are the sensitivity of the sensor to the overlapping of the community's song, proximity or remoteness of the sound source, background noise, etc. These can lead to amplitude variations that should not be interpreted as relevant differences to the study.

Although there is great interest in comparing sounds to identify species or individuals, the methods used are adapted to closely related sounds (such as vocalisations produced by a single individual) and are generally very simple indices that need improvement. In most cases they would not be optimal for comparing sounds coming from complete and complex communities and landscapes where it is difficult to define similarities.

The experimental work carried out in the second part covers firstly the analysis of the general soundscape without focusing on any specific sound community in the area. However, recordings of another soundscape with different characteristics are also available. Therefore, although precise results are not assured, the Acoustic Dissimilarity Index (D) is defined below to compare the two environments and to see how their differences are reflected.

4.1. Acoustic Dissimilarity Index

The Acoustic Dissimilarity Index (D) [9], as its name indicates, evaluates the dissimilarity between two different communities or soundscapes from temporal and spectral data. Similarly, to H, it is the product of a multiplication, in this case, between spectral dissimilarity (D_f) and temporal dissimilarity (D_t).

$$D = D_t \times D_f \tag{16}$$

The dissimilarity between two different signals, which must have the same duration and sampling frequency, is evaluated by calculating the difference between the amplitude envelope (Eq. 8) of each recording divided by 2.

$$D_t = -\frac{\sum_{t=1}^{n} |A_1(t) - A_2(t)|}{2} \tag{17}$$

In turn, the spectral dissimilarity (D_f) is calculated in the same way by replacing the amplitude envelope by its Short-Time Fourier Transform (STFT), obtaining the signal *S(f)* (mean spectrum) as it was calculated to find H_f. The equation is then as follows.

$$D_f = -\frac{\sum_{f=1}^{N} |S_1(f) - S_2(f)|}{2} \tag{18}$$

The values between which D oscillates vary from 0 to 1. To test its reliability, Sueur and colleagues [9] simulated recordings with vocalisations of different species. These were randomly mixed to obtain recordings to be compared that were different in a greater or lesser number of species. In the simulation, those recordings that shared a smaller number of species increased the value of D, while those recordings with more homogeneous species provided a lower value. Therefore, it can be said that the D index could be used to argue for differences between sound communities, with higher values being obtained the greater the dissimilarity.

5. CONCLUSIONS AND FUTURE LINES OF ACTION

In this first part of the work, the main ecoacoustic indices used in underwater environments have been described for informative and theoretical purposes. The experimental part is based on recordings taken in the deep sea. The application of these indices to the underwater environment is initially complex, as seas and the ocean are noise-filled environments. All these indices were created for the analysis of terrestrial habitats, just as it has been observed in this work. Some of them even started by being tested in simulated environments, where the recordings contained the exact results that made the theory for each index accurate and rigorous.

On the other hand, there is a need to create underwater studies to serve as a baseline in different environments, both polluted with anthropogenic sound and filled with biological presence. Each ecosystem is unique and has varied sounds that make it peculiar and different from the rest. The scarcity of reference data related to the study of underwater habitat makes it difficult to compare results, patterns, trends, etc.

Therefore, after the theoretical compilation, the aim is to transfer these indices to water, even though the acoustic and propagation characteristics change from one medium to another. Further experimental studies will be carried out identifying common patterns and differences, making a final discussion of the viability of the studied indices in marine environments, contributing to increase the reference for future works.

ACKNOWLEDGEMENTS

This publication is part of the project PCI22022-135081-2, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR", where MCIN stands for the Ministry of Science and Innovation; AEI for the State Research Agency; 10.13039/501100011033 for the DOI (Digital Object Identifier) of the Agency; and PRTR for the acronym of the Plan for Recovery, Transformation and Resilience.

REFERENCES

[1] N. Pieretti, and R. Danovaro, "Acoustic indexes for marine biodiversity trends and ecosystem health," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 375, no. 1814, 20190447, 2020.

[2] G. Rose, "On the Need to Ask How, Exactly, Is Geography "Visual"?," *Antipode*, vol. 35, no. 2, pp. 212-221, 2003.

[3] B. Krause, *The Great Animal Orchestra: Finding the Origins of Music in the World's Wild Places*, Little, Brown, United States, March 2012.

[4] A. Farina, and S. H. Gage, *Ecoacoustics: The Ecological Role of Sounds*, Wiley, United Kingdom, July 2017.

[5] C. Desjonquères, T. Gifford, and S. Linke, "Passive acoustic monitoring as a potential tool to survey animal and ecosystem processes in freshwater environments", *Freshwater Biology*, vol. 65, no. 1, pp. 7-19, 2020.

[6] J. Sueur, A. Farina, A. Gasc, N. Pieretti, and S. Pavoine, "Acoustic Indices for Biodiversity Assessment and Landscape Investigation," *Acta Acustica United With Acustica*, vol. 100, no. 4, pp. 772-781, 2014.

[7] N. Pieretti, A. Farina, and D. Morri, "A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI)," *Ecological Indicators*, vol. 11, no. 3, pp. 868-873, 2011.

[8] E. P. Kasten, S. H. Gage, J. Fox, and W. Joo, "The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology," *Ecological Informatics*, vol. 12, pp. 50-67, 2012.

[9] J. Sueur, S. Pavoine, O. Hamerlynck, and S. Duvail, "Rapid Acoustic Survey for Biodiversity Appraisal," *PLOS ONE*, vol. 3, no. 12, pp. 1-9, 2008.

[10] M. Depraetere, S. Pavoine, F. Jiguet, A. Gasc, S. Duvail, and J. Sueur, "Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland," *Ecological Indicators*, vol. 13, no. 1, pp. 46-54, 2012.

[11] L. J. Villanueva-Rivera, B. C. Pijanowski, J. Doucette, and B. Pekin, "A primer of acoustic analysis for landscape ecologists," *Landscape Ecology*, vol. 26, no. 9, pp. 1233-1246, 2011.

[12] A. Farina, and D. Morri, "Source-sink e eco-field: ipotesi ed evidenze sperimentali. Atti del X congresso nazionale della SIEP-IALE. Ecologia e governance del paesaggio: esperienze e prospettive," *Bari*, pp. 365-372, 2008.

[13] S. H. Gage, and A. C. Axel, "Visualization of temporal change in soundscape power of a Michigan lake habitat over a 4-year period," *Ecological Informatics*, vol. 21, pp. 100-109, 2014.

[14] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *in IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70-73, 1967.

[15] S. H. Gage, and B. M. Napoletano, "Envirosonics equipment and operations manual," *Computational Ecology and Visualization Laboratory*, Michigan State University, 2004.

[16] S. H. Gage, B. M. Napoletano, and M. C. Cooper, "Assessment of ecosystem biodiversity by acoustic diversity indices," *The Journal of the Acoustical Society of America*, vol. 109, no. 5, pp. 2430-2430, 2001.

[17] C. E. Shannon, and W. Weaver, *The mathematical theory of communication*, Urbana, Illinois University Press, United States, 1949.

[18] B. K. Pekin, J. Jung, L. J. Villanueva-Rivera, B. C. Pijanowski, and J. A. Ahumada, "Modeling acoustic diversity using soundscape recordings and LIDAR-derived metrics of vertical forest structure in a neotropical rainforest," *Landscape Ecology*, vol. 27, pp. 1513-1522, 2012.

[19] A. Farina, and P. Li, *Methods in Ecoacoustics: The Acoustic Complexity Indices*, Springer Nature, Berlin, 2021.