# Predicting and Visualising City Noise Levels to Support Tinnitus Sufferers

William Hurst[1], Graham Davis[2], Abdennour El Rhalibi[1], David Tully[1], Zhigeng Pan[3]

[1]*School of Computing and Mathematical Sciences, Liverpool John Moores University, Liverpool, UK*
[2] *WYG Environmental, UK*
[3]*Digital Media and HCI Research Center, Hangzhou Normal University, Hangzhou, PRC*
W.Hurst, A.Elrhalibi {@ljmu.ac.uk}, graham.davis@wyg.com, d.tully@2008.ljmu.ac.uk, zgpan@hznu.edu.cn

**Abstract:** On a daily basis, urban residents are unconsciously exposed to hazardous noise levels. This has a detrimental effect on the ear-drum, with symptoms often not apparent till later in life. The impact of harmful noises levels has a damaging impact on wellbeing. It is estimated that 10 million people suffer from damaged hearing in the UK alone, with 6.4million of retirement age or above. With this number expected to increase significantly by 2031, the demand and cost for healthcare providers is expected to intensify. Tinnitus affects about 10 percent of the UK population, with the condition ranging from mild to severe. The effects can have psychological impact on the patient. Often communication becomes difficult, and the sufferer may also be unable to use a hearing aid due to buzzing, ringing or monotonous sounds in the ear. Action on Hearing Loss states that sufferers of hearing related illnesses are more likely to withdraw from social activities. Tinnitus sufferers are known to avoid noisy environments and busy urban areas, as exposure to excessive noise levels exacerbates the symptoms. In this paper, an approach for evaluating and predicting urban noise levels is put forward. The system performs a data classification process to identify and predict harmful noise areas at diverse periods. The goal is to provide Tinnitus sufferers with a real-time tool, which can be used as a guide to find quieter routes to work; identify harmful areas to avoid or predict when noise levels on certain roads will be dangerous to the ear-drum. Our system also performs a visualisation function, which overlays real-time noise levels onto an interactive 3D map.

**Keywords**: Hazardous Noise Levels, Data Classification, Tinnitus, Visualisation, Hearing Loss, Prediction, Real-Time

## 1. INTRODUCTION

Hearing loss in the elderly has a significant cost impact on the UK's National Health Service every year, where it is estimated that 10 million people suffer from damaged hearing [5]. This specifically covers 6.4 million who are retirement age or above and 3.7 million individuals of working age. It is estimated that by 2031 there will be 14.5 million people suffering from some form of hearing loss [5]. This will have a major cost bearing on the health service provision, more so than conditions such as diabetes or cataracts [5].

The current UK health and safety guideline state that, when exposed to a sound of 85 dB(A) or higher in the work place, the employee must wear some form of hearing protection. The European Control of Noise at Work Regulations, however, states that noise exposure of 80dB(A) or higher should be counteracted with protective gear. We use the European guidelines for our research. The impact of harmful noise levels on an individual's hearing can differ. However, the more time spent exposed to a sound over 80db, the more damage is caused to the ear-drum. Tinnitus is a condition, for which, currently, there is no cure. The symptoms include buzzing, ringing or monotonous sounds in the ear, which are permanent and constant throughout the day and night. Action on Hearing Loss (formerly RNID) state that in 2011, 10% of adults in the UK had some form of Tinnitus. The condition is brought about by damage to the ear drum and often caused by expose to frequent loud noises on a daily basis. Sufferers of Tinnitus find that the symptoms are exacerbated by regular exposure to loud noises encountered during their daily activities.

As our research shows, the level of noise in city centres and public places can have a significant impact on an individual's hearing. In this paper, an approach for assessing and predicting noises levels at specific urban locations during the day is presented. The system performs a data classification process to identify and predict harmful noise areas at diverse periods. The approach also visualises the noise levels in public places, such as city centres. The analysis is achieved by using machine learning classifiers to detect, and subsequently predict, trends in high noise levels at specific times of the day. A visualisation of the results allows the user to view the best times to avoid certain areas of a city through an interactive tool.

The rest of the paper is as follows. Section 2 presents a background and motivation behind the research. The dataset used in our research is detailed in Section 3 which includes an account of our system design and data classification techniques. Section 4 provides a discussion and account of the results. Section 5 presents an overview of the visualisation process for a real-time graphical display of the harmful levels in a city. The paper is concluded in Section 6, which provides a discussion on the work presented and details how the work will be taken further in the future.

## 2. BACKGROUND

In the UK, in 2010, £1.34 was spent on care for each individual affected by hearing damage. This is significantly lower than the nation spends on diabetes (£21.21) and sight loss (£14.21) wellbeing. In Australia, it is estimated that the disease burden on the economy associated with hearing loss was $11.3 billion, during a study taken in 2005 [5]. The World Health Organisation anticipates that hearing loss will have an incremental drain on health care providers in the next 15 years. For example, it is projected that hearing damage will be the most common disease in the UK by 2031 [4]. There is a significant cost benefit to governments in the prevention of hearing damage. The aim of this project is to illustrate how noise levels in metropolitan areas impact hearing. A greater awareness of harmful sounds, and where and when they occur, may reduce the costs of health care and alleviate symptoms in conditions such as Tinnitus. By providing urban residents with detailed real-time noise maps, specific to their location, enables individuals to safeguard their hearing and be aware of the potential damage to the regular exposure to levels over 80dBs in their daily lives.

### 2.1 Hearing Loss and Noise Prediction

The most common cause of hearing damage is age related. However, the cause can be prolonged exposure to sounds which are over 80db(A). There are known to be four different levels of hearing loss:

(1) Mild hearing loss, where individuals have difficulty hearing outside the mean range of 25 to 39 decibels.
(2) Moderate hearing loss relates to individuals who find it difficult to follow speech and hear between 40 and 69 decibels.
(3) Severe hearing loss refers to individuals who require hearing aids and the use of sign language to communicate.
(4) Profound deafness, refers to individuals who are able to only hear 95 decibels or higher.

Hearing loss is a non-life-threatening condition and, for that reason, is often overlooked in a healthcare environment, particularly in developing markets [17]. This is enforced by the research put forward by Figueira *et al*., which details the creation of humanitarian apps, for audiometric hearing tests in affordable format [17]. The idea is to make hearing-loss healthcare more available in emerging markets by employing existing mobile technology. Their proposed App evaluates an individual's hearing ability through their mobile device and, subsequently, detects if the user has hearing damage. The project, however, is reliant on the availability of smart phones and the results have not been compared with an audiometer to test the successfulness of the project. Research, such as this, paves the way for making hearing loss treatment more widely available, particularly in emerging markets, however it does little towards the prevention of hearing-loss.

The system proposed in this paper consists of three topographies: intelligent noise evaluation; sound prediction and interactive visualisation. Noise visualisation is the ability to digitally record or assess sound and present a conception to a user. The conception of sounds provides an ideal way to communicate data which is invisible to the human eye and where the effects are not visible. The process can be made possible through use of an acoustic camera [19] to directly visualise sound in real-time. Alternatively, class 1 integrating sound level meters can monitor specific noise levels and construct datasets, which can be used for post-analysis.

Noise level prediction focuses on intelligent forecasting down to the street level in an urban environment. No research has covered the use of intelligently predicting noise levels in urban areas or the generation of interactive noise maps as a guide to hearing-impairment sufferers. Current sound prediction models rely on forecasting through use of simple calculations to estimate future sound levels. This technique is employed when new transportation or development projects are planned. No systems present a real-time hour by hour visualisation to the street level.

### 2.2 Tinnitus

Tinnitus is a condition which results in the perception of monotonous sounds resonating in the ear. The resonances are the result of absence of corresponding external sounds. The condition is permanent and there is no cure and can be made worse by expose to frequent loud noises. For that reason, the condition requires management. The British Tinnitus Association estimated that around 10% of the UK population have some form of Tinnitus; with a further 1% having a condition which can affect their lifestyle [6]. It is a condition which is unseen and the level of suffering is only known by the patient as there are no visible symptoms. Having a strong form of Tinnitus is also linked to and can cause depression. Holmes et al stated that in 2009, the level of depression is higher amongst sufferers of Tinnitus than it is in the general population [7]. Tinnitus can have psychological effects on the patient. Often communication becomes difficult, and the effects may also mean that the individual is unable to use a hearing aid as the buzzing, ringing or monotonous sounds in the ear are enhanced. Action on Hearing Loss state that sufferers of hearing loss are more likely to withdraw from social activities.

A specific goal, of the research put forward in this paper, is to aid with the prevention of Tinnitus and reduce its symptoms. This is achieved by providing an approach for sufferers to avoid exposure to excessive noise levels during the day. Frequent and prolonged encounters of loud noises in the daytime can trigger the ringing and buzzing sounds which are associated with Tinnitus and often last well into the evening and night.

### 2.3 Noise Visualisation

Data is not fixed and is a changing entity [1]. This is particularly true for sound data, which is dynamic and has a varied level of granularity. Creating sound visualisations for the hearing impaired is developing research area [8]. As technology advances, access to smart equipment has been made easier [9]. For example, Brophy *et al.*, focus on the visualisation of loud sounds [9] in real-time to aid the hearing impaired visualise the environment around them. Their approach provides greater interactivity for an individual with a hearing impairment and the surroundings. The project works by capturing the environment using a camera when a loud sound is detected. The image is then displayed to the individual with the hearing impairment via the use of virtual reality glasses. Using visualisation techniques can help project unseen environmental characteristics in real-time. This is of particular benefit to sufferers of Tinnitus, and the reducing of hearing damage caused by urban noise. By providing a visual guide about noise levels down to individual street level, areas can be avoided if necessary.

In this paper, we present an approach and tool for the analysis, prediction and visualisation of noise data in public places. Specifically, the focus on the sound data sets from Leicester city centre in the UK. The dataset is provided by WYG Environmental.

### 2.4 Noise Assessment of an Urban Area

The approach put forward involves interacting with data to find hidden information and view if trends in noise patterns develop over time. The environmental noise monitoring was undertaken using Rion NL-52 class 1 integrating sound level meters to establish baseline ambient, background and specific source noise levels. Measurements were taken in accordance with BS 7445-1:2003 The Description and Measurement of Environmental Noise: Guide to Quantities and Procedures. The measurement equipment was checked against the appropriate calibrator at the beginning and end of the measurements, and no drift was observed. The following statistical parameters were recorded at a variety of logging periods, including: $L_{Aeq}$, $L_{Amax}$, $L_{Amin}$, $L_{A10}$, $L_{A90}$ and linear $L_{eq}$ values. All the values are sound pressure levels in dB (re: 2 x $10^{-5}$ Pa). Sound levels can be measured in frequency bands to provide detailed information about the spectral content of the noise. These measurements are usually undertaken in octave or third octave frequency bands. If these values are summed logarithmically, a single figure value can be calculated. This describes the total amount of acoustic energy measured but does not take any account of the ear's ability to hear certain frequencies more readily than others.

Instead, the dB(A) figure is used. This is found to relate better to the loudness of the sound heard. The dBA figure is obtained by subtracting an appropriate correction, which represents the variation in the ear's ability to hear different frequencies, from the individual octave or third octave band values, before summing them logarithmically. As a result, the dB(A) value provides a depiction of how loud a sound is in reality. The 'A' is used to state average, whereas 'C' would be the peak noise, i.e. dB(C). Consequently, the dataset includes 10 features, each is accounted below:

- $L_{Aeq}$: Sounds vary and fluctuate with time. Instead of having an instantaneous value to describe the noise event, an average of the total acoustic energy experienced over its duration provides a more accurate account. The LAeq, 07:00 – 23:00 for example, describes the equivalent continuous noise level over the 12 hour period between 7 am and 11 pm. LAeq is calculated using the formula:

$$L_{Aeq} = SEL - 10Log(t) + 10Log(n)$$

T is Time and n is the amount of events within a given time. SEL is the sound level over one second. This would typically have the same energy content as the whole event.

- $L_{Amin}$: The quietest instantaneous noise level recorded, specifically the quietest 125 milliseconds measured during any given period of time, is given the LAmin annotation.
- $L_{Amax}$: The $L_{Amax}$ is the loudest instantaneous noise level. Again, this is usually the loudest 125 milliseconds measured during a given time block.
- $L_E$: The $L_E$ feature provides an assessment of impact sounds and blast noises, used for actions such as train passes. So that for a given number of passes an overall average can be calculated. It consists of the sound exposure level. The value represents the energy rate for the measurement range that is replaced by the energy value for one second. In other words, it is essentially a one second equivalent of the overall measurement.

- $L_y$: $L_y$ is the peak 'C' weighted sound pressure level used for occupational noise assessments to determine requirements for hearing protection.
- LN1-5: LN1 to LN5 consist of the percentile levels ($5^{th}$, $10^{th}$, $50^{th}$, $90^{th}$ and $95^{th}$). The most common ones to use are $10^{th}$ and $90^{th}$, referred to as LA10 and LA90.

A sample of this data captured is displayed in Table 1. The table shows a section of the data recorded for 5 of the features. The total dataset used for this research involves 2386 records or data with 10 features.

Table I. Sample Dataset (dB)

| Record | Start Time | Leq | LE | Lmax | Ly | LN1 |
|---|---|---|---|---|---|---|
| **1** | 11:31:19 | 69.0 | 93.8 | 79.8 | 101.8 | 77.1 |
| **2** | 11:36:19 | 69.9 | 94.7 | 92.0 | 111.3 | 78.6 |
| **3** | 11:41:19 | 69.3 | 94.1 | 84.3 | 100.1 | 82.3 |
| **4** | 11:46:19 | 69.9 | 94.7 | 82.2 | 99.3 | 78.0 |
| **5** | 11:51:19 | 68.8 | 93.6 | 81.4 | 103.9 | 77.8 |
| **6** | 11:56:19 | 67.7 | 92.5 | 83.6 | 102.2 | 76.3 |
| **7** | 12:01:19 | 80.0 | 104.8 | 102.1 | 110.8 | 93.1 |
| **8** | 12:06:19 | 71.3 | 96.1 | 85.2 | 100.8 | 82.6 |
| **9** | 12:11:19 | 70.8 | 95.6 | 82.6 | 100.0 | 80.0 |
| **10** | 12:16:19 | 67.5 | 92.3 | 83.6 | 95.8 | 74.4 |

## 3. APPROACH

Tinnitus sufferers often avoid going into public places as they fear their symptoms will be exacerbated by exposure to loud noises throughout the day [7]. The tool presented in this paper helps patients identify times of the day when it may be preferable to go outside. More specifically, it details areas they should avoid in order to ensure the symptoms are not triggered or made worse [8]. We also predict that the tool will be cost beneficial to governments. The aim is to help with the action on hearing loss by making people more aware of the damage which can be caused, unawares, in busy urban areas. The approach conducts the following steps.

(1) City noise collection as defined in section 2.4.
(2) Noise data classification and noise level prediction: The focus is on the identification of trends in data and a detection of when the most harmful levels occur so that the information can be overlaid on a map. First a visualisation of the entire data set is presented using a scatter graph format. The idea is to display trends in data in occur and that noise levels have characteristics at different times of day. Secondly a classification process forms a predictive model to identify the noise models anticipated at specific times of the day.
(3) Sound visualisation: The third stage involves the development of a visualisation tool which provides an interactive guide to Tinnitus sufferers. The tool details which areas of a city should be avoided and suggest what times of day they should best go out in public. This involves two stages:
    a. Develop map interface of city or urban environment.
    b. Overlay results in real-time in an understandable and interactive format.

### 3.1 Data Examination

In this section a visualisation of the noise collection, which took place in Leicester city centre (UK) on 21/08/2013 to 29/08/2013, is presented. The visualisations show a scatter plot of the decibel levels throughout the 2386 rows of raw data. The idea is to demonstrate that visual trends in the data occur at regular intervals. Figure 1 displays an overview of the maximum noises levels recorded at five minute intervals over a period of 9 days.

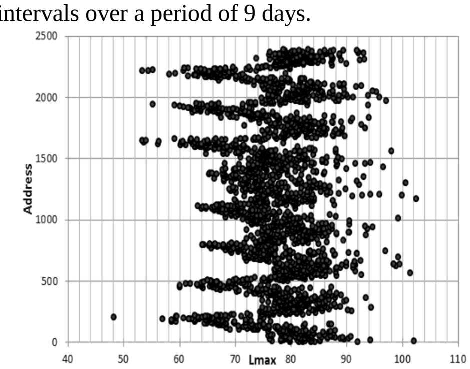


Fig. 1. $L_{max}$ Noise Levels

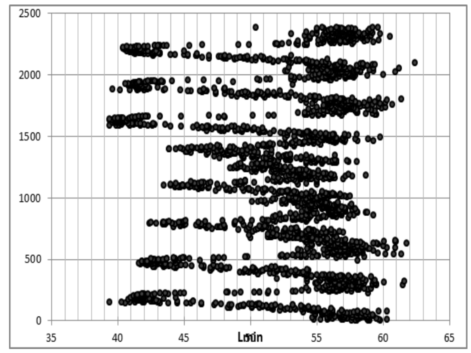


Fig. 2. $L_{min}$ Noise Levels

Noise levels are displayed along the x-axis while the reading labels are displayed along the y-axis. The numbers 0 to 2500 refer to each of the noise levels samples starting at 11:31:19 on Day 1 and ending at 18:16:19 on Day 9. The graph shows a clear repetitive trend in the data. Each-day has a similar spike in noise levels. This is caused by rush hour traffic at regular intervals. Similarly, in figure 2, a trend in the min noise levels over the 9 days is visible. There are identifiable times of day, which are regularly louder than other times of day. However, the main cluster on noise is around the 55dB to 60dB level.

The approach of this work lies in being able to predict nose levels and present the results in an interactive map. Increases in noise levels during rush hours are predictable, but the exact noise levels and the impacts on Tinnitus sufferers are unknown. Often subtle changes in noise levels can have a negative impact on the sufferer. This work also aims to present an approach for recognising the unexpected noise levels and identify if harmful noise levels occur outside of rush hour periods.

### 3.2 Case Study

Breaking down the graph into individual days we see identifiable trends in noise data. Figure 3 displays a scatter plot of the maximum noise level readings taken at a sampling rate of 5 minute intervals over 24 hours.

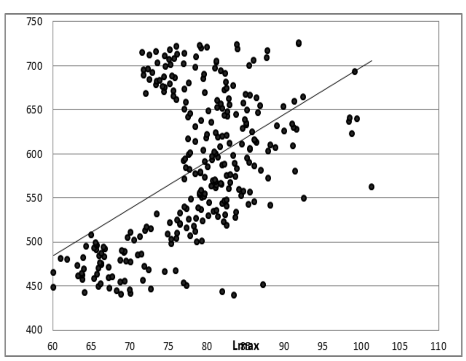


Fig. 3. $L_{max}$ 24 Noise Levels over 24 Hours

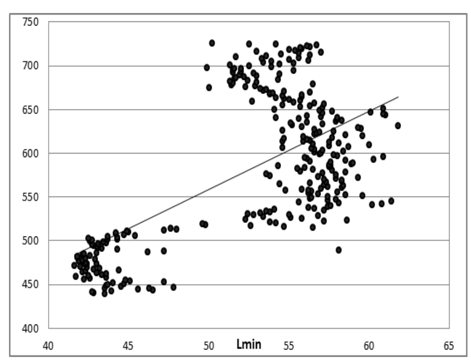


Fig. 4. $L_{min}$ Noise Levels over 24 Hours

Whereas, figure 4 displays a scatter plot of the minimum noise level readings taken, again at a sampling rate of 5 minute intervals over 24 hours.

A linear trend line is included in both Figures 3 and 4 to show the linear division of the data samples recorded during the 24 hour time blocks.

Figures 5 and 6 display the Max and Min noise levels over day 2. A similarity in the pattern of noise behaviour can be seen at comparable times of day.

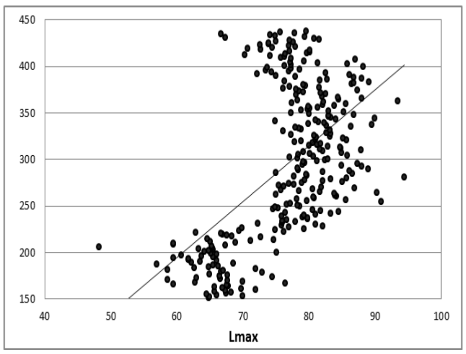


Fig. 5. $L_{max}$ Noise Levels Day 2

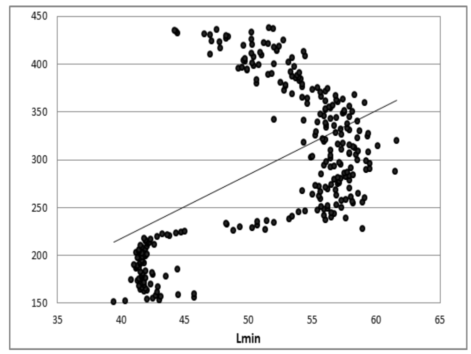


Fig. 6. $L_{min}$ Noise Levels Day 2

However, it can be a challenge to identify these patterns by eye when the data set is considerably large and containing data from different areas of the city. For the reason the use of data classification techniques is adopted to identify the trends in noise levels. The results are used for the visualisation process. The results are used to show a projection of noise levels at specific times of day and allocate a warning level so that the users will know when to avoid the area and how great the risk is to their hearing.

## 4. DATA CLASSIFICATION

The methodology involves employing a selection of supervised learning to predict trends in the noise levels. Specifically, the classifiers used include: Uncorrelated Normal Density based Classifier (UDC), Quadratic Discriminant Classifier (QDC), Linear Discriminant Classifier (LDC), Decision Tree (TREEC), and Parzen Classifier (PARZENC).

Linear Discriminant Classifier (LDC), is a technique which works by sorting or dividing data into groups based on characteristics to create a classification [11]. A discriminant function is obtained by monotonic transformation of posterior probabilities [12]. In other words, it performs an ordered transformation of unknown quantities, which are separated by a linear vector. Quadratic Discriminant Classifier (QDC) works in a similar way to LDC by dividing the data into groups based on given characteristics. However, by using QDC the data is divided using a quadratic surface rather than a one-dimensional one. QDC makes no assumptions that covariance are alike. In other words, it assumes that the changing of two random variables will not be the same [13].

Uncorrelated Normal Density based Classifier (UDC) also operates comparably to the QDC classifier but computation of a quadratic classifier, between the classes in the dataset, is done by assuming normal densities with uncorrelated features. Quadratic Bayes takes decisions by assuming different normal distribution of data [14]. LDC, QDC and UDC are density based classifiers. Decision Tree (TREEC) is a classifier which uses decision rules to divide the classes of data [12]. It operates by using criterion functions (the sum of squared errors), stopping rules (criteria for appropriate number of splits in a decision tree) or pruning techniques (the removal of unwanted tree sections).

Using decision tree is a particularly ideal choice of classifier because it is well-known as one of the most effective supervised classification techniques [13]. Parzen Classifier (PARZENC) functions by including aspects of the training data when the classifier is built up. It is a non-linear classifier and it has the benefit that its parameters can be user supplied or optimised [12, 14].

### 4.1 Data Processing

The entire dataset consists of 2385 data recordings sampled at 5 minute intervals over an eight and a half day period.

Table 2. Total Sample Dataset (dB)

| Recording time | Number of Hours | Number of days | Number of Recording |
|---|---|---|---|
| **Total recordings** | 204.5 | 8.52 | 2385 |

The raw dataset was divided into different hours of the day. In the analysis, rush hours are compared against off peak travel times to show how different noise levels can be identified. 3 days are randomly selected from the dataset for the comparison. The dataset was then divided into on-peak and off-peak hours. Table 3 details the number of total on-peak and off peak recordings over the three random days selected for a case study.

Table 3. On-Peak/Off-Peak Recordings

| Recording time | Number of Hours | Number of days | Number of Recording |
|---|---|---|---|
| **On-Peak** | 27 | 1.125 | 324 |
| **Off-Peak** | 45 | 1.875 | 540 |
| **Totals** | 72 | 3 | 864 |

The data set for on-peak and off-peak is unbalanced, as the number of on-peak hours in the day is lower than off-peak. In this research, off-peak is defined as between 10 pm – 6am and 9 am to 4 pm. On-Peak is defined as between 6 am 9 am and 4 pm to 7 pm. Using this model, a balanced dataset for the classification process is developed. Table 4 displays the number of samples used for the classification process.

Table 4. Noise Samples for Classification

| Recording Time | Total Hours | Time of recording | Number of Recordings |
|---|---|---|---|
| **On-Peak** | 9 | 3:00 – 5:59 | 108 |
| **Off-Peak** | 9 | 6:00 – 8:59 | 108 |
| **Total** | 18 | ------- | 216 |

### 4.2 Data Analysis

Similarly to figures 1 to 6, Figure 7 displays a scatter plot of the off-peak and on-peak noise levels in Leicester city centre for two features selected from the dataset.

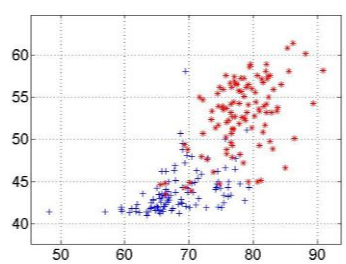

Fig. 7. Scatter Plot of the Off-Peak and On-Peak

Off-peak is represented by blue crosses, while On-Peak is displayed as red dots. There is a clear visible difference in behaviour when comparing usage levels. The visualisation relates to two features, $L_{max}$ and $L_{min}$ noise levels. The values along the x-axis (feature 1) denote the maximum noise level for the On-Peak usage. The values along the y-axis refer to the maximum noise levels for Off-Peak. Both features cover a three day period.

Visualisation techniques may not always provide a straightforward approach to distinguishing between types of behaviour. The difference in noise levels between on-peak and off-peak hours may not always be easy to identify. Often the changes in behaviour may be subtle and a challenge to detect. The visualisations displays that there is cross-over between 65dB and 80dB, for example. However small the differences, the effects are felt by tinnitus sufferers [17].

For that reason, the approach put forward in this paper involves a machine learning classification technique to detect subtle trends in data and predict when noise patterns occur. A mathematical comparison enables detection and predicative model for identifying trends in noise levels in real-time. The results of the classification techniques used are displayed in Table 5. Each of the experiments were conducted 30 times in order to account for errors. This also enables consistency in the results [15, 16].

Table 5. Classification Results

| Classifiers | AUC (%) | Sensitivity | Specificity | Error |
|---|---|---|---|---|
| **PolyC** | 94.44 | 0.9074 | 0.9259 | 0.0556 |
| **UDC** | 92.59 | 0.9259 | 0.9259 | 0.0741 |
| **KNNC** | 92.59 | 0.9259 | 0.9259 | 0.0741 |
| **ParzenC** | 92.59 | 0.9259 | 0.9259 | 0.0741 |
| **NaivebC** | 91.67 | 0.9074 | 0.9259 | 0.0833 |
| **LDC** | 90.74 | 0.9259 | 0.8889 | 0.0926 |
| **SVC** | 90.74 | 0.9074 | 0.9074 | 0.0926 |
| **TreeC** | 90.74 | 0.8704 | 0.9444 | 0.0926 |
| **QDC** | 89.81 | 0.9259 | 0.8704 | 0.1019 |

Overall, the classifiers were able to identify the difference between on-peak and off-peak hours which a high success rate. The results are provided by a confusion matrix evaluation. Table 6 displays a sample confusion matrix provided by the Naivebc classification results.

Table 6. Naivebc Confusion Matrix

| True Labels | Estimated Labels | | |
|---|---|---|---|
| | **1** | **2** | **Totals** |
| **1** | 49 | 5 | 54 |
| **2** | 9 | 45 | 54 |
| **Totals** | 58 | 50 | 108 |

The confusion matrix determines the distribution of errors across all classes. The estimate of the classifier is calculated as the trace of the matrix divided by the total number of entries. Additionally, a confusion matrix provides the point where miss-classification occurs. In other words, it shows true positive (TP), false positive (FP), true negative (TN) and false negative (FN) values. Diagonal elements show the performance of the classifier, while off diagonal presents errors. In this case the matrix shows that the classifier detects 49 of 54 off-peak sounds accurately. 45 of 54 on-peak sounds are correct.

Figures 8 to 10 display a histogram distribution of the classification performance results. As figure 8 displays, there is a divide in the classifiers' performance, however, most are able to perform consistently, with four able to achieve between 92% and 94.44%.

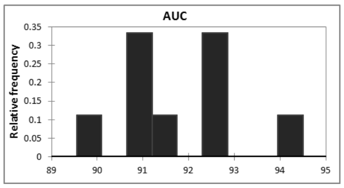


Fig. 8. AUC Histogram

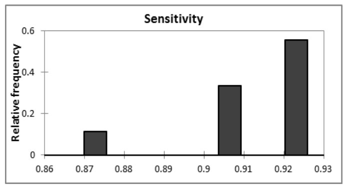


Fig. 9. Sensitivity Histogram

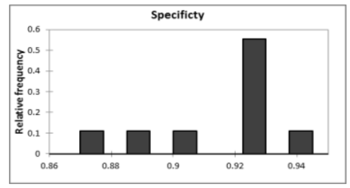


Fig. 10. Specificity Histogram

The results obtained support the findings that our methodology can be used to accurately classify and predict noise levels. In this case, off-peak and on-peak noise is identified accurately.

### 4.3 Discussion

In this section, a discussion on the classification results is presented. A visualisation of the High range (between 92.59 and 94.44%); the medium range (90.74% to 91.67%) and the low level range (89%) is presented. Figure 11 displays a visualisation of the results for the UDC classification (high level).

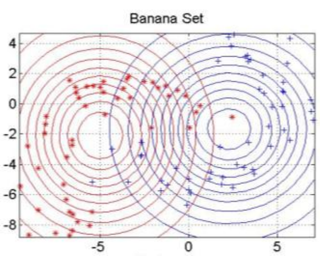


Fig. 11. UDC Visualisation

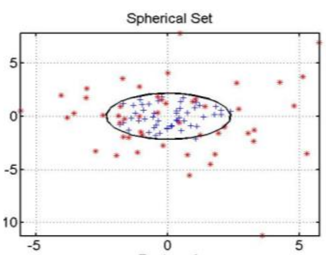


Fig. 12. UDC Sphere Plot


Fig. 13. Naivebc Banana Set Plot

As in figure 7, off-peak is represented by blue crosses, while on-peak is displayed as red dots. The data is in the form of a banana set and separated into groups by likelihood contours. The data inside the contours is most likely to belong to the group. Figure 12, again shows the UDC classification, in this case in the form of a spherical set plot. The graph shows how the classifier is able to accurately group 92.59% of the data accurately, and separates the on-peak from the off-peak noise. Both graphs display how the classifier is able to sort the data mathematically.

Figure 13 displays a visualisation of the naivebc classifier (mid-level). Contour-lines, formed by the naivebc process, encase groups of data, which are assed to belong to either the on-peak or off-peak grouping. As before, off-peak is represented by blue crosses, while on-peak is displayed as red dots and the scatter plot is in the form of a banana set. The graph displays a visualisation of the AUC of the classifier's performance. 91.67% of the data is grouped accurately, however misclassification can be identified by the red dots and blue crosses which are grouped incorrectly.

Figures 14 and 15 present a visualisation of the QDC classification process (low range). The linear and spherical line, generated by the QDC analysis, displays the division between the two sets of data.

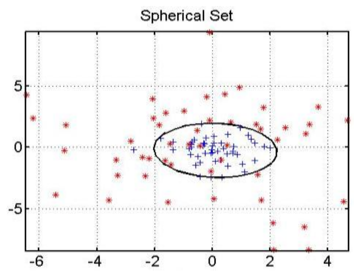


Fig. 14. QDC Spherical Plot

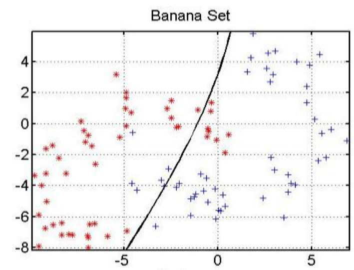


Fig. 15. QDC Linear Plot

As both graphs display, misclassification occurs and blue (off-peak) data can be seen outside of the spherical plot or on the incorrect side of the linear line (figure 15).

## 5. DATA VISUALISATION

The classification process offers the approach for noise level analysis. The visualisation employs the use of a game engine to visualise the results. By means of this technique, an interface, which can present a directional sound map for the target user, is devised. The results are fed into the visualisation software, which provides the user with an account of areas of a city to avoid at certain times of day. Figure 16 displays a map interface which is part of Project Vision Support (PVS) [10]. PVS is framework test bed used for the visualisation of real world environments by parsing data obtained from OpenStreetMap.com to procedurally generate virtual, navigable environments [10].

Fig. 16. Map Interface

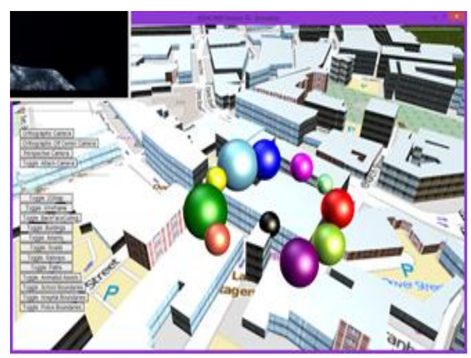

Fig. 17. Map Interface Leicester City Centre (UK)

The image shows a map with spheres relating to sound levels. Each of the spheres relates to one of the features from our dataset. The size of the sphere is denoted by the levels of sound types in the area visualised. Figure 17 displays a map interface of Leicester city centre (UK). The conception functions by incorporating real-world map data incorporating map, which is converted in the data manager. Features, initially sent to a temporary data store are used to create feature vectors and train classifiers for the sound prediction. The results are then projected into the map.

The system provides a predictive function to allow users to view expected noise levels in the area they wish to travel to.

## 6. CONCLUSION AND FUTURE WORK

This paper covers a classification technique for distinguishing between on-peak and off-peak hours. However, the techniques employed are a demonstration of how noise levels can be classified in real-time to present an analysis of where and when is the best time to travel or enter a certain area of the city. Tinnitus sufferers can simply choose to avoid leaving the house in peak times. However, our work provides a method for sufferers to be able to leave the house whenever they want by being able to identify in real-time when to avoid specific areas. In the future, we will extend this research to perform an evaluation of other areas such as motorways and different cities around the UK. We will use the data to both present visualisations of the noise patterns in cities and to show how noise pollution levels vary across different environments. We will incorporate further datasets such as holiday periods and weather patterns to create a more holistic noise level model. The future aim of the work will be to conduct a user case study assess the effectiveness of the app.

## ACKNOWLEDGMENTS

This research is co-supported by the NSF key project with grant no (61332017) and the National project with grant no (2013BAH24F00).

## REFERENCES

1. B. Fry, 2008. Visualizing Data, O'Reilly Media; 1 edition (28 Dec. 2007), ISBN-10: 0596514557.
2. W. Hurst., M. Merabti and P. Fergus, 2013. Behavioural Observation for Critical Infrastructure Security Support, 7th European Modelling Symposium on Mathematical Modelling and Computer Simulation, pp 36-41.
3. W. Hurst., M., Merabti, S. Iram and P. Fergus, 2014. Protecting Critical Infrastructures through Behavioural Observation. Inderscience International Journal of Critical Infrastructures, Vol. 10, No. 2, pp 174-192.
4. C. D. Mathers., D. Loncar, 2006. Updated projections of global mortality and burden of disease, 2002-2030: data sources, methods and results. WP on Evidence and Information for Policy World Health Organization.
5. Action on Hearing Loss, [www.actiononhearingloss.org.uk] retrieved May 2015.
6. British Tinnitus Association (BTA) 2015, [www.tinnitus.org.uk/, facts-and-figures]. retrieved May 2015.
7. S. Holmes and N. Padgham, 2009. More than ringing in the ears: a review of tinnitus and its psychosocial impact. Journal of Clinical Nursing, vol 18(21) pp 2927-37.
8. RNID, 2010, What's That Noise? A profile of personal and professional experience of tinnitus in Northern Ireland. RNID, Belfast.
9. J. Brophy, 2013. The Visual Representation of Sound for the Hearing Impaired, Senior Project Electrical Engineering Department California Polytechnic State University San Luis Obispo.
10. D. Tully, A. El Rhalibi, M. Merabti, Y. Shen and C. Carter, 2014, Game Based Decision Support System and Visualisation for Crisis Management and Response. 15th Annual Post-Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting pp 283-288.
11. E. Kuncheva, 2004. Combining Pattern Classifiers: Methods and Algorithms.
12. P. Fergus., P. Cheung ., A. Hussain., D. Al-Jumeily., C. Dobbins and S. Iram, 2013. Prediction of Preterm Deliveries from EHG Signals Using Machine Learning, PLoS One, vol. 8, no. 10, p. e77154.
13. R. P. Duin., P. Juszczak., P. Paclik., P. Pakalska., D. De Ridder., D. M. Tax, and S. Verzakov, 2007. A Matlab Toolbox for Pattern Recognition, Version 4. Delft Pattern Recognition Research.
14. F. Lotte, 2009. Study of Electroencephalographic Signal Processing and Classification Techniques towards the use of Brain-Computer Interfaces in Virtual Reality Applications.
15. N. J. Salkind, 2008. Statistics for people who (think they) hate statistics, Third Edition, Sage Publications.
16. I. Idowu, P. Fergus, A. Hussain, C. Dobbins and H. Al-Askar, Advance Artificial Neural Network Classification Techniques Using EHG for Detecting Preterm Births.
17. S, Figueira., K Nguyen., S Panditrao, 2014. HearThat? - An app for diagnosing hearing loss, Global Humanitarian Technology Conference (GHTC).
18. J. Lang, 1986. Assessment of Noise Impact on the Urban Environment, A Study on Noise Prediction Models, national Institute for Research on Heat and Noise Technology, Austria.
19. E, Lumnitzer., M. Behun., M. Bilova 2011. Implementation of method for dynamic noise visualisation into educational process, 9th International Conference on Emerging eLearning Technologies and Applications (ICETA), pp 17-19.