DESIGN, IMPLEMENTATION AND EVALUATION OF AN AIR QUALITY SONIFICATION SYSTEM

Carla Terboven

Hasso Plattner Institute
Prof.-Dr.-Helmert-Str. 2-3
14482 Potsdam, Germany
`carla.terboven@student.hpi.de`

Nico Steckhan

Digital Health Cluster, Hasso Plattner Institute
Prof.-Dr.-Helmert-Str. 2-3
14482 Potsdam, Germany
`nico.steckhan@hpi.de`

Eleni-Ira Panourgia

Film University Babelsberg KONRAD WOLF
Marlene-Dietrich-Allee 11
14482 Potsdam, Germany
`e.panourgia@filmuniversitaet.de`

Bert Arnrich

Digital Health Cluster, Hasso Plattner Institute
Prof.-Dr.-Helmert-Str. 2-3
14482 Potsdam, Germany
`bert.arnrich@hpi.de`

ABSTRACT

Air pollution is the world's deadliest environmental risk factor. Yet there is little effort to educate the public about personal exposure to pollutants such as particulate matter (PM). This paper presents the design and implementation of a portable sensor box (PSB) to collect local, spatially highly resolved particulate matter data. To counteract common inaccuracies in mobile particulate matter measurements, data were aggregated and cleaned according to their location in a real-world investigation. We employed parameter mapping to develop a real-time, interactive and intuitive, yet scientifically accurate sonification of the data. The sonification was made accessible to listeners within a test area in Potsdam using a physical prototype. Through the implementation of a scripted exposure design, we investigated realistic, comparable statements and evaluation by the participants about the sonification device and the sonification.

1. INTRODUCTION

No environmental risk factor causes more premature deaths than air pollution [1]. Yet, 99% of the world's population lives in areas where air pollution is considered too high by the World Health Organization (WHO) [2]. To date, policies and regulations have failed to protect the public from excessively polluted air. At the same time, few people are aware of the air pollution levels in their neighborhoods and the corresponding health risks. Moreover, the public often dismisses air pollution levels, assuming that the problem is more severe in other parts of the earth [3]. Therefore, the question arises how air quality can be communicated and thus a sustainable action of the public be achieved. This communication is typically done with visual and textual interpretations of the few official measuring stations. However, the public is already saturated with this information media, and relatively new approaches like sonification likely attract more attention [4]. Listening to sonified data facilitates the communication and interpretation of the information contained in the data.

We envision that an auditory display could be used to communicate collected on-site sensor data to listeners in real time. Within this scenario, we ask to what extent the sonified data can be understood and classified by the listeners. Related to the live sonification of the data, we then examine whether expectations of the data can falsely influence what is heard. In the context of human-computer interaction, it has already been found that the same sounds are perceived differently when presented through different sound sources [5]. Possibly, the context in which the sound of particulate matter data is presented such as near a street or in a park, may also influence the perception of sonification.

The content of the paper is organized as follows. Section 2 introduces the essential terminology and related projects. Afterward the content of the paper is organized as illustrated in Figure 1. Section 3 describes the used methods, i.e., how we obtain robust data for real-time sonification, the sonification design, and the chosen study design. The resulting artifacts are then presented in section 4. Finally, the results are discussed and concluded in section 5.

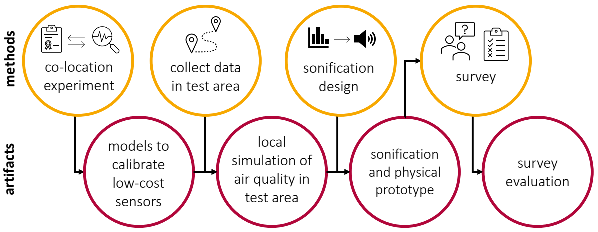


Figure 1: Methodological alignment of project

2. BACKGROUND AND RELATED WORK

Air pollution can be measured with different pollutants, e.g. with Particulate Matter (PM). PM is a mixture of the smallest particles in the air. The particles can be solid or liquid and consist of different chemical components. Distinctions are made in the size of the particles and the most common measurements are PM2.5 and PM10. They describe the mass concentration (in µg/m³) of particles smaller than 2.5 µm or 10 µm. We focus on PM as long-term exposure to ambient particulate matter (PM2.5), with 4.14 million deaths caused in 2019, has a particularly large impact on human health [1].

There are a number of projects that communicate air quality data with sound. St Pierre and Droumeva [6] scale and map pollutants (CO, O3, SO2, and NO2) to individual frequencies.The frequencies are complemented by a clicking Geiger counter sound based on PM2.5 data. They found that listeners understand which pollutant changes at any time once the mapping is known. Arango introduces three projects called *AirQ Jacket*, *Esmog Data*, and *BREATHE!* [7]. Esmog Data is an art installation using audio and motion graphics. Each collected sensor data point (CO, CO2, SO2, and PM10) is connected to multiple parameters of a synthesizer. The artists of BREATHE! consider breathing as a communication medium that can be universally understood. Listeners can identify levels of toxic air solely by listening to a breathing sound based on measured toxic gases in the air from different speakers in space. The auditory display by Skov [8] focuses on dimensions like temperature, light, humidity, and noise. It also uses voice to classify the used urban environmental data of different cities in categories like “high” or “medium”.

However, these projects rely on historically collected data for their sonification design. We experiment with real-time sonification in local environments and hypothesize that the connection to the data is more substantial when using on-site data. Real-time sonification can allow listeners to explore their surroundings in an interactive manner, while improving the usability of the auditory display and the learning effect of the sonified data [9, 10, 11].

Other projects provide live feedback with the sonification of real-time air quality data. The AirQ Jacket [7] consists of an eye-catching white jacket with multiple built-in sensors, LEDs, and small, lightweight speakers. The person wearing the jacket can visually and acoustically experience the gathered air quality data and temperature. The project AirCase [12] uses small earbuds and the corresponding charging case. The batteries in the charging case are not only used to charge the earbuds but also to supply environmental sensors with power. Thus, the small charging case can collect mobile environmental data (e.g., CO2 and temperature). This data is directly sonified and can be perceived by the user with the earbuds that initially belong to the case. Another project, called Sonic Bike [13], uses an PM sensor to gather live data on a bike. The data is then processed at the back of the bike so the bike rider can finally experience the sonified air pollution via two speakers attached to the bicycle handlebar and a sub-woofer behind the bike seat.

All of the aforementioned projects use directly collected live data from their own sensors but do not specify how these sensors are calibrated and maintained. There is also no indication of the complexity of mobile air measurements or how associated outliers, caused, e.g., by vibrations, are handled. To use sonification as a scientific method, the underlying data must also be collected scientifically. Therefore, compared to [7, 12, 13], we focus more intensively on obtaining robust, local PM data.

3. METHODS

3.1. Co-location of reference sensors with portable sensor box

Co-locations are used to calibrate optical PM sensors for specific meteorological conditions [14, 15, 16, 17]. The sensors to be calibrated are placed in the immediate vicinity of official, calibrated reference systems for several days. The data of the reference system are then compared with the data of the sensor to be calibrated. Including meteorological data makes it possible to develop a model that can predict the data of the reference system reasonably accurately. This model can subsequently be used in comparable meteorological conditions to calibrate sensor data. Calibration is essential for the reliable use of our Portable Sensor Box (PSB, Figure 2). The sensors used provide raw data only in terms of particle concentrations, however, PM is usually reported using mass concentrations. For this reason, we conducted a co-location experiment in Potsdam at the measuring station DEBB054 Potsdam-Zeppelinstraße. The co-location took place from November 21 to December 5, 2022.

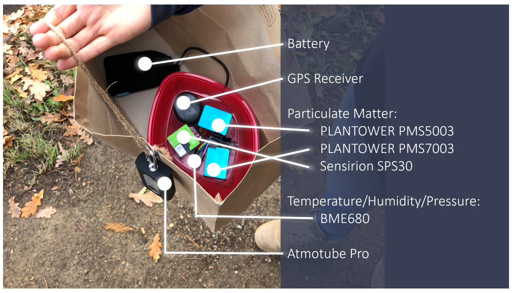


Figure 2: Components of the portable sensor box. Four different particulate matter sensors, a GPS receiver and temperature/humidity/pressure sensor were combined and attached to a Raspberry Pi.

3.2. Data collection

Official air quality values are only available from permanently installed, widely separated measuring stations. Only limited statements can be made for the areas between the measuring stations. Using mobile measuring devices is one possibility to still gather air quality data extensively over wide areas. However, these are controversial in accuracy because sensor performance is assumed to decrease with speed and vibration [15, 18].

Since our objective is to make statements about the communication of particulate matter values in local areas, we need a comprehensive data basis. We aim to obtain this with mobile measurements. However, we generate robust values with repeated measurements in always the same area in order to counteract measurement inaccuracies due to movement and vibration. We assume that the average particulate matter load can be determined site-specifically for enough repeatedly collected data for a defined test area. We call this final determination of particulate matter values per location hereinafter *simulation of the test area*.

All data used for the simulation is collected using the PSB presented in section 3.1. Each data point is stored together with a GPS position and a time stamp. As particulate matter values are subject to diurnal variations, we collect data at the recurring time on weekdays during the onset of commuters’ evening traffic between 4 and 6 pm. Therefore, our simulation can be used only in this period, e.g., for studies.

The test area is an approximately 1.9 km long track in Potsdam, next to the air monitoring station Zeppelinstraße. The route leads through different types of development and landscaping along roads with varying traffic levels. A total of 18 data sets were collected from October 10 to December 15, 2022. We recorded data approximately every 2.5 m. In total, 12,253 data records are available to simulate the track.

3.3. Sonification design

The basic idea of the auditory display is to hear the air quality of the area one is located in. For example, if one stands in an area with high air pollution, an alerting, insistent sonification should be heard. The sound should change when one moves to an area with less air pollution. The sonification design is interactive in the sense that spatial movements of the listener can change the sonified data and, therefore, the sound. The listener thus has a certain control loop and can learn the different levels of PM2.5 and PM10 at different locations by using the sonification. Interactive sonification is used to enhance the learning of the sonified data, it is suitable for data exploration, and reduces fatigue for listeners interacting with the sonification in a control loop [11].

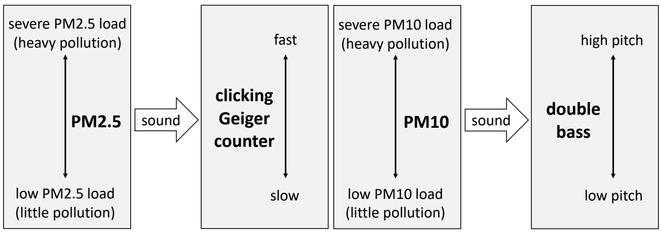
In terms of content, the PM values represent the mass concentration of the smallest particles in the air. If higher PM values are assigned to an area, this means an increased health risk in this area [1, 2]. From a conceptual perspective, the sonification of the data, therefore, should communicate the danger and urgency of high PM levels.

We decided to create individual mapping functions for PM2.5 and PM10. Figure 3 provides a summary of the two mappings. The two PM categories can be distinguished using two different timbres. PM2.5 is represented by the sound of a Geiger counter, and PM10 by the sound of a stringed double bass. The data manipulate different sound parameters. PM2.5 influences the duration of the clicking sound, and PM10 the frequency of the bass sound. We decided to use this method of sonification because we thought that it would be easier for untrained listeners to distinguish between the two parameters. An alternative method would certainly be multivariate sonification on one auditory stream only.

The PM value is queried every two seconds, and in the next two seconds, the queried value is sonified. Increasing PM2.5 values are represented by faster clicking of a Geiger counter sound (see Figure 3). The PM2.5 data are rounded to the nearest integer, and the Geiger counter sound clicks once for every 2 µg/m³ of PM2.5 mass concentration.

The listeners can experience the PM10 data through the changing pitch of a double bass (see Figure 3). If the PM10 values increase, the pitch becomes higher. We manipulate the pitch by playing a recorded sound sample of a double bass at different speeds. The faster the sample is played, the higher all frequencies and the finally perceived pitch. We use a linear mapping and have subjectively set the pitch boundaries so that the designer perceives them as aesthetically pleasing, and the double bass sound remains recognizable. At a PM10 value of 0 µg/m³, we play the sound sample in 700 ms (it is repeated until after 2 seconds another sound is played). At a PM10 value of 36 µg/m³, the sound sample is played in 2100 ms (and cut off after 2 seconds). The length of the sound sample in ms is therefore given by 2100-38.89*PM10.

The sonification thrives when it is explored independently in our test area. However, an exemplary recording of the sonification while a listener is walking through the test area can be heard on SoundCloud [19].



Sonification of PM2.5 Sonification of PM10

Figure 3: Sonification of particulate matter as dual representation of PM fractions below 2.5 and 10 micrometer mapped to clicking Geiger counter sound and pitched double bass accordingly. Both sonifications are played at the same time (mono).

3.4. Survey

The evaluation study was approved by the ethics board of University Potsdam on 9th of January 2023 under no. 79/2022.

A *scripted exposure study* is a beneficial study design to monitor multiple levels of real-world exposure [20]. The study participants walk a predefined route with route segments of varying degrees of particulate matter pollution. The participants are equipped with GPS receivers. This allows to obtain spatially comparable measurements of the desired effects before, during, and after exposure. In order to expose all participants to comparable levels of particulate matter, all participants of our study walk the same route at a fixed time (weekdays between 4 and 6 pm). An auditory display informs the participant in real-time of the PM exposure he or she is exposed to at the various locations along the test route. In order to make this exposure to the sonified PM data comparable among all participants, we use a location-based, uniform simulation. Overall, the study can thus be described as a pilot study using a simulated scripted exposure design.

To conduct the study, participants walk a distance of about 1.9 km with the sonification device on their ears. The auditory display of the sonification device is replayed location-based. The route followed is the simulated test track introduced in section 3.2. For road safety and to ensure uniform tempo and that the correct route is chosen, an instructor accompanies all participants on the test track. The instructor asks the participants to answer questionnaires at various points along the route. These questionnaires ask for the assessment and opinion on the air quality of the different route sections based on the expectations and visual impressions of the subjects on the one hand and the heard auditory display on the other. In particular, we are interested whether expectations about air quality (e.g., based on visual impressions) create biased interpretations of the auditory display. For example, whether what is heard is interpreted differently in parks than at major street intersections. In the context of human-computer interaction, it has already been found that the same sounds are perceived differently when presented from different sound sources [5]. Possibly, the context in which the sound of particulate matter data is presented (e.g., on a street or in a park) may also influence the perception of the sonification.

4. RESULTS

4.1. Models to calibrate low-cost sensors

For calibrating particulate matter sensors, Multiple linear regression (MLR) models are most suitable [14]. According to Ockham's razor, we choose the simplest of the models that explain the target variable as adequately as possible.

Following the calculation of different models, we combined two PM sensors and the meteorological sensor and to use the LinearRegression function in the Python package *scikit learn* [21] to train the model. With a coefficient of determination (R^2) of 0.97, a Mean Squared Error (MSE) of 0.92, and a BIC of 27.81, our trained model adequately explains PM2.5 mass concentrations. In accordance with related literature calibration models for PM10 perform less well [14]. We decided to use an MLR model with an R^2 of 0.93, an MSE of 2.26, and a BIC of 89.2.

4.2. Local simulation of air quality in test area

In order to use the data for the simulation, data preprocessing and the determination of the PM values with the calibration model (see section 3.1) is necessary. The detailed Python code can be found on GitHub[1].

A challenge is posed by enormous peaks in particle concentration such as those caused by passing smokers. These peaks are not representative of the location's primary PM concentration and are therefore removed. We remove all data records with a z-score greater than 3 for one of the parameters used in the calibration model (temperature, relative humidity, pressure, and number concentrations from PM sensors PLANTOWER PMS7003 and Sensirion SPS30). The check for a normal distribution can be neglected since the dataset is large enough [22]. After the outlier treatment, 12,005 from 12,253 data records remain. For these data records, we calculate the PM2.5 and PM10 values according to our MLR calibration model. All these calibrated data points are shown with blue dots in Figure 4 on the right side. It can be seen that data records cannot be collected for every millimeter of the test track. For positions without measured values, the most probable value must be derived from the surrounding measured values. The grid density depends on the measurement accuracy of our GPS receiver. We used a grid with field size $0.00015\,^\circ \times 0.00015\,^\circ$ respectively $16.67\,\text{m} \times 10.17\,\text{m}$. The values of the grid fields can be determined using different interpolation methods. We have chosen the median as a simple variant that requires little computing power. Eventually, for each field, we can calculate the median of all data points to obtain the region's simulated particulate matter value. Figure 4 shows how the data points are assigned to the fields. On average, each field contains 30 data points.

4.3. Sonification and prototype

An interactive design requires feedback as close to real-time as possible. The movement of the listener to another location should therefore be accompanied by a change of the sound in real-time. In our sonification design, we have deliberately chosen short sections of two seconds, after which the location data is retrieved again and updated if necessary. However, the grid size of approximately 10 × 10 meters defined in section 4.2 also plays a certain part. The GPS accuracy prevents a reduction of this grid size, and thus also the feedback loop. Only when the listener enters a new grid field, the sonification can be updated. If the listener moves within the field, no interaction with the sonification can be experienced.

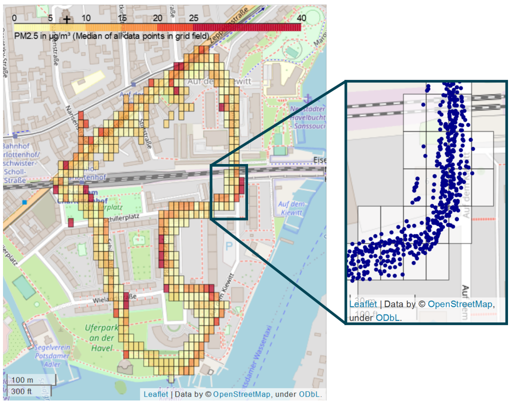


Figure 4: The test track had a distance of 1.9km and an average of 30 samples per grid.

To make the sonification design accessible to listeners, we decided to use headphones to transmit the sound in a targeted and individual way only to the people who interact with the sonification. The headphones are part of our prototype, which is hereinafter called *sonification device*. In addition to the technical sound display, several other components of an auditory display are included in the device: it enables user and data interaction, stores the PM data and further application context, and finally renders the sonification. The sonification device was explicitly assembled for our chosen study design and therefore provides a comparable position-based sonification in a specified test area in Potsdam.

The assembled sonification device is illustrated in Figure 5. All individual parts were mounted directly on the headphones to allow easy handling during the study. We decided to operate the Raspberry Pi headless for the quick and inclusive use of the device.

After starting the program, the GPS position of the sonification device is queried every two seconds. Subsequently, it is calculated in which grid field of the simulated area the position is located, and the historically collected PM2.5 and PM10 data are obtained for this grid field. This sonification can only be used in the prepared, surveyed test area. The feedback loop is two seconds long as the data are read every two seconds, and then the data is sonified during the next two seconds.

The translation from data to sound can be divided into three sections: (1) data processing with Python, (2) data transfer via Open Sound Control (OSC), and (3) sound generation and playback in Pure Data (Pd). The related code is publicly available in the GitHub repository of the project.

The sonification device requires two types of data processing. One is processing real-time GPS data to simulated PM values and, furthermore, converting PM values to sound information. The mapping of data to sound could also be implemented in Pd but we deliberately decided to use our experience with Python and choose the easy-to-read programming language for all calculations and decisions over the visual, quickly cluttered programming environment Pd. We propose a software architecture that does all the computing steps in Python and only switches to Pd for the final stage of sound generation.

---

[1] `https://github.com/carlaterboven/simulated_scripted_exposure_study` (visited on 2023-01-15).

Figure 5: Components of the sonification device

The duration of the sounds and possible pauses between the sounds are calculated in Python based on the PM data. We use Python's multiprocessing package to simultaneously play multiple sound layers with different audio samples. Thus, for each data value, length and pauses are calculated simultaneously for the respective audio files. After the calculation, the information for each of the two sounds is transmitted individually, via OSC to the Pd application. Finally, the processes are rejoined, and the messages sent are further processed on the Pd side. Pd was chosen because of low entry barriers and because it is also available for the Raspberry Pi in version Pd Vanilla. Since all the logic has already been calculated in Python, the only remaining task in Pd is interpreting the calculated information. All OSC messages are split into their individual parts within the Pd patch, and the received information is used to play prerecorded samples at a certain speed. The samples used are freely available sound recordings of a Geiger counter clicking and a double bass sound. For the Geiger counter, the sample file is always played back at the same speed, and the pause between the clicks is regulated with the Pd metronome. For the double bass sound, different playback speeds of the sample produce different pitch variations. Individual incoming OSC messages allow specific, even simultaneous, playback of the sounds. Following this experimentation, further sounds could extend the setup.

4.4. Survey evaluation

The study involved 21 persons (16 male and 5 female). The participants were between 18 and 39 years old, with an average age of about 25 years. All studies were conducted between 4 and 6 pm on weekdays during the period from January 10 to 27, 2023. The question about listening skills was answered by most participants with very good or good. Only two persons rated their hearing ability as moderate but stated that they did not have a medically diagnosed hearing disorder.

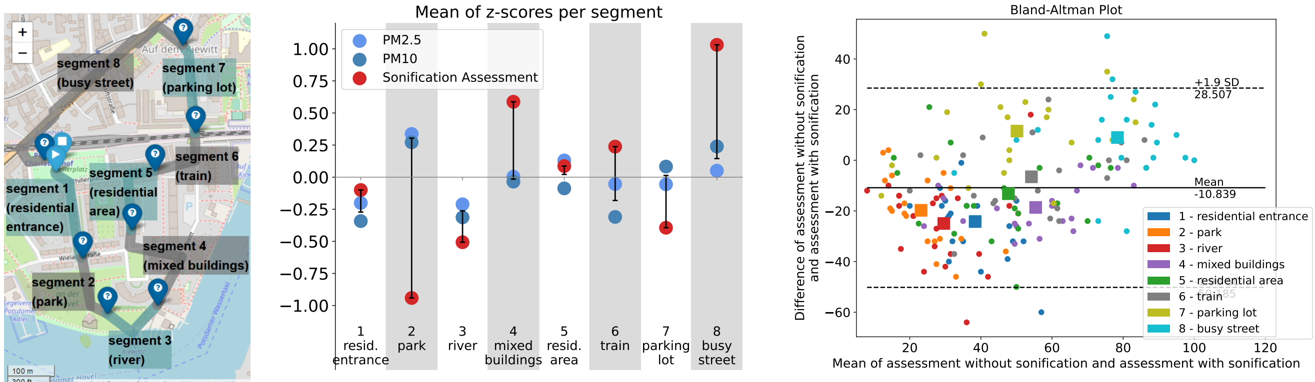
We can qualitatively evaluate the sonification design by using study participants who answered open-ended questions in the test area after using the device for about one hour. Interestingly, there is no clearly favored mapping (Geiger counter duration or double bass frequencies) among the study participants. A correlation between prior musical knowledge and preferred mapping could not be found either. Many of the participants stated that they paid equal attention to both mappings. Study participants who liked the use of rhythm as an experience of PM data better respond to the question *What do you consider when interpreting or evaluating the sonification?* with sentences like “Mainly the interval and the volume of the ticking. The pitch of the double bass was a bit more difficult to interpret.”, “Especially the Geiger counter, the pitch is difficult to estimate over a long period of time.”, or “Above all, spikes in intensity, but also the frequency of the Geiger counter.”.When asked about situations in which interpretation was difficult, rhythm-favoring participants stated, for example, that “differences of the notes of the double bass were more difficult to hear than the differences of the ticking”. In addition, the continuous tone of the double bass was criticized with sentences like “Due to the fact that I heard the tone permanently, it was very monotonous, and I had to concentrate on it quite a bit to still be aware of it”. On the other hand, there were also proponents of mapping PM10 to frequency. When interpreting the sonification, they state that “In the end, [I relied] more on the bass than the clicking. [And rather focused on] the lasting impression than individual peaks”. Another person stated that the focus was “rather on the deeper sounds. In fact, over time, I got used to the clicking. Only when there was a change I became attentive”. However, this could also be interpreted in favor of a selective-attentive listening behavior of the Geiger counter mapping.

Overall, the design seems to trigger the desired ideas in the participants. Participants stated that they would pay attention to “whether the sound stresses me or not” or “on the threatening nature of the sounds”. The reflection and the personal response to the sonified data can be found in answers such as “If my personal impression differs from what I hear, I think about the discrepancy”. Or the statement that judgments are made based on “[rhythm] and pitch in relation to previous experience”. Difficulties in interpretation were mainly described for sections in which there were many changes of high and low PM data, and “it was difficult to form an average”.

The selected test track consists of several segments. Each segment is numbered and given an individual descriptive name for better referencing (see Figure 6a). For each route segment, PM exposure can be categorized in three different ways. Based on the (1) measured and calibrated objective data of the PSB, based on (2) the subjective perception of these sonified values by study participants, and based on the (3) subjective personal assessment of air pollution by study participants (e.g., by visual characteristics of the route segments). In the following, the three measurement variants are presented consecutively and then compared with each other.

4.4.1. *Objective data of the PSB*

The PM values are grouped according to their location. We found slight variations in the medians for the different route segments. Route segment 4 has the lowest PM2.5 median (5,704 µg/m³), while route segment 2 (park) has the highest PM2.5 median (7,815 µg/m³). For the mass concentration of PM10, the median is lowest in road section 6 (11,303 µg/m³) and highest in road section 8 at the busy road (13,21 µg/m³).



(a) Real world test track

(b) Comparison of z-scores from objectively measured PM values and perceived sonification

(c) Bland Altman Plot showing differences within user ratings with and without sonification

Figure 6: Comparison of objectively measured PM levels, perceived sonification and visually assessed PM exposure per route segment.

4.4.2. *Subjective perception of the sonified PM data*

We queried the sonification assessment with the question “How would you rate PM loads for the last route segment regarding the sonification?” and had the participants estimate the PM mass concentration using a visual analog scale from “no PM load” to “heavy PM load”. For the evaluation, the visual scales were assigned to a value range from 0 to 100. The Friedman test confirms significant differences (p<0.001) in the sonification assessment for the different route segments. The post hoc test (Nemenyi test) shows that each section is significantly different from at least one other section (p<0.01). Route segments 2 and 8 have the most significant pairwise differences from the other route segments. Moreover, it can be observed that the interpretation of the sonification does not necessarily reflect the underlying PM data. For example, section 2 (park) is rated as the cleanest with a median rating of 30. However, at the same time, the underlying PM readings for the section are actually quite severe (highest PM2.5 median and 4th highest PM10 median). With ideal conditions and a perfect understanding of the sonified data, the assessment of air quality based on sonification should reflect the PM data.

4.4.3. *Subjective PM assessment without sonification*

The participants were also asked to assess the PM load without sonification. They indicated that this estimation was given mainly based on visual characteristics and olfactory stimuli. For example, the number of stationary and moving cars, the presence of green areas or water, and unpleasant scents such as exhaust fumes. Individual participants also referred to the feeling of how freely one can breathe, the paving of the paths (earth or stones), and noise levels. The participants were asked again to give their assessment on a visual analog scale ranging from “no PM load” to “very high PM load”, with a mapping to values from 0 to 100 in the evaluation. On average, the lowest levels of PM pollution are expected in the park (segment 2) and at the waterfront (segment 3). In contrast, the most pollution is expected at the busy Zeppelinstraße (segment 8). We perform a Friedman test for repeated samples to test whether the “assessments of PM load without sonification” in these segments significantly differ from the other segments. The p-value of the test is clearly below the significance level of 0.001, which is why at least one population median is significantly different from the rest. With a post hoc test (Nemenyi test), we can determine that each route segment was rated significantly differently to at least two other route segments. Route segments 2 (park) and 8 (busy street) were rated particularly differently from the rest of the route

4.4.4. *Comparison of PM measuring methods*

To better compare the different scales of PM mass concentration and visual analog scale, we standardize all data records with their z-score. The higher the absolute value of the z-score, the further the raw measurement data deviates from the attribute’s mean. Further interpretations should be handled cautiously, as the underlying data are not normally distributed for our attributes. However, the data sets are sufficiently large (168 data records of questionnaires, 335 values of each PM2.5 and PM10) to use the z-score as a rough comparative measure of the different attributes. Results are visualized in Figure 6b. The objective PM values in route section 2 (park) are above average but were assessed lowest by the listeners. In route segment 8 (busy street), on the other hand, PM values close to the average were estimated by the listeners to be well above average. We compare the distribution of the normalized z-scores of both attributes for each of the track segments with a Mann-Whitney-U test. The Mann-Whitney U tests we perform for the distributions of the sonification assessment and the PM2.5 data reveal significant differences between the distributions for route segments 2 (park) and 8 (busy street) with p-values < 0.001. No significant differences in the distributions are found for all other route segments. The results are similar for the PM10 data communicated with double bass and the corresponding sonification assessments. The Mann-Whitney U tests find significant differences between the two populations in route segments 2 and 8 (p-values<0.001).

Overall, the results in this section indicate statistically significant differences between the PM data and the sonification interpretation in route sections 2 (park) and 8 (busy street). Since these are precisely the road sections with prominent personal expectations (see section 4.4.1), we hypothesize that the strong expectations may be the cause that the sonification assessment in the two sections does not seem to have a clear connection to the actual sonified data.

Even though in an ideal sonification environment, the listener's interpretation of the heard data should depend only on the data itself, we hypothesize that there are factors, like the listener's expectation regarding the sonified data, that can influence the interpretation in real-world applications. Knowing that the distribution of the underlying PM data in track segments 2 (park) and 8 (busy street) shows significant differences compared to the interpreted sonifications, we test whether these differences might be related to listeners' expectations. Generally, for the whole test track a positive correlation can be observed between the assessment of PM load without sonification and the sonification assessment (Spearman's correlation coefficient 0.697, and p-value<0.001). If we calculate the Spearman's correlation only for route sections 2 and 8, we get a correlation coefficient of 0.792 (p<0.001). However, in the remaining route segments (segments 1 and 3-6), the interpretation of the sonification seems to be less closely related to the PM assessment without sonification (Spearman's correlation coefficient of 0.606 and p<0.001). These results indicate that the sonification perception seems to be not independent of the overall expectation. However, this hypothesis should be further explored using larger sample sizes.

To assess the influence of the study participants, we compare the differences between the visual analog scales of the assessment without sonification and the assessment with sonification in a Bland-Altman plot (see Figure 6c). The different route segments are shown within the plot in different colors. The squares mark the comparisons of the means of both measurement methods for each route segment. Most of the points in the plot are located in the bottom left corner. Contextually, this means that the study participants considered these route segments to have a PM load below average. Moreover, the sonified PM values assessment was higher than the assessment without the sonification device (negative difference). Only the last two route segments (parking lot and busy street) showed a positive difference between the two measurement methods for the majority of the participants. Thus, the sonification was rated to communicate a lower PM load compared to the assessment of the PM load without sonification. Therefore, in the last route segment (busy street), the majority of the study participants understood that the sonification communicated lower values than expected but nevertheless tended to rate the sonification higher than the sonified data actually reported (see Figure 6c). For route segment 2 in the park (see Figure 6c in orange), participants classify the PM load as below average for both measures with and without sonification. Still, the differences between the measurements are relatively small for most participants, and there are no values outside the tolerance range. This indicates an agreement between the measurement methods in the park without the users noticing that the actual sonified data were significantly higher.

4.4.5. User experience

Overall, the sonification device could contribute to raising awareness of particulate matter exposure, as can be gathered from statements made by study participants. They stated, for example: "If my personal impression differs from what I hear, I think about the deviation." Overall, the perception of the environment appears to change as a result of the immediate sonification of environmental data. For example, one person described "I started paying more attention to my surroundings, what I saw (river, nature, cars, etc.). [...] Without the sound, I probably would have just walked without really noticing where I was." Another person described that the sonification "makes you aware of things that you otherwise don't even notice/fade out in everyday life." However, further qualitative evaluation of the study shows that this increased awareness strongly promotes the desire to discover the sources of particulate matter. Numerous statements were made such as "one is constantly looking for the reasons for sudden exposure", "sometimes the frequency suddenly went up or down, I would have liked to know why exactly". However, it is difficult to give correct information on PM emitters due to primary and secondary pollution and broad areas where PM can be dispersed, e.g., by winds.

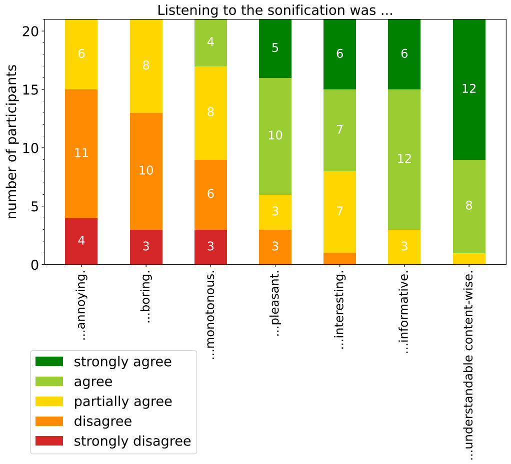


Figure 7: User feedback related to sonification

Nevertheless, the sonification device contributed to the "special appreciation of the less polluted natural areas." One person described the feelings during sonification as follows: "When it was low, I appreciated the surroundings more. When it was high, I got annoyed with every passing car." Annoyance at every peak of PM pollution probably becomes stressful in the long run. However, 48% of the users stated that they would use the device during their leisure time. 67% of users could imagine using the device to choose less polluted routes when walking.

Figure 7 describes possible adjectives to describe the sonification. The study participants were asked how much they would agree with these adjectives as a characterization of the sonification. None of the study participants felt that listening to the sonification was boring or annoying. 4 of the 21 users agreed, and 8 additional users partially agreed with the statement that the sonification was monotonous to listen to. At the same time, however, the vast majority of users also agreed that they perceived listening to the sonification as pleasant. Concerning the context of the sonified data, about 60% of the users found the sonification interesting. Even more users found it informative, and almost all users stated that they understood the sonification in terms of its content.

Overall, the sonification device could help users in route segments with weaker expectations regarding PM exposure to assess air quality or at least raise awareness on the topic of PM exposure.

5. DISCUSSION AND CONCLUSIONS

Our sonification provides a way to perceive air quality, which is an unfamiliar input for a holistic interpretation. The presence of visual stimuli has the ability to capture one's attention and impact their overall sensory experience, ultimately shaping their expectations. If we combine the previous evaluations, it is striking that the focus is always on the same two route segments. Our analysis of the personal assessments of the route segments reveals that route segments 2 (park) and 8 (busy street) show the greatest expectations. These same expectations also significantly correlate with the sonification assessment in the park and along the busy street, while at the same time, the distribution of the actual underlying sonified data shows significant differences from the sonification assessments in the same route sections (2 and 8). The results indicate that when the sonification-independent expectations (e.g., based on visual impressions) of the real-time sonified data are strong, the interpretation of the sonification is linked to the expectations. However, it is doubtful whether changes to the design of sonification would make a significant difference here; rather, it is perhaps a general fact that visual stimuli often dominate or at least influence perception and thus expectation. Furthermore, we hypothesize that when expectations are weaker, the evaluation of the sonification is more likely to be connected to the actual sonified data. The accuracy of the user data would be interesting if visual or olfactory stimuli could be excluded. Sonification is a possible way to influence the behavior. For example, the change of walking paths to healthier conditions. Sonification would be an ideal way of communicating the surrounding environmental influences, since these do not limit the visual capacity.

6. FUNDING

The project was funded by the European Union under Horizon 2020, under grant agreement No. 101003799.

7. REFERENCES

[1] Health Effects Institute, “State of global air 2020. Special report.” MA:Health Effects Institute, Tech. Rep., 2020.

[2] World Health Organization, “Ambient (outdoor) air pollution,” 2021, (visited on 2022-05-17). [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

[3] R. Riley, L. de Preux, P. Capella, C. Mejia, Y. Kajikawa, and A. de Nazelle, “How do we effectively communicate air pollution to change public attitudes and behaviours? a review,” *Sustainability Science*, vol. 16, pp. 2027–2047, 11 2021.

[4] K. Groß-Vogt, T. Hermann, M. W. Jury, A. K. Steiner, and S. Kartadinata, “Klimalanlage - performing climate data,” in *Addressing the Challenges in Communicating Climate Change Across Various Audiences*, W. Leal Filho, B. Lackner, and H. McGhie, Eds. Springer, 2019, pp. 339–355.

[5] T. Komatsu and S. Yamada, “How does the agents’ appearance affect users’ interpretation of the agents’ attitudes: Experimental investigation on expressing the same artificial sounds from agents with different appearances,” *Intl. Journal of Human–Computer Interaction*, vol. 27, no. 3, 2011.

[6] M. S. Pierre and M. Droumeva, “Sonifying for public engagement: A context-based model for sonifying air pollution data,” in *The 22nd International Conference on Auditory Display (ICAD–2016).* International Community for Auditory Display, 5 2017, pp. 60–64.

[7] J. J. Arango, “Airq sonification as a context for mutual contribution between science and music,” *Revista Música Hodie*, vol. 18, no. 1, pp. 92–102, 2018.

[8] K. F. Skov, “Sonification excerpt #4: Rio de janeiro,” 2015, (visited on 2022-06-07). [Online]. Available: https://soundcloud.com/kasper-skov/sonification-excerpt-4-rio-de

[9] T. Hermann and A. Hunt, “An introduction to interactive sonification,” *IEEE multimedia*, vol. 12, no. 2, 2005.

[10] S. Pauletto and A. Hunt, “Interactive sonification of complex data,” *International Journal of Human-Computer Studies*, vol. 67, no. 11, pp. 923–933, 2009.

[11] A. Hunt and T. Hermann, “Interactive sonification,” in *The Sonification Handbook*, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Logos Verlag, 2011, ch. 11, pp. 273–298.

[12] H. Zhao, T. Röddiger, and M. Beigl, “Aircase: Earable charging case with air quality monitoring and soundscape sonification,” in *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, 2021, pp. 180–184.

[13] Bicrophonic Research Institute, “Enviro bike,” 2020, (visited on 2022-06-07). [Online]. Available: https://sonicbikes.net/enviro-bike/

[14] F. Karagulian, M. Barbiere, A. Kotsev, L. Spinelle, M. Gerboles, F. Lagler, N. Redon, S. Crunaire, and A. Borowiak, “Review of the performance of low-cost sensors for air quality monitoring,” *Atmosphere*, vol. 10, no. 9, p. 506, 2019.

[15] A. Lewis, W. R. Peltier, and E. von Schneidemesser, “Low-cost sensors for the measurement of atmospheric composition: overview of topic and future applications,” *World Meteorological Organization (WMO). WMO-No. 1215.*, 2018.

[16] T. Sayahi, A. Butterfield, and K. Kelly, “Long-term field evaluation of the plantower pms low-cost particulate matter sensors,” *Environmental pollution*, vol. 245, 2019.

[17] M. Zusman, C. S. Schumacher, A. J. Gassett, E. W. Spalt, E. Austin, T. V. Larson, G. Carvlin, E. Seto, J. D. Kaufman, and L. Sheppard, “Calibration of low-cost particulate matter sensors: Model development for a multi-city epidemiological study,” *Environment international*, vol. 134, 2020.

[18] S. Schmitz, A. Caseiro, A. Kerschbaumer, and E. von Schneidemesser, “Do new bike lanes impact air pollution exposure for cyclists?—a case study from berlin,” *Environmental Research Letters*, vol. 16, no. 8, p. 084031, 2021.

[19] C. Terboven, “Particulate Matter (PM) Sonification Potsdam Germany,” https://on.soundcloud.com/svQdK, 2023.

[20] S. Jarjour, M. Jerrett, D. Westerdahl, A. de Nazelle, C. Hanning, L. Daly, J. Lipsitt, and J. Balmes, “Cyclist route choice, traffic-related air pollution, and lung function: a scripted exposure study,” *Environmental Health*, vol. 12, no. 1, 2013.

[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[22] A. Ghasemi and S. Zahediasl, “Normality tests for statistical analysis: a guide for non-statisticians,” *International journal of endocrinology and metabolism*, vol. 10, no. 2, 2012.