MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Big Data Analysis and Machine Learning Techniques to Answer New Demands on Long Term Monitoring Noise Analysis

Bañuelos-Arteagoitia, Oier[1]; Giraldo-Valencia, José Omar[1]; Bañuelos-Irusta, Alberto[1]; Gómez-Arteagoitia, Aritz[1]
AAC Centro de Acústica Aplicada S.L.
Leonardo da Vinci 14-7B, Parque Tecnológico de Álava; 01240 Vitoria-Gasteiz (Spain)

ABSTRACT

This paper presents the combination of different noise monitoring systems (conventional and low-cost), combined with data analysis procedures to manage the big amount of data coming from long term acoustic monitoring in real world cases. Signal processing and machine learning techniques are used to solve higher requirements in noise monitoring. Noise is getting a more relevant role in the environmental urban management and awareness is growing to include actions to reduce its annoyance, demanding a better description of noise problems and the identification of sources with their contribution to a specific situation. Direct measurements are difficult and expensive for many situations due to non-permanent sources in urban areas and conventional monitoring systems just giving noise levels information are not enough to identify the noise sources along time.
To respond to variable situations, flexible methods will be presented using environments that mix noise from recreational, transportation and other sources. Those sonic environments will be studied combining analysis of SPL data in short intervals with audio data processed through a pre-trained tuned model from Audioset and own recorded databases, trying ultimately to self-identify the contribution of specific noise sources related with annoyance in specific sounds environments.

Keywords: Noise, Monitoring, Machine learning
I-INCE Classification of Subject Number: 72, 74

1. INTRODUCTION

The analysis of urban noise levels have been carried out through the use of long term acoustic monitoring studies. Having in mind the uncertainties and fluctuations that exhibit noise levels in the city, a statistical approach had been the common technique to evaluate the results coming from monitoring stations. The first systems [1] were capable of computing probability distribution histograms of sound levels and its cumulative plots, which was useful to study phenomena occurring in short sampling times (15 min, 30 min, 2 hours) in part by the memory limit.

However, noise monitoring studies for general environmental purpose as urban noise monitoring, try to cover larger time scales and commonly run for several days and

[1]aac@aacacustica.com

even months to look for patterns at different time scales [2]. The strategy used to overcome memory limits have been the storage of noise descriptors in relative short periods, 10 to 15 minutes intervals, or frequently 1 hour intervals; such studies have found typical evolution patterns for the urban noise levels correlated with traffic peaks in the morning and in the evening, a steady state during the day and a decreased trend in the night [3].

It's also common to store percentile based descriptors as the environmental noise levels in a certain location can be decomposed in a mixture of distant or permanent sources correlated with lower percentiles(residual sound) and a mix of closer or sporadic sources correlated with higher percentiles(specific sounds) [4].

To describe specific events as train pass-by or an aircraft overflight, the monitoring system provides specific tools to analyse short events over a defined noise level, in order to have additional information for these events that could be relevant to describe the noise environment in a specific situation.

Nowadays, sound level meters have greater storage capacity which is useful to save high-resolution time history data and at the same time data and signal processing methodologies are offering new possibilities to go towards a new noise monitoring concept. During the last years AAC Centro de Acústica Aplicada has been working in the development of new methodologies to exploit the features of this new scenario.

The noise action plans in urban areas requires more detailed information to have a more detailed knowledge of the sources that contribute to the noise levels and to have enough data for taking decisions about possible solutions to control the noise source that are generating annoyance on the population.

As a result, we present some more detailed analysis based on $L_{Aeq,1s}$ levels and audio signal processing, that allow us to search for patterns in several time scales and detect specific sound events that could be related with annoyance.

2. METHODOLOGY

2.1 Data Processing with data science tools

Measurements results from sound level meters are commonly stored in csv or txt formats for which the default analysis tool is a spreadsheet. However, making measurements each second during relative long periods can generate considerable amount of information, moreover if there are various measurements locations, as can be seen on Table 1.

Table 1. Data size for different measuring periods with 1 second samples

Measurement Period	**Data size (only $L_{Aeq,1s}$)**
1 hour	3,600
1 day	86,400
1 week	604,800

Data sizes from typical long term monitoring scenarios are on the million scale of records. Pandas, a Python data analysis library can handle easily such data sizes without additional configuration as needed in Excel with the use of Power Pivot and Power Query. Using Pandas we load the entire measurement data, retrieve the information about size and data type present in the file with a few lines of code as shown on Figure 1.

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 601617 entries, 2018-09-12 12:30:09 to 2018-09-19 11:37:03
Data columns (total 9 columns):
Record #        601617 non-null int64
Record Type     2 non-null object
Date            601617 non-null datetime64[ns]
LAeq            601615 non-null float64
LAFmax          601615 non-null float64
LAFmin          601615 non-null float64
OVLD            601615 non-null object
OBA OVLD        601615 non-null object
Marker          0 non-null float64
dtypes: datetime64[ns](1), float64(4), int64(1), object(3)
```

Figure 1. Information about data loaded by pandas

Later, we assign as index of the data frame (tabular data structure in Pandas) the timestamp information. After that, we can slice through the measurement data specifying start and stop dates of the analysis period, making straightforward the study of different time scales (Figure. 2) in the data.

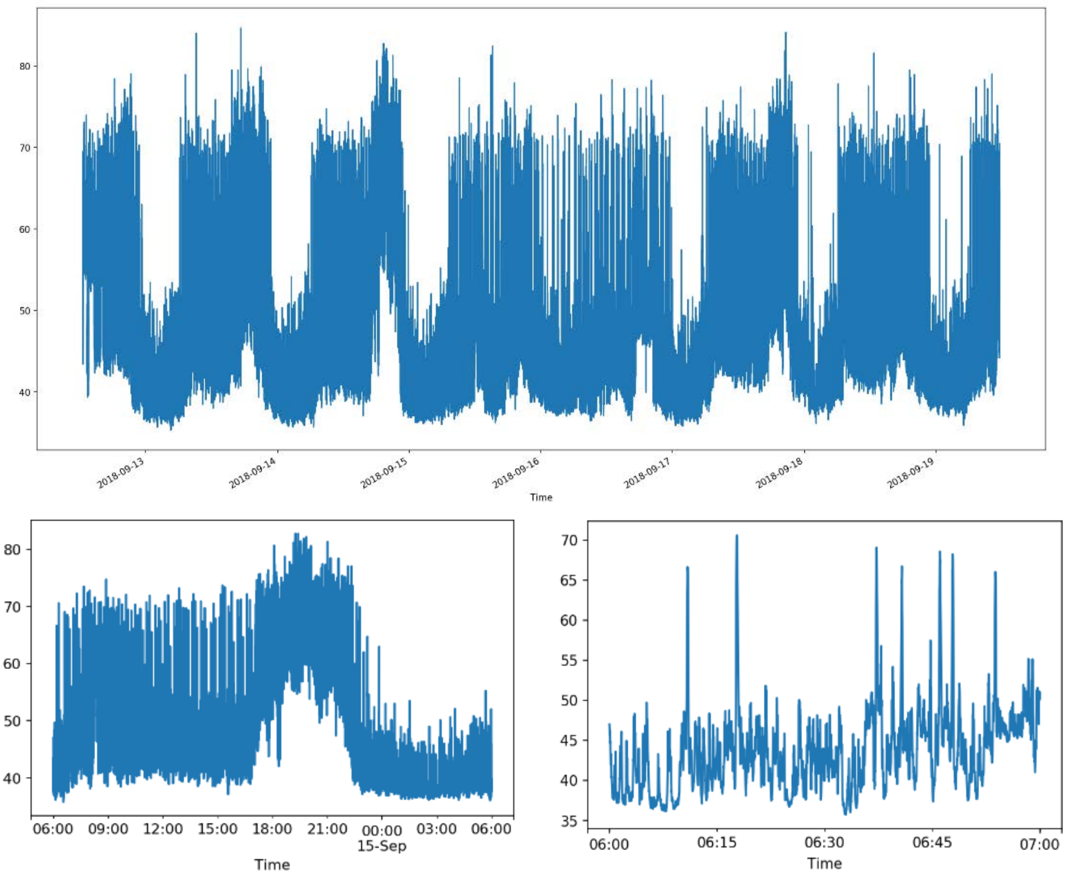

Figure. 2. Different time scale analysis, week, day and hour

One of the advantages of using $L_{Aeq,1s}$ data is that we can focus on specific sound events, for example by doing a peak detection on all the data and then performing a grouping of sound events by week, day or hour. This is useful specifically with irregular noise sources that can cause annoyance but could be masked on L_{Aeq} data from longer periods.

2.2 Linking audio and slm data

After detecting the prominent sound events from sound level meter (slm) data, the next step is to identify the type of noise source. This task is complex in long term studies as the presence of an annotator would be almost impossible and the identification by 1/3 octave levels is difficult in complex sonic environments. Audio data is the logic choice to perform this task, therefore, we must find a way to link audio with slm data in order to hear those sound events.

Our approach is to use low cost audio recorders, besides of the audio recording functionalities provided by some slm, continuous recording through long term periods (week, months) is rarely supported because of the data sizes (Table 2.).

Table 2. File size for different recording durations

Duration of recording	File size
1 Hour	317.5 MB
1 Day	7.62 GB
1 Week	53.34 GB

AAC's strategy synchronizes the information from audio recorders and slm measurements, using the cross-correlation function to find the delay between devices as shown on Figure 3. All of this processing is done in Python with the aid of scipy and numpy libraries.

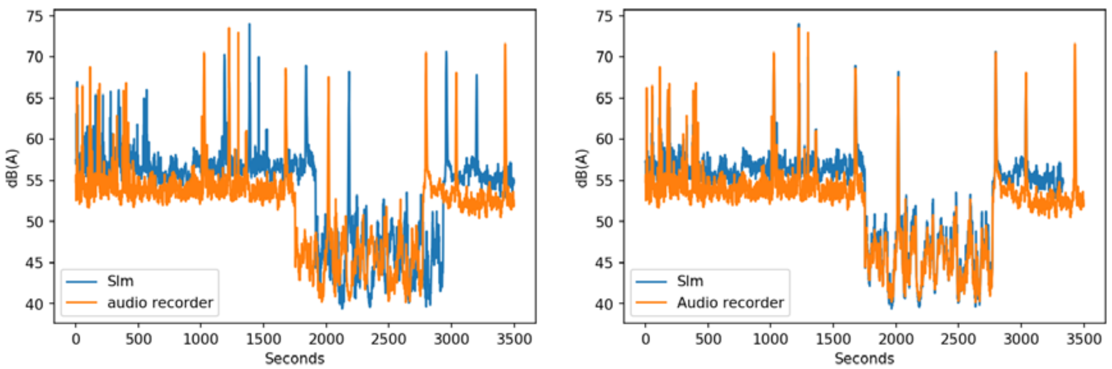

Figure. 3 Results before and after synchronization process

Having now the information synchronized, we extract small audio clips (5s to 10s) that correspond to the sound event timestamps retrieved from the time history analysis of $L_{Aeq,1s}$. We can now hear the sound event and get a sense of the annoyance caused by the noise source in a specific moment, however this process can generate thousands of small clips which are still a lot of information to be processed by a human. As a consequence, the next step is to use machine learning techniques to automatically classify the small audio clips.

2.3 Sound event classification using machine learning models

There are several studies on environmental sound event classification, ranging from those that use MFCC representation with machine learning classifiers, bag of words and lastly deep learning models. Currently the state of the art are CNN and RNN models that operate on MEL spectrogram data according to the DCASE challenge.

On this case we choose to use the Audioset model [5] called VGGish because it has been released with open source license and also because it’s trained on the biggest dataset [6] of environmental audio clips. The input for the Audioset model is a 96 frames x 64 bins matrix representing log MEL spectrograms of 0.96 seconds of audio. As can be seen on Figure 4, the final layer provides a 128 vector embedding that compress audio information and gives a useful feature to train custom audio classification models. Because of the lack of labelled and big datasets on the environmental audio domain we use the Audioset embeddings and then test on our owned small datasets that are carefully labelled. Hence, the small audio clips generated on the previous step could be classified on a pre-defined set of classes.

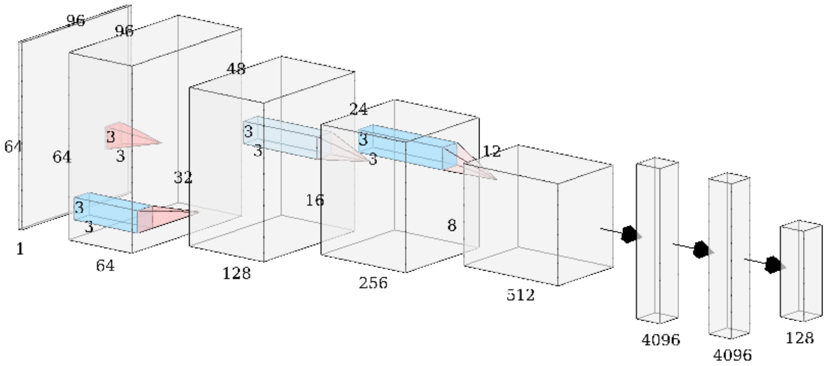

Figure 4. VGGish model architecture from Audioset (graph made using NN-SVG[7])

3. RESULTS FROM REAL MONITORING SCENARIOS

In this section we present some example results obtained in real world cases after having applied the methodology previously presented. As a first example, the peak detection results are illustrated on Figure. 5 for a 3 minute period on $L_{Aeq,1s}$ data. In this particular case a threshold of 70 dB(A) was chosen, however it should be carefully taken when choosing that parameter, as it is highly dependent of the situation under analysis. Another options are the use of adaptative thresholds that exploit statistical properties of the signal.

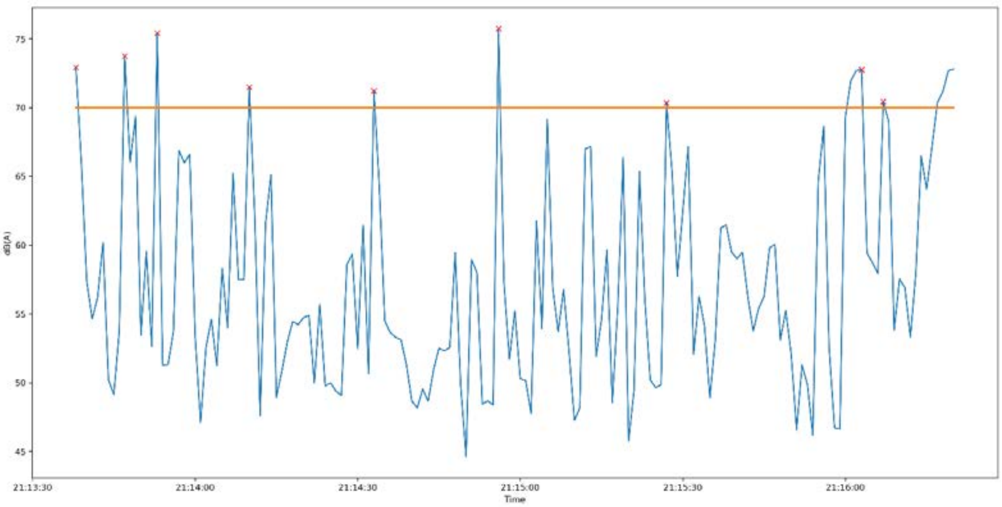

Figure. 5. Peaks detected on $L_{Aeq,1s}$ data.

The Figure. 6 shows the histograms of sound events grouped by hour on specific days of the measurement period. It can be seen that from 18:00 to 21:00 the number of events rises and it was effectively correlated with the activity of a skate park near to the residential area under study.

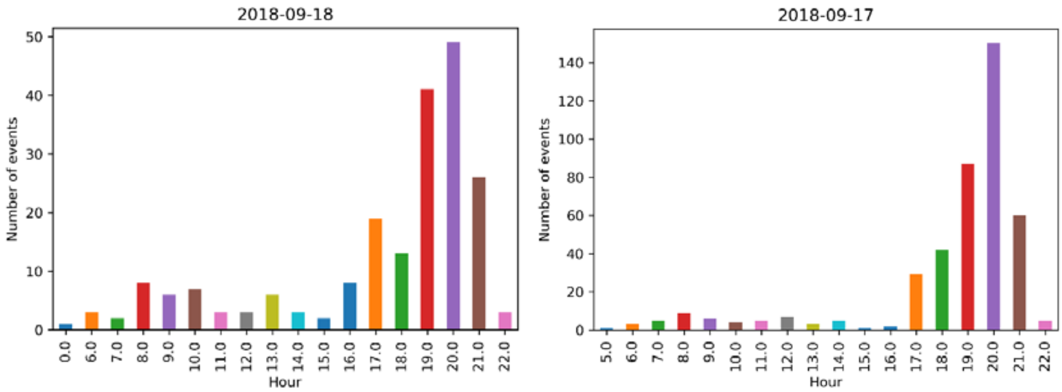

Figure. 6 Sound events on different days

A summary (Figure. 7) of the long term monitoring study can be presented with the aid the histograms that allow to see patterns of sound events grouped by hours and days, or a different window analysis relevant to the study.

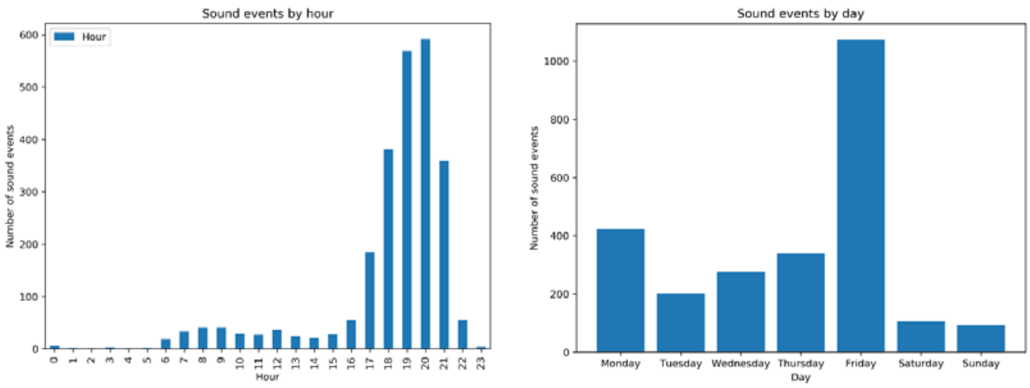

Figure 7. Example of summary results obtained

After the extraction of the small audio clips related with the sound events of slm data, we noticed that there were some train pass-by sounds mixed with skate park sounds. Then, we train an audio classifier on top of the embeddings from the VGGish model. The Figure 8 shows how the embedding compress around 1 second of audio (64x96 MEL spectrogram) into a 128 vector. Audioset dataset has 527 audio classes, it’s useful to reduce the training to a few of sound classes that are likely to be present on the recordings in order to get a better classification performance.

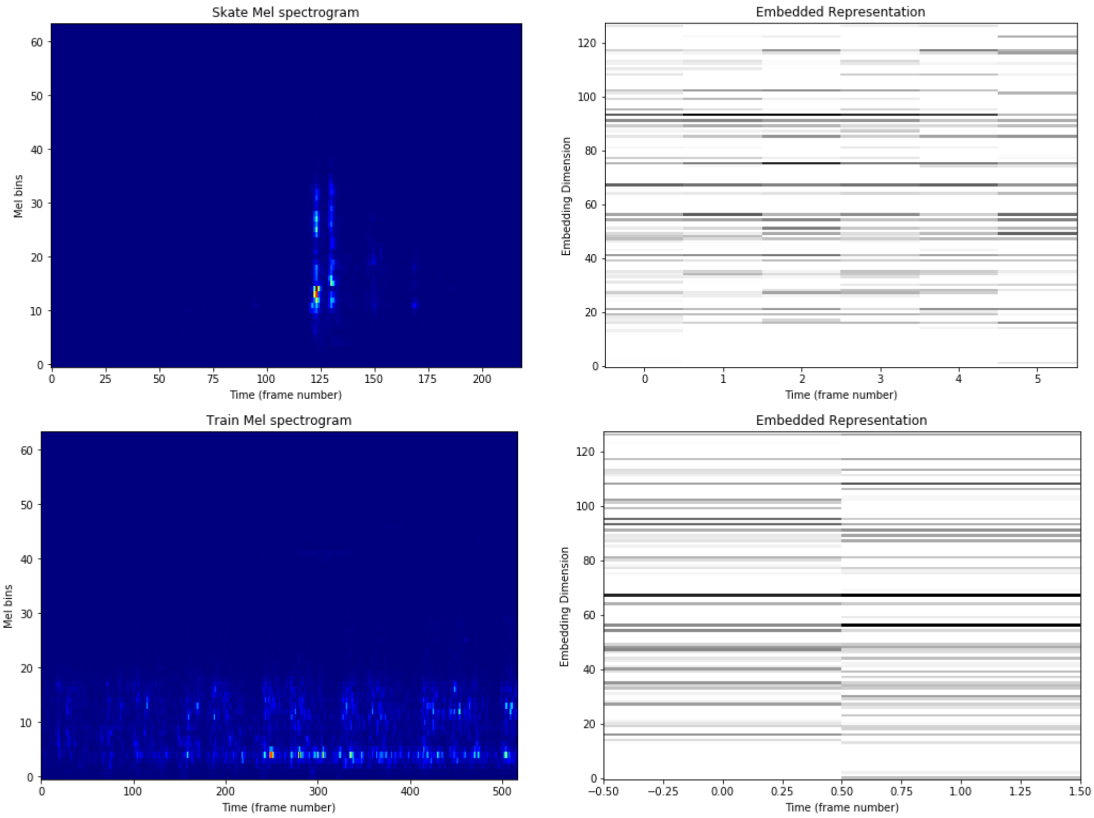

Figure 8. MEL spectrograms and embedded representation of skate and train sounds.

4. CONCLUSIONS

The methodologies developed by AAC for data and signal processing are able to offer new interesting opportunities to get information that will help to solve many noise situations complex to solve with conventional methods.

Long term monitoring using intervals for data collection each second requires the use of methodologies based on specific processing methods able to deal with really high volume of data, able also to include sound event detection processes to identify specific sources that generates annoyance or to assess the specific impact of those events on the long term noise levels.

Noise level data is not enough to describe many situations and audio recordings must support the analysis. For long term monitoring, it is needed to have synchronization methods that could manage the small deviations in time for different signals.

Complementary, the audio recorder itself is a powerful tool to assess annoyance and perform source identification that could be a great value for the monitoring system, even using low-cost solutions. To do that, methods based on machine learning techniques are needed, and the work already developed by AAC evidence the high possibility to include them in daily acoustics studies to have a practical response to their analysis and diagnostic.

Of course the R&D in this field is opening new features that should be developed in the near future that will contribute a new way to deal with practical and effective monitoring methodologies.

5. ACKNOWLEDGEMENTS

We acknowledge to CDTI (Spanish Government) to finance the research project ARZA-TIK where AAC developed the methodologies of the studies that support this paper.

6. REFERENCES

1. K. E. Kittelsen and C. Poulsen. Vér, “Statistical Analysis of Sound Levels”, in “Technical review No. 1” edited by Bruel & Kjaer, Denmark (1964)

2. D Geraghty a , Margaret O’Mahony, “Investigating the temporal variability of noise in an urban environment” in “International Journal of Sustainable Built Environment (2016) 5, 34–45

3. J. Song and V. Lenchine, “Statistical properties of urban noise – results of a long term monitoring program” in “Proceedings of ACOUSTICS 2017”, Perth, Australia(2017)

4. ISO 1996-1:2016 and ISO 1996-2:2017

5. S. Hershey et. al,”CNN Architectures for Large-Scale Audio Classification” (https://research.google.com/pubs/pub45611.html), ICASSP 2017

6. Gemmeke, J. et. al., “AudioSet: An ontology and human-labelled dataset for audio events” (https://research.google.com/pubs/pub45857.html), ICASSP 2017

7. LeNail, (2019). “NN-SVG: Publication-Ready Neural Network Architecture Schematics” In “Journal of Open Source Software”, 4(33), 747, https://doi.org/10.21105/joss.00747