CONSTRUCTION SITE ENVIRONMENTAL POLLUTION MANAGEMENT – INITIATIVES AND INNOVATION IN ASIA

Ting Cai Co-founder of Novox Limited, Hong Kong
SiewSun Chan Regional sales manager, Acoem, Malaysia
Jean-Florent Cros Product manager, Acoem, Limonest, France
Julien Preuilh Data science engineer, Acoem, Limonest, France

1 INTRODUCTION

This paper describes how recent innovations using AI can support site managers in complying with legislation and limiting the impact caused by excessive noise. It's a fact that noise pollution is one of the major issues of this century, affecting every region of the world. Today's noisy environments are many and varied, including road traffic, aircraft, industry, sporting and cultural events and construction activities.

Prolonged exposure to high levels of noise can have harmful effects on human health, wildlife and the environment in general. This can lead to health problems such as stress, sleep disorders, hearing loss, cardiovascular problems and poor concentration.

A number of measures are being put in place to combat noise pollution, such as soundproofing buildings, regulating vehicle noise levels, urban planning to reduce noise sources, and raising public awareness of the effects of noise pollution. It is essential to promote policies and practices that encourage a healthier noise environment and to develop quieter technologies to reduce noise levels.

Monitoring site noise is essential for preventing noise nuisance, ensuring the safety of workers and managing complaints from residents. This helps to create a healthier and more respectful working environment. For this reason, site managers need tools that enable them to manage excess noise on their sites, including the date on which it occurred, and to identify the reason(s) for the noise levels being exceeded.

This document describes an acoustic monitoring solution developed by ACOEM France. This system triggers an alert corresponding to an excess of noise (exceeding a sound level) with the following information: the date, time and direction of the event. This system has been installed initially to monitor a construction site since April 2023 under the supervision of Novox Limited, a partner of ACOEM France. Following this first installation, other system was deployed in Europe.

The objective of this installation is to be able to enrich a sufficiently representative database of audios to be able to train a model based on artificial intelligence capable of discriminating between construction noise and any other environmental noise.

2 CONSTRUCTION NOISE MONITORING

2.1 Architecture solution

This section describes the various components of the system used to monitor the worksite. For the remainder of this document, the term "alert" will be used when a noise level threshold is exceeded. The architecture of the solution is made up of several devices shown in Figure 1:

- A class 1 FUSION sound level meter, whose role is to certify the sound level recorded in the field;
- An ATD-300 sensor comprising 4 microphones and with powerful on edge computation capability. This sensor continuously records the sound environment around it and triggers an alert whenever the sound level exceeds a predefined threshold. It then determines a direction

and, in a future version of the system, will be able to identify the nature of the source using artificial intelligence algorithms.
- A 360° camera or PTZ (Pan Tilt Zoom) camera to provide additional information to the alert to help better identify the alert in addition to the audio.
- A NAS (Network Attached Storage) server to centralise the various alert data (audio, sound level, location results, photos, etc.) from the various devices.
- A Cadence environment for storing data and displaying alerts in real time.

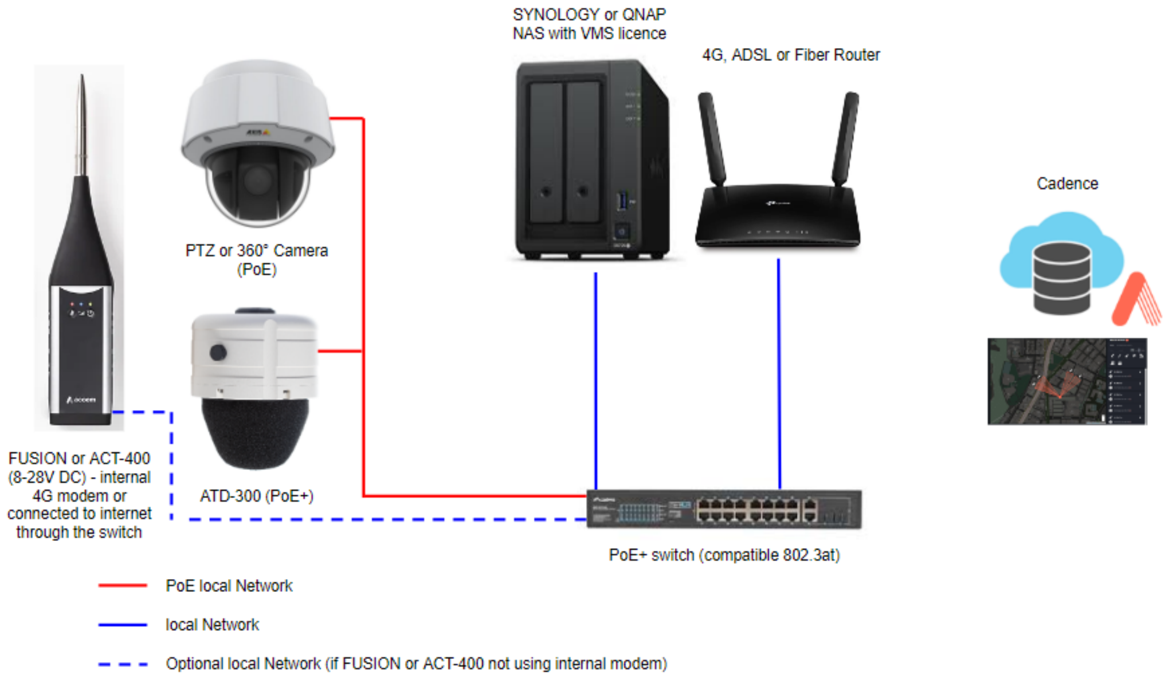

Figure 1: architecture solution

2.2 Construction site

An initial installation of the complete system was carried out in April 2023 on a construction site in Hong Kong. This trial demonstrated the difficulties that can be encountered in an urban environment. The emergence of a sound level can be characterized by a complex mixture of several different sources. This can compromise the main objective, which is to determine whether the construction site is responsible for the noise emission. Added to this are the various forms of reverberation which will tend to destabilize the source location algorithm.

The trigger threshold was set at 85 dBA. Over a 4-month period, there were more than 10,000 alerts, i.e. an average of around one hundred detections per day. Most of these alerts were triggered during the site's working hours. Given that each detection is associated with a direction, analysis shows that around 80% of alerts come from the site area. At present, there is nothing to indicate whether the calculated location is correct, so a detailed analysis must be carried out on each audio in order to analyses and label each of the alerts. So, in this document, we only talk about trends.

It is therefore essential to set up a tool for viewing and retrieving the various data in order to be able to confirm whether or not the alert originates from the worksite, by listening to the audio and images from the camera, and at best to determine the source responsible.

A representation of the site plan is shown in Figure 2, and a photograph of the complete system on the same site in Figure 3.

Figure 2: Aerial plan of the construction site

Figure 3: complete system photography

2.3 An example of an alert

An alert is triggered by the ATD-300 device which, in addition to providing an instant of detection, determines the direction of the event and sends all this information back to the Cadence platform. A representation of an alert on Cadence is shown in Figure 4 below.

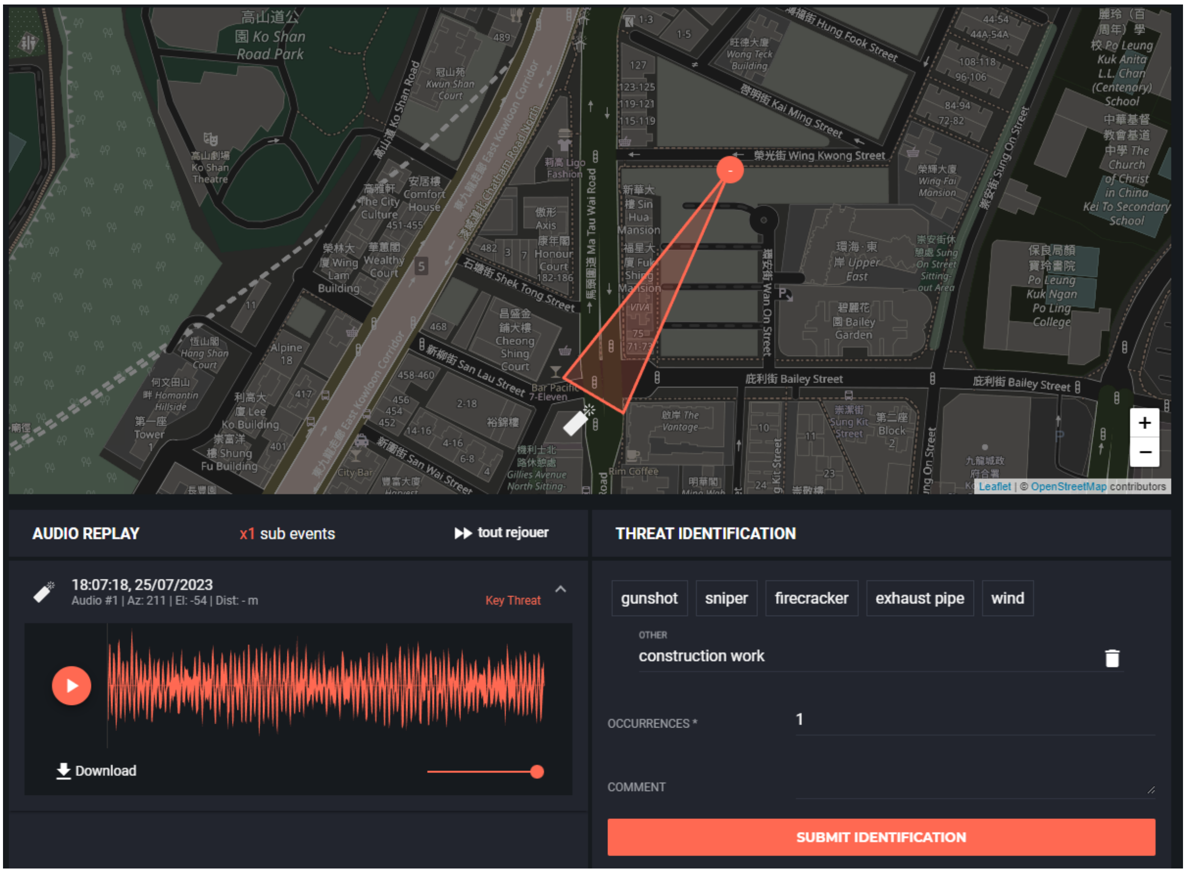

Figure 4: example of an alert

There are three information panels concerning the alert:

- a map of the site with information on the direction of the alert (represented by the orange triangle). In this way, it is possible to check that the direction of the alert is observed if the threshold exceedance is due to a source originating from the site area or not.
- an interface for viewing, listening to and downloading the audio corresponding to the alert, enabling the user to analyses the alert and determine the source responsible for the alert.
- an identification field enabling the user to enter, if they have the right to do so, the identification of the alert, in this case "construction work". To do this, they will also analyze the camera images, which will help them to identify the alert more correctly.

3 DETECT AND LOCALISE

Detection is defined as exceeding a sound level. In the system, detection is performed by the ATD-300. Acquisition is performed on the device's 4 microphones with a sampling frequency of 44100 Hz. An analysis is performed for each 1s frame. The analysis consists of performing an A-weighting filter on the frame, then determining the LAeq value, and generating an alarm if this value is above a predefined threshold.

The A-weighting filter reproduces the human ear's perception of noise. The transfer function of the applied filter is given by the following formula:

$$H_A(f) = \frac{7.39705 \times 10^9 \times f^4}{(f + 129.4)^2 (f + 676.7)(f + 4636)(f + 76655)^2}$$

A representation of the filter gain is shown in Figure 5:

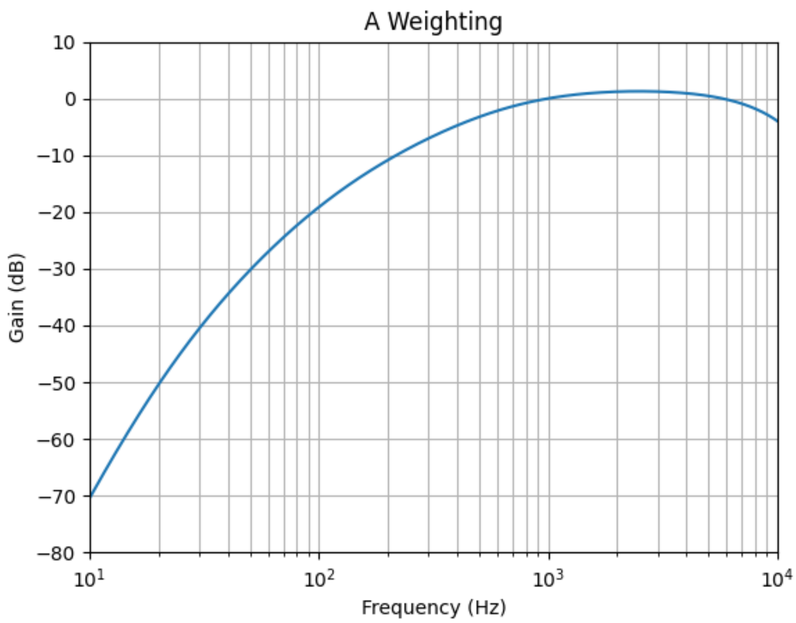

Figure 5: A-weighted filter

The value of LAeq(T) is defined as the average energy received over the period T. In the study, T=1s.

Considering the frame signal of size N samples: $\{x[i]\}_{1 \le i \le N}$, the signal power is defined by the formula:

$$P[i] = x[i]x^*[i] = |x[i]|^2$$

with $x^*[i]$ conjugate of a vector $x[i]$

The formula used to calculate the $LAeq(T)$ is defined by:

$$LAeq(T) = 10\log_{10}\left(\frac{1}{T}\sum_{i=0}^{T}\frac{P^2[i]}{P_0}\right)$$

with $P[i]$ the instant sound pressure and $P[0]$ the equal reference pressure $P_0 = 2 \times 10^{-5}\ Pa$.

For each alert detected, a localisation is performed to determine a direction in relation to the ATD-300 device. Localisation is based on the beamforming technique, which makes it possible to determine the direction of the emergent sound source. This technique is widely used to locate and isolate sound sources in noisy environments, thus improving the detection and analysis of acoustic signals in a given direction. For this reason, this method will use the 4 microphones of the device and the result will give a result in azimuth and elevation.

Initially, the expected localisation performance is based on the case where an alert triggered by the construction site gives a localisation in the direction of the construction site.

4 IDENTIFY THE SOURCE

In this section, we will present the theories that will be used to identify the sound sources that trigger alerts when thresholds are exceeded. We have two objectives: firstly, to be able to recognize construction site noise among other classes of noise. The second is to distinguish the type of construction site event (alarm, rotating machine, impulse machine, etc.). These two objectives are complementary, i.e., we can first detect construction site noise and then identify the noise category.

In this paper, no results will be presented on the identification of construction site noise, because initially, the objective of these initiative in Asia and Europe is to collect sufficient audio data sample (duration of 1,5s every minutes triggered on high noise level to exclude all personal data). These audios will be labelled and prepared in databases, which will then be used to create models using artificial intelligence algorithms such as machine learning or deep learning.

However, a preliminary study on the recognition of sound sources has been carried out by Acoem France. The noise sources studied are represented by taxonomy in Figure 6 below. These sources include classes of sounds associated with construction sites: "construction machinery", "continuous construction sites" and "impulse construction sites".

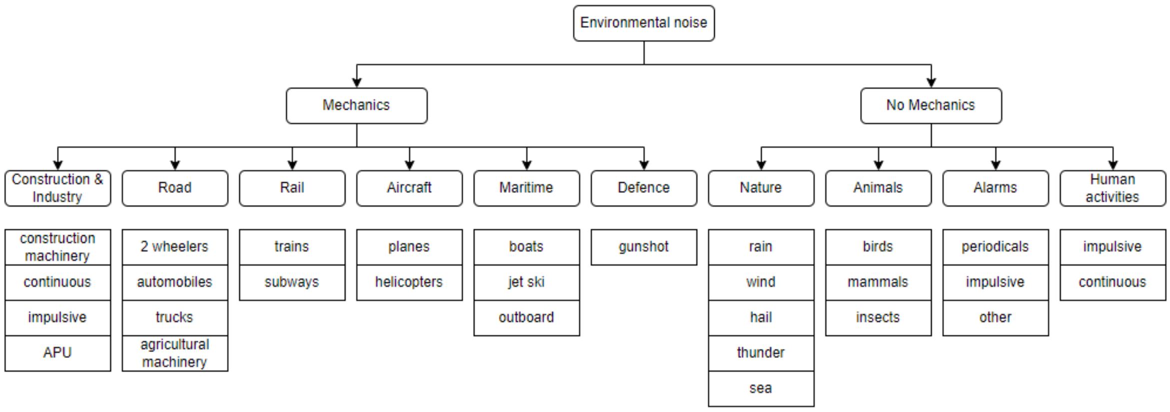

Figure 6: Taxonomy of the 29 classes of environmental noise

Recognition of environmental noise is achieved using classifier-type models previously learned from a large amount of data and using machine learning algorithms to identify and distinguish between sounds. The general approach to creating such a model is shown in Figure 7 below.

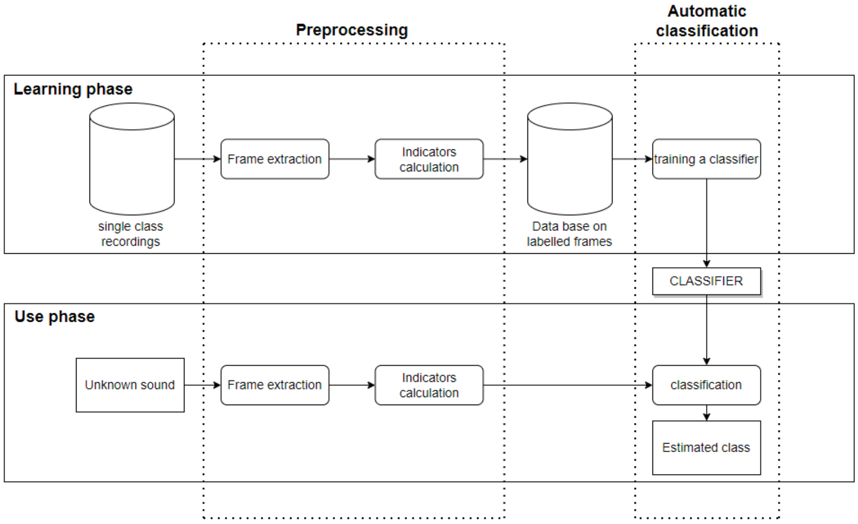

Figure 7: Overview of the environmental noise classification project

This structure is divided into two phases: a learning phase and a use phase.

In the first phase, we must first collect single-class sound recordings to establish a database of sounds categorised into the 29 sound classes shown in Figure 6. This data will be pre-processed, i.e., the maximum amount of information relevant to the sound classification will be extracted from the signal. In our case, we will first extract the signal frames. Not everything in an audio signal can be used; only a few parts of the audio actually contain the sound information. For example, in the case of a jackhammer, what really characterises the tool are the impulsive blows of the impacts as it operates. Once the frames have been extracted, each one is summarised by a vector of indicators. The indicators used come mainly from signal processing theory[1], and include temporal and frequency indicators such as skewness, octave band energies, spectral centroid and others of a similar nature. The use of indicator vectors instead of the signal directly allows us to control the information retained from the signal, we extract relevant and non-redundant indicators[2] in order to better control the inputs to the classifier, we thus bring acoustic knowledge to our model, which can be interpreted as an aid to classification. On the other hand, there is a disadvantage in terms of calculation time: the indicators have to be calculated for each extracted frame, which can lead to conflicts when controlling a system in real time.

Once the base of labelled frames is ready, the classifier can be created. To do this, we will use a neural network model whose input will represent the pre-calculated indicators and whose output will represent the probabilities of belonging to the 29 proposed sound classes, with the highest probability defining the audio class. In this project, because of the large number of sound classes and their heterogeneity, several neural networks will be used, forming a hierarchical structure like a decision tree. An example of architecture is proposed in Figure 8:

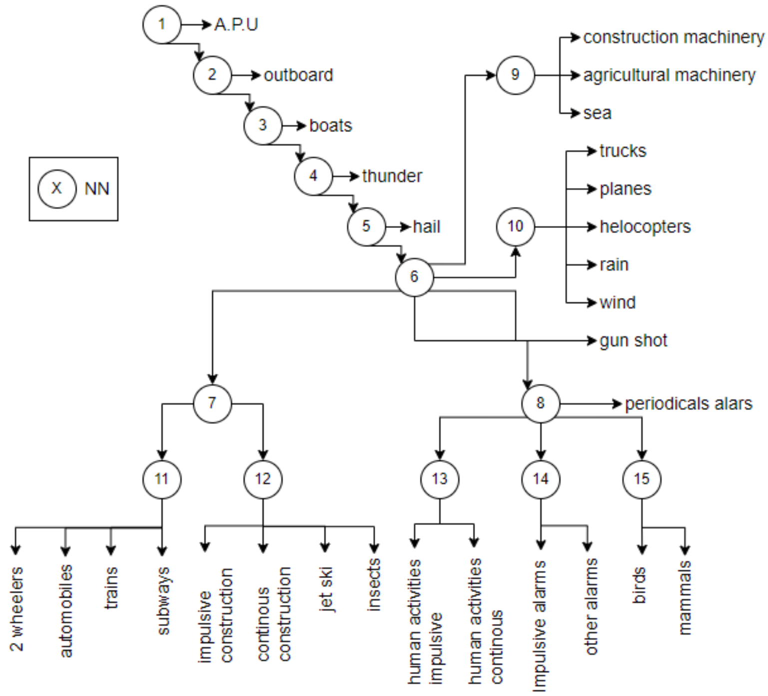

Figure 8: hierarchical classification

We note that to classify and distinguish the 29 classes of noise, it is necessary to use 15 neural networks. The placement of these networks is established by statistically comparing the separation difficulty of each class. Thus, the easiest classes to distinguish appear in the first networks, and the most complicated in the last networks.

In the second phase, once the classifier is ready for use, it is used on unknown data, in other words data that the classifier has not seen in training, and the different performances of the complete system are observed.

The results obtained are very encouraging, and some classification results are shown below. The performance represents the rate of correct identification of the sound class.

Best performances		Worst performances	
Classes	**Performance**	**Classes**	**Performance**
Periodic alarms	99.4 %	2 wheels	76.1 %
Gunshot	99.2 %	Insects	77.3 %
Thunder	98.0 %	Human activity	78.9 %
A.P.U	97.8 %	Mammals	80.7 %
Boats	97.7 %	Train	82.2 %
Hail	97.7 %	Helicopters	82.8 %

For the sound classes relating to construction noise, the results obtained are presented below:

Classes	Performance
Construction machinery	90.1 %
Continuous construction sites	83.0 %
Impulse construction sites	95.3 %

5 CONCLUSION

In conclusion, the monitoring and control of noise pollution on construction sites is an important issue that needs to be brought under control. The initiative carried out in Asia and Europe have made it possible to set up an autonomous system for collecting a large quantity of audio data characterizing the construction site and its surroundings, and also to observe the problems encountered due to the urban environment, which is a complex environment due to the large number of sound sources. The data collected will be used to set up a classifier to detect noise from the site and identify the sound source responsible for exceeding the threshold. In view of the initial studies carried out in the field of sound source recognition, the use of indicators and neural networks may represent a very suitable classification solution. The prospects for the current system are therefore the implementation of a sound source classifier to help identify alerts, ensuring good quality monitoring of worksite areas.

6 REFERENCES

1. Ben. D. Fulcher, highly comparative time-series analysis: The empirical structure of time series and their methods. Department of Physics, University of Oxford (2013)
2. Patent: Method and system for the automated creation of recognition models of content, publication: 2923043, inventor: DEFREVILLE BORIS (2007)