DOI: 10.61782/fa.2025.0491

INTRODUCING NORBAERT– AN OPEN SOURCE SOFTWARE TO FACILITATE THE RESEARCH ON COMPLEX NETWORKS FROM THE ACOUSTIC ENVIRONMENT.

Timo Haselhoff[1*] **Mikel Jedrusiak**[2] **Frank Weichert**[2] **Susanne Moebus**[1]
[1] Institute for Urban Public Health, University Hospital Essen, Essen, Germany
[2] Department of Computer Science, Computergraphics, TU Dortmund, Dortmund, Germany

ABSTRACT

The urban acoustic environment (AE) encapsulates information of systems like traffic, the built environment, or biodiversity. Reduced costs of storage capacities and computational power enabled the collection of extensive acoustical datasets in previous years. However, few established approaches exist that are tailored to capture the longitudinal time-frequency dynamics of the urban AE.

Recently, complex networks were suggested as a methodology to analyse the urban AE. Metrics such as Link Density can quantify e.g. acoustic dominance, offering insights into the extent that single sound sources influence time-frequency dynamics. Studies have demonstrated associations between these metrics and urban land-use patterns, sound sources, and human perception. Additionally, complex network metrics do not require precise calibrations of sound pressure levels and provide a consistent visual representation of acoustic dynamics that is independent of recording length.

To advance research in this field we introduce the open-source software NORBAERT, which calculates frequency correlation matrices (FCMs), complex networks and selected measures for single or multiple audio recordings and provides a visual representation of the respective FCMs and adjacency matrices. Its aim is to facilitate broader adoption of this methodology to extent research and to provide the basis for advances in complex network research of the AE.

Keywords: *acoustic environment, soundscape, complex networks, acoustic indices, software*

[*]***Corresponding author***: *timo.haselhoff@uk-essen.de*

1. INTRODUCTION

The recent decline in the cost of audio recordings has made it feasible to conduct high-resolution temporal and spatial sampling of the acoustic environment (AE). However, the reduced costs often lead to the generation of vast amounts of data, reaching several terabytes, while effective analysis methods for large-scale audio data, particularly in urban environments, remain limited. In addition to traditional noise measurements using different variations of sound pressure levels, psycho- and ecoacoustic indices are currently being investigated for their potential to provide information on the urban AE [1, 2]. However, as those were not designed to consider specific characteristics of the urban AE, many indices do not provide the intended information [1] or even produce contradictory results [3]. Recently, frequency correlation matrices (FCMs) and complex networks derived from those, were suggested as a promising approach to investigate the urban AE [4].

A detailed description of the methodological aspects of FCMs and complex networks corresponding to the AE can be found elsewhere [4]. In its approach to not rely on absolute amplitude measures, but rather investigate the interrelationship between relative frequency amplitudes over time, FCMs represent a different approach to many other acoustic measures. Without relying on precisely calibrated sound pressure levels, FCMs are a promising tool for low-cost acoustic measures, as devices used are often of lower overall measurement accuracy. In addition, an advantage of FCMs is that their dimension depends solely on the number of frequency bins. This consistency simplifies follow-up analyses like image classification and network analysis, avoiding the issue of varying input dimensions due to different recording lengths. Furthermore, it thus offers a solution for visualizing large longitudinal audio datasets [5].

Regarding its applications, FCMs already delivered insights into source contributions to low frequency ocean sound [6] or improved the detection of bird sounds in noisy settings

[7]. Jedrusiak et al. demonstrated that FCMs can be used as a "fingerprint" to classify and identify the AEs of different urban recording locations, using convolutional neural networks [8]. Furthermore, FCM-based network measures like Link Density have been identified as key descriptors for the AE of health-related urban green spaces [9] as well as being associated to human perception of, e.g., pleasant or annoying soundscapes [10].
However, despite the already promising results, this method is still in its infancy. To facilitate research in this direction, we introduce NORBAERT (Networks Of Recordings Based on the urban Acoustic EnviRonmenT) as a framework, offering a user interface-based program for examining FCMs, complex networks and selected measures for single or multiple audio recordings. In the following, we provide information on its calculations and functionality. Python-code and desktop applications can be found on https://github.com/THaselhoff/NORBAERT.

2. METHODS

In general, there are two ways in which FCMs, complex networks and selected network measures can be calculated:

1. Within each recording (default state), producing output for each input recording separately;
2. Between recordings (enabled by checking the box "Between all recordings"), producing one output for all input recordings (see Fig. 1).

2.1 Frequency power

The calculation procedure for frequency power differs slightly for both approaches. A requirement for recordings analysed with NORBAERT is a minimum sampling frequency of 44.1 kHz.

2.1.1 Within each recording (one output for each recording)

Soundscapes can be understood as a temporal composition $\boldsymbol{x}(t) \in \mathbb{R}^n$ of various sound sources, as outlined in our developed formalization [8]. To analyse the frequency range $u \in \{0, \dots, 44.1\ kHz\}$ over time of the recorded composition, only the first (usually the left) channel (in case of stereo) is used. The signal is split into $n \in \mathbb{N}$ parts, where n is determined by the length of the recording in seconds divided by $s \in \mathbb{N}$ seconds, rounded down. Here, s is defined by the parameter "Snippet Length" from the NORBAERT user interface (UI). To examine the spectral composition of the urban sound environment over time, a Fast Fourier Transform (FFT) is performed for each of the n segments of the signal, transforming the time-domain signal $\boldsymbol{x}(t)$ into the frequency domain. If the recording has a sampling frequency $u > 44.1$ kHz, it will be limited to $u_{max} = 44.1$ kHz:

$$F_k(u) = \frac{1}{u_{max}} \sum_{t=0}^{u_{max}} \boldsymbol{x}_k(t) e^{-2\pi i \frac{tu}{u_{max}}} \in \mathbb{C} \tag{1}$$

The parameter $k \in \{1, \dots, n\}$ represents the split number and depends on the recording split. Following this approach, we extract the amplitude spectrum $|F_k(u)|$ for each of the n segments for each frequency u. The data is then segmented into 1024 uniformly spaced frequency bins. After computing the root mean square (RMS) for each bin, a logarithmic transformation is applied, yielding a $n \times 1024$ matrix $\boldsymbol{M}$ with entries

2.1.2 Between recordings (one output for multiple recordings)

Using this mode, recordings are not split into parts, rather, the FFT is calculated for the entire recording. The following procedure is the same as described in section 2.1.1. The result is a $n \times 1024$ matrix, but here, n is the number of recordings.

2.2 Frequency correlation matrix

The FCM for both approaches is calculated by correlating the frequency power of each frequency bin with all other bins over time using Pearson correlation (r), resulting in a 1024×1024 matrix. To ease interpretation, R^2 is reported. By default, the matrix is then limited to 13 kHz, as previous research showed no substantial power above that threshold in the urban environment [11]. However, this can be adjusted by the parameter "Frequency Limit" from the UI.

2.3 Adjacency matrix

The adjacency matrix is calculated by thresholding each value of the FCM by r, where r is the correlation threshold defined by the parameter "Threshold" from the NORBAERT UI. For each value of the FCM, if it is smaller than r, it is set to zero, else it is set to 1. In addition, the diagonal is set to 0. The result is a binary adjacency matrix (for unweighted networks and networks without multi- and self-edges), which indicates which frequency bins are related over time.

2.4 Network measures

NORBAERT provides the possibility to directly calculate the Link Density. As Link Density has proven to be an effective measure of FCM based complex networks [9] it is included in this version of NORBAERT.

2.4.1 Link Density

Link Density is a measure of connections between frequency bins in the FCM-based network that are actually present in relation to the maximum of possible connections [12]. It ranges between 0 and 1, where 0 indicates no connections and 1 a fully connected network. It represents a measure of so-called acoustic dominance. AEs of high acoustic dominance (e.g. streets) consist of single sound sources that occupy a majority of the frequency range over time and thus result in a Link Density closer to 1. AEs of low acoustic dominance (e.g. forests) consist of multiple sound sources that occupy different frequency ranges over time and thus result in a Link Density closer to 0 [12].

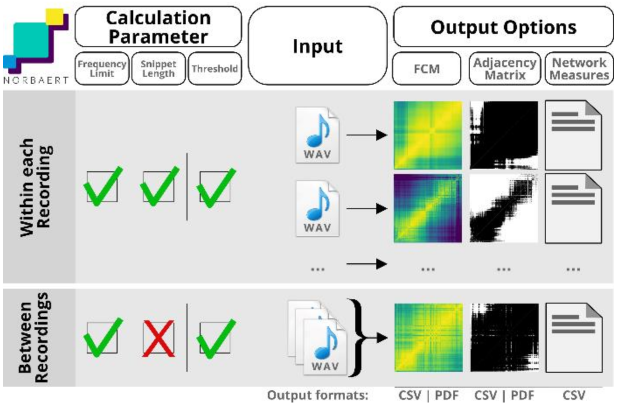

Figure 1. Overview on required calculation parameters (indicated by the green tick), input and output options for both approaches. "Within each Recording" produces one output for each recording, "Between Recordings" produces one output for all recordings in the respective input directory (FCM=Frequency Correlation Matrix).

3. HOW TO USE NORBAERT

3.1 Input

The "Input Directory" from the NORBAERT UI specifies the directory of the audio files for the analysis. Considered are all ".wav" files in the specified directory and its subdirectories. Results from studies mentioned above are based on a minimum recording length of 180 s [4, 8, 9, 10, 11]. Shorter recordings will work, but the shorter the length, the greater the chance of picking up only single sound signals.
The "Output Directory" specifies the directory where the analysis results shall be stored. Depending on the selected output options, folders will be created containing the respective output. By default, outputs will be produced for each file separately. By checking the box "Between all recordings", one output will be produced for all files in the specified input directory.
Under "Input Options", the sampling frequency for all audio files needs to be specified. All audio files need to have the same sampling frequency and a minimum sampling frequency of 44.1 kHz.

3.2 Calculation parameter

There are three parameters that need to be specified. The "Frequency Limit" defines the upper frequency limit (in Hz) of the FCM and its respective network. Note that this does not change the calculation of the frequency power. Rather, it limits the 1024 × 1024 matrix up to the frequency bin, in which the defined frequency limit is contained.
The "Snippet Length" is essential only for the "Within each recording" mode. It defines the length for each part a single audio file is divided into. As the FFT is performed on these parts, it cannot be smaller than 1 s without compromising the frequency resolution. Previous research in the urban environment is based on a minimum snippet length of 10 s, therefore, it is given as the default option. Different settings might be feasible, depending on the specific research questions. However, the recording length divided by the snippet length has to be $n > 2$ to provide correlation values smaller than 1. From previous research, we recommend n to be at least 15.
The "Threshold" defines the correlation threshold for the FCM to obtain the adjacency matrix and is essential for both modi. As with the "Snippet Length", the default setting of 0.8 is derived from previous research [4], but different settings might be feasible, especially for shorter snippet lengths.

3.3 Output

There are three main categories of outputs produced by NORBAERT. First, the FCM as well as the Adjacency Matrix can be exported as ".csv". Respectively, folders named "FCMs" and "Adj_Matrices" will be created in the output directions, containing a ".csv" file for each recording, named by the recording. If the mode "Between

Recordings” is selected, the file will be named by the specified input directory.
With the “.csv” output, additional analysis or visualizations can be carried out (e.g. image classification, calculation of additional network measures, visualizations of the networks). The first column and row of the “.csv” correspond to the frequency mids of the respective frequency bins.
The second output option is the calculation of Link Density. If checked, a “.csv” file will be created, including the Link Density of the respective adjacency matrix. The file contains a list of all audio files in the input directory and their corresponding Link Density. If the mode “Between recordings” is selected, the file will contain the input directory as row name and the respective Link Density based on the adjacency matrix calculated between all audio files.
The third output option is a visualization of the respective FCMs and adjacency matrices as “.pdf” files (see Fig. 2). Respectively, folders named “FCMs_pdf” and “Adj_Matrices_pdf” will be created in the output directions, containing a “.pdf” file for each recording, named by the recording. If the mode “Between Recordings” is selected, the file will be named by the specified input directory. The given frequency labels correspond to frequency mids of the respective frequency bins.

4. FUTURE DEVELOPMENT AND CONCLUSION.

FCMs and complex networks based on them already provide valuable insights into AE. Previous research has shown promising results in capturing location-specific frequency dynamics of AE, especially for the urban environment. Fewer studies have investigated FCMs outside the urban environment. In order to promote the wider use of this methodology, to extend the research, and to lay the foundations for advances in the study of complex AE networks, we have introduced NORBAERT. It is intended to serve as a basic framework for examining the AE and related variables using FCMs and complex networks.
At the time, holistic approaches that consider a wide frequency range or the use of rather simple complex network measures like Link Density already provide useful insights. In the future, adjusted adaptions of these approaches could lead to more sophisticated results. For instance, the focus on selected frequency ranges might be more suitable to differentiate between bird calls or the use of additional network measures could yield more detailed insights into the properties of the underlying AE. We aim to integrate the option to focus on specific frequency ranges in the future. In addition, we intend to include other network measures if they are found to be useful in the literature.

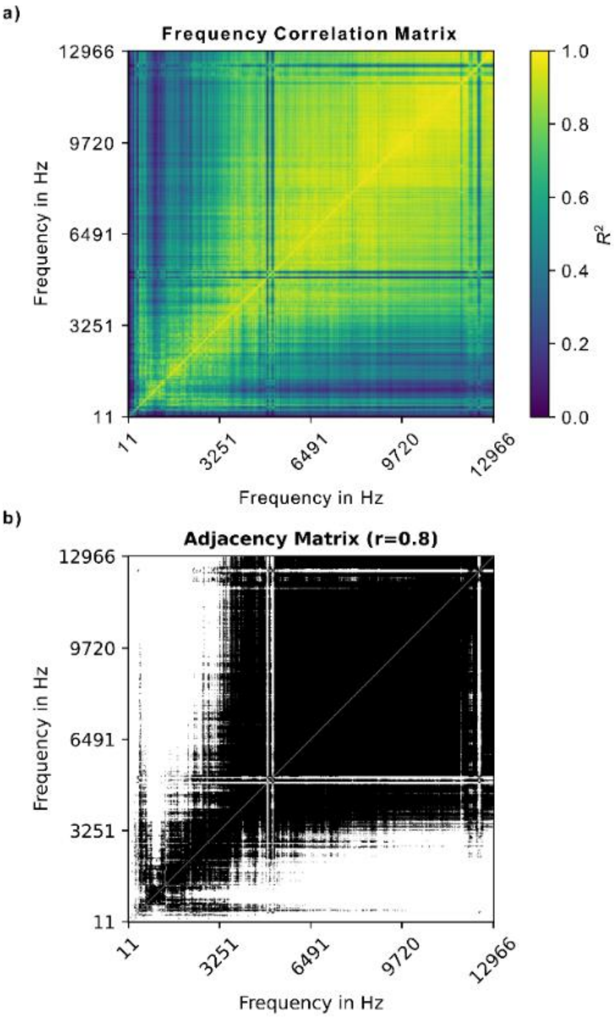

Figure 2. Two examples of the output figures for a FCM (a) and its corresponding adjacency matrix (b), generated for a single recording.

5. REFERENCES

[1] A. J. Fairbrass, P. Rennert, C. Williams, H. Titheridge, and K.E. Jones: “Biases of acoustic indices measuring biodiversity in urban areas”, *Ecological Indicators*, 83, 169-177, 2017.

[2] M. S. Engel, A. Fiebig, C. Pfaffenbach, and J. Fels: “A review of the use of psychoacoustic indicators on

soundscape studies", *Current Pollution Reports*, 1-20, 2021.

[3] T. Haselhoff, J. Hornberg, J. L. Fischer, B. T. Lawrence, S. Ahmed, D. Gruehn, and S. Moebus: "The acoustic environment before and during the SARS-CoV-2 lockdown in a major German city as measured by ecoacoustic indices", *The Journal of the Acoustical Society of America*, 152(2), 1192-1200, 2022.

[4] T. Haselhoff, T. Braun, A. Fiebig, J. Hornberg, B. T. Lawrence, N. Marwan, and S. Moebus: "Complex networks for analyzing the urban acoustic environment", *Ecological Informatics*, 78, 102326, 2023.

[5] A. Farina and S. H. Gage: *Ecoacoustics: The ecological role of sounds*. Oxford: John Wiley & Sons, 2017.

[6] J. L. Miksis-Olds, and S. M. Nichols: "Is low frequency ocean sound increasing globally?", *The Journal of the Acoustical Society of America*, 139(1), 501-511, 2016.

[7] Y. Wang, L. Zhang, J. Xie, J. Zhang, and R. Zhu: „DFEFM: Fusing frequency correlation and mel features for robust edge bird audio detection", *Avian Research*, 100232, 2025.

[8] M. D. Jedrusiak, T. Harweg, T. Haselhoff, B. T. Lawrence, S. Moebus, and F. Weichert: "Towards an interdisciplinary formalization of soundscapes", *The Journal of the Acoustical Society of America*, 155(4), 2549-2560, 2024.

[9] T. Haselhoff, M. Schuck, B. T. Lawrence, A. Fiebig, and S. Moebus: "Characterizing acoustic dimensions of health-related urban greenspace", *Ecological Indicators*, 166, 112547, 2024.

[10] T. Haselhoff, K. Schröer, J. Schröder, and S. Moebus, „Associations between acoustic indices and the evaluation of the urban soundscape according to ISO 12913", in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Nantes, France)*, pp. 8209-8214, 2024.

[11] T. Haselhoff, T. Braun, J. Hornberg, B. T. Lawrence, S. Ahmed, D. Gruehn, and S. Moebus: "Analysing interlinked frequency dynamics of the urban acoustic environment", *International Journal of Environmental Research and Public Health*, 19(22), 15014, 2022.

[12] A. L. Barabási: *Network science*. Cambridge: Cambridge University Press, 2016.