DOI: 10.61782/fa.2023.0302

THE URBAN ACOUSTIC ENVIRONMENT AS A COMPLEX NETWORK

Timo Haselhoff[1*] **Tobias Braun**[2] **André Fiebig**[3]
Bryce T. Lawrence[4] **Susanne Moebus**[1]
[1] Institute for Urban Public Health, University Hospital Essen, Germany
[2] Complexity Science, Potsdam Institute for Climate Impact Research, Germany
[3] Engineering Acoustics, Technical University of Berlin, Germany
[4] Department of Spatial Planning, TU Dortmund University, Germany

ABSTRACT

The urban acoustic environment (AE) contains valuable information on complex sub-systems of urban areas, such as traffic or biodiversity. As the availability of cheap sensors and computational power increases, so does the need for methods to process high-dimensional audio data. We take this as an opportunity to introduce complex networks (CNs) to the field of urban acoustics. CNs have proved effective in capturing the complexity of e.g. climate dynamics or brain structures, and thus, represent a promising tool for the high-dimensional urban AE. We present how CNs are constructed based on frequency correlation matrices and show their behavior for various time periods and sound sources. To demonstrate their use on audio data from the urban environment, we apply them to the dataset from the SALVE study to systematically characterize the urban AE. Here, we use subsets of SALVE with two different temporal resolutions: 1 s over three minutes and three minutes over 24 h. Measures such as the average shortest path length identify urban sites with similar AEs, indicating the utility of CNs to identify non-random patterns in large datasets of the urban AE.

Keywords: *soundscape, complex networks, passive acoustic monitoring, urban environment*

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

1. INTRODUCTION

The Acoustic environment (AE) matters for a plethora of fields. For instance, it can be utilized to measure biodiversity in ecoacoustics [1] and the pleasantness of public spaces in urban planning [2]. Accordingly, due to the decreased cost of audio recordings in recent years, the temporal and spatially high-resolution sampling of the AE becomes feasible. This results in the availability of spectral information (e.g. sound sources) over time and at various locations. Thus, passive acoustic monitoring (PAM) represents an emerging source of valuable information for many research fields. However, the reduction in costs can easily result in several terabytes of data, but evaluated analysis methods for large-scale audio data are scarce – especially in the urban environment. In more natural environments, ecoacoustic indices [3] are deployed as a method to quantify the AE since approx. one decade. Still, the exact interpretations from these indices are under an ongoing debate, even for the more natural areas they were developed for [4]. Consequently, their use for more urban environments is not yet resolved either [5]. We take this as an opportunity to introduce complex networks (CNs) as a new framework for the characterization of the urban AE. In recent decades, CNs have emerged as a powerful tool for accurately representing real-world system (e.g. climate dynamics and brain activities [6]). One of the main advantages of CNs is their ability to capture the emergent properties and behaviors of the system as a whole, rather than focusing solely on individual components [6]. Here, we apply CNs to quantify frequency correlation matrices (FCMs), which have shown to be a promising tool to depict environment specific interrelationships between power spectra in high spatial and temporal resolution [7,8]. The aim of this work is to give an intuitive understanding on how FCM based CNs in the urban AE work. For this, we briefly introduce the methodology and give examples for CN behavior on two different time-scales: (i) for three minutes and (ii) for 24 hours.

2. METHODS

The basis of our CN analysis are FCMs. We use FCMs to capture linear relations between frequency bins. Here, we calculate the correlation between all frequency bins over time (Fig. 1). In the urban acoustic environment, high correlations between frequency bins characterize the prevalence of particular sound sources and therefore provide valuable information to distinguish between the AEs of different urban settings [8]. We calculate the FCMs using Pearson correlation [7]:

$$r(f_1, f_2) = \frac{\sum_{i=1}^{n}(x(f_1)_i - \overline{x(f_1)})(x(f_2)_i - \overline{x(f_2)})}{\sqrt{\sum_{i=1}^{n}(x(f_1)_i - \overline{x(f_1)})^2 \sum_{i=1}^{n}(x(f_2)_i - \overline{x(f_2)})^2}} \quad (1)$$

Where $x(f_1)$ is the set of amplitudes at frequency f_1. The calculated correlation coefficients form a symmetric matrix, where the i-by-j-th element represents the correlation coefficient between power levels at f_i and f_j.

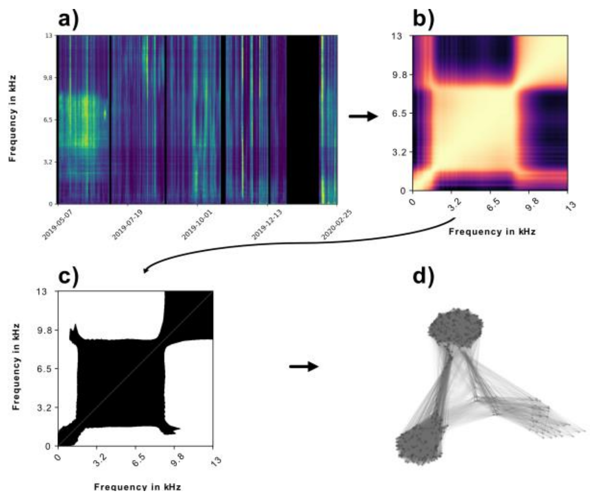

Figure 1. Procedure to build a simple graph from a spectrogram. First, a spectrogram is created (a) from which the FCM is derived (b), which is then thresholded to achieve the adjacency matrix (c) and the corresponding network representation (d).

Subsequently, the coefficient of determination (R2) is used to measure the proportion of explained variance between two frequency bins. R2 quantifies the strength of the relationship between frequency bins and ranges from 0 to 1. As a next step, it is necessary to establish a threshold to dichotomize the FCM, as adjacency matrices for simple graphs have to be binary. This involves assigning a value of one to all R2 values above the threshold, indicating connected frequency bins, and zero to those below. Figure 1 depicts the whole process.

There is a plethora of measurements available from CN theory, once the network is created. In this work, we focus on the average shortest path length (ASPL) of the network, as preliminary analyses using ASPL indicated promising results. ASPL represents the average number of "steps" required to go from any given node to every other node in the network, taking into account the total number of nodes present. Thus, its value is dependent on the number of connections and the topology of those connections (i.e. a measure of connectivity). In our context, low ASPL indicates a greater number of high correlations between multiple frequency bins, resulting in fewer distinct communities. When there is a single dominant sound source, such as traffic, there are many high-frequency correlations across the spectrum, resulting in a lower ASPL. Conversely, when there are multiple sound sources that form distinct correlation communities, the ASPL will be higher. Therefore, the ASPL of the urban AE can be interpreted as an index of "acoustic dominance". Finally, it should be noted that the network should be fully connected to calculate the ASPL [6], which is why different R2 thresholds are used for the following analysis. For the sake of conciseness, the identification of optimal thresholds will be discussed in future works.

3. DATA

The audio data used in this work originates from the SALVE study [9]. Briefly, 50 3-min audio files have been recorded daily at more than 82 locations in Bochum since 2019. Recordings were made using Wildlife Acoustics SM4 recorders with a SMMA2 microphone [10]. The devices were programmed to record 3-minute recordings every 26 minutes at a sampling frequency of 44.1 kHz and 16 bit depth. Here, we use subsets of SALVE for the two different temporal resolutions:

i. One second windows of two 3-minute recordings from the urban forest and a main street,
ii. 3-minute windows of 50 daily recordings from 23 urban sites.

For (i) we derived the spectral content using Welchs method with a “tukey” window of a block size of 2048 and an overlap of 25%. FCMs, CNs and their respective ASPL were calculated for each second of the respective recording. For (ii) a power spectrum for each recording was calculated using Fast Fourier Transform and then averaged for 1024 bins with equal bandwidth. FCMs, CNs and their respective ASPL were then calculated between all frequency bins for each respective day. The frequency range for all data was limited to 13 kHz, as we found no substantial power above 13 kHz in the urban environment [8].

4. RESULTS

In the following, the behavior of complex networks of the urban AE is illustrated, using two different temporal resolutions.

4.1 Complex Networks on 1 s in 3-minute recordings

Figure 2, shows that the ASPL can capture the occurrence of distinct sound sources quite well. First off, the average ASPL of background noise is around 1.8, which drops remarkably during the time frames were distinct sound sources are present (e.g. bird vocalizations or single cars passing by). However, as soon as multiple sound sources begin to overlay, the ASPL drops become substantially less pronounced (Fig 2b: 120 s - 180 s). This suggests that ASPL is highly sensitive to the presence of sound sources on very short time scales and could be a useful tool to differentiate between polyphonic and monophonic sounds in the urban AE.

4.2 Complex Networks on one day of recordings

Figure 3 depicts the FCM and the corresponding CN of two days of recordings at the main street (a) and the urban forest (b). The ASPL for the main street is 1.01 and the ASPL for urban forest is 2.32. It shows that CNs between 3-minute recordings exhibit distinct patterns from those on a 1 s resolution. From the latter (Fig. 2), overlapping sound sources show higher ASPL, thus, a busy main street could be expected to have a higher ASPL than a rural urban forest. However, this assumption does not transfer to the ASPL behavior on a 3-minute resolution. This is most likely the case, because power spectra over three minutes are more similar to each other than the spectral content of single seconds [8].
However, another interesting pattern can be derived from the ASPL for a 3-minute resolution. As virtually all frequency bins for the main street correlate with each other, the AE can be assumed to be “dominated” by only a few factors. In contrast, multiple rectangular communities form for the urban forest, which indicates several factors causing the communities. This is further validated by Fig 3c, which represents a mixture of (a) and (b), where the main street spectrogram from 05:00 to 08:59 was replaced with the bird chorus from the urban forest during that time. Here it shows that the integration of bird vocalizations favors the emergence of rectangular structures in the FCM, which will increase ASPL: (a) is 1.01 and (c) is 1.72.

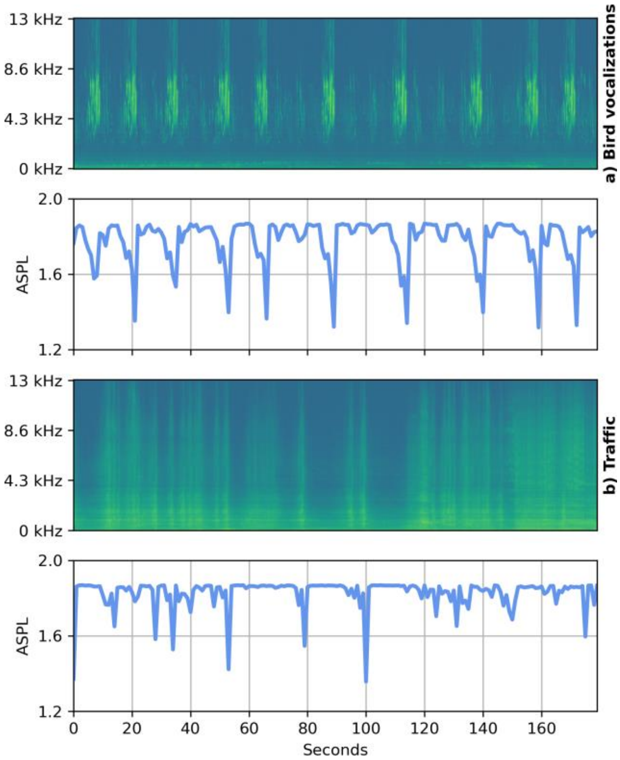

Figure 2. Two 3-minute recordings of bird vocalizations (a) and passing cars (b). For each second of the signal a FCM, the CN (threshold: R2=0.1) and the respective ASPL were calculated, which is plotted over time. It can be seen that distinct sound events are recognized, but that polyphonic signals (e.g. (b) from 120 s to 180 s) result in a lower ASPL than monophonic signals.

5. DISCUSSION & COCNLUSION

The decrease in cost for PAM and the resulting increase in the availability of large-scale audio data from the urban environment calls for an elaboration of methods tailored towards its special features. Here, we build CNs based on FCMs to capture frequency dynamics of the urban AE. We use the ASPL to quantify the topology of those networks and interpret it as a measure of "acoustic dominance".

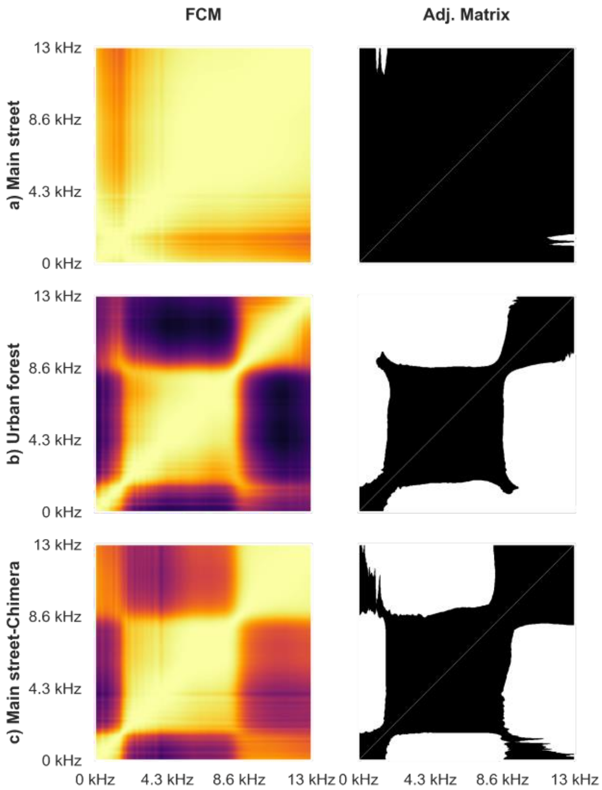

Figure 3. Depicted are FCMs (using R2) and the respective Adjacency Matrix (threshold: R2=0.7) for the respective land use over 24 h for one day. The "Main Street-Chimera" is the FCM of a spectrogram-mix from a) and b), where the data between 00:00 to 04:59 & 09:00 to 23:59 originates from "Main street" and data from 05:00 to 08:59 originates from "Urban forest".

We show application examples on two time-scales, which indicate that (i) the ASPL identifies sound signals on a 1 s time resolution and could be of use to differentiate between polyphonic and monophonic signals; and (ii) that the ASPL on a 3-minute time resolution shows promising results to differentiate between urban environments with higher/lower "acoustic dominance". Altogether, CNs are a new framework that could be used to characterize sound source constellations in the urban environment to e.g. link them to human perception. To do so, future works should validate the behavior identified in this work for larger datasets and the mechanisms, which influence the emergence of frequency communities need to be further understood.

6. ACKNOWLEDGMENTS

This work was supported by the HEAD-Genuit-Foundation [P-21/02-W].

7. REFERENCES

[1] B.C. Pijanowski, A. Farina, S.H. Gage, S.L. Dumyhan and B.L. Krause: "What is soundscape ecology? An introduction and overview of an emerging new science", *Landsc. Ecol.*, 2011;26:1213-32.

[2] J. Kang and B. Schulte-Fortkamp: *Soundscape and the built environment*. City: Boca Raton, FL, USA, CRC press, 2016.

[3] J. Sueur: *Sound analysis and synthesis with R*. City: Culemborg, Netherlands, Springer, 2018.

[4] I. Alcocer, H. Lima, L. Sugai and D. Llusia: "Acoustic indices as proxies for biodiversity: a meta-analysis". *Biol. Rev.*, 2022;97:2209-36.

[5] A.J. Fairbrass, P. Rennert, C. Williams, H. Titheridge and K.E. Jones: "Biases of acoustic indices measuring biodiversity in urban areas". *Ecol. Indic.*, 83, 169-177, 2017.

[6] A-L. Barabási: *Network science*. City: Cambridge, Cambridge University Press, 2016.

[7] S.M. Nichols and D.L. Bradley. "Use of noise correlation matrices to interpret ocean ambient noise". *JASA*. 2019;145:2337-49.

[8] 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". *IJERPH*, 19(22), 15014, 2022.

[9] T. Haselhoff, B.T. Lawrence, J. Hornberg, S. Ahmed, R. Sutcliffe, D. Gruehn and S. Moebus. "The acoustic quality and health in urban environments (SALVE) project: Study design, rationale and methodology". *Appl. Acoust.*, 188, 108538, 2022.

[10] Wildlife Acoustics. "Song Meter SM4 Acoustic Recorder". https://www.wildlifeacoustics.com/products/song-meter-sm4, (accessed on 15.04.2023).