# A widened array of metrics (WAM) approach to characterize the urban acoustic environment

**Bryce Lawrence[1*], Jonas Poppen[2], Timo Haselhoff[2], Robynne Sutcliffe[2], Salman Ahmed[2], Susanne Moebus[2], Dietwald Gruehn[1]**

[1]School of Spatial Planning, Department of Landscape Ecology and Landscape Planning, TU Dortmund University, Dortmund, Germany

[2]Institute Urban Public Health, University Hospital Essen, University Duisburg-Essen, Essen, Germany

*** Correspondence:**
Bryce Lawrence
bryce.lawrence@tu-dortmund.de

## 1 Supplemental Material Description

This supplemental information provides a detailed description of the origin of alpha and beta indices, the packages necessary for their calculation in R, and the parameters of the R scripts used in calculation of WAM in R.  It serves as documentation of the parameters used in WAM for the project SALVE and as supplemental material for alpha and beta R scripts available via github at bt-lawrence/Alpha-and-Beta-Indices-R.

## 2 Overview of Packages, Functions, and Key Terminology

In the following, we shortly describe key terminology used in the descriptions of the indices that are critical to understand the index, and provide an overview of packages and functions in R necessary for their calculation.

### 2.1 R and R Packages

The computer language R (Ihaka and Gentleman, 1996) calculates most alpha and beta indices as functions within a package.  A function (denoted here with the symbol () after the function name) can be equated to a formula that has variable parameters called arguments.  Thus, all the original formulas for acoustic indices have been embedded into specific functions in R.  Functions from multiple packages are necessary to prepare or transform data before calculating a specific alpha or beta index. Two primary packages for sound index analyses are seewave (Sueur et al., 2008a) and soundecology (Villanueva-Rivera and Pijanowski, 2018). Calculation of functions in these packages require the supporting packages parallel (R Core Team 2019), pracma (Borchers, 2019), oce (Kelley and Richards, 2019), ineq (Zeileis, 2014), gsw (Kelley et al., 2017), data.table (Dowle and Srinivasan, 2019), tuneR (Ligges et al., 2018), and vegan (Oksanen et al., 2019).

Beta Indices rely on all of the same packages as the alpha indices with the addition of the package openxlsx (Schauburger and Walker, 2019).

## 2.2 Key Terms and Transformations

Alpha and beta indices must transform a Windows Audio File Format (WAV) recording into a useable numeric transformation before carrying out functions. Here we explain some of the most commonly used transformations and functions as well as their adjustable parameters to eliminate the need for repetition in section 4 below.

### 2.2.1 Amplitude and Hilbert Envelopes

In an absolute amplitude envelope, positive oscillating waves are isolated but left in their raw absolute state. The Hilbert envelope interpolates the peaks of the oscillating amplitude envelope to create a smoothed amplitude profile (Seuer, 2018 pgs. 127-129). Both methods quantify the amplitude of a signal over time. From either of these envelopes, one can calculate the maximum, minimum, mean, median amplitude as single values that are very useful for comparison.

### 2.2.2 Fourier Transformation

The Fourier transformation (FT) is used for frequency analysis that converts an acoustic observation from a time domain to a frequency domain. This transformation creates a distribution of amplitude on the y-axis and frequency ranges on the x-axis, called a *frequency spectrum*, that allows for identification and quantification of the frequency ranges present in an acoustic environment. The Discrete Fourier transformation (DFT) is carried out for time-limited (discrete) lengths of an acoustic recording to return a single time 'discrete' frequency spectrum. A short time (discrete) Fourier transformation (STFT, STDFT) computes this transformation for time-limited sections of a recording (windows) to compare frequency spectra between discrete time windows (Sueur, 2018 pg. 216; 309-311).

The Fourier transformation can be computed with the function spec() from package seewave from any WAV file. There are several adjustable arguments of this transformation, including:

- the window length (wl) describing the time and frequency resolution of the transformation (128, 256, 512, 1024), which is sufficient at the default 512 for the urban acoustic environment (Sueur, 2018 pg. 314);
- the window name or shape (wn) which controls the taper of the transformation at the high, middle or low frequency ranges (bins) of the frequency spectrum (hanning, hamming, bartlett, blackman, flattop, rectangle) for which the usually default 'hanning' is used (Sueur, 2018 pg. 239);
- the overlap (ovlp) of windows in percent when STFT or STDFT is computed, for which the default is 0% overlap. We did not include any overlap to reduce processing time, which increases with overlap percentage, and to eliminate uncertainty stemming from

window overlap that we may not be able to account for statistically (Sueur, 2018 pg. 132; 314).

### 2.2.3 Mean Frequency Spectrum

The mean (average) frequency spectrum is a summary of energy distributed amongst frequency bands in a signal for a specified time window. For the indices based on a mean frequency spectrum, this value is calculated in R with the functions meanspec() or soundscapespec() from a WAV file using the package seewave. In meanspec(), the argument spectro makes a 2-D spectrogram out of the WAV using the default fftw parameters (3.2.2) to calculate the mathematical mean from the STDFT matrix, where the sum of energy in each frequency row is divided by the total number of STDFT columns and then averaged to produce the mean frequency spectrum. In this method, longer sounds are weighted more heavily than shorter sounds (Sueur, 2018 pg. 377). This means that continuous lower frequency sound common in the urban acoustic environment may bias the mean frequency spectrum. In the NDSI, however, the function soundscapespec() in the seewave packages is used to bin the frequency spectrum, computing the STDFT for 1Khz bins using Welch's method (Welch, 1967; Sueur, 2018 pg. 377). This calculation method is useful for the NDSI because it does not summarize 1Khz bins relative to all 1Khz bins, meaning continuous sounds have less effect on the NDSI value.

### 2.2.4 Cumulative Distribution Function

Frequency spectra can be transformed into cumulative distribution functions (CDF) by summarizing the amplitude across all frequency bins in a DFT frequency spectrum. This calculation is done with the function specprop() in seewave (3.2.3) and is used in the beta indices for the purposes of comparing two spectral distributions with each other (Sueur et al., 2008a; Sueur, 2018 pg. 495-6).

### 2.2.5 Diversity and Entropy Index

The Shannon index of general diversity (also called Shannon's H) is used in several of the diversity or entropy indices and has been used in ecological studies for at least the past 50 years (Odum, 1971 pg. 144). When the Shannon index is used without normalization, allowing values ranging from zero to infinity, it is simply called the Shannon index and is considered a measure of diversity. When the Shannon index is normalized, constraining the index value range between 0 and 1, it is called an evenness index and considered a measure of entropy.

## 3 Description of Selected Indices

In order to facilitate an understanding of the general meaning of each index and to discuss which parameters are adjustable, we summarize each of the alpha, beta and SPL indices individually, including:

- the source of the mathematical background of each index and an explanation of the original intended use of the index.

- a mention of the functions in R used to calculate each selected index, and their relevant adjustable arguments, if any, and the range of each index from high to low as determined by its mathematical formula.
- a simple description of the index to explain what it means practically for analysis of the urban acoustic environment based in part on the mathematically determined range, and supported by the given performance samples. Indices that do not perform are addressed in the discussion.
- an overview of the SPL indices used and their original sources or calculation platforms.

## 3.1 Alpha Indices

### 3.1.1 Bioacoustic Index

The Bioacoustic Index (Bio) was developed to estimate relative avian abundance in native and exotic vegetated land cover types in Hawaii (Boelman et al., 2007). The index uses the STFT, then calculates the mean spectrum of each 1Khz frequency bin, plots them, then calculates the area under the curve (Boelman et al., 2007 App. 1). This index can be useful to determine the presence of biophonies between seasons or between land cover / land use types. A simplified interpretation of what this index means in the acoustic environment is as follows:

- Bio is a measure of amplitude in the default (assumed biophonic) frequency range of 2000 to 8000 Hz. The range is adjustable.
- The higher the Bio index, the greater the amplitude of biophonic signals.
- In the urban acoustic environment, high Bio values may indicate avian abundance, but may also indicate machine noises that can have overlapping Hz ranges.

The index is calculated in R with the function bioacoustic_index() within the soundecology package where the index is calculated as an 'area value' representing the total amplitude within each 1Khz bin between 2000-8000Hz (Villanueva-Rivera and Pijinowski, 2018). Frequency bandwidth maximum and minimum can be changed within the arguments min_freq and max_freq with a default range of 2000 to 8000 Hz. The index range is 0 to infinity, where zero represents no amplitude 'area' between 2000 to 8000 Hz in a recording, and values greater than zero represent an increasing 'area' between 2000 and 8000 Hz.

### 3.1.2 Acoustic Complexity Index (ACI)

The acoustic complexity index focuses on amplitude modulation between frequency bins with the assumption that anthrophonic sounds have low modulation over all bins and biophonic sounds have higher amplitude modulation over frequency bins. The ACI, discussed extensively in Farina and Morri (2008) and Pieretti et al. (2011) has a multi-step calculation procedure that essentially sums the differences in amplitude by frequency bins between user defined time ranges of a WAV recording. The logic behind the ACI is that when the differences between frequencies within time bins are consistently different, then the acoustic environment is complex; when the differences in frequencies between time bins is small then the acoustic environment is more uniform and thus less complex. However, since the ACI is cumulative over time, longer sound samples will be more heavily weighted, possibly resulting in lower values for urban acoustic environments that often contain continuous sound sources. A simplified interpretation of what this index means in the acoustic environment is as follows:

- The ACI indicates modulation of amplitude distribution across frequency ranges over time.

The index is calculated in R with the function ACI() from seewave (Sueur et al., 2008a) or acoustic_complexity() from soundecology (Villanueva-Rivera and Pijinowski, 2018). Several parameters can be set, including parameters of the Fourier time window (3.2.2) and the limitation of a defined frequency range for analysis (Sueur, 2018 pg. 489). The index ranges from zero to infinity where zero indicates no modulation in amplitude between frequency bins over time and higher values indicate greater modulation in amplitude between frequency bins over time.

### 3.1.3 Acoustic Diversity Index (ADI)

The acoustic diversity index (Pijinowski et al, 2011), also discussed extensively in Sueur (2014; 2018), was developed to analyze species call diversity amongst different land cover types in Tippecanoe County, Indiana, USA, including mature and mixed forest types, agricultural areas in different successional stages, and on Purdue University campus itself. The ADI uses the short time discrete Fourier transformation (3.2.2) to split an observation into frequency bins and then uses the relative amplitude above a defined dB threshold to calculate the Shannon index (Sueur, 2018 pg. 486). This index is useful because it is based on hypothesized differences in the acoustic environment amongst various land use types, both urban and non-urban. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of the diversity of frequency and amplitude distribution between discrete time windows.

The index is calculated in R with the function acoustic_diversity() in the package soundecology (Villanueva-Rivera and Pijinowski, 2018). Adjustable parameters in the STDFT function can be left in their default values (3.2.2), but dB threshold and width of frequency bins need to be defined. The index ranges from 0 to the natural log of the number of Hz bins, that for the default settings of ten bins it is ln(10) = 2.303. In the index, larger values indicate a high diversity of amplitude distribution amongst frequency bins and values closer to zero indicates that amplitude is distributed evenly throughout all frequency bins.

### 3.1.4 Acoustic Evenness Index (AEI)

The acoustic evenness proceeds with the same first step as the ADI, but computes the Gini coefficient (Gini, 1912) rather than the Shannon index (Sueur, 2018 pg. 488). The Gini coefficient is a measure of statistical dispersion that measures the inequality among values of a frequency distribution (Dixon et al. 1987). The measure is useful because as the value of the Gini coefficient goes up, the concentration of the energy of the overall recording is concentrated in fewer frequency bins. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of the concentration or dispersion of energy amongst frequency ranges.

The formula is calculated with the function acoustic_evenness() in the soundecology package (Villanueva-Rivera et al. 2011) with adjustable parameters of the STDFT as described above (3.2.2) with dB threshold and frequency steps that need to be defined. The use of more than the default ten frequency bins (of 2000Hz each) may improve accuracy of the Gini-coefficient if sound sources existing within ranges of less than 2000Hz, which is likely the case in the 0 to 4000Hz range in the urban environment where biophonies, geophonies and anthrophonies overlap. The index ranges from 0 to 1, where zero represents perfect Gini-equality of amplitude distribution between frequency bins and one represents maximal Gini-inequality of amplitude between frequency bins where all energy would be distributed to only one frequency range.

### 3.1.5 Normalized Difference Soundscape Index (NDSI)

The NDSI (Kasten et al., 2012) is an often-used index in ecoacoustics and soundscape ecology studies. The NDSI, built on the principle of the NDVI (Rouse et al., 1973) which is widely applied in remote sensing, estimates the difference between signals from two different ranges within an observation, normalized by the sum of signals in the two ranges. The NDSI uses the range comparison approach to calculate the proportion of anthrophonic sounds (1000 to 2000 Hz) to biophonic sounds (2000 to 8000 Hz) in an acoustic environment. The NDSI is useful to test the assumption that increases in vegetation will reflect a larger ratio of biophonic to anthrophonic sounds, estimate the impact that urbanization or seasonality have on biophonies or anthrophonies, and to investigate the effects of spatial structure on the acoustic environment. A simplified interpretation of what this index means in the acoustic environment is as follows:

- The NDSI is a ratio of how much of the amplitude of an acoustic observation is contained within the range of biophony and how much within the range of anthrophony.

The index is calculated in R with the function ndsi() in the soundecology package (Villanueva-Rivera and Pijanowski, 2018) or NDSI in seewave (Sueur et al., 2008a). The adjustable parameters in the NDSI include those of the Fourier transformation (3.2.2) and the default anthrophonic and biophonic minimum and maximum ranges of 1000-2000 (anthro) and 2000 to 8000 (bio). Usually all arguments can be left at default values since birds do appear in the 2000 to 8000 Hz range in the urban environment. In areas with high avifauna diversity, such as the tropical or Mediterranean plant communities, this range may need to be increased to account for birds with higher frequency calls. The index ranges from -1 to +1, where the closer the value to positive one, the more influence biophony has in an observation and the closer to minus one the more influence anthrophony has in an observation.

### 3.1.6 Number of Frequency Peaks (NP)

The NP index was developed as a rapid biodiversity assessment tool to estimate the abundance of acoustic activity in tropical and subtropical mountain habitats where habitat ranges of some species are extremely spatially limited. Peaks are determined based on the steepness of amplitude slopes with specified time distance between slopes (Gasc et al., 2013b). A simplified interpretation of what this index means in the acoustic environment is as follows:

- The higher the value, the greater the number of frequency peak events within specified amplitude and frequency thresholds.

The index is calculated in R with the function fpeaks() in the package seewave (Sueur et al., 2008a), that computes the mean frequency spectrum with the function meanspec (3.2.3). The fpeaks() function has adjustable amplitude, frequency and threshold parameters (Sueur, 2018 pg. 265) that should be considered based on the type of acoustic environment sampled the source from which peaks are to be quantified. The index ranges from zero to infinity where higher values indicate more frequency peaks.

### 3.1.7 Spectral Entropy ($H_f$)

The spectral entropy index ($H_f$) estimates entropy over time in the frequency domain by calculating Shannon evenness of the probability mass function derived from the (mean) frequency spectrum (3.2.1) (Sueur, 2018 pg. 485). The index is valuable because acoustic environments of greater frequency entropy may indicate non-uniform or chaotic acoustic environments, whereas acoustic environments with low entropy may indicate frequency uniformity. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of frequency evenness in an acoustic environment.

The index is automatically calculated with the functions sh() in seewave (Sueur et al., 2008a) on a mean frequency spectrum (3.2.3) generated from function meanspec (3.2.3). The alpha argument in sh() allows the selection of Shannon evenness or Simpson entropy index, with Shannon evenness as the default value. The index range is from 0 to 1, with higher values indicating greater evenness within a frequency spectrum and lower values indicating increasing entropy within a frequency spectrum (Sueur et al., 2008a).

### 3.1.8 Temporal Entropy Index ($H_t$)

The temporal entropy index ($H_t$) (Sueur et al., 2008b) estimates the Shannon evenness of the Hilbert amplitude envelope (3.2.1) scaled by its sum (Sueur, 2018 pg. 483). The index is useful because it measures the evenness of amplitude over the recording period, which could indicate areas that are evenly loud or evenly quiet when cross-referenced with $LA_{eq}$ or SPL measures –a consideration in noise or annoyance studies (Lercher and Schulte-Fortkamp, 2003). A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of amplitude evenness in an acoustic environment.

The index is calculated in a two-step process in R with the function env() in seewave to calculate the Hilbert envelope and then th() in seewave to calculate the index (Sueur et al., 2008a; Sueur, 2018 pg. 483). The range of the index is 0 to 1, with one equating to complete evenness of the Hilbert amplitude envelope (no entropy) and zero equating to complete unevenness of the Hilbert amplitude envelope (complete entropy).

### 3.1.9 Acoustic Entropy Index (H)

The acoustic entropy index (H) is the product of $H_t$ and $H_f$, thereby considering the entropy of both the frequency and amplitude domains together (Sueur et al., 2008b; Sueur, 2018 pg. 486). The index is useful because it clarifies whether an acoustic environment is consistently loud or quiet and if the amplitude is evenly or unevenly distributed across the frequency spectrum. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of frequency and amplitude evenness in an acoustic environment.

The index is calculated in R with the function H() in the package seewave (Sueur et al., 2008a) with default parameters of Hilbert envelope (3.2.1) for $H_t$ and $H_f$ selected. The index value ranges from 0 to 1 where higher values indicate low entropy where consistently even frequency ranges at a consistent amplitude occur over time in an observation and low values indicate high entropy where variation of both frequency and amplitude occur throughout an observation.

### 3.1.10 Amplitude Index (M)

The amplitude index (M) was developed to estimate the overall amplitude of an acoustic environment (Depraetere et al., 2012) in temperate woodlands. The index is calculated as the median of the amplitude envelope (3.2.1) scaled by the digitization depth[1] of the recording (Sueur, 2018 pg. 482-483). The index is useful because it informs about the consistent loudness of an acoustic environment. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of how close the median loudness is to the maximum loudness.
- An absolute measure of the median amplitude of a recording scaled by the maximum amplitude.

[1] Also referred to as bit-depth or quantization with common units of 8-bit, 16-bit, and 24-bit (https://www.presonus.com/learn/technical-articles/Sample-Rate-and-Bit-Depth).

The formula is calculated in R with the function M() in package seewave (Sueur et al., 2008a) with the absolute or Hilbert envelope (3.2.1). There are no variable arguments in M other than digitation depth, which is embedded in a WAV already. The index range is 0 to 1, where a value of one indicates that the median amplitude of the recording is identical to the maximum amplitude over the entire duration of the recording and a values closer to zero indicates that the median amplitude is almost never the same as the maximum amplitude over the entire duration of a recording.

#### 3.1.11 Acoustic Richness (AR)

Acoustic richness (AR) was developed to assess species richness in temperate climates where fewer bird calls and higher background noises, such as wind or rain, can bias the signal to noise ratio of $H_f$ (Depraetere et al., 2012 pg. 48). The index is the product of the ranked M and $H_t$ values from a group of observations, divided by the square of the observations (Sueur, 2018 pg. 484). Thus, the index is a measure of relative acoustic richness of one observation compared to an entire group of observations. Since AR is a ranked function, the value is relative to the group of observations with which it was calculated. The index can be calculated for land uses, times of day, or seasons for which one wants to understand the ranked amplitude differences. A simplified interpretation of what this index means in the acoustic environment is as follows:

- A measure of persistent loudness of the acoustic environment.
- A measure of amplitude saturation (richness) within a single acoustic environment observation compared to a group of observations.

The index is calculated in R with the function AR() from the package seewave (Sueur et al., 2008a). Since the index is composited from M and $H_t$, there are no adjustable arguments in the index. The index ranges from 0 to 1, where higher values indicate a higher consistent amplitude (richness) and lower values indicate less consistent amplitude characterized as higher amplitude entropy environments.

### 3.2 Beta Indices

The beta indices focus on comparisons of the frequency spectra in pair-wise analysis. The preliminary step to calculate most of the beta indices is to transform the WAV into a cumulative distribution function (3.2.4) or cumulative probability mass functions (Sueur, 2018 pg. 495). After transformation, frequency spectra can be compared using a number of different mathematical arguments.

In all cases, the beta indices return a matrix of pair-wise comparisons between observations. This means that unlike the alpha indices where there is one independent value calculated for each index in an observation, the beta indices change based on their pair-wise comparison. One way to reduce this complexity is to use the beta indices in pair-wise comparisons in a spatially or temporally restrictive hypothesis testing approach. In this approach, observations that are expected to be similar or dissimilar based on their placement within similar or different land uses, land cover types, times of day, seasons, or distance to roadways, forests, etc., can be grouped and compared to validate or refute a similarity or dissimilarity hypothesis between acoustic environments.

### 3.2.1 Spectral Dissimilarity ($D_f$)

The spectral dissimilarity index ($D_f$) calculates the dissimilarity between two frequency spectra then sums the absolute difference in amplitude for all frequency bins (Sueur, 2018 pg. 496). The index was developed for rapid avifaunal appraisal in tropical regions to identify and differentiate different bird habitats (Sueur et al., 2008b). The index is useful because the comparison of frequency spectra with $D_f$ compares the entire Hz range and thus should identify differences within the urban acoustic environment as well. A simplified interpretation of what this index means is as follows:

- A comparative measure of the total differences of amplitude distribution amongst frequency bins between two observations.

The index is calculated with the function diffspec() in the package seewave (Sueur et al., 2008a) which uses meanspec() to compute the mean frequency spectrum (3.2.3). The index ranges from 0 to 1 where values of one indicate that the pairwise comparison has completely different frequency composition in all frequency bins over the time window and values of zero indicate that the pairwise comparison has a completely similar frequency composition between bins over the time window.

### 3.2.2 Cumulative Spectral Dissimilarity ($D_{cf}$)

The cumulative spectral dissimilarity ($D_{cf}$) is similar to the spectral dissimilarity index but instead of comparing mean frequency spectra, the index compares cumulative frequency spectra (3.2.4) (Sueur et al., 2008b). The index is useful because comparisons of mean values over time can weight results for long continuous or single powerful sounds in an acoustic environment, thereby missing outlier values or misrepresenting non-normally distributed frequency spectra. The cumulative comparison can indicate total differences between observations and highlight cases where mean values may not provide a full understanding of the acoustic environment. A simplified interpretation of what this index means is as follows:

- A measure of the total difference between two spectra by comparing the cumulated amplitude in all frequency ranges over the entire time of the observation.

The function diffcumspec() in seewave is used to calculate $D_{cf}$ which in turn relies on meanspec() (3.2.3). The index ranges from 0 to 1, where zero represents total similarity of cumulative frequency between two spectra and one represents total dissimilarity cumulative frequency between two spectra (Sueur, 2018 pg. 497).

### 3.2.3 Kolmogorov-Smirnov Distance ($D_{KS}$)

The Kolomogorov-Smirnov Distance ($D_{KS}$) calculates the frequency location of the maximum difference between two cumulative frequency spectra using the Kolomogorov-Smirnov distance, returning both the maximum difference and the frequency location of the maximum difference (Massey, 1951). The index was developed to identify differences in species richness and assemblages in avian habitat (Gasc et al., 2013a). The index is useful for the urban acoustic environment because it will clarify which land use types have the greatest maximal statistical differences and in which frequency range that maximum distance occurs. A simplified interpretation of what this index means is as follows:

- The KS distance is the maximum distance between cumulative frequency spectra from two different WAVs and indicates the frequency range of the maximum difference.
- A measure of statistical or probabilistic maximum difference between two observations.

The $D_{KS}$ is calculated with the function ks.dist() in seewave (Sueur et al., 2008a) that includes adjustable arguments as discussed within meanspec() (3.2.3). All arguments in ks.dist() can be left at default values when starting with a WAV recording. The index range is from 0 to 1, where lower values indicate a smaller maximum K-S distances and higher values indicate larger maximum K-S distances between frequency spectra.

### 3.2.4 Log-Spectral Distance ($D_{LS}$)

The log-spectral distance ($D_{LS}$) is the log of the difference between two frequency spectra, which can be scaled by the length of the recording (Sueur 2018 pg., 502). The benefit to using log is that it returns a positive number quantifying the power to which one observation (base) must be raised in order to equal a second (larger) observation (Shirali, 2002). An increasing log, translates into an exponentially increasing difference between two observations. Since dB is also a logarithmic ratio, it makes sense to compare frequency spectra using a logarithmic scale. The $D_{LS}$ is a counterpoint to the non-logarithmic KS distance, and it is useful to compare day or night recordings and testing differences between land use types. A simplified interpretation of what this index means is as follows:

- The power in amplitude across a frequency spectrum to which one observation must be raised to equal a second observation, quantified as a positive ‘distance’.

The $D_{LS}$ is calculated in R with the function logspec.dist() in seewave (Sueur et al., 2008a) that has the adjustable argument scale, where the distance can be scaled by time. The frequency spectrum distribution is calculated with the function meanspec() (3.2.3). The value ranges from zero to the maximum possible log of the spectral distances, with lower values indicating a lower log difference between observations and larger values indicating a larger log difference between observations.

### 3.2.5 Mutual Information (I)

This Mutual Information index (I) transforms the frequency spectra of two different recordings into five symbols representing an a) increase, b) decrease, c) peak, d) trough, e) flat region of each spectrum. The absolute frequency of each symbol is computed and then used to calculate a level of entropy by applying a normalized Shannon's index on the re-occurrence of the five symbols. The entropy values are combined for the five symbols in both spectra and calculates the overlap, or mutual information, between both spectra (Cazelles, 2004; Sueur, 2018 pg. 501). A simplified interpretation of what this index means is as follows:

- Measures the differences in shape between two cumulative frequency spectra.
- A multi-dimensional difference in the frequency distribution between two recordings, measuring differences in increases, decreases, peaks, troughs and length of flat segments between two spectra.

The index is calculated with the function symba() in seewave (Sueur et al., 2008a), further discussed in Ligges et al. (2018), for which all parameters are left in their default values and the index is calculated as a dissimilarity index using the expression 1-symba() (Sueur, 2018 pg. 501). The index range is 0 to 1, where the closer the value to one, the greater the dissimilarity (entropy) between all 5 symbols in the two spectra, and the closer the value to zero then the more similar (less entropy) in all five symbols between the two spectra.

### 3.2.6 Relative Frequency Dissimilarity (S)

The Relative frequency dissimilarity index (S) calculates the relative dissimilarity between minimum and maximum frequency of two frequency spectra, expressed as a percentage (Deecke and Janik, 2006 pg. 648). The index is useful because it compares the total range of difference between frequency spectra rather than just the maximum difference ($D_{KS}$) between spectra. Ranges in the urban acoustic environment are useful to identify land uses with relatively homogenous and heterogeneous frequency and amplitude distributions. A simplified interpretation of what this index means is as follows:

- The total minimum and maximum amplitude difference amongst all frequency bins, normalized by time.

The index is calculated with the function simspec() in seewave (Sueur et al., 2008a) where the WAV is converted to a mean frequency spectrum using default parameters in the function meanspec() (3.2.3). The function simspec() is a similarity index, but is converted to a dissimilarity index in WAM by applying the simple argument 100 – S (Sueur, 2018 pg. 502) to test hypotheses of differences between land uses and seasons. The index ranges from 0% to 100%, where 100% would indicate complete dissimilarity and 0% would indicate complete similarity between two frequency spectra.

### 3.2.7 Correlation Based Dissimilarity (R)

This index uses a correlation coefficient to estimate how much two frequency spectra covary with each other using Pearson's (Pearson, 1895), Spearman's (Spearman, 1904) or Kendal (Kendal, 1938) tests, depending on the scale level and distribution of the acoustic data. The index can be calculated in inverse to generate a correlation-based dissimilarity index, which is applied when the study aim is to find dissimilarities in the acoustic environment between land uses, seasons, and time (Sueur, 2018). A simplified interpretation of what this index means is as follows:

- The similarity or dissimilarity between mean frequency spectra of two observations using positive correlation as the determination for similarity.

The index is calculated in R with the function cor() from the package stats (Lellouch et al., 2014) and is made a dissimilarity index by calculating 1- cor (Lellouch et al., 2014; Sueur 2018 pg. 503). The index ranges from 0 to 1, where 0 indicates highly positive values of Spearman's r between two mean frequency spectra (similarity) and 1 indicates highly negative values of Spearman's r between mean frequency spectra (dissimilarity).

### 3.2.8 Temporal Dissimilarity ($D_t$)

The temporal dissimilarity index ($D_t$) compares the difference in absolute or Hilbert envelopes (3.2.1) between two recordings to determine the similarity in distribution of amplitude amongst frequency ranges between two frequency spectra (Sueur et al., 2008a; Sueur, 2018 pg. 503). The index is useful to identify areas where the temporal distribution of frequency are not the same, thereby inferring differences in evenness, entropy or diversity between acoustic environments. A simplified interpretation of what this index means is as follows:

- The index compares similarity and dissimilarity in amplitude distribution amongst frequency bins between two observations.

The $D_t$ index is calculated in R with the function diffenv() from seewave (Sueur et al., 2008a) with adjustable smoothing parameters. The index ranges from 0 to 1, where values of one indicate that the pairwise comparison has completely dissimilar Hilbert envelope composition in all frequency bins over the time window and values of zero indicate that the pairwise comparison has a completely similar Hilbert envelope composition between bins over the time window.

### 3.2.9 Acoustic Dissimilarity (D)

Acoustic dissimilarity is the product of temporal ($D_t$) and frequency ($D_f$) dissimilarity indices and can be used to estimate the relative mathematic distance (dissimilarity) between two WAVs for either the absolute or the Hilbert envelope (3.2.1) (Sueur et al., 2008b pg. e4065). The D index is useful as a composite index of $D_t$ and $D_f$ which may pinpoint acoustic environments where differences in both amplitude and frequency occur and possibly be more useful than $D_t$ or $D_f$ alone which can only identify one or the other. We assume that there will be high correlation between $D_t$, $D_f$ and D in urban acoustic environments. A simplified interpretation of what this index means is as follows:

- The index is the product of the dissimilarities in temporal and frequency vectors.

The D index can be calculated in R using the function diffwav() from package seewave (Sueur et al., 2008a). The index ranges from 0 to 1, where one equates to total frequency dissimilarity over all time steps and zero equates to total frequency similarity between every time step.

## 3.3 Sound Pressure Level (SPL) Indices

In addition to the alpha and beta indices, WAM includes a selection of well-understood and applied SPL-based measures that have been used in a wide range of past studies. The inclusion of decibel indices is critical, because they provide a central pivot common to many studies that can be used to put the alpha and beta indices into a known context.

### 3.3.1 Equivalent Continuous Sound pressure level ($LA_{eq}$)

$LA_{eq}$ is an A-weighted dB measure of equivalent continuous SPL, where A-weighting refers to the adjustment of dB measures necessary to account for the perceived loudness of an acoustic environment by humans (Fletcher and Munson. 1933; IEC 2013, 2003; Kang, 2007; TS, 2019) SPL meters provide $LA_{eq}$ values.

### 3.3.2 A-Weighted Mean Sound Pressure Level (SPL dBA)

This measure is calculated from the WAV using Wildlife Acoustics Noise Analysis software (Wildlife Acoustics, 2019). The SPL in dB is calculated for every second with A weighting and then the mean value of all 1 second dB(A) values is calculated to produce a value equivalent to an A-weighted $L_{eq}$ with +/- 4 dB accuracy.  Recognizing the limits of this value for precise SPL measures, this index is still useful to understand daily, weekly, monthly, seasonal, and annual SPL variations in the urban environment.

### 3.3.3 Distribution-Based Statistical Noise Levels ($L_n$)

We included statistical measures of loudness in WAM to provide a measure of ‘amplitude saturation’ that gives the idea of what the maximum, average and minimum A-weighted dB levels are in an acoustic environment (Kang, 2007 pg. 27). The $L_{10}$ is the A-weighted dB value exceed for 10% of the recording time and has a high correlation with continuous high volume traffic noise (Georgiadou et al., 2004) and loud foreground sound incursions. The $L_{50}$ is the A-weighted dB value exceeded for 50% of the recording time representing the median or middle ground SPL level. The $L_{90}$ is the A-weighted dB value exceed for 90% of the recording time generally corresponding to background SPL exposure levels (Kang, 2007; Margaritis et al., 2018). The $L_{10}$, $L_{50}$, and $L_{90}$ measures are usually provided by SPL meters.

### 3.3.4 Day-evening-night level ($L_{den}$)

Temporal variations of the day-evening-night level, $L_{den}$, is calculated using mean SPL dB(A) values from day (7:00 to 19:00), evening (19:01 to 23:00) and night (23:01 to 6:59) using the equations in Kang (2007 pg. 29). One might also calculate the often-used day-night level ($L_{dn}$), however, since the $L_{den}$ already calculates a direct comparison of day and night recordings we conclude that it alone is sufficient.  To include this metric, one must have sound data for all hours of the day, as can be collected with automatic recorders.

# 4 Bibliography


Boelman N. T., Asner G. P, Hart P. J., and Martin R. E. (2007). Multitrophic invasion resistance in Hawaii: bioacoustics, field surveys, and airborne remote sensins. Ecological Applications. 17: 2137–2144. Appendix 1 accessed on March 3, 2020 from: http://www.esapubs.org/archive/appl/A017/086/appendix-A.htm

Borchers, H.W. (2019). pracma: Practical Numerical Math Functions. R package version 2.2.9. https://CRAN.R-project.org/package=pracma

Cazelles, B. (2004). Symbolic dynamics for identifying similarity between rhythms of ecological time series. *Ecology Letters*, *7*(9), 755-763.

Deecke, V. B., and Janik, V. M. (2006). Automated categorization of bioacoustic signals: avoiding perceptual pitfalls. The Journal of the Acoustical Society of America, 119(1), 645-653.

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

Dowle M. and Srinivasan A. (2019). data.table: Extension of `data.frame`. R package version 1.12.8. https://CRAN.R-project.org/package=data.table

Farina, A., and Morri, D. (2008). 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, 365–372.

Flanagan, J. L. (2013). *Speech analysis synthesis and perception* (Vol. 3). Springer Science & Business Media.

Fletcher, H., and Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, *12*(4), 377-430.

Gasc, A., Sueur, J., Jiguet, F., Devictor, V., Grandcolas, P., Burrow, C., and Pavoine, S. (2013a). Assessing biodiversity with sound: Do acoustic diversity indices reflect phylogenetic and functional diversities of bird communities?. *Ecological Indicators*, *25*, 279-287.

Gasc, A., Sueur, J., Pavoine, S., Pellens, R., and Grandcolas, P. (2013b). Biodiversity sampling using a global acoustic approach: contrasting sites with microendemics in New Caledonia. *PloS one*, *8*(5).

Georgiadou, E., Kourtidis, K., & Ziomas, I. (2004). Exploratory traffic noise measurements at five main streets of Thessaloniki, Greece. *Global NestI International Journal*, *6*(1), 53-61.

Gini, C. (1909). "Concentration and dependency ratios" (in Italian). English translation in *Rivista di Politica Economica*, **87** (1997), 769–789.

International Electrotechnical Commission (IEC) (2013). Electroacoustics—Sound level meters—Part 1: Specifications (IEC 61672-1). *Geneva, Switzerland*.

Ihaka, R.R. and Gentleman, R. (1996). R: A language for data analysis and graphics. J. Comput Graph Stat.; 5: 299-314.

Kang, J. (2007). Urban sound environment. CRC Press.

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

Kelley D. and Richards C. (2019). oce: Analysis of Oceanographic Data. R package version 1.1-1. https://CRAN.R-project.org/package=oce

Kelley D., Richards C. and WG127 SCOR/IAPSO (2017). gsw: Gibbs Sea Water Functions. R package version 1.0-5. https://CRAN.R-project.org/package=gsw

Kendall, M. (1938). "A New Measure of Rank Correlation". *Biometrika*. **30** (1–2): 81–89. *doi*:*10.1093/biomet/30.1-2.81*. *JSTOR* *2332226*.

Lercher, P., and Schulte-Fortkamp, B. (2003). The relevance of soundscape research to the assessment of noise annoyance at the community level. In *Proceedings of the Eighth International Congress on Noise as a Public Health Problem* (pp. 225-231).

Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H., and Sueur, J. (2014). Monitoring temporal change of bird communities with dissimilarity acoustic indices. *Methods in Ecology and Evolution*, *5*(6), 495-505.

Ligges U., Krey S., Mersmann O., and Schnackenberg S. (2018). tuneR: Analysis of Music and Speech. URL: https://CRAN.R-project.org/package=tuneR

Margaritis, E., Kang, J., Filipan, K., and Botteldooren, D. (2018). The influence of vegetation and surrounding traffic noise parameters on the sound environment of urban parks. In Applied Geography; 94, pp. 199–212. DOI: 10.1016/j.apgeog.2018.02.017.

Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, *46*(253), 68-78.

Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P.R.,O'Hara R. B., Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E. and Wagner H. (2019). vegan:Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

Odum, E. P., & Barrett, G. W. (1971). *Fundamentals of ecology* (Vol. 3, p. 5). Philadelphia: Saunders.

Park, T. H., Lee, J. H., You, J., Yoo, M. J., and Turner, J. (2014). Towards soundscape information retrieval (SIR). In *ICMC*.

Pearson, K. (1895). VII. Note on regression and inheritance in the case of two parents. *proceedings of the royal society of London*, *58*(347-352), 240-242.

Pieretti, N., Morri, D., and Farina, A. (2011). A new methodology to infer the singing activity of an avian community. The Acoustic Complexity Index (ACI). In *Ecological Indicators* 11 (3), pp. 868–873. DOI: 10.1016/j.ecolind.2010.11.005.

Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., and Pieretti, N. (2011). Soundscape ecology: the science of sound in the landscape. BioScience; 61(3), 203-216.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rouse Jr, J. W. (1973). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation.

Schauberger P. and Walker A. (2019). openxlsx: Read, Write and Edit xlsx Files. R package version 4.1.4. https://CRAN.R-project.org/package=openxlsx

Shirali, S. (2002), *A Primer on Logarithms*, Hyderabad: Universities Press, *ISBN 978-81-7371-414-6, esp. section 2*

Spearman C. (1904). The proof and measurement of association between two things. American Journal of Psychology. **15** (1): 72–101. *doi*:*10.2307/1412159*. *JSTOR* *1412159*.

TS III. DIN ISO 12913-3 (E) (2019). —Acoustics—Soundscape Part 3: Data Analysis. *ISO: Geneva, Switzerland*.

Weihs, C., Jannach, D., Vatolkin, I., and Rudolph, G. (2016). *Music data analysis: Foundations and applications*. Chapman and Hall/CRC.

Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, *15*(2), 70-73.

Wildlife Acoustics (2019). Kaleidoscope Pro Analysis Software. Wildlife Acoustics, Inc, Maynard MA. https://www.wildlifeacoustics.com/products/kaleidoscope-pro.

Zeileis A. (2014). ineq: Measuring Inequality, Concentration, and Poverty. R package version 0.2-13 https://CRAN.R-project.org/package=ineq