Soundless: a Citizen Science platform to monitor and analyze noise pollution and its health effects

Pedro García López
pedro.garcia@urv.cat
Universitat Rovira i Virgili
Tarragona, Spain

Daniel Alejandro Coll Tejeda
danielalejandro.coll@urv.cat
Universitat Rovira i Virgili
Tarragona, Spain

Aleksander Slominski
aslom@us.ibm.com
IBM Watson
New York, US

ABSTRACT

Noise pollution is a serious environmental issue that harms the health and well-being of millions of people worldwide. In particular, night noise can seriously disturb sleep patterns and cause heart rate anomalies. To address this problem, we developed Soundless, a citizen science project that leverages mobile phones and wristbands to monitor noise and sleep patterns during the night.

Soundless was deployed in different neighbourhoods of the city of Tarragona (Spain) that suffer night noise due to their proximity to train tracks. Our platform enabled the collection and analysis of acoustic (measured decibels) and health (heart rate, sleep modes) data from 50 volunteers over two phases, one in winter and one in summer. The project provides scientific evidence for policy makers and stakeholders to design and implement noise reduction strategies and actions.

This paper presents Soundless distributed architecture and the results of our data analysis, focusing on three main aspects: the identification of noise and health data incidents, the detection and analysis of individual sensitivity and its response to noise exposure, and a comparison of noisy and silent nights for different users. As expected, we discovered that noise pollution significantly impacts the sleep quality and heart rate variability of the participants. It leads to awakenings, disturbances, and anomalous fluctuations in heart rates.

KEYWORDS

Acoustic pollution, sleep monitoring, mobile devices, Cloud Computing

1 INTRODUCTION

Since 2011, the World Health Organization (OMS) began warning about a growing problem and possible effects of noise on human health [5]. This was followed by a study published in 2019, confirming the severity of this issue and its impact on European citizens [7]. This is a global problem affecting millions of people around the world.

Previous studies [1, 8, 11] have demonstrated that constant exposure to environmental noise during sleep stages can lead to hearing loss, sleep disruption, cardiovascular disease, social handicaps, reduced productivity, impaired teaching and learning, absenteeism, increased drug use, and accidents. These studies introduced a reference value of 45 dB inside homes as the limit for safe noise levels. Above this value, health effects increased significantly.

In this paper, we present the Soundless project, a citizen science initiative exploring the effects of noise pollution in the city of Tarragona, Spain. Soundless comprises a mobile app and a cloud platform enabling the collection and analysis of acoustic and health data from 50 volunteers across two phases, one in winter and one in summer.

The goal of this work is to create a citizen science platform capable of measuring sleep disturbances caused by nocturnal noise. We aim to provide a scientific demonstration of sleep disturbances caused by noise and to analyze the correlations between noise, sleep, and heart rates. We also aim to measure the noise sensitivity of different citizens and to provide them with personalized information.

An important goal of this citizen science project is to raise public awareness about a severe problem that affect citizens. Citizen science can be used as a powerful tool for citizen mobilization and empowerment. In our case, we leveraged active citizen engagement from existing neighborhood associations affected by nocturnal train noise in the city of Tarragona.

The major contribution of this work are:

(1) Development of a distributed citizen science platform involving a Cloud backend service for data capture and analysis, and an Android mobile application (soundless.app) capturing noise levels (decibels) and health variables such as heart rate and sleep levels from Fitbit wearables.
(2) Real-world deployment of the platform for five months in 50 citizens of Tarragona in neighborhoods close to railway tracks, resulting in a dataset with thousands of hours recording the different parameters.

(3) Data collection on noise and health every night of the study, allowing for a more comprehensive understanding compared to previous studies.
(4) Analysis and correlation between noise and health variables, demonstrating a clear relationship between ambient noise and its effects on individuals. We also demonstrate excessive (greater than 45dB) nocturnal noise levels and analyze user sensitivity to noise during different sleep patterns.
(5) Development of several methods capable of calculating each person's sensitivity to noise, and the decibel level necessary for each individual to start experiencing negative health effects due to nocturnal noise.
(6) Significant public impact and visibility of the problem in the city of Tarragona due to our experiment (13 articles in newspapers, 4 interviews on radio, 3 interviews on TV). The study garnered the attention of mass media (TVs, newspapers, radios) and is driving citizen mobilization around this problem to push policymakers to take corrective actions.

2 RELATED WORK

Soundless is a pioneering study in the analysis of nocturnal noise pollution and its health effects. This study allows for the active involvement of citizens in a mobile-cloud technology project to highlight and corroborate noise levels and its health implications during the night. In the preparation of this study, we have reviewed previous works addressing the following aspects: environmental noise measurement, acoustic and health data analysis, and citizen participation in environmental research.

Environmental noise measurement. Various studies have utilized mobile devices to measure noise levels in different urban environments [13, 16–18]. These works have demonstrated the viability and utility of smartphones as acoustic sensors, capitalizing on their portability, connectivity, and processing capability. Some of these studies collect information such as people's sensations at different scales, and even the generation of different noise maps.

However, they have also identified some challenges and limitations, such as microphone calibration, data privacy, measurement quality, and sample representativeness.

An important limitation of these studies is the quality and heterogeneity of microphones, that may question the results of the study. In our case, we compared our measurements with legal certified sound level meter devices showing high accuracy and less than 10% errors. In any case, the goal of our study is to raise awareness of the problem in the society to push policy makers to solve citizen problems.

Acoustic and health data analysis. Some works have analyzed the relationship between environmental noise and human health, using different data sources, such as medical devices, surveys, or statistical data [6, 9, 11]. These works have found evidence that nocturnal noise can affect sleep quality, increase awakenings, and reduce subjective well-being.

Our work extends these findings by collecting and analyzing acoustic and health data from 50 citizens, using mobile devices and fitness trackers. This allows us to identify the individual sensitivity and response to noise exposure.

Citizen participation in environmental research. Lastly, some works have explored citizen participation in the collection and analysis of environmental data, using citizen science platforms [9, 12, 14]. These works have discussed the benefits and potential of involving citizens in science, providing more perspectives and data sources for better and more representative results.

However, they focused on the challenges and issues of citizen science, such as data quality, participant motivation, and ethical considerations. They have also proposed some design guidelines and best practices for citizen science platforms, such as providing clear instructions, feedback, and recognition. Furthermore, they have proposed a framework for evaluating the outcomes and impacts of citizen science, considering the scientific, educational, social, and environmental aspects. Our work is based on these projects but focuses on the problem of noise pollution and aims to involve citizens in identifying and solving this problem, in collaboration with neighborhood associations and public administrations.

3 CONTEXT

Tarragona is a port city located in northeast Spain (Catalonia) by the Mediterranean Sea. A major problem is that railway tracks create a barrier between the city and the sea (beach, port). Apartment blocks are located just at 2-5 meters of distance from the tracks without any acoustic barriers, as we can see on the Figure 1.

The situation has been further exacerbated by a governmental decision to reroute passenger high-speed trains to the interior line, assigning the coastal line to freight trains. This has resulted in long, heavy freight trains passing through different cities in the Tarragona region, generating noise levels exceeding 80 decibels throughout the day and night. With the impending completion of the Mediterranean train corridor, which will connect major cities such as Barcelona, Valencia, Murcia, and Cartagena with Europe, freight traffic is projected to increase further.

Adding to the city's noise pollution, Tarragona's port operates its own rail tracks for cargo movement between docks,

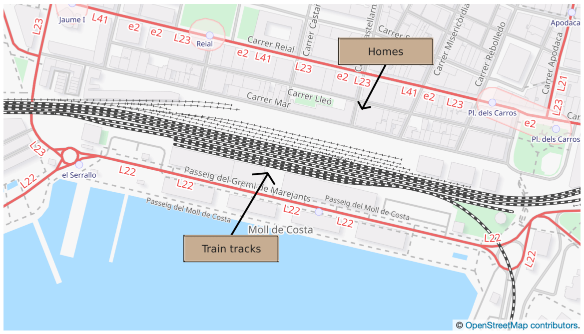

Figure 1: Map of the railway lines of Tarragona

tracks that traverse through several of Tarragona's residential neighborhoods, including Serrallo and Barri del Port. Numerous residents have reported that their buildings vibrate noticeably when freight train pass.

In response to these issues, several concerned citizens approached our research group at Universitat Rovira i Virgili for assistance and potential solutions. Consequently, we initiated "Soundless," a citizen science project designed to identify and address these community problems.

4 SOUNDLESS ARCHITECTURE

Soundless follows a classical client-server architecture where clients interact with an Android mobile application that obtains data from a Fitbit wristband (heart rate, sleep mode) and mobile microphone (decibel level). The client application is executed every night by users before they go to sleep, and it is halted in the morning. All data gathered in the night is submitted to a Cloud service that collects and stores data in an anonymous way.

Moreover, the system incorporates multiple services from the Google Cloud Platform, which furnish the necessary tools for the handling, storage, processing, generation, and analysis of all gathered data.

4.1 Architecture processes

As we can see on the Figure 2, the architecture life cycle start on the data acquisition, this process takes place during the night, as these are the hours focused on the study. This process is carried out on the Android application, which runs on each user's mobile device.

Once the recording is started, the application automatically connects to the Fitbit platform, which, through an individual verification process, provides the health data. These data are stored on each user's mobile device, along with their sound level data.

This information always remains encrypted, and once the user voluntarily decides to share the data with the study researchers, it is sent anonymously and privately to the cloud.

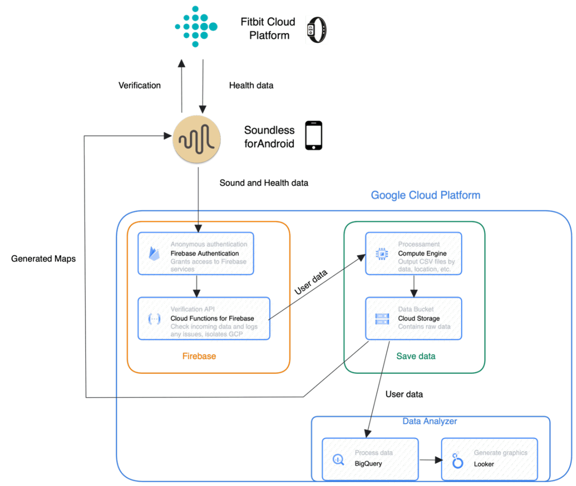

Figure 2: Project architecture

```
uuid,timestamp,dB,heartRate,sleepStage,user
6667f000-a60d-40da-8094-d8e04d70f02d,1670121343732,32,61,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
6667f000-a60d-40da-8094-d8e04d70f02d,1670121354751,33,67,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
6667f000-a60d-40da-8094-d8e04d70f02d,1670121365751,33,67,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
6667f000-a60d-40da-8094-d8e04d70f02d,1670121376770,33,67,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
6667f000-a60d-40da-8094-d8e04d70f02d,1670121387788,33,68,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
6667f000-a60d-40da-8094-d8e04d70f02d,1670121398808,34,62,2,48vvF2k9Ktfg8MMxUUy6Djp0iK73
```

Figure 3: Example of CSV format

4.2 Data Management in the Cloud

These data, once received, are stored in an object bucket on the remote Storage, where all data are stored in an aggregated manner. The format of this data is a CSV format like the one shown in the following image, Figure 3.

This image shows several columns that make up each of the files in our storage.

- uuid: unique and random identifier for each user's recording.
- timestamp: value that allows knowing the time of the night of the record.
- dB: numeric value of the sound level at the time of capture.
- heartRate: numeric value of the person's heart rate obtained through the Fitbit wear.
- sleepStage: the sleep stage in which the person is, this variable can have several values:
  - -1 / 0: the person is completely awake.
  - 1: corresponding to a light sleep stage.
  - 2: deep sleep stage, one of the most important stages of rest.
  - 3: REM sleep.
- user: and finally, a minimal anonymous identifier, which allows knowing the recordings that belong to each of the users, but at no time allows obtaining any

kind of information about the person as expressed in the privacy section.

After this data acquisition and storage process, the data are processed using a series of services that allow the filtering of invalid data, such as: recordings of a few minutes, recordings with impossible sound levels, recordings with data loss, etc.

This filtering process allows us to recognize outliers and remove them so that the data analysis is much more accurate. Once the filtered data are obtained, they are sent to a series of processes such as Google Big Query and analyzed to try to obtain the conclusions that are shown throughout this document.

Finally, we tested the mobile-cloud architecture in production with real users, capable of obtaining, storing, and processing all the data under investigation. This makes the data analysis presented subsequently possible.

5 DATA ANALYSIS

This section explains and details the entire data analysis process. It describes the methods we have used, the reasons for choosing them, and the challenges encountered during the research process to determine the best way to handle the data analysis.

Additionally, the concept of 'incident' is specified and defined, explaining how anomalous events were detected throughout the study and outlining the differences between the types of events that are assessed. All this is to ensure a better understanding of the results obtained from this analysis process.

5.1 Identification of an Incident

In order to correlate noise with health variables, we must first define the concept of a noise incident. an incident is defined as an abrupt and abnormal surge in a variable. For example, if during a 2-hour recording, with an average of 30 dB, there is a moment in the night where this value scale to 50 dB in a few seconds, it is considered an incident and a violation of current night time noise regulations.

To address this, an initial study was carried out to explore the various existing methods for detecting anomalous values in a time series recordings. Following an extensive review, we opted to employ an algorithm based on the principle of dispersion, known as z-scores.

Specifically, this algorithm analyses the data and detects as an incident when any point in the data is at a distance greater than 3 times the standard deviation of the moving average quoted [10, 15, 20]. This moving average is calculated in a time window of 30 seconds, since it is the average time it usually takes for a train to pass, as can be seen in the Tarragona's city council measurement of train noise [3].

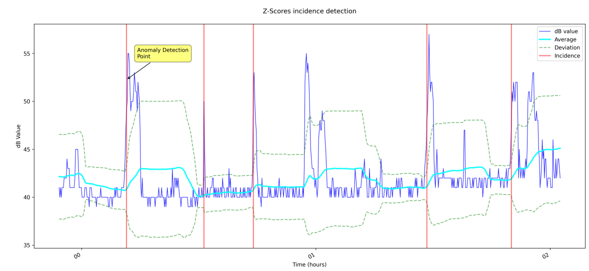

Figure 4: Graphic representation of the functioning of z-scores

These values ensure that an abrupt rise in the data is detected in the time windows marked by the passage of trains.

Z-scores measure how many standard deviations a data point is away from the mean of its dataset. This allows us to identify data points that deviate significantly from the expected range, effectively pinpointing moments of anomalous spikes in sound levels during the night. This algorithm is ideal for our dataset, as it can detect changes over time.

This choice was driven by several key factors:

- Sensitivity to Change Over Time: Z-scores is used to detect deviations from the mean within a time series, making it ideal for identifying sudden spikes in noise levels that might indicate a disruptive event. As illustrated in Figure 4, z-scores effectively highlight these deviations, allowing us to pinpoint potential noise incidents throughout the night.
- Adaptability and Configuration: The z-scores algorithm offers flexibility through various configurations, enabling us to adapt it to the specific characteristics of our dataset and the different variables being analyzed (e.g., decibel levels, heart rate). This adaptability ensures that the algorithm remains effective and reliable across diverse data points.
- Proven Effectiveness: Our initial investigations and tests demonstrated that the z-score algorithm can generate a data set with these new incidents that yielded results that were consistently faithful to what is experienced daily in citizens' homes.

Before applying the z-scores algorithm, we performed data cleaning to remove invalid recordings or impossible data points. This pre-processing step ensured the accuracy and representativeness of our results.

5.2 Health Incidents

While detecting sound incidents is crucial, understanding their impact on health is equally important. Soundless tracked two health metrics: heart rate and sleep stages, providing real-time insights into participants' states during the thousands of hours of recording.

Initially, we explored establishing a direct mathematical correlation between sound levels and heart rate using existing algorithms. Some of this algorithms were Pearson correlation, Spearman, Kendall, and Mutual Information. However, the unique characteristics of our dataset presented a challenge. Heart rate variability during sleep, influenced by natural sleep cycles and individual characteristics, created significant "noise" in the data, making it difficult to isolate the impact of sound incidents. While we observed strong correlations (up to 0.75 using Pearson's correlation) in isolated recordings, the average correlation remained weak (around 0.20 using Pearson's correlation) due to several interfering factors:

- Individual Variability: Participants exhibited diverse heart rate patterns and sensitivities to noise.
- Data Complexity: Recordings contained inherent noise even after filtering, adding to the challenge of isolating the influence of sound incidents.
- Reaction Time Differences: Individual response times to noise varied considerably, ranging from 10 to 40 seconds, further complicating correlation analysis.
- Sensitivity Fluctuations: Individual sensitivity to noise appeared to fluctuate, with inconsistent reactions even within the same participant.
- Sleep Stage Influences: Natural heart rate variations during different sleep stages added another layer of complexity.

Despite these challenges, we observed clear correlations between significant sound incidents and subsequent heart rate fluctuations when analyzing specific events. This led us to adopt the z-scores algorithm for its effectiveness in identifying anomalies within time series data.

After the data pre-processing process, they are analyzed by the z-scores algorithm. This algorithm analyzes each of the recordings individually, first analyzing the sound and then the heart rate. Each of these analysis are carried out by configuring the algorithm parameters for the different types of data that we find (sound and heart rate). Finally a dataset is obtained with the noise and heart rate incidents.

Once the localized incidents were obtained using this algorithm, a search is conducted to find a possible sound incident related with a heart incident within a window time of 10 to 40 seconds. This approach allow us to detect instances where a noise incident likely triggered a subsequent heart rate incident.

Regarding sleep modes correlated with noise, we define a sleep incident as a moment when, due to a previous sound incident, the person has transitioned from a deep sleep or REM phase to being fully awake within a few seconds. This is an event that should not occur under normal circumstances.

6 RESULTS OF THE ANALYSIS

In this section, we provide details about the results obtained in this citizen science project. The first of these is the introduction of the problem, i.e., individual, and aggregate data on the number of incidents detected, the frequency with which the legal limit is exceeded each night, among other important data. Also, we will look on the potential hours when most incidents occur.

First of all, we present the ability to detect and define each user's sensitivity to a series of noises. This sensitivity is defined as a value where a person's health variable (heart, sleep) considerably surges as a consequence of a sudden noise or noticeable increase in sound. That allows us to make a comparison between users and different sensitivity levels. As a result, we can see how each user is completely different from the others.

We also try to compare user behaviour under different noise scenarios. To perform this task, we locate the night with the highest number of noise incidents for each user and compare it with their quietest night. This comparison showcases the significant influence that noises and external incidents have on people's health variables by comparing both cases.

After the analysis of the entire dataset, we can assert that the legal limit set by the World Health Organization (WHO) of 45 decibels was exceeded on 98% of nights with recordings. These data unveil a worrying reality: the legal limit is significantly exceeded during the citizens' rest hours, a limit that should not be permitted by the relevant authorities.

In addition to these alarming figures, during the study, each and every one of our users, on average, **has detected 12 noise incidents every night** where there has been a recording.

These figures are also reflected in the health incidents, since **half of the nights** where there has been a recording, each of the users has experienced an **incident in their heart rate** caused by excessive noise.

In addition, **one in five nights**, each user **has been woken up** due to a high level of ambient noise.

These data do not correspond to a single individual or a small portion of the data but are aggregated values. These values have been obtained by grouping and analyzing the entire set of participants, that is, they are values that affect each and every one of our citizens. These values rely on thousands of hours of recordings and data that allow us to have a much more accurate and realistic picture of what is truly experienced every night in the city.

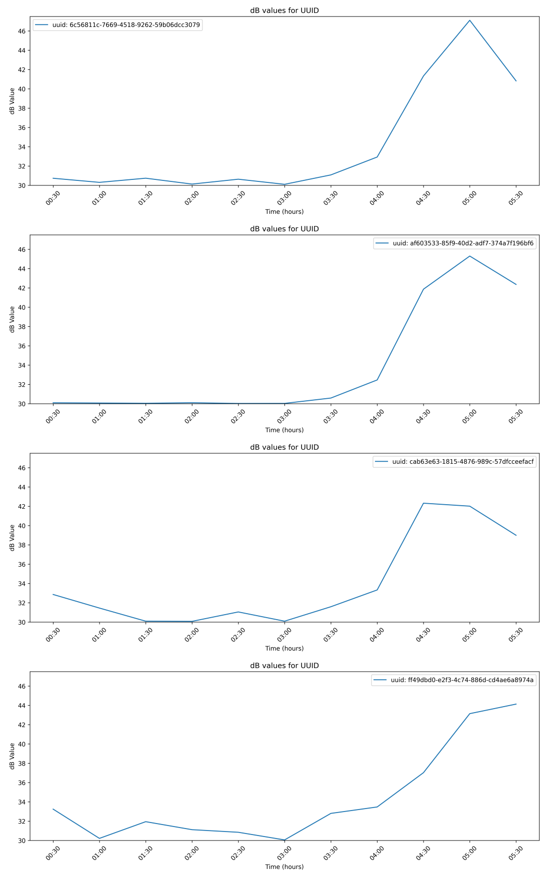

Figure 5: Decibel levels across multiple nights for a single user

Following these alarming data, we began to investigate what could be causing all these noises and the consequences they have on our users. After a long period of research and potential origins of these problems, we detected a specific range of hours when people experienced more noise incidents than at other times.

6.1 Studying night noise levels

There is a certain range of hours during the night when people experience a significant increase in the noise level. This sudden increase consequently results in more health impacts, leading to poorer rest of our citizens.

To illustrate this, we examined decibel levels during several nights randomly chosen and checking the noise level in these recordings. These graphs provide a more realistic and less aggregated view of the noise levels that users encounter during their rest hours.

As can be observed from these recordings Figure 5, decibel level remains relatively constant for most of the night, hovering between 30 and 35 dB. However, between approximately 4:00 AM and 5:00 AM, there is a significant increase in these values.

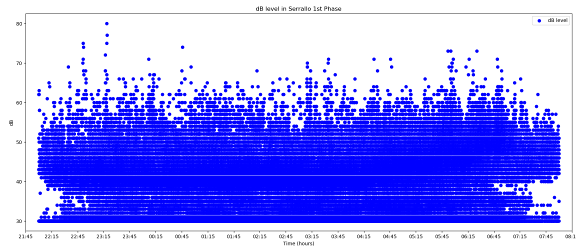

Figure 6: Decibel Level in Serrallo in the 1st phase

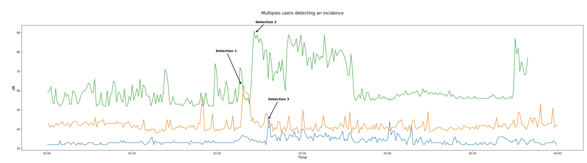

Figure 7: Incident detected by multiples users

This increase once again underscores the sudden surge in noise that users experience at a specific time during the night. This leads to serious health disturbances such as irregular heart rhythms, abrupt awakenings, poor sleep quality, among many other factors caused by abnormal noise levels.

According to data observations, these increases and large variations during this specific range of hours are not an isolated event of a single user, also occur in the majority of our users consistently.

As can be seen in the image, Figure 6, the average decibel level experienced throughout the night during the entire first phase in the Serrallo neighborhood, which is one of the most affected by noise. This image shows how practically all nights the noise level is beyond 45dB and reaching 80dB.

Not only do we have this data, we also observe how multiple incidents are often detected by several users at the same time, as can be seen in the image, Figure 7.

This example shows that these incidents are not an isolated case, but that multiple users detect the same incident at the same time. This shows another example of aggregated data where it can be seen that the problem is not individual, is collective.

Next, to further corroborate the previous information about these hours with greater activity, we have external data provided by official entities, independent of our study, which allow us to identify a possible origin of this trend.

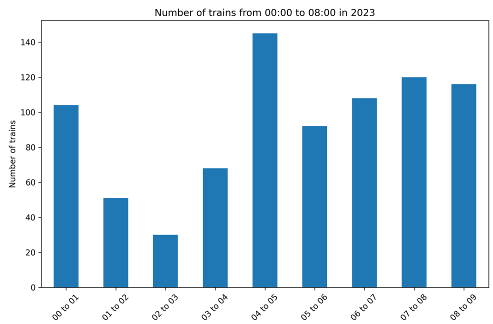

Figure 8: Number of Trains in 2023 and their Actual Times of Passage through Tarragona Station

6.2 Hourly Train Traffic Through Tarragona

As mentioned earlier, this range of extremely noisy hours can not only be observed in our data obtained thanks to the active participation of our citizens, also is observed in the data obtained from official sources that collaborated in our research.

In this case, after reaching out to multiple railway entities, the Port of Tarragona, and the Institut d'Estadística de Catalunya (IDESCAT), among others, we only received responses from two of them: Adif, the Spanish entity responsible for all infrastructure and its management, and IDESCAT.

Both institutions provided data showing the amount of goods transported in different months of the year, maritime traffic at Spanish ports, among many other data. After studying and managing all these data, we were able to filter and obtain the desired values to show, but this time with external data. We extracted the time range during which a larger number of trains are circulating to and from the Port of Tarragona thanks to these external datasets.

This graph, Figure 8, shows the number of trains that passed through the tracks of Tarragona throughout of all the 2023, all grouped by hours. These values represent the actual arrival or departure times from Tarragona provided by Adif.

Examining these data, we can once again demonstrate that the time with the highest number of circulating trains, and therefore, the most problems caused by our main noise source, is from 4:00 AM onward. This is the time with the highest peak of train traffic, with a total of 143 trains.

Once again and starting from official data entirely external to our study, it is shown that these hours are the ones with the highest incident throughout the night. They represent a significant source of disturbances in the heart rhythm and sleep of our citizens.

7 SENSITIVITY OF EACH PERSON

Building upon the next objective of this study, we examined individuals' sensitivity to environmental noise. This is one of the highlights of the study as it introduces an innovation compared to other studies [1, 13, 18] that assess each person's environmental sensitivity.

In this case, we conducted a study and a series of analyses that allowed us to determine a person's noise sensitivity threshold. In other words, we could identify a specific decibel level at which an individual begins to experience health effects such as changes in heart rate, sleep disturbances, or both simultaneously.

To do this study, we used the statistical analysis model known as z-scores, mentioned earlier. This model allows us to detect a substantial increase in a specific variable, be it decibels or heart rate, over a short period, also called incidents.

After identifying these incidents, we combined this data with the entire dataset to find the moment when the incident occurred. This allowed us to understand the decibel level, heart rate, and sleep stage at that exact moment. We then averaged the decibel data registered from the 20 seconds before and after the incident to provide an approximate value for each incident.

Once we obtained this data, we performed a series of averages to ultimately obtain the specific decibel value at which the individual most often experienced a health disturbance. In this case, for the study of sensitivity, we used data from the 11 most representative users. The range of cases varied, from highly sensitive individuals with values close to 38 dB, to individuals with low sensitivity who require louder noises to experience a health disturbance. The data from these users represent 621 recordings used to determine the sensitivities mentioned.

Finally, this study not only allows us to determine the individualized noise sensitivity threshold for different individuals but also enables us to compare sensitivity levels between different individuals, as we will see next.

The previous graph, Figure 9, shows the decibel level at which individuals began to experience health disturbances, with the x-axis representing different users. Moreover, it includes three lines for each user: the decibel threshold necessary to affect the heart rate, the value necessary to affect sleep, and finally, the noise level that begins to affect both simultaneously.

To obtain these values, we considered all recordings and the health effects the users experienced throughout their history. As expected, the more recordings and data we have from a single user, the value at which noise begins to affect the individual's health will be more precise.

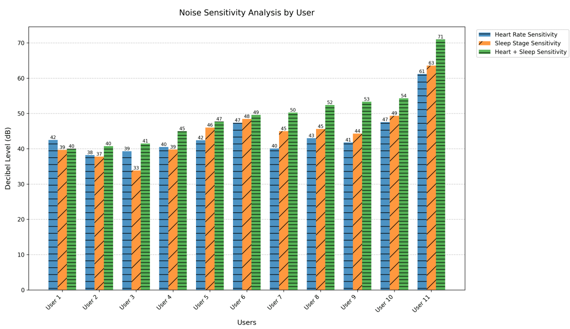

Figure 9: Sensitivity levels of heart rate, sleep, and both combined, for different users

As can be seen in this graph, there are significant differences between individuals, showing that some people are extremely sensitive, with a sensitivity of 39 decibels to start affecting health, while others have a much higher tolerance to the same noise levels.

7.1 General sensitivity

Upon examining this graph in depth, in addition to the differences between individuals, we can also observe a nearly universal trend of noise affecting heart rate before sleep. That is, a lower intensity noise is required to affect an individual's heart rate during sleep than to awaken that individual. Consequently, a higher decibel level is needed to simultaneously affect these two values, demonstrating the sensitivity people have during their sleep phases and the different health effects that can occur during this stage.

Finally, by collectively evaluating all users, their incidents, and the thousands of hours of data we have, we can extrapolate the decibel levels needed to affect each of these variables generally. In other words, the decibel level most likely to affect the majority of users.

These values show us that between **40 and 50 decibels** our users have begun to show effects on their heart, sleep and both at the same time. This range is due to the limitation of the mobile devices used in the study. Even so, it is worth noting that these same values correspond to what is mentioned throughout the document.

7.2 Sleep Interruptions in Sensitive vs Non-Sensitive Individuals

With this information about the sensitivity of each person to different noises during the night, it is possible to aggregate and combine these data with our dataset of incidents and reach new conclusions, thanks to this new variable.

In this case, special emphasis has been placed on investigating the number of sleep incidents detected throughout the nights based on whether a person is sensitive or less sensitive to noise. Remember that we define a sleep incident as when a person moves from a deep sleep phase to being completely awake due to excessive noise.

After investigating these data, we obtained the percentage of times that a noise, exceeding the legal limit of 45 decibels, has caused an abrupt awakening in each people in the study. Here we observe that among all the data we have, it is clearly those more sensitive individuals who register more sleep incidents, due to their lower tolerance to different noises. Specifically, these results show the following:

- 70.59% of the times a sleep disturbance has been detected, it has occurred in people with a higher sensitivity to noise, in this case, a sensitivity less than 45 decibels.
- On the other hand, people with lower sensitivity to this type of situation have only registered 29.41% of the sleep disturbances caused by noise.

In this section, several conclusions were drawn that allowed us to gain a new perspective for the study, enabling us to demonstrate and get a level of user sensitivity.

This level of sensitivity was determined through the incidents detected by each user, and the decibel level at which these incidents occurred. Thus, the more data we had for each user, and the more incidents were detected, more precise this highly individual and personal value would be.

This sensitivity allows us to make comparisons between the different users in the study. It demonstrates that as there are more sensitive and less sensitive individuals, each of them experiences the same noises in entirely different ways. This results in people with lower noise tolerance being more easily affected, leading to a poorer quality of rest.

In addition, we also obtained values for general sensitivities, that is, values at which the majority of people in the study began to experience heart disturbances. We found that this general sensitivity corresponded to the legal limit set by the World Health Organization (WHO) of 45 decibels.

Finally, we also gathered some data on the frequency with which sensitive individuals reacted to noises and experienced a disturbance in their sleep, causing them to wake up, compared to less sensitive individuals. The former accounted for 70% of the total sleep disturbances.

8 SELF-COMPARISON ANALYSIS

The complexity and diversity of personal heart rates and sleep modes makes very difficult to create a control group that can be compared to the users of this study. We decided that it is more adequate to compare the same user at different nights showing different levels of noise. That way we can observe how noise affects every user comparing noisy and silent nights.

Figure 10: Sound levels and heart rate during a quiet and very noisy night for User 1

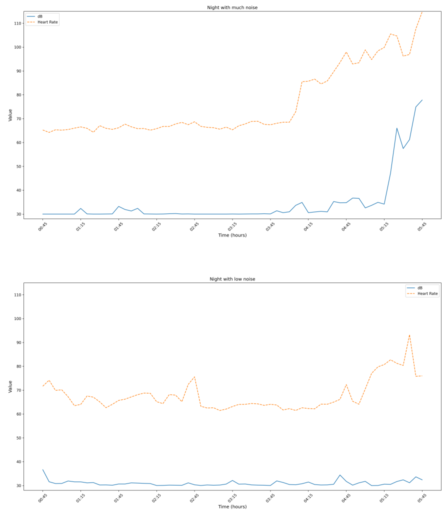

Figure 11: Sound levels and heart rate during a quiet and very noisy night for User 2

We compare, for each of the users, nights with a higher rate of environmental noise, and consequently, a greater health disturbance, with quieter nights where there were recordings. This approach allows us to individually compare the implications that external factors during sleep have on a person's health and rest.

In this case, we have conducted the study from different perspectives, one more individualized and detailed, and another showing the same reality, but from a general perspective. In the case of the more individualized view, we have chosen three users selected at random, along with their noisiest nights and their quietest nights. Observing in each of the graphs the sound values, where in many cases, the noise level on the noisiest nights exceeds values above 80 decibels, a value practically double what is allowed by law in Spain regarding this type of nocturnal noise inside homes.

It should be emphasized that **the trend of the selection of random users of the following graphs, are observed in practically all the users** and their different recordings, showing a general trend when comparing their quietest nights against the noisiest ones.

8.1 Individual view

In these graphs, Figures 10, 11, 12, two prominent lines can be observed. The blue line represents the recorded decibel level, and the orange line represents the user's heart rate at each moment. As can be seen, there is a significant difference between the noisiest and least noisy nights for all random users, with differences of tens of decibels and a much wider heart rate range.

Figure 12: Sound levels and heart rate during a quiet and very noisy night for User 3

In the first user, we observe in the first graph, Figure 10, how the sound values remain constant throughout almost the entire night, close to 30 dB, which is absolute silence. On the other hand, in terms of heart rate, we see the person fluctuating between 50 and 60 BPM (beats per minute), being these heart rate variations a completely normal and natural process over sleep cycles.

However, when we compare to the opposite graph, Figure 11, where there is a high level of noise, we observe how

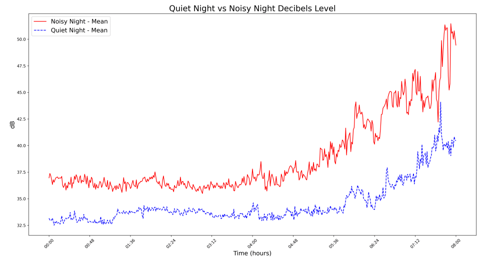

Figure 13: Average decibels on the noisiest and quietest nights

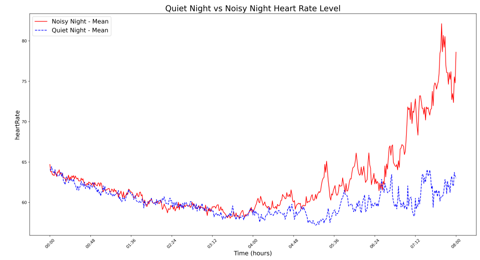

Figure 14: Average heart rate level on the noisiest and quietest nights

the sound remains constant until a certain point in the night where it exceeds 55 dB, thus surpassing the legal limit. With this sudden increase, we also observe a significant increase and deviation from the normal heart rate.

This trend is observed in each of the studied users, where we see in the first graph of these, a normal heart rate, fluctuating as the different sleep phases change, and a sound practically in silence. Compared to the second graph where it is observed that, at a moment in the night, a great noise occurs, significantly increasing the noise produced by environmental factors, and consequently causing a significant alteration in people's heart rates.

8.2 Sound and Health Relationship

We observe that during a night with virtually no noise and constant silence, individuals' heart rates remain steady throughout the night. However, when there are significant variations in sound, the participant's heart rate reflects these changes almost immediately.

This observation shows a relationship between these variables, where certain levels of noise can almost immediately disrupt a person's heart rate, especially during the sleep hours of the participants.

8.3 Grouped View

After this individualized view of some randomly selected users, it's time to see if this trend is isolated or if it is a pattern that is observed throughout the entire study. In this case, to make the graph simpler and more visual, it was decided to display the average decibel level throughout an entire night with each of the data sets, the noisy nights and the quiet nights. This presents a graph with aggregated and general data, thereby allowing us to assess different trends in the entire study.

In this case, Figure 13, as was observed before, there is a significant difference in the noise level on the noisiest nights and the quietest nights. We observe that on the quieter nights, with the blue line, the noise level remains practically constant, with a maximum deviation of 7 points, in addition to an average sound level clearly lower than the second data set.

On the other hand, on the noisiest nights, we observe that the average sound level is much higher, reaching values above 50 decibels, remember this is an average, and a deviation of more than 10 points.

In this case, Figure 14, we observe that the average heart rate of these individuals during the night is also a constant value. Initially, there is a drop as people start to enter deep sleep, causing the heart rate to slow down as part of the human sleep stages. However, once the general noise level begins to rise, the heart rate also increases exponentially.

We see a transition from an average heart rate of less than 65 beats per minute (bpm), with minimal deviation, to values above 80 bpm at the highest peak, and deviations of more than 10 points. All of this illustrates how this relationship between the two variables affects how the human body responds to rest, which should be a peaceful and disturbance-free time.

Finally, comparing these data, we see that, on one hand, we obtain a much lower average in the case of the quieter nights, and on the other hand, we observe that the deviation, that is, how much the average varies throughout the night, is much smaller. It is also worth noting the peak in decibels exceeding the legal limit that is recorded on the noisiest nights.

Finally, to conclude this section, we once again observe what has already been mentioned in previous sections, 6.1 and 6.2, the hours during which the highest noises and most frequent disturbances are observed.

In this case, as in the previous ones, and observing the hours when there is a significant increase in noise levels, it is the range of hours from 4:00 AM to 5:00 AM, approximately. This trend is observed throughout all the graphs shown in this self-comparison.

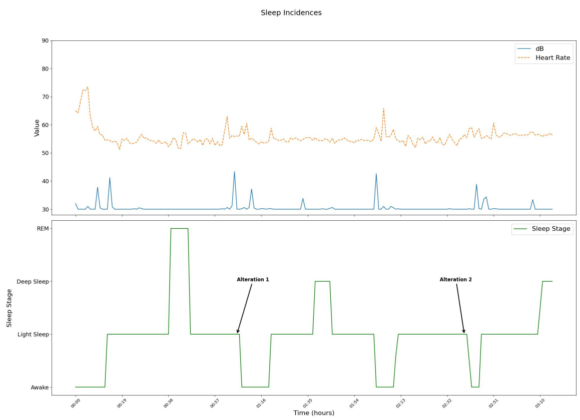

Figure 15: Sleep incident produced by noise on User 1

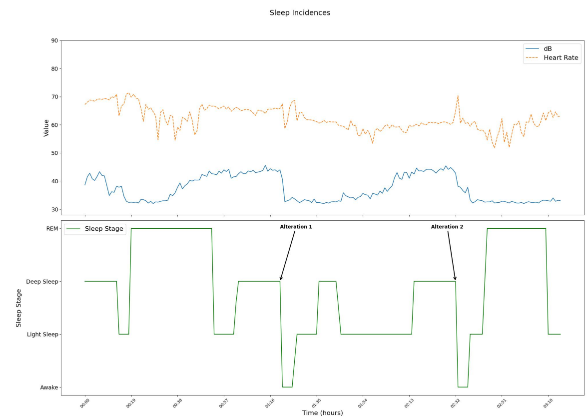

Figure 16: Sleep incident produced by noise on User 2

This reaffirms, once again, the hours already detected by this study, where a greater number of incidents in both aspects, sound and health, occur. In many cases, these incidents involve sounds that exceed the limit set by government institutions multiple times, and which should not be permitted during rest hours.

Furthermore, we observe both individually and collectively, the differences between the records of our users' quietest nights compared to their noisiest nights are quite significant. This difference is evident in the average level and deviation of both sound and heart rate. We see that nights with more disturbances are the nights where the heart rate level is much higher, corresponding with the noise level experienced by our users. All of this demonstrates a certain dependency between these two variables.

9 SLEEP INCIDENTS

The study has highlighted the importance of detecting incidents, the problem that noise represents and how it significantly affects the heart at night. However, in this section we will focus on its direct impact on sleep.

After conducting a specific study, it is observed that sleep is affected in one on five nights for each user. This means that, due to an abnormal noise, citizens wake up completely, which is a worrying fact.

In addition, it has been recorded that 20% of the sleep incidents caused by noise have been caused by levels higher than 60 dB, being values much higher than expected during the night [2, 4].

This values are alarming because they indicate that people's sleep is interrupted by sounds whose decibel levels exceed the limits established by the WHO.

This problem is clearly illustrated in Figures 15 and 16, which clearly shows how on several occasions during the night, users go from deep sleep, one of the most important and restorative phases of sleep, to being completely awake due to an incident caused by noise.

As seen in these graphs, these noise incidents, again, exceed 45 dB in both users, even reaching levels close to 50 dB. After this increase, users are suddenly awakened.

These graphs provide a representative view of what happens to the study users during the night, taking into account the intrinsic differences of individual, as discussed in the section 5.2.

10 HEALTH AND SOCIAL IMPACT

This is not the first study that reports how night noise is causing sleep disturbance and health problems. Many studies [19] confirm WHO analysis [5] where above 40dB there are strong impacts in health. Awakenings and sleep disturbance clearly preclude restorative sleep, but the chronic exposure to night noise produces different diseases in the long term [19].

We identified a severe health impacts in Tarragona's neighborhoods close to the train tracks. Apartments are located at less than 5 meters of the tracks without any noise barrier, which means that they are exposed to extraordinary noise levels reaching 80dB in our study. All involved neighbors suffered heart and sleep disturbances when noise increased suddenly over 40-50 dB.

Concerning the validity of our noise measurements, numerous previous studies also using mobile devices report less than 10% accuracy difference with legal specialized acoustic sensors. On the other hand, the city of Tarragona has already executed a legal noise study in the same neighborhoods [3], which showed high noise during day and night (around 70 and 80 dB) that coincided with our measurements.

The problem will get even worse in 2025-2026, when the freight train traffic will be multiplied by three as it has been officially recognized by the Mediterranean corridor office.

Furthermore, the length of train convoys will grow from 500 meters to 750 meters. This means much more noise and more prolonged in time during all nights. This will obviously aggravate the health impacts and further disturb the sleep of citizens.

This study has received public attention since it appeared a lot in mass media like newspapers, radios, and TVs. Our citizen science project and the credibility of the University scientific work has attracted the attention of the public and was useful for increasing awareness about this severe problem.

Furthermore, the experiment is boosting user mobilization and engagement to solve this problem. In fact, we are collaborating closely with an existing citizen association "Mercaderies per l'Interior" that aims to convince policy makers to move all freight trains traffic to the interior line. This association has organized public demonstrations and has extended their reach to more citizens that did not find a way to organize their complaints.

11 CONCLUSION

Soundless is a citizen science problem that aims to provide scientific evidence of nocturnal noise disturbance to raise public awareness. The project involved 50 citizens that used an android mobile application and fitbit wristband in the night to monitor acoustic pollution and their health data (heart rate, sleep) during five months.

After an exhaustive study of hundreds of data points with thousands of hours of recording, we can conclude that the study has fulfilled its purpose. We report that the night noise produced by freight trains in the seaside of Tarragona is far beyond legal limits (>45dB) –and up to 80dB– during all nights. Furthermore, we consistently detected the impact of noise in health variables like heart rates and sleep modes, which affected most users in our study during all nights. This disturbances during sleep are causing irregular heart rates, sudden awakenings, and noticeable impacts in sleep quality in the citizens.

We have shown that there is a serious noise problem caused by the passage of freight trains in the night close to apartments violating noise legal levels. This problem has been detected on multiple users, on multiple nights and finally corroborated by the data provided by official acoustic studies from the city of Tarragona[3].

REFERENCES

[1] Stansfeld-Stephen A and Matheson-P Mark. 2003. Noise pollution: non-auditory effects on health. *Oxford University Press* 68, 1 (Dec. 2003), 243–257. https://doi.org/10.1093/bmb/ldg033

[2] IAC ACOUSTICS. 2024. Comparative Examples of Noise Levels. https://www.iacacoustics.com/blog-full/comparative-examples-of-noise-levels

[3] Enginyeria Industrial JMC Ayuntamiento de Tarragona. 2021. *Informe sonométrico del Ayuntamiento de Tarragona*. Technical Report. Ayuntamiento de Tarragona. Informe sonométrico elaborado por Enginyeria Industrial JMC.

[4] Royal National Institute for Deaf People. 2024. How loud is too loud? https://rnid.org.uk/information-and-support/ear-health/protect-your-hearing/how-loud-is-too-loud/

[5] World Health Organization. Regional Office for Europe. 2011. Burden of disease from environmental noise. Quantification of healthy life years lost in Europe *(WHO EURO)*. World Health Organization. Regional Office for Europe, 106 pages. https://www.who.int/publications/i/item/9789289002295

[6] World Health Organization. Regional Office for Europe. 2019. Burden of disease from environmental noise: quantification of healthy life years lost in Europe. World Health Organization. Regional Office for Europe, 106 pages. https://iris.who.int/handle/10665/326424

[7] World Health Organization. Regional Office for Europe. 2019. Night noise guidelines for Europe *(WHO EURO)*. World Health Organization. Regional Office for Europe, 162 pages. https://iris.who.int/handle/10665/326486

[8] Lisa Goines and Louis Hagler. 2007. Noise pollution: a modem plague. *Wolters Kluwer Health, Inc.* 100, 3 (Jan. 2007), 287–294. https://doi.org/10.1097/smj.0b013e3180318be5

[9] Rainer Guski, Dirk Schreckenberg, and Rudolf Schuemer. 2017. Environmental noise guidelines for the European Region. *International Journal of Environmental Research and Public Health* 14, 12 (Dec. 2017), 1539. https://doi.org/10.3390/ijerph14121539

[10] Data Heroes. 2023. What is the Z-Score for Anomaly Detection? https://dataheroes.ai/glossary/z-score-for-anomaly-detection/

[11] J. A. Horne, F. L. Pankhurs, L. A. Reyner, K. Hume, and I. D. Diamond. 1994. A field study of sleep disturbance: effects of aircraft noise and other factors on 5,742 nights of actimetrically monitored sleep in a large subject sample. *Oxford University Press* 17, 2 (March 1994), 146–159. https://doi.org/10.1093/sleep/17.2.146

[12] Hiromi Kobori, Janis L. Dickinson, Izumi Washitani, Ryo Sakurai, Tatsuya Amano, Naoya Komatsu, Wataru Kitamura, Shinichi Takagawa, Kazuo Koyama, Takao Ogawara, and A. J. Miller-Rushing. 2015. Citizen science: a new approach to advance ecology, education, and conservation. *John Wiley and Sons* 31, 1 (Nov. 2015), 1–19. https://doi.org/10.1007/s11284-015-1314-y

[13] Zipf L, Primack RB, and Rothendler M. 2020. Citizen scientists and university students monitor noise pollution in cities and protected areas with smartphones. *PLOS One* 15, 9 (Sept. 2020), 15. https://doi.org/10.1371/journal.pone.0236785

[14] Nicolas Maisonneuvea, Matthias Stevens, Maria E. Niessen, Peter Hanappe, and Luc Steels. 2009. Citizen noise pollution monitoring. Association for Computing Machinery, 96-103 pages. http://hdl.handle.net/10261/127898

[15] Oracle. 2014. Z-Score. https://docs.oracle.com/en/cloud/saas/planning-budgeting-cloud/pfusu/insights_metrics_Z-Score.html

[16] Carlo Andrea Rozzi, Francesco Frigerio, Luca Balletti, Silvia Mattoni, Daniele Grasso, and Jacopo Fogola. 2022. Indoor noise level measurements and subjective comfort: Feasibility of smartphone-based participatory experiments. *PLOS ONE* 17, 1 (Jan. 2022), e0262835. https://doi.org/10.1371/journal.pone.0262835

[17] Eunyoung Shim, Dohyeong Kim, Hyekyung Woo, and Youngtae Cho. 2016. Designing a Sustainable Noise Mapping System Based on Citizen Scientists Smartphone Sensor Data. *PLOS ONE* 11, 9 (Sept. 2016), e0161835. https://doi.org/10.1371/journal.pone.0161835

[18] Artemis Skarlatidou, Muki Haklay, and Antonella Radicchi. 2021. Chapter 6. Citizen science mobile apps for soundscape research and

public spaces studies: lessons from the Hush City project. In *Geographic Citizen Science Design: No one left behind*. University College London Press, 130–148. https://doi.org/10.2307/j.ctv15d8174

[19] Michael G Smith, Makayla Cordoza, and Mathias Basner. 2022. Environmental noise and effects on sleep: an update to the WHO systematic review and meta-analysis. *Environmental health perspectives* 130, 7 (2022), 076001.

[20] Akash Srivastava. 2023. Detecting Anomalies with Z-Scores: A Practical Approach. https://medium.com/@akashsri306/detecting-anomalies-with-z-scores-a-practical-approach-2f9a0f27458d