The Amazon rainforest soundscape characterized through Information Theory quantifiers

Juan G. Colonna[1,†,*], José Reginaldo H. Carvalho[1,‡], Osvaldo A. Rosso[2,††]

1 Instituto de Computação (IComp), Universidade Federal do Amazonas (UFAM), Manaus, Amazonas, Brasil.
2 Instituto de Física, Universidade Federal de Alagoas (UFAL), Brasil.

†Writing - Original Draft Preparation, Conceptualization, Investigation, Methodology, Software, Visualization.
‡Writing - Review & Editing, Project Administration, Supervision, Validation, Investigation.
††Writing - Review & Editing, Validation.
* juancolonna@icomp.ufam.edu.br (JGC)

Abstract

Automatic monitoring of biodiversity by acoustic sensors has become an indispensable tool to assess ecological stress at an early stage. Due to the difficulty in recognizing the Amazon's high acoustic diversity and the large amounts of raw audio data recorded by the sensors, the labeling and manual inspection are not feasible. Therefore, we propose a new ecoacoustic index that allows us to quantify the complexity of an audio segment and correlate such complexity with the soundscape's biodiversity. The approach uses unsupervised methods to avoid the problem of labeling each species individually. The proposed index is called the Ecoacoustic Complexity Index (ECI) and makes use of Statistical Complexity, a measure of Information Theory Quantifiers. An interesting feature of this index is that each audio segment, even of variable length, can be characterized as a single two-dimensional point in the Causal plane, helping us understand the ecoacoustic dynamics of the rainforest. The main results show regularity in the ecoacoustic richness of a floodplain considering different temporal granularities, be it between hours of the day or between consecutive days of the monitoring program. We found that this regularity does a good job of characterizing the soundscape of the environmental protection area of Mamirauá, in the Amazon, differentiating between species richness, environmental phenomena, and insects.

Introduction

Recent research shows that climate change modifies the Earth's natural soundscapes [1], for animal species are sensitive to their environmental conditions. This observation has driven many researchers to monitor variations of animal populations through time, and use them as an indicator of environmental degradation [2]. There are currently two ways to achieve this goal: by acoustic surveys or by acoustic diversity indices. Acoustics surveys are the most widely used method to monitor animal populations, and it takes advantage of the animal vocalization capability [3]. However, a solid survey in remote tropical areas, such as the Amazon rainforest, demands significant investment of both human and financial resources as well as expert knowledge. Diversity indices, on the

other hand, belongs to a broader class of methods capable of quantifying diversity without needing to recognize each particular sound [4–6].

Our goal is to propose a new acoustic index capable of tracking the ecoacoustic status of a given landscape, promoting the understanding of the variation in animal populations. The proposed Ecoacoustic Complexity Index (ECI) is unsupervised, meaning that no human expertise is needed to label and classify long audio records. Information Theory provides powerful and elegant methods for signal analysis, such as the Entropy-Complexity causality plane theory to map generalized statistical complexity. This new ECI index benefits from two classical Information Theory quantifiers: the Shannon entropy and the Jensen-Shannon divergence, which, when combined, are able to differentiate between a broad range of signal features, in both temporal and spectral domains. Besides, the dissimilarity between two recordings can be naturally quantified, since our index was defined in terms of the Jensen-Shannon divergence.

The proposed index can be applied to acoustic signals with variable temporal length, only requiring the adjustment of a single parameter (τ). To compose the ECI, we combined the *Von Neumann* entropy, calculated from the eigenvalues of the autocorrelation matrix, with the *Statistical Complexity*. It is worth mentioning that this new index encapsulates the representation of any acoustic signal as a two-dimensional point in the Entropy-Complexity (HxC) causality plane [7]. In addition to its low-dimensional representation, each point on the HxC plane has useful interpretations regarding the underlying physical system. Therefore, the ECI characterizes nontrivial sound correlations.

The usefulness of the ECI was evaluated on a set of continuous recordings, generated by a sensor node positioned in a preserved floodplain area called *Mamirauá*, located in the central Amazon rainforest. In this remote, hard-to-reach location, biodiversity remains virtually unchanged. Furthermore, the region is under the influence of the flood cycle of the Amazon river, reaching its maximum typically in June. The available records are from July (water level is high) and September (water level is close to the minimum), representing the two characteristic seasons of this region well. These recordings allow us to obtain a baseline to compare with acoustic landscapes modified by human intervention in future studies.

In summary, the main contribution of this paper is a new ecoacoustic data analysis tool from the perspective of Information Theory Quantifiers and statistical measures. If we consider a soundscape as a complex physical system, with ECI we can study and interpret the acoustic dynamics of this system as a whole. Note that we are not referring to the dynamics of a particular species. Additionally, we provide the source code for experiment replication together with the figures script at `http://bit.ly/2m12PWc`.

Related works

Most ecoacoustic recognition methods rely on supervised classifiers to identify and catalog species. Among the main species that have been recognized with these methods, we can highlight frogs, birds, whales, dolphins, elephants, mosquitoes, gibbon, and others [8–14]. Fully automatic multi-species monitoring systems have also been proposed, integrating hardware and software on a single platform [15]. All of these works faced the same challenge, a manual labeling of long audio recordings. The difficulty of labeling large databases and the decreasing of classifier performance when the number of classes increases, is still an open issue.

An alternative to deal with this problem is to use unsupervised methods [16]. These methods do not require labeled data and are particularly well suited to analyze acoustic diversity from a given soundscape, such as tropical environments [17]. However, determining the number of clusters or validating the content of each recording group,

also requires expert knowledge, being time consuming and subjective. Beside that, clusters could be heterogeneous in different ways, so assigning a biodiversity score to each group is not a trivial task.

Another type of unsupervised approach has emerged in recent years. Most of them define an acoustic index, usually linked to the acoustic richness of the audio recordings. Some examples of these are the Vocal Activity Index (VAI) [16], the Acoustic Complexity Indices (ACI) [6, 18] and the Acoustic Entropy Index (AEI or simply H) [4].

The VAI is not completely unsupervised [16]. The authors proposed training a binary classifier to predict whenever there is a bird call in a recording segment. After that, the number of segments predicted as the positive class is divided by the total segments recorded over the same time period. The authors also validated an alternative to estimate this index using an unsupervised clustering method. After determining the optimal number of clusters, Principal Component Analysis (PCA) should still be applied to reduce the dimensionality of the segments. All these processing steps, along with the manual tuning of parameters, make it difficult to deploy this method to a low-cost acoustic sensor.

Another alternative is the ACI [18]. This index is calculated from a spectrogram using the Short-Time Fourier Transform (STFT) algorithm. The spectrogram is a matrix with the number of rows as the number of frequency bins obtained with the STFT and the number of columns as the signal length divided by an overlap factor. Most modern hardware can perform STFT without any difficulty. The final ACI calculation is based on the aggregated values of the spectrogram. Therefore, the number of rows and columns of the spectrogram change as we change both the STFT resolution and the signal length. This mutual dependence between resolution and overlap makes it difficult to compare results when the rules of the monitoring program change due to hardware updates or even when the hardware fails. Furthermore, this index was developed with the purpose of being used as a feature for classification methods and not as a measure of acoustic richness.

Finally, Sueur et al. [18] proposed to multiply the Spectral Entropy (H_f) and the Temporal Entropy (H_t) to compute the AEI. To calculate H_f it is assumed that the normalized frequency spectrum, obtained by the Fast Fourier Transform (FFT), is a histogram or a Probability Density Function (PDF). Similarly, one should assumed that oscillations in the amplitude envelope of the signal can be used as a histogram to obtain H_t. The final index is $H = H_f \times H_t$, so the closer H to one, the greater the acoustic diversity. One may notice that short-term impulsive noises can drastically alter the value of this index. Unfortunately, the multiplication of two entropy values lacks physical interpretation. However, this index inspired our proposal, for we believe that entropy, or measures from Information Theory, are natural choices to quantifying diversity.

The AEI and ACI indices have two drawbacks. First, the increase of quasi-white random noise increases entropy, hence H values close to one cannot be always interpreted as high acoustic diversity produced by vocalization patterns of different species. Second, these indices depend directly on the length of the audio recorded by the sensor. Therefore, signals with variable lengths change the index ranges, making the values no longer comparable. We, then, concluded that there is room for improvement and we propose an index that is not affected by the issues mentioned above.

Methods

In this section we present the three fundamental concepts used to calculate the proposed index: (a) the autocorrelation matrix, (b) the *Von Neumann* entropy, and (c) the Statistical Complexity measure.

Autocorrelation Matrix

The Pearson's correlation coefficient (r_{xy}) is a measure of the intensity and direction of the linear relationship between two signals x and y. The autocorrelation coefficient r_{xx} has a similar interpretation, but instead of using two different signals, it uses a version of the same signal shifted by τ units. For instance, if $\tau = 1$ then r_{xx} quantifies the strength of the association between x_i and x_{i+1} [19].

Let $X = \{x_1, x_2, \dots x_N\}$ be the acoustic signal of length N at the sensor input; the unbiased autocorrelation coefficient r_{xx} is defined as:

$$r_{xx}(\tau) = \frac{1}{(N-\tau)s^2} \sum_{i=1}^{N-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x}), \tag{1}$$

where N, $\bar{x}$, and s are the length of x, the sample mean, and the sample standard deviation, respectively. Here, the maximum value of τ must satisfy the condition $\tau_{max} \ll {}^{N}/_{2}$. This equation may also be referred to as the Autocorrelation Function (ACF).

Given a maximum τ_{max} value, the autocorrelation matrix R_{xx} can be formed as a Toeplitz matrix, or diagonal-constant matrix, with shape [20]:

$$R_{xx}(\tau) = \begin{bmatrix} 1 & r_{xx}(1) & r_{xx}(2) & \cdots & r_{xx}(\tau_{max}) \\ r_{xx}(1) & 1 & r_{xx}(1) & \cdots & r_{xx}(\tau_{max}-1) \\ r_{xx}(2) & r_{xx}(1) & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & r_{xx}(1) \\ r_{xx}(\tau_{max}) & r_{xx}(\tau-1) & r_{xx}(\tau-2) & \cdots & 1 \end{bmatrix}, \tag{2}$$

where $r_{xx}(0) = 1$.

These autocorrelation coefficients are efficiently calculated by the Fast Fourier Transform (FFT) algorithm. Regardless of how they are obtained, it is well known that these coefficients carry information about the signal's main frequencies. Therefore, we can consider them spectral features that accurately describe ecoacoustic signals. For more details on this, please refer to S1 Appendix.

Entropy methodology

The Von Neumann entropy was defined by 1927 for quantum measurement processes [21]. It has a fundamental role in studying correlated systems. This Information Theory quantifier is defined as the normalized Shannon entropy (H) of the singular spectrum as:

$$H[P] = \frac{-1}{\log(\tau_{max})} \sum_{i=1}^{\tau_{max}} \left(\frac{\lambda_i}{\sum_i^{\tau_{max}} \lambda_i} \right) \log \left(\frac{\lambda_i}{\sum_i^{\tau_{max}} \lambda_i} \right), \tag{3}$$

where λ_i are the eigenvalues of a given R_{xx} matrix. It is worth mentioning that, the denominator term $\sum_i^{\tau_{max}} \lambda_i$ normalizes the eigenvalues values between $0 \leq \lambda_i \leq 1$. Hence, the whole term ${}^{\lambda_i}/_{\sum_i^{\tau_{max}} \lambda_i}$ can be interpreted as an histogram. Lastly, the term ${}^{-1}/_{\log(\tau_{max})}$ normalizes the entropy within $0 \leq H[P] \leq 1$.

Generalized Statistical Complexity measure

The original proposal of López-Ruiz *et al.* [22] and the extended work of Rosso *et al.* [23] define the Generalized Statistical Complexity Measure as the functional product:

$$C[P] = Q[P, P_e]H[P] \tag{4}$$

where $H[P]$ is the normalized entropy (equation 3), P is a normalized histogram obtained from the eigenvalues of R_{xx}, P_e is a reference histogram with uniform distribution, and the disequilibrium $Q[P, P_e]$ is defined in terms of the Jensen-Shannon divergence $J[P, P_e]$. That is:

$$Q[P, P_e] = Q_0 J[P, P_e] \tag{5}$$

with:

$$J[P, P_e] = H\left[\frac{(P + P_e)}{2}\right] - \frac{H[P]}{2} - \frac{H[P_e]}{2} \tag{6}$$

where Q_0 is a normalization constant used to maintain $0 \leq Q \leq 1$. For more details about Q_0, please refer to S2 Appendix. Note that $Q[P, P_e]$ depends on two different probability distributions. The first one, P, is related to the signal under analysis, and the second, P_e, is an uniform distribution, which represents a white noise signal. This reference histogram with uniform distribution is considered the equilibrium point of any physical systems.

The proposed Ecoacoustic Complexity Index

Given the fundamental concepts in the previous section, we can now summarize the proposed ECI calculation. For any ecoacoustic signal obtained by a sensor node, these steps must be followed:

1. from a signal X, apply the autocorrelation (equation 1) choosing a maximum value of τ_{max} to obtain a Toeplitz matrix R_{xx} (equation 2);
2. apply the Singular Value Decomposition (SVD) on R_{xx} to recover its singular spectrum (*i.e.*, the eigenvalues $\lambda_{1:\tau_{max}}$ of R_{xx}); then,
3. normalize each eigenvalue λ_i by the sum of all eigenvalues to get the histogram P and calculate the normalized entropy $H[P]$ according to the equation 3;
4. compute the Jensen-Shannon divergence using the histogram P and a uniform histogram P_e (equation 5 and equation 6); and finally,
5. apply equation 4 to estimate the ECI, the complexity index of the ecoacoustic signal.

This procedure maps each signal to a unique point, with H and C coordinates, in the Entropy-Complexity (HxC) plane. Note that, R_{xx} is a full rank matrix, therefore all its eigenvalues, as well as their normalized values, are positive real numbers. This allows us to interpret the singular spectrum as a Probability Density Function (PDF) of a signal recorded by any acoustics sensor. Moreover, the only free parameter of our method is τ_{max}. This methodology holds useful properties, from a practical standpoint. For instance, when an ecoacoustic signal has no deterministic patterns, lacking significant correlations, its singular spectrum tends to be flat, causing $H[P] \approx 1$. On the contrary, when the signal is purely deterministic, the entropy tends to its minimum $H[P] \approx 0$. As the complexity function is concave, we find the maximum complexity value between the two extremes of $H[P]$, as illustrated in Fig. 1a.

In some specific situations, two or more different signals may have singular spectra with the same entropy. For instance, consider the three points p_1, p_2, and p_3 depicted in Fig. 1b. In these cases, divergence plays a key role in helping to separate their complexities. In other words, divergence is useful to separate histograms with equal entropy, as in the Entropy-Complexity plane depicted in Fig. 1b. One may note that entropy weighted by the divergence causes a large range of possible C values for each

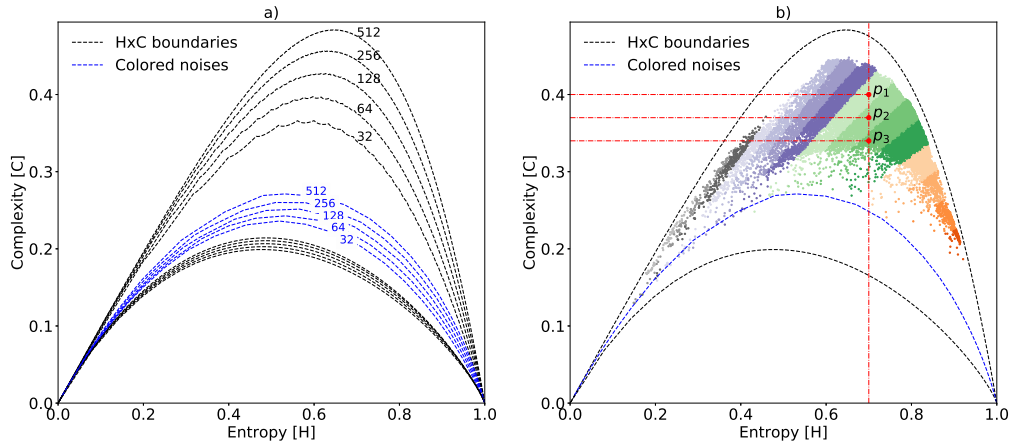

Fig 1. Boundaries of the HxC plane and the color noises reference line. (a) Maximum and minimum boundaries of the Generalized Statistical Complexity represented according to the chosen time lag τ_{max}. Here, the blue dashed line shows the variation of the correlated noises according to α. (b) Distribution of some acoustic samples from our dataset when $\tau = 512$. Points p_1, p_2 and p_3 have a singular spectrum with equal entropy but different complexity. The color gradient illustrates regions of the plane in which the samples have similar divergence.

entropy value. Such values are contained between the upper and lower limits, shown in the same figure by the black dashed lines. In this same figure, we illustrate the distribution of the audio samples from our dataset using color gradients to highlight the effect caused by the Jensen-Shannon divergence term in equation 4.

Lastly, the dashed blue line in Fig. 1 is the position of simulated color noise signals by a function with Power Spectral Density (PSD) that obeys a power law of the form $\xi(f) = \frac{1}{|f|^{\alpha}}$ [24, 25], where f denotes frequency. Thus, we vary the alpha parameter between $0 \leq \alpha \leq 2$ in small increments generating several random time series, each of them with a sampling frequency of $f_s = 44.1$kHz and a Signal-to-Noise Ratio of SNR $= 0$dB. The position on the HxC plane of each one of these series draws the blue dashed reference line shown in Fig. 1. For more details about colored noises, please refer to S3 Appendix.

Results

In this section, we will analyse the ECI computed from our experiments in the field.

Characterization of reference samples

Sueur *et al.* [4] made publicly available seven signal records specifying the acoustic richness of each one according to their index. Thus, we can use them as references and investigate the characteristics of these signals through the ECI. This allows us to validate and compare the proposed index.

Fig. 2a shows the complexity of these signals. According to Sueur *et al.*, the recorded signals s_{19} and s_{18} have an elevated biodiversity richness, for their higher entropy values. The ECI provided deeper information, as we realized that the increase in entropy may be a consequence of both: the cricket's chirping (or other insects) or a raise in environmental noises. Thus, ECI will not just label those samples with high

biodiversity richness, but it will provide more information about them. Take, for instance, the following examples.

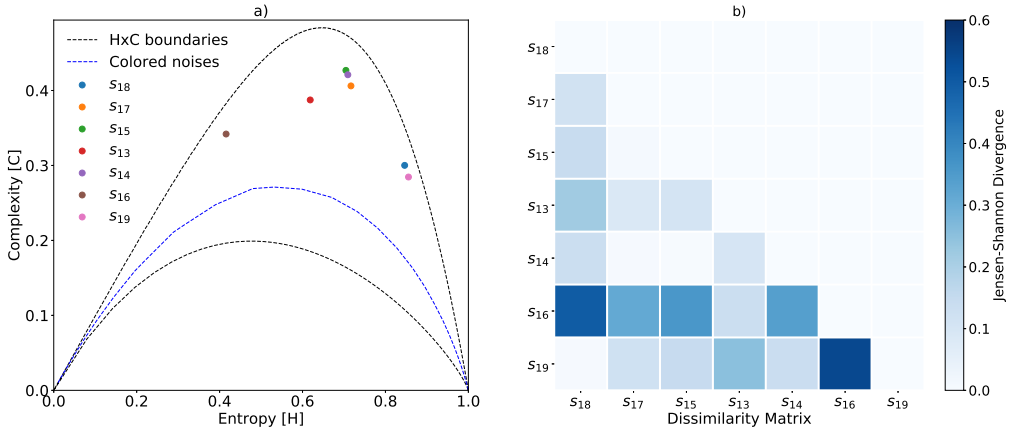

Fig 2. Characterization of reference signal samples using $\tau_{max} = 512$. Every point in subfigure a) represents a ecoacoustic signal. Subfigure b) shows the divergence matrix between signal pairs. We keep the original numbering of the signals to help the interpretation and comparison with Sueur *et al.* [4]. These records are available online at `https://doi.org/10.1371/journal.pone.0004065`

By employing the ECI, we figure out that signals s_{14}, s_{15} and s_{17} carry a larger number of different deterministic patterns, increasing complexity. This indicates higher ecoacoustic richness. The characterization of s_{13} shows that this sample has slightly more noise with few additional vocalization patterns compared to s_{16}, but keeping less diversity than, for instance, s_{14}. Lastly, s_{16} has the smallest diversity of acoustic patterns, and therefore the smallest complexity and entropy. One has to understand the role of τ in the HxC plane. A small value of τ results in a low-discriminant plane, being difficult to visualize the differences among signals. A high value of τ, on the other hand, will divide the main frequencies captured by the auto-correlation matrix into several groups. Thus, there is a trade-off when tuning τ.

One of the advantages of using Jensen-Shannon divergence is that we can quantify the difference between two arbitrary signals. Thus, we can make peer-to-peer comparisons between a given signal and a reference signal from another soundscape or between signal variations of the same soundscape at different times. The matrix of Fig. 2b shows the divergence between all pairs of reference signals. The colorbar indicates that the greater the divergence between the signals, the stronger the blue color. It is worth noting that there is a correspondence between the distance of the points in the complexity plane and their divergences. For instance, the greatest divergence is obtained by comparing the singular spectra of s_{16} and s_{19}, which are the farthest points in the HxC plane. A similar observation holds when comparing points s_{16} and s_{18}. Although the position of the points in the HxC plane uses the uniform distribution as a reference, the dissimilarity between them has a correspondence with the divergences of the signals; that is, the closer the points are in the HxC plane, the greater is their ecoacoustic similarity.

Characterization of samples from *in situ* monitoring

Our second investigation used a dataset with signals recorded in the Mamirauá conservation area, in the Brazilian Amazon rainforest. Our acoustic sensor captured

approximately 40Gb of raw audio signals, during thirteen days of the monitoring program. The audios were stored in segments with 22 seconds of duration, with a sampling frequency of $f_s = 44.1$kHz, summing up 43348 audio samples. The monitoring program, called *Providence*, was divided into two phases. The former, consisting of five consecutive days of recording in July 2016, and the latter, consisting of eight consecutive days at the beginning of September in the same year. All collected data compose a single data set. Authorization for the use of this sensitive biological data has been released by The Biodiversity Authorization and Information System (SISBio) through number 72722. The raw signals cannot be made available due to legal restrictions. A copy of the original data can be requested through the Providence website `http://projectprovidence.org`.

Fig. 3 depicts the relationship between the main elements of our analysis. The key step of our methodology is how to obtain the probability distributions to calculate entropy. As previously pointed out, the histogram of equation 3 comes from the normalized eigenvalues of the autocorrelation matrix. Fig. 3a shows the complexity plane for $\tau_{max} = 512$, where the 43348 segments were represented as ECI points. According to our methodology, points with greater complexity should exhibit a greater acoustic richness. Therefore, to illustrate the behavior of the proposed index, three extreme points were arbitrarily chosen: s_1, which has low entropy and low complexity; s_2, with medium entropy and high complexity; and s_3, with high entropy and complexity lower than s_2 but greater than s_1. For each of these points, we plot their singular spectrum (Fig. 3b), the ACF given by the equation 1 (Fig. 3c) and their PSD spectrograms (Fig. 3d). From top to bottom we have s_1, s_2, and s_3, respectively.

We can verify in Fig. 3b that ecoacoustic signals with few components and long-range correlations, as s_1, tend to have a concentrated singular spectrum, decreasing $H[P]$, while signals with uncorrelated noises (ie. tending to a white noise), such as s_3, have a flat singular spectrum, raising up $H[P]$. In the intermediate case, when the signals have different deterministic patterns, singular spectra similar to that of s_2 are generated, causing an intermediate entropy value. The ACF of sample s_1 has long-range correlations (Fig. 3c), implying an environmental colored noise, and a spectrogram with energy accumulated mainly at low frequencies (top spectrogram in Fig. 3d). We verified that this recording corresponds to rain sound without disturbance of animals. As we expected, s_1 is very close to the curve of environmental colored noises with $\alpha \approx 2$. Additionally, the ACF of s_1 produces the singular spectrum shown in the upper plot of Fig. 3b, which justifies its low entropy value.

A similar analysis can be made about s_2 and s_3. In the case of s_2, the ACF plot shows short- and medium-range correlations, with a spectrogram richer in different acoustic patterns (middle plot in Fig. 3c). Such ACF produces the singular spectrum shown in the third plot of Fig. 3b. The distribution of the eigenvalues of this singular spectrum returns an average entropy value, but its divergence tends to be maximum, thus justifying the high ECI. From its spectrogram, we also note that there is a low-energy ambient noise spread in a few frequency bands. We verified that this recording has high acoustic richness, containing calls of at least four different bird species, two frog species, and some insects.

Finally, s_3 shows high entropy and decreased complexity. Its ACF ($R_{xx}(s_3)$) plot shows only short-range correlations. This may be due to the lack of repeated deterministic patterns and the presence of noises spreading out energy at low-, medium-, and high-frequencies. In this case, the ACF plot has values only for low τ_i indices, emphasizing the presence of approximately white noise, which helps break weak correlations of signal components. As we know, uncorrelated noise signals tend to have an approximated flat singular spectra increasing their entropy value, a fact that can be seen in the lower plot of Fig. 3b. In the bottom spectrogram of Fig. 3d, we observed

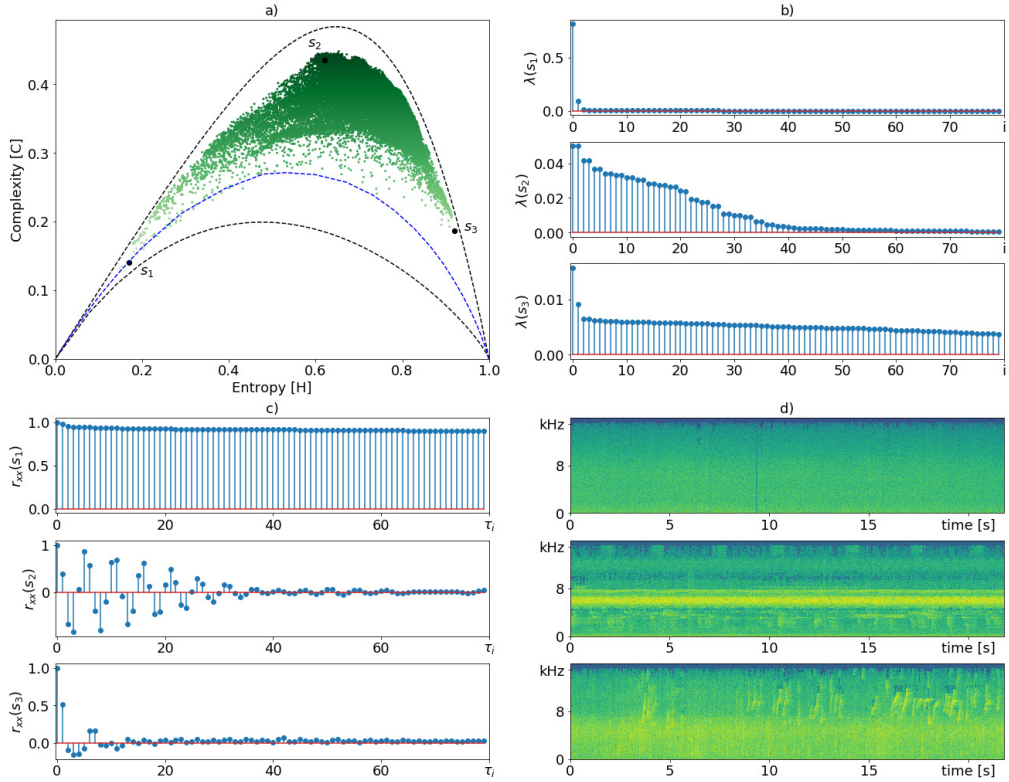

Fig 3. ECI characterization using $\tau_{max} = 512$ of samples recorded at Mamirauá protection area in the Amazon rainforest. Subfigure (a) depicts the HxC complexity plane, where each point is a signal segment of 22 seconds. Subfigures (a-d) show the singular spectrum, the autocorrelation function (ACF) and the spectrograms, respectively, of points s_1, s_2, and s_3, from top to down.

scattered noise contaminating most frequency bands and also a call of a single bird species. The presence of this quasi-deterministic pattern prevented the ECI from achieving even greater entropy and lower complexity.

So far we have discussed how the HxC plane characterization of ecoacoustic samples relates to signal processing concepts, such as ACF and PSD. We also demonstrate, in a practical way, how the proposed index relates to the acoustic richness of a signal segment. The effect of τ parameter variation on the distribution of samples in the HxC plane can be seen in S1 Fig. In the following sections, we present a detailed data analysis on how the proposed index varies in relation to the time of the day and along the monitoring days, characterizing temporal soundscapes patterns.

Temporal characterization through ECI

Temporal variations of the proposed index allow us to characterize the monitored location considering variations of the acoustic richness through time. Fig. 4a shows the complexity of each segment within three half-hour periods in a plane with $\tau_{max} = 512$, from where we can verify that there was an entropy increase in the period between 03:00 and 03:30 AM (green dots). This may be related to an increase in environmental noises, the intense acoustic activity of insects, and other environmental factors, and also a decrease of birds, amphibians, and other animal calls with daytime habits, causing a

lower acoustic diversity. Also, the interval corresponding to 12:00 and 12:30 pm (blue dots), presents few high ECI values and at the same time shows extreme dots with low entropy, this is probably due to two factors: 1) the acoustic activity of some animals with daytime habits combined with insects, such as cicadas, which increases entropy, and 2) the regular rainfall at this time of day during the monitoring period, which decreases entropy. Lastly, we have the interval between 07:00 and 07:30 AM (red dots), which presents the highest ECI values and few dots shifted slightly to the left of the plane, this may be a consequence of a higher acoustic activity of birds at dawn.

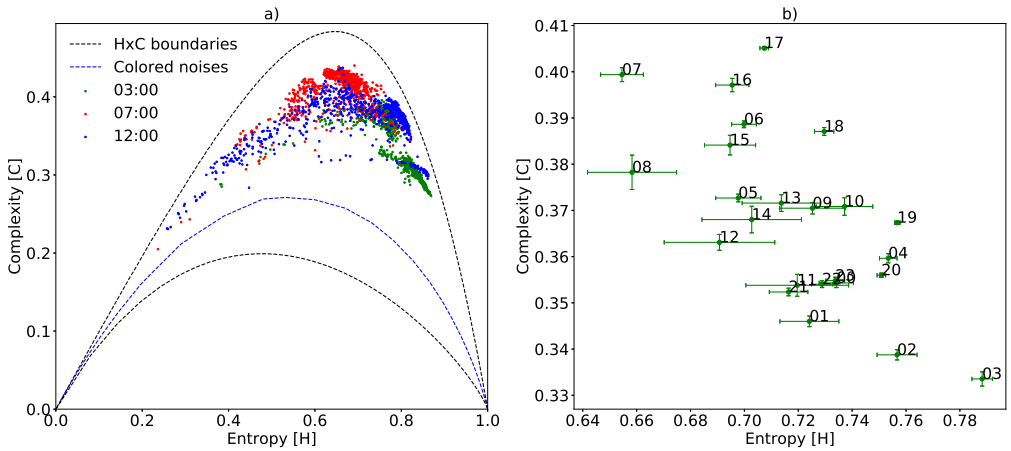

Fig 4. Temporal grouping of samples using $\tau_{max} = 512$ during the seven days of monitoring. Subfigure a) shows the spatial distribution of the samples at three half-hour intervals. Subfigure b) shows the cluster centroids with complexity and entropy variance bars at one-hour intervals. The proximity between these centroids highlight hours of the day with similar ecoacoustic richness.

The complete characterization of acoustic richness variations every one hour of the day can be better appreciated in Fig. 4b. This figure shows the centroids and their respective scatter bars for each sample group at each hour of the day. As we expected, the horizontal bars are always larger than the vertical bars due to the shape of the HxC plane. Comparing the different groups, we can see which periods of the day present the greatest ecoacoustic dispersion (e.g. 12 pm). Interestingly, at sunset (e.g. 5 pm) the highest ECI values are achieved with a smaller dispersion, which suggests that there is an intense acoustic activity with less perturbation of environmental phenomena at this time. The second key time is dawn (e.g. 7 am) in which it is known that there is a greater acoustic activity of morning birds. Points 21, 22 and 23 are relatively close, and as described in the previous sections, near points in the plane have histograms with low divergence caused by similar acoustic patterns.

Looking at the spatial centroid's distribution in Fig. 5, we also noticed that before dawn (e.g. 3 am) the acoustic activity of the birds, anurans and other species is lower, giving rise to greater insect activity such as cicadas, a fact that increases entropy. Finally, we can see groups of centroids with similar characterizations (near location and comparable dispersion), for example, the set $\{06h, 07h, 08h, 15h, 16h, 17h, 18\}$ or $\{05h, 09h, 10h, 12h, 13h, 14h\}$. Such centroid variations are better depicted in Fig. 5. This figure presents the ECI variation with a half-hour resolution, where we can notice two peaks of maximum acoustic activity, at 07:00 and 17:00, a fact that matches the knowledge of the experts about soundscapes of the Mamirauá region.

Environmental phenomena may change over the days, however, we expect that this fingerprint of acoustic complexity must be repeated with some frequency, except when

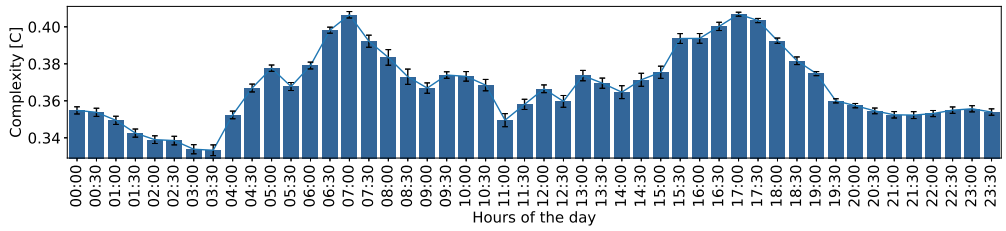

Fig 5. Ecoacoustic complexity with its confidence interval at every half hour using $\tau_{max} = 512$. Average variation considering the seven days of monitoring. This variation can be considered the soundscape fingerprint of the region within the sensor microphone range.

biodiversity changes. To verify such consistency through time we included Fig. 6. In other words, the daily pattern shown in Fig. 5 is almost regular across the six consecutive days of the monitoring program, shown in Fig. 6. We observed that the temporal variation of ECI is little affected by the τ parameter variation. This parameter increases the scale between the maximum and minimum peaks of acoustic complexity but keeps the shape of the curve. During a few days of the monitoring program, the sensor experienced technical difficulties that caused the loss of some minutes of recording. However, the proposed index proved to be resilient to these issues. The soundscape fingerprint of daily variation including the Entropy and Divergence quantifiers are additional resources presented in S2 Fig.

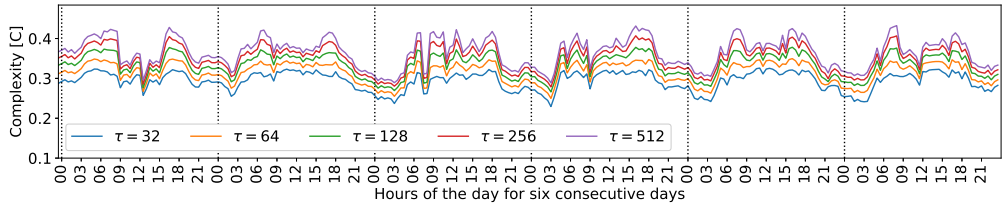

Fig 6. ECI regular patterns varying through six days. Behavior patterns of the ECI index indicate greater activity during the day and lower activity at night. Daytime variations are more irregular compared to nighttime variations.

ECI variability in Mamirauá floodplain during drought and flood events

The Mamirauá Sustainable Development Reserve is mostly floodplain, where the different phases of the hydrological cycle influence the ecosystem [26, 27]. In Mamirauá, the seasons are strongly marked by the water level. During the flood season the land is completely underwater, while during the dry season, there is plenty of drylands available. These changes impose mobility restrictions for some species, which directly modifies the ecoacoustic landscape. As mentioned at the beginning, the monitoring period was divided into two months, approximately one week in July and another week in September during flood and drought event peaks, as shown in S3 Fig. These two months are characterized by the high water phase and the low water phase, directly affecting the landscape. Stratifying our dataset by month and plotting the ECI for each week, we observed a shift over the spatial distribution of points in Fig. 7). The effect of this seasonal variation can best be observed by the centroid displacement of each month.

This comparison shows the usefulness of ECI in capturing the variation of the ecoacoustic soundscape due to the change of season.

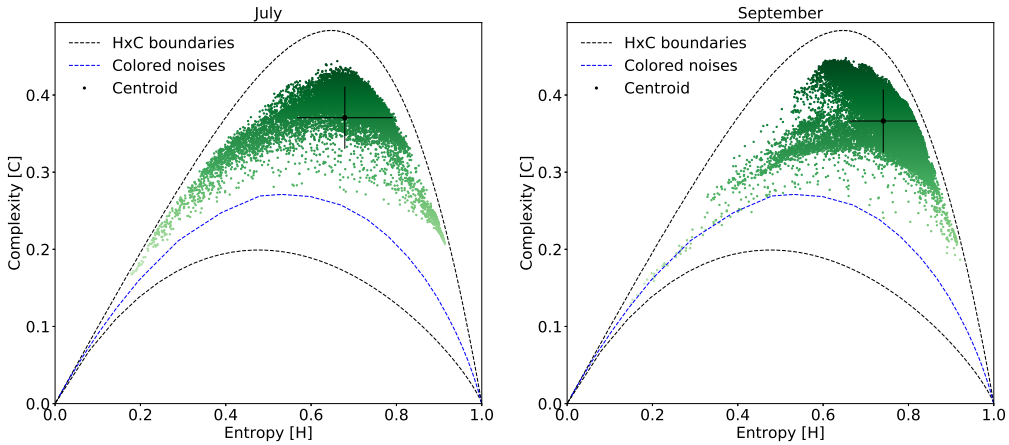

Fig 7. Seasonal characterization of the soundscape using tau = 512 $\tau_{\mathbf{max}}$. This possibility of characterization can be extended to long term monitoring programs.

Rainfall sound characterization

Environmental phenomena are also captured by the sensor microphone. Consequently, environmental sounds can also be characterized in the HxC plane. We know that these phenomena, for instance, the sound of the rain, have a frequency spectrum known as colored noises. This is the main reason why we included the simulation of this type of noise, generating the blue reference curve in Fig. 8a.

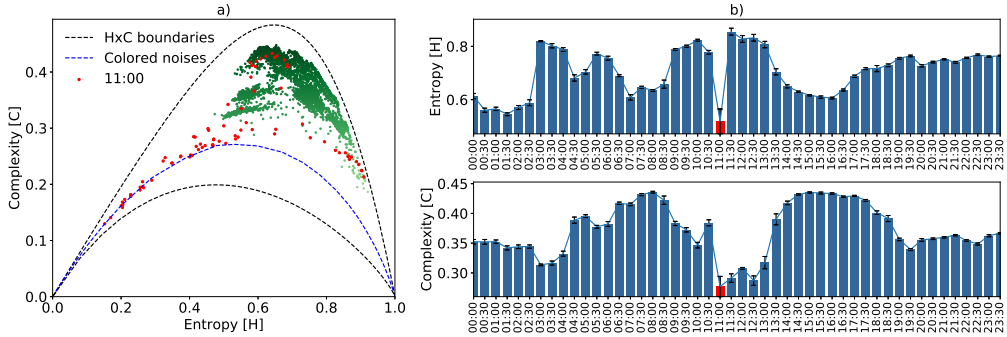

Fig 8. Rainfall characterization using $\tau_{\mathbf{max}}$. Subfigure a) shows the ECI from September 2, 2016. The red dots correspond to a half hour interval when it rained. Subfigure b) shows the average variation, along with their respective confidence intervals, of the quantifiers during that day every half hour.

The period of days in which the monitoring was performed is considered to be within the dry season of the Amazon [28]. However, on September 2, 2016, there was a slight amount of rainfall during the interval between 11:00 and 11:30 am. Only the acoustic samples corresponding to this day are illustrated in the HxC plane of Fig. 8a. Here, the red dots correspond to a half-hour interval starting at 11:00 am. During this interval there was rainfall, and the resulting approximation of the samples to the color noise

curve. In this same time interval, both entropy and complexity decreased abruptly, as can be seen in Fig. 8b. The left shift of the red dots in the HxC plane causes a considerable increase in the variance of the information quantifiers, as indicated by the confidence interval on the red bar. After the rain stopped, the entropy quickly recovered a high value, while the complexity still took time to increase. For those with field experience, it can be seen that after a storm the forest is silent for a few minutes, recovering acoustic activity first by insects, which increases entropy, and then with birds and other animal species, ultimately increasing ecoacoustic complexity. Unfortunately, we do not have detailed rainfall records of the monitored location on these dates. However, the recorded audios were hand-inspected by a specialist who heard and confirmed rain sound at the mentioned time period.

Discussion

In previous sections, we presented the theoretical bases and some evidences that support the composite index of ecoacoustic complexity. Our main challenge was to make this index sensitive to bioacoustic diversity while differentiating among environmental phenomena. Thus, the physical interpretations characterized by the HxC plane contribute to the understanding of the system dynamics of the Amazon soundscape. It is worth noting that the ECI, in our specific experiments, characterized the region covered by the sensitivity of the sensing device (or microphone).

With characterization analysis using the complexity plane, we found that the greater the acoustic diversity, the greater the complexity. In the HxC plane, vocalizations of species that have regular spectral-temporal patterns balance entropy, whereas species with irregular calls - with a greater number of different patterns - tend to increase entropy. The presence of insects can also be observed through the ECI. We realize that insect choral singing, like cicadas, increases entropy for they are similar to high-frequency uncorrelated noise, typically between 5 and 8 kHz. Thus, the increase in entropy caused by insects of the same species generally decreases the complexity indicating less diversity, regardless of how many individuals the coral has, if they are of the same species then the diversity is smaller and consequently the ECI should decrease as noted. Although insects' sound increases entropy, their characterization is very important, for they support the life of other species.

The temporal analysis presented above shows that there is a regularity of the soundscape between hours of the day and between monitoring days captured by ECI. As expected, from Fig. 5 and Fig. 6 we noticed a greater daytime acoustic activity. Fig. 5 highlights two high ECI peaks at dawn and dusk, characterizing the hours of the day with greater intensity and acoustic diversity, which is a characteristic behavior of birds and frogs in the Amazon rainforest. As shown in Fig. 6, this characteristic behavior is almost regular between consecutive days. Short- and long-term environmental phenomena are also well characterized by ECI. Short-term sporadic phenomena, as verified by the rain example, can be recognized with our approach. In this example, the low-frequency rain sound - characterized as colored noises - decrease entropy and complexity, whenever not disturbed by other sounds, generating HxC points close to the color noises reference curve (Fig. 8). With regard to the long-term environmental phenomena, such as the hydrological cycle, there is also a different characterization of samples in the HxC plane for different seasons. Mamirauá is a floodplain area with a season of the year completely flooded. This change in landscape alters the dynamics of the species and hence, the soundscape, as was characterized in Fig. 7.

Conclusion

We presented a new Ecoacoustic Index called ECI. This index was based on the Entropy-Complexity causal plane theory, which in turn is a predictor of generalized statistical complexity. The calculation of ECI is obtained from information theory quantifiers, which are the Shannon entropy and the Jensen-Shannon divergence using a uniform distribution as a reference. We emphasize that our main contribution is a new index to convert acoustic signals recorded by microphones (or environmental sensors) into probability distributions, allowing us to characterize the soundscape of a given location. Besides being unsupervised, experiments have shown that, different from existing approaches, ECI is not susceptible to the variable length signals.

Intrinsically, our methodology is based on the autocorrection matrix, allowing us to relate the spectral content of signals to their spatial characterization in the HxC plane. Moreover, the low computational complexity of the SVD algorithm used to obtain the eigenvalues makes the method interesting for contexts with limited hardware resources. There is a trade-off between a small τ (and a low-discriminant HxC plane), and a large one (and a HxC broken into several frequency clusters).

In addition to these contributions, we detailed the dynamics of the HxC plane, its upper and lower bounds, the effect of divergence on the spatial characterization of the samples, the effect of the tau autocorrelation parameter, and also add a curve simulating the colored noises possibly found in nature. These features allow the differentiation of signals where natural patterns, such as the singing of birds, amphibians, insects or other animals, and environmental phenomena are combined. Finally, simulations and figures are available at `http://bit.ly/2m12PWc`, facilitating the reproduction of results and possible comparisons with other methods.

A direct application of ECI is the characterization of periodic seasons. As shown in the Mamirauá reserve case study, seasonal variations of the water level are the most important factor in determining the presence of different species communities present in flooded areas in the Amazon. The proposed index proved sensitive enough to capture the variations in the ecoacoustic landscape due to the change from flood season to drought. However, further investigation is needed, extending monitoring time to more than two seasons over a year, and also expanding the monitoring program to other areas with different soundscapes, such as the Brazilian Cerrado or the Atlantic Forest.

Supporting information

S1 Appendix. Efficient autocorrelation calculation. The relation between the Power Spectral Density (PSD) of a signal and its autocorrelation function is described by the Wiener-Khintchine theorem [29] as:

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} S_{xx}(f) e^{i2\pi f\tau} \mathrm{d}f \tag{7}$$

where $S_{xx}(f)$ is the PSD. Therefore, the autocorrelation coefficients expressed by equation 1 can be efficiently calculated using the Fast Fourier Transform algorithm (FFT) applying the three following steps:

1. apply the FFT on $X(t)$ to obtain its frequency spectrum, $F(f) = \mathrm{FFT}[X(t)]$;
2. multiply $F(f)$ by its complex conjugate to get the PSD, $S_{xx}(f) = F(f)F^*(f)$; and finally
3. obtain $R_{xx}(\tau)$ using the Inverse FFT, $R(\tau) = \mathrm{IFFT}[S(f)]$.

This procedure allows computing the autocorrelation from the raw data $X(t)$ with two FFT's. It is worth noting that, a naive procedure using equation 1 has a computational complexity order $O(n^2)$, whereas using the FFT has complexity $O(n \log_2 n)$. Its reduced cost makes this procedure attractive to be embedded into small sensor motes. Moreover, the minimum computational complexity is reached when the length of $R_{xx}(\tau)$ is a power of 2, justifying why we chose the values 32, 64, 128, 256 and 512 in our experiments.

S2 Appendix. Normalized Jensen-Shannon divergence. The quantity expressed by equation 5, known as "disequilibrium", can be rewritten as:

$$Q[P, P_e] = \frac{J[P, P_e]}{Q_{max}}, \tag{8}$$

where $Q_0 = {}^1/_{Q_{max}}$, then

$$Q_0 = -2 \left\{ \frac{\tau_{max} + 1}{\tau_{max}} \log(\tau_{max} + 1) - 2 \log(2\tau_{max}) + \log(\tau_{max}) \right\}^{-1} \tag{9}$$

is the maximum possible value of $Q[P, P_e]$, obtained when only one component of P is equal to one and all the others become zero [30]. This means $\lambda_1 = 1$ and $\lambda_{2:\tau_{max}} = 0$ in our definition.

S3 Appendix. Correlated stochastic noises Colored noises, also known as correlated stochastic noises, are present in almost every acoustic signal recorded in a rainforest environment [31, 32]. This class of noise typically represents natural phenomena. The following α values determine some common types of noise:

1. $\alpha = 0$ models the white noise containing an equal amount of energy in all frequency bands;
2. $\alpha = 1$ models the pink noise with equal sound pressure level in each octave band decreasing the energy as the frequency increases; and
3. $\alpha = 2$ models the red (or brown) noise, which is common in oceanographic recordings, it describes the ambient underwater noise from distant sources

Values of $1 \leq \alpha \leq 2$ include several low-frequency natural phenomena recorded by the microphone, for example, rain or wind to name a few. Here, illustrating these noises becomes a reference curve to determine which ecoacoustic signal samples have characteristics of colored environmental noise.

S1 Fig. The relationship between the τ parameter and the HxC plane. We can notice, in Fig. 9, how the spatial distribution of points changes in relation to τ. The higher the τ, the greater the distance between the lower and upper boundaries, allowing new grouping patterns to appear.

S2 Fig. Radar chart plot of 24h. Each chart in Fig. 10 shows the quantifier's variation over the hours of the day. These charts can be adopted as the fingerprint of the monitored soundscape.

S3 Fig. Hydrological cycle of Mamirauá floodplain. Fig. 11 shows the water level phases according to the months of the year 2016 measured in meters above sea level (MASL or m.a.s.l.). This data is publicly accessible through the link `https://www.mamiraua.org.br/fluviometrico-na-reserva`.

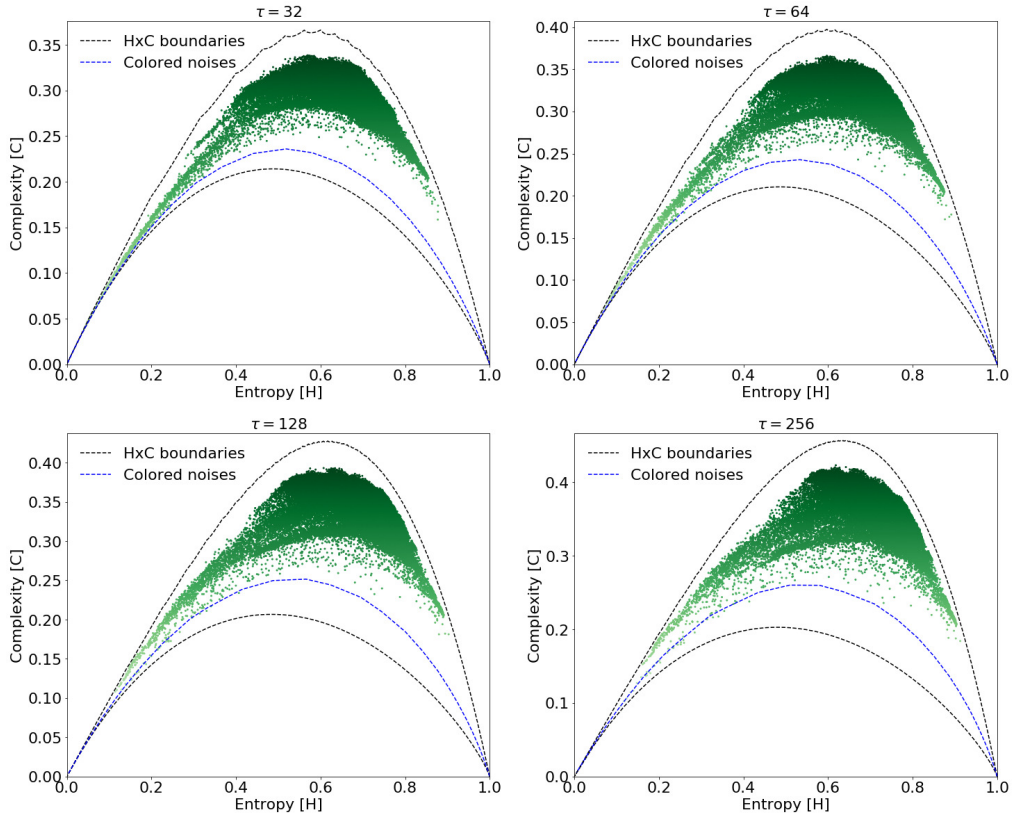

Fig 9. HxC planes generated from several values of τ. The separation between the upper and lower limits makes it possible to better discern the distribution of points in the HxC plane.

Acknowledgments

The authors thank the Instituto de Desenvolvimento Sustentável Mamirauá for their support in coordinating the Project Providence. The authors thank the Gordon and Betty Moore Foundation for their grant support and scholarships. We thank all the institutions that were part of the Providence project, including the Commonwealth Scientific and Industrial Research Organization (CSIRO), The Sense of Silence Foundation and the Technical University of Catalonia. We also thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the institutional support.

References

1. Sueur J, Krause B, Farina A. Climate Change Is Breaking Earth's Beat. Trends in Ecology & Evolution. 2019;doi:10.1016/j.tree.2019.07.014.

2. Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL, Farina A, Krause BL, Napoletano BM, et al. Soundscape ecology: the science of sound in the landscape. BioScience. 2011;61(3):203–216.

3. Gasc A, Sueur J, Pavoine S, Pellens R, Grandcolas P. Biodiversity sampling using a global acoustic approach: contrasting sites with microendemics in New Caledonia. PLoS One. 2013;8(5):e65311.

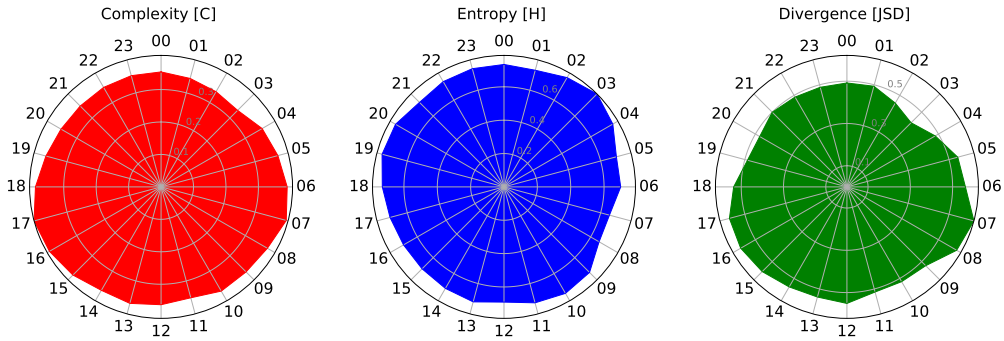

Fig 10. Radar chart plot using $\tau_{\mathbf{max}} = 512$. From left to right the three information quantifiers used in our methodology.

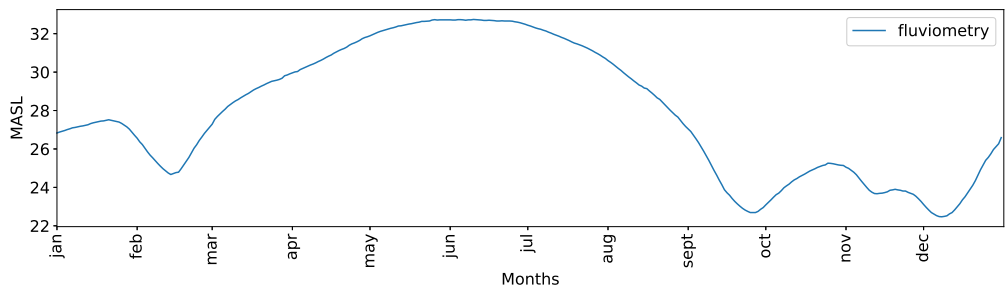

Fig 11. Seasonal variation of water level in Mamirauá during 2016. July and September are the months where acoustic data was collected.

4. Sueur J, Pavoine S, Hamerlynck O, Duvail S. Rapid acoustic survey for biodiversity appraisal. PloS one. 2009;3(12):1–9. doi:10.1371/journal.pone.0004065.

5. Farina A, Gage SH. 1. In: Ecoacoustics: A New Science. John Wiley & Sons, Ltd; 2017. p. 1 – 11.

6. Farina A. Ecoacoustics: A Quantitative Approach to Investigate the Ecological Role of Environmental Sounds. Mathematics. 2018;7(1). doi:10.3390/math7010021.

7. Rosso OA, Carpi LC, Saco PM, Ravetti MG, Plastino A, Larrondo HA. Causality and the entropy-complexity plane: Robustness and missing ordinal patterns. Physica A: Statistical Mechanics and its Applications. 2012;391(1-2):42–55.

8. Alonso JB, Cabrera J, Shyamnani R, Travieso CM, Bolaños F, García A, et al. Automatic anuran identification using noise removal and audio activity detection. Expert Systems with Applications. 2017;72:83 – 92. doi:10.1016/j.eswa.2016.12.019.

9. Xie J, Zhu M. Handcrafted features and late fusion with deep learning for bird sound classification. Ecological Informatics. 2019;52:74 – 81. doi:10.1016/j.ecoinf.2019.05.007.

10. Bermant PC, Bronstein MM, Wood RJ, Gero S, Gruber DF. Deep Machine Learning Techniques for the Detection and Classification of Sperm Whale

Bioacoustics. Scientific Reports. 2019;9(1):1 – 10. doi:10.1038/s41598-019-48909-4.

11. Melo-Santos G, Figueiredo Rodrigues AL, Tardin RH, de Sá Maciel I, Marmontel M, Da Silva ML, et al. The newly described Araguaian river dolphins, *Inia araguaiaensis* (Cetartiodactyla, Iniidae), produce a diverse repertoire of acoustic signals. PeerJ. 2019;7:e6670. doi:10.7717/peerj.6670.

12. Bjorck J, Rappazzo BH, Chen D, Bernstein R, Wrege PH, Gomes CP. Automatic Detection and Compression for Passive Acoustic Monitoring of the African Forest Elephant; 2019.

13. Ivan K, Davide Z, Yunpeng L, Marianne S, Kathy W, Stephen R. Bioacoustic detection with wavelet-conditioned convolutional neural networks. Neural Computing and Applications. 2018;doi:10.1007/s00521-018-3626-7.

14. Vu TT, Tran LM. An Application of Autonomous Recorders for Gibbon Monitoring. International Journal of Primatology. 2019;40(2):169 – 186. doi:10.1007/s10764-018-0073-3.

15. Aide TM, Corrada-Bravo C, Campos-Cerqueira M, Milan C, Vega G, Alvarez R. Real-time bioacoustics monitoring and automated species identification. PeerJ. 2013;1:e103. doi:10.7717/peerj.103.

16. Oliver RY, Ellis DPW, Chmura HE, Krause JS, Pérez JH, Sweet SK, et al. Eavesdropping on the Arctic: Automated bioacoustics reveal dynamics in songbird breeding phenology. Science Advances. 2018;4(6). doi:10.1126/sciadv.aaq1084.

17. Rama Rao KVSN, Garg S, Montgomery J. Investigation of Unsupervised Models for Biodiversity Assessment. In: AI 2018: Advances in Artificial Intelligence. Springer International Publishing; 2018. p. 160 – 171.

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

19. Wilcox RR. Introduction to Robust Estimation and Hypothesis Testing. 3rd ed. Academic Press; 2012.

20. Tomé AM, Teixeira AR, Figueiredo N, Santos IM, Georgieva P, Lang EW. SSA of biomedical signals: A linear invariant systems approach. Statistics and its Interface. 2010;3(1):345–355.

21. Passerini F, Severini S. The von Neumann entropy of networks. arXiv:08122597v2. 2012;doi:10.4018/978-1-60960-171-3.ch005.

22. López-Ruiz R, Mancini HL, Calbet X. A statistical measure of complexity. Physics Letters A. 1995;209(5-6):321–326.

23. Rosso OA, De Micco L, Larrondo HA, Martin MT, Plastino A. Generalized statistical complexity measure. International Journal of Bifurcation and Chaos. 2010;20(03):775–785.

24. Rosso OA, Larrondo HA, Martin MT, Plastino A, Fuentes MA. Distinguishing Noise from Chaos. Phys Rev Lett. 2007;99:154102. doi:10.1103/PhysRevLett.99.154102.

25. Ravetti MG, Carpi LC, Gonçalves BA, Frery AC, Rosso OA. Distinguishing Noise from Chaos: Objective versus Subjective Criteria Using Horizontal Visibility Graph. PLOS ONE. 2014;9(9):1–15. doi:10.1371/journal.pone.0108004.

26. Affonso AG, Queiroz HL, Novo EMLM. Abiotic variability among different aquatic systems of the central Amazon floodplain during drought and flood events. Brazilian Journal of Biology. 2015;75(4):60–69.

27. Ferreira-Ferreira J, Silva TSF, Streher AS, Affonso AG, de Almeida Furtado LF, Forsberg BR, et al. Combining ALOS/PALSAR derived vegetation structure and inundation patterns to characterize major vegetation types in the Mamirauá Sustainable Development Reserve, Central Amazon floodplain, Brazil. Wetlands Ecology and Management. 2015;23(1):41–59.

28. Yoon JH, Zeng N. An Atlantic influence on Amazon rainfall. Climate Dynamics. 2010;34(2):249–264. doi:10.1007/s00382-009-0551-6.

29. Ricker DW. Echo Signal Processing. The Springer International Series in Engineering and Computer Science. Springer US; 2003.

30. Ribeiro HV, Zunino L, Lenzi EK, Santoro PA, Mendes RS. Complexity-Entropy Causality Plane as a Complexity Measure for Two-Dimensional Patterns. PLOS ONE. 2012;7(8):1–9. doi:10.1371/journal.pone.0040689.

31. Colonna JG, Nakamura EF, Rosso OA. Feature evaluation for unsupervised bioacoustic signal segmentation of anuran calls. Expert Systems with Applications. 2018;106:107 – 120. doi:10.1016/j.eswa.2018.03.062.

32. Colonna JG, Nakamura EF. Unsupervised selection of the singular spectrum components based on information theory for bioacoustic signal filtering. Digital Signal Processing. 2018;82:64 – 79. doi:10.1016/j.dsp.2018.07.009.