TRAINING AN ANOMALOUS NOISE EVENT DETECTION ALGORITHM FOR DYNAMIC ROAD TRAFFIC NOISE MAPPING: ENVIRONMENTAL NOISE RECORDING CAMPAIGN

Referencia PACS: 43.50.Rq,

Francesc Alías[1], Joan Claudi Socoró[1], Xavier Sevillano[1], Luca Nencini[2]
[1]GTM - Grup de Recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull
Quatre Camins 30, 08022 Barcelona, Spain, Tel. +34932902440, E-mail: falias@salleurl.edu; jclaudi@salleurl.edu ; xavis@salleurl.edu
[2]Bluewave Acoustics, E-mail: nencini@blue-wave.com

ABSTRACT
The LIFE+ DYNAMAP project aims at creating dynamic road traffic noise maps automatically upon the levels measured by a low cost sensors network. To ensure these maps reflect the acoustic impact of road infrastructures, it is necessary to exclude other acoustic sources (e.g. aircrafts, industries, etc.) from the noise level computation. To that end, an anomalous noise event detection algorithm based on supervised learning techniques is designed. This paper describes the recording campaign conducted in the two pilot areas of the DYNAMAP project in Rome and Milan, aimed at collecting road traffic and anomalous noise event samples to train the algorithm.

INTRODUCTION
Given the negative effects of continued exposure to high traffic noise levels [1], European authorities have driven several initiatives to study, prevent and reduce the effects of exposure of population to traffic noise. An example is the European Noise Directive (END 2002/49/EC), focused on the creation of noise level maps [2]. In order to take into account changes in environmental conditions, these maps have to be updated with a 5-year periodicity, which entails a time and cost consuming process that is undertaken by local and regional bodies of government.

In an attempt to simplify and reduce the cost of updating road traffic noise maps, the LIFE+ DYNAMAP project (http://www.life-dynamap.eu/) aims at automating the whole process. To that end, the main goal of the project is to develop a network of low cost acoustic sensors and an integrated system for data acquisition and processing able to detect, report and map the acoustic impact caused by road infrastructures in real time.

To validate the proposed approach, during the LIFE+ DYNAMAP project two demonstrative pilot systems will be implemented and tested for at least one year in the cities of Milan and Rome. The first one will be located inside the city of Milan, thus allowing to test the system in an urban scenario, while the second one will be located along a major road surrounding the city of Rome (the A90 highway), making it possible to validate the performance of the system in a suburban environment. More specifically, in previous stages of the project, a support tool was designed to choose the specific pilot areas where the demonstrative systems will be implemented [3]. Taking into account several environmental and infrastructural factors (e.g. noise levels, population density, number of dwellings, etc.), the candidate areas were ranked based on scores dependent on those attributes, selecting Milan’s district 9 as the urban pilot area [4], and a total of 17 critical areas located along the A90 highway in Rome for the suburban setting [3].

However, the automation of the road traffic noise data gathering and analysis processes entails several consequences. One of them has to do with the fact that acoustic events produced by nontraffic sources that could alter the measured noise levels (e.g. an air-craft flying over, nearby industries or railways, road works, church bells, animals, etc.) should be detected and eliminated from the noise map computation to provide a reliable picture of the actual road traffic impact. For this reason, it is necessary to devise strategies to automatically identify anomalous noise events captured by the network of sensors.

To that end, the LIFE+ DYNAMAP project includes the development of an Anomalous Noise Event Detection (ANED) algorithm to detect such events and ensure that the noise levels represented on the dynamic maps only reflect road traffic noise. In its current version, the ANED algorithm follows a semi-supervised machine learning approach that requires the construction of a reliable acoustic model of road traffic noise [5].

Given the diversity of operating scenarios (i.e. urban and suburban), it is necessary to build acoustic models that faithfully reflect the characteristics of road traffic noise in both types of settings. For this reason, an environmental noise recording campaign was conducted on the pilot areas where the two demonstrative versions of the DYNAMAP system will be implemented.

As a result of the recording campaign, nearly 10 hours of audio were collected, labeled and processed to train the mentioned acoustic model for subsequent development stages of the ANED algorithm. The primary goal of this paper is to describe all the procedures related to the recording campaign in the following terms. First, section 2 presents the main technical aspects of the recording campaign. Next, section 3 is devoted to the presentation of the audio post-processing tasks conducted on the recorded audio. Finally, section 4 outlines the conclusions and future steps towards the implementation of the ANED algorithm.

RECORDING CAMPAIGN

The main goal of the recording campaign was to collect widely diverse road traffic noise samples in their actual environment conditions to train, validate and test the current version of the ANED algorithm (see [5] for more details). To that effect, several recordings were conducted between the 18th and 21st of May 2015 in specific locations of the two pilot areas of the LIFE+ DYNAMAP project. The selection of these locations was based on obtaining representative samples of the traffic conditions and acoustic characteristics of the pilot areas.

Moreover, it is worth noting that the monitoring network of sensors of the LIFE+ DYNAMAP system will be composed of low cost acoustic sensors. The recordings were conducted with two measuring devices simultaneously: one low cost sensor from Bluewave [6] connected to a ZOOM H4n digital recorder, (see Figure 1.b) and a Bruel&Kjaer 2250 sonometer (see Figure 1.a), used as a reference. These dual recordings were conducted to allow the validation of the low cost acoustic sensor performance with respect to the sonometer in the near future. The recording setup was the following:

- Situation of both measuring devices: 50 cm distance between them.
- Sampling: 48KHz sampling rate with 24 bits/sample,
- Sensitivity verification using a 94 dB_{SPL},1 KHz calibration tone.
- Clapping: in order to align the audio recordings from both measuring devices, a sequence of 5 sec. of clapping was performed between both sensors with a separation that assured a very good signal to noise ratio despite the environmental noise.
- Gain adjustment: the input gain of each recorder was selected to guarantee enough room for in site audio dynamics (no saturation).
- Installation: both recording systems were installed on a tripod and included a windscreen to protect the sensor from wind.

- Orientation: the final orientation of the DYNAMAP low cost sensors with respect to the traffic flow is still undefined. For this reason, recordings were made with three orientations: putting the sensor in the direction of the traffic –forward orientation–, in the opposite direction –backward–, or orthogonal to the vehicles flow. Moreover, three elevation angles of the sensors positions were also employed: 0º, 45º and -45º.

(a) Bruel&Kjaer 2250 sonometer.　　(b) Low-cost measuring device.

Figure 1 – Recording equipment.

During the 18th and 19th of May 2015, recordings were conducted in 6 sites along the A90 highway in Rome (see Figure 2), which constituted a representative subset of the 17 sites in this pilot area according to the following four classes [3]: single road; additional crossing or parallel roads; railway lines running parallel or crossing the A90 motorway; and a complex scenario including multiple connections. Concretely, the recording equipment were installed in 6 highway portals owned by ANAS S.p.A (see Figure 3), a partner of the DYNAMAP project that is a government-owned company under the control of the Ministry of Infrastructure and Transport in Italy. During these recordings, the weather conditions were sunny, without rain and with an average temperature of 19ºC.

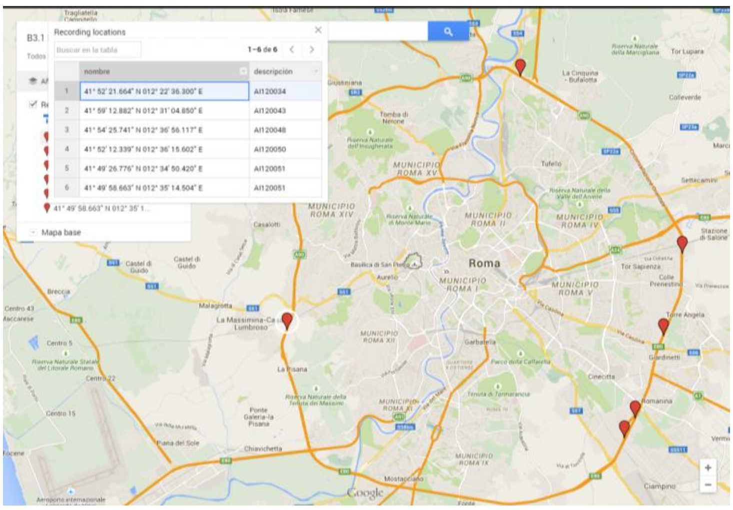


Figure 2 – Locations of the suburban recordings in the A90 highway surrounding Rome.

Figure 3 –Examples of the recording setup installed in the ANAS S.p.A. portals situated on the A90 highway surrounding Rome.

From the 20th to the 21st of May 2015, we moved to the Milan's district 9 pilot area to collect urban road traffic noise samples in 12 locations at different times of day and night (see Figure 4).

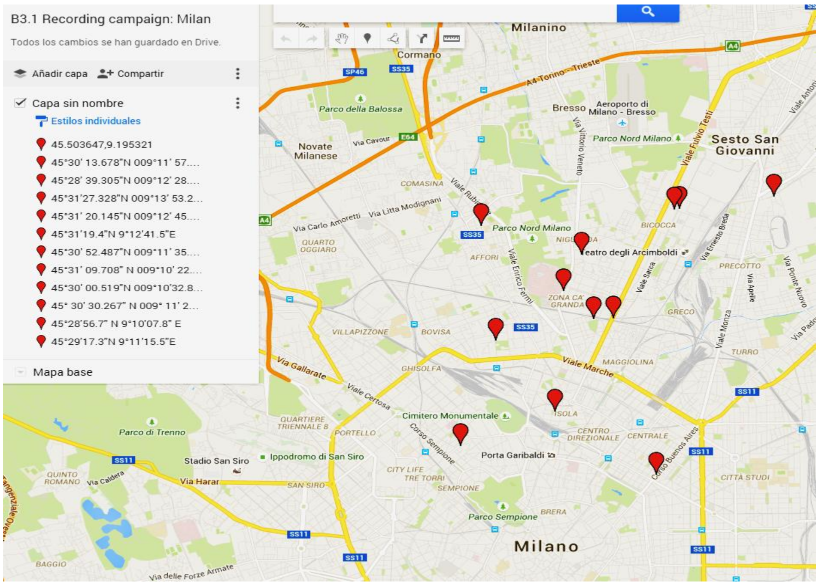


Figure 4 – Locations of the urban recordings within the Milan municipality.

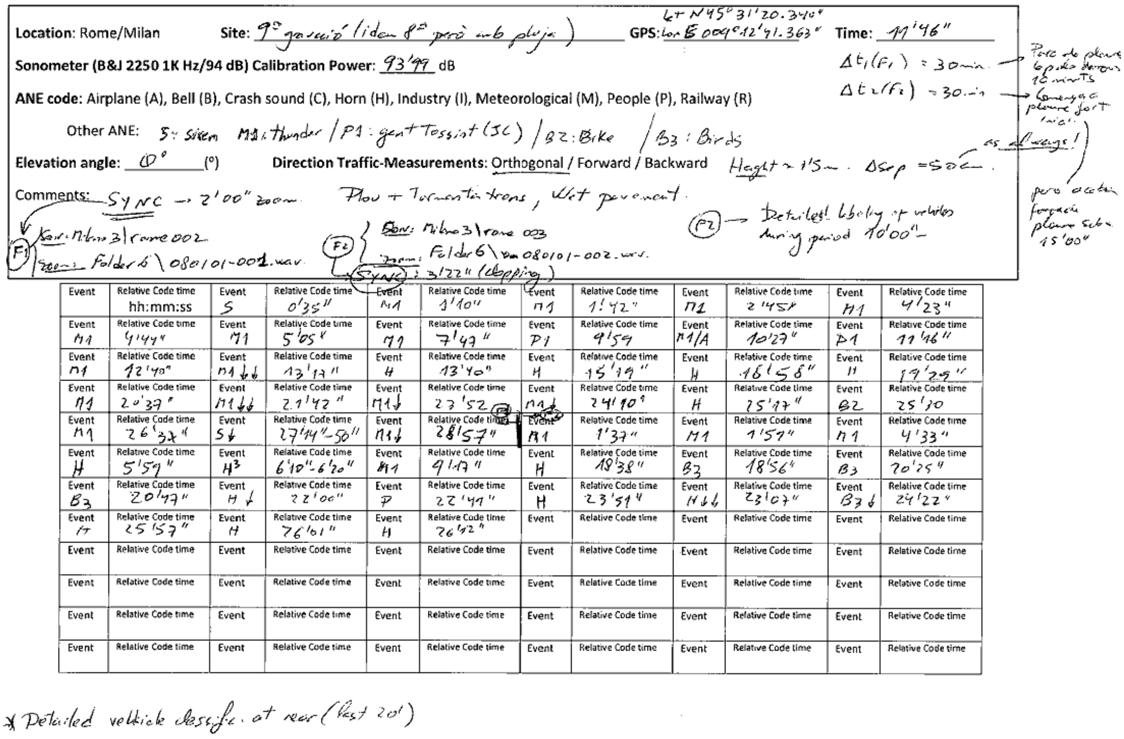
Concretely, the twelve locations correspond to the following specific locations:

1. Near hospital location, including tramways and low traffic.
2. Very low traffic one-way road.
3. Highly dense but slow traffic, with tramways, stone pavement, traffic lights and retentions.
4. Railways, very low traffic.
5. Tram and railways, fluid fast traffic multilane.
6. City center, shopping road, crossroad with traffic lights. Wet pavement.
7. Night very low fluid traffic two-way road (multilane).
8. Fluid traffic two-way road near university (multilane).
9. The same location as location number 8 but with wet pavement.
10. Residential area, fluid traffic, narrow two-way road.
11. Near school narrow two-way road with very low density traffic.
12. Low traffic, one-way narrow street near city council building.

During the Milan recordings, the weather was quite sunny, except in the second day when thunderstorms occurred during one of the recordings, making it possible to record the noise of thunders as well as road traffic noise with wet pavement.

Furthermore, in order to help the subsequent labeling and post-processing, the recordings were accompanied with a measurement annotation sheet including the following information (see Figure 5): the site location and time, the calibration tone power measured by the sensors, the elevation angle, and the direction of traffic measurements (facing forwards or backwards). Moreover, the annotation sheet allowed the manual pre-labeling of anomalous events occurring during the recording periods, together with some other particular annotations of the recording context (e.g., new types of acoustic events, or weather conditions).

As a result of the four-day recording campaign between Rome and Milan, a total 9 hours and 51 minutes of audio were collected and prepared for the subsequent labelling and post-processing phase, which is described in the next Section.



Location: Rome/Milan Site: 9º gravació (idem 8ª però amb pluja) GPS: Lat N 45º31'20.340" Lon E 009º12'41.363" Time: 11'46"

Sonometer (B&J 2250 1K Hz/94 dB) Calibration Power: 93'79 dB

ANE code: Airplane (A), Bell (B), Crash sound (C), Horn (H), Industry (I), Meteorological (M), People (P), Railway (R)

Other ANE: S: Siren M1: thunder / P1: gent Tossint (SC) / B2: Bike / B3: Birds

Elevation angle: 0º (º) Direction Traffic-Measurements: Orthogonal / Forward / Backward

Comments: SYNC → 2'00" zoom. Plou + Tormenta trons, Wet pavement.

Event	Relative Code time hh:mm:ss	Event S	Relative Code time 0'35"	Event M1	Relative Code time 1'10"	Event M1	Relative Code time 1'42"	Event M1	Relative Code time 2'45"	Event M1	Relative Code time 4'23"
Event M1	Relative Code time 4'44"	Event M1	Relative Code time 5'05"	Event M1	Relative Code time 7'43"	Event P1	Relative Code time 9'59"	Event M1/A	Relative Code time 10'27"	Event P1	Relative Code time 11'46"
Event M1	Relative Code time 12'40"	Event M1↓↓	Relative Code time 13'13"	Event H	Relative Code time 13'40"	Event H	Relative Code time 15'19"	Event H	Relative Code time 16'58"	Event H	Relative Code time 19'29"
Event M1	Relative Code time 20'37"	Event M1↓↓	Relative Code time 21'42"	Event M1↓	Relative Code time 23'52	Event M1↓	Relative Code time 24'10"	Event H	Relative Code time 25'47"	Event B2	Relative Code time 25'30
Event M1	Relative Code time 26'37"	Event S↓	Relative Code time 27'14"-50"	Event M1↓	Relative Code time 28'57"	Event M1	Relative Code time 1'37"	Event M1	Relative Code time 1'57"	Event M1	Relative Code time 4'33"
Event H	Relative Code time 5'59"	Event H3	Relative Code time 6'10"-6'20"	Event M1	Relative Code time 9'17"	Event H	Relative Code time 18'38"	Event B3	Relative Code time 18'56"	Event B3	Relative Code time 20'25"
Event B3	Relative Code time 20'47"	Event H↓	Relative Code time 22'00"	Event P	Relative Code time 22'41"	Event H	Relative Code time 23'51"	Event H↓↓	Relative Code time 23'07"	Event B3↓	Relative Code time 24'22"
Event FT	Relative Code time 25'57"	Event H	Relative Code time 26'01"	Event H	Relative Code time 26'42"	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time
Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time
Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time
Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time
Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time	Event	Relative Code time

* Detailed vehicle classific. at rear (last 20')

Figure 5 – Example of the measurement annotation sheet.

AUDIO DATABASE GENERATION

After finishing the recording campaign, a post-processing phase was conducted in order to normalize and label all the recorded audio files, and export the audio clips according to a designed taxonomy of the acoustic database specifically conceived for the ANED. To that effect, we used the Audacity freeware software.

A total of 18 audio projects were generated, one for each session during the recording campaign. Six projects were related to the Rome recordings, and twelve to the Milan ones. In the next paragraphs, the two main procedures performed during the post-processing of the audio recordings are described, divided into two steps: normalization, and labeling plus audio clip export.

Normalization step: For each audio project, the audio files gathered from the two recording devices were imported. Next, the amplitude of the audio files was normalized to make all the noise recordings uniform. This is a key process to avoid biasing the performance of the ANED algorithm. To this end, the amplitude gain was set to adjust the amplitude of the 1 kHz calibration tone of $94dB_{SPL}$ to -30 dB full scale in the audio signal spectrum. For that, we used the Audacity spectrogram analysis tool with a 512-points FFT and a Hanning window of 10.6 ms. On the one hand, this setting avoids the clipping of the regions of interest along the recordings, and, on the other hand, it allows obtaining a relevant audio signal amplitude to perform the posterior subjective labeling comfortably. Secondly, both parallel audios were manually aligned thanks to the clapping passages at the beginning of each recording session, using the clap impulsive signals (of 5 ms length each in average) to reliably align the parallel recordings through simple visual inspection. This process was followed by a subsequent perceptual validation stage based on assigning each audio to a different stereo channel (i.e., the signal from the low cost sensor to the left channel and the signal recorded by the sonometer to the right channel) to ensure there was no perceived delay in the audio.

Figure 6 shows an example of an Audacity project to illustrate the processes just described. The two tracks correspond to the signals simultaneously recorded by the two devices in the same location. The amplitude normalization adjusts the amplitude of the calibration tone (corresponding to the selected area in the first track) to -30dB full scale with the help of the spectrogram in the bottom-right region of the figure. The time alignment between both tracks is conducted by zooming the clapping area (interval between seconds 35 and 38) and delaying the second track with respect the first one, as shown in the bottom-left part of the figure.

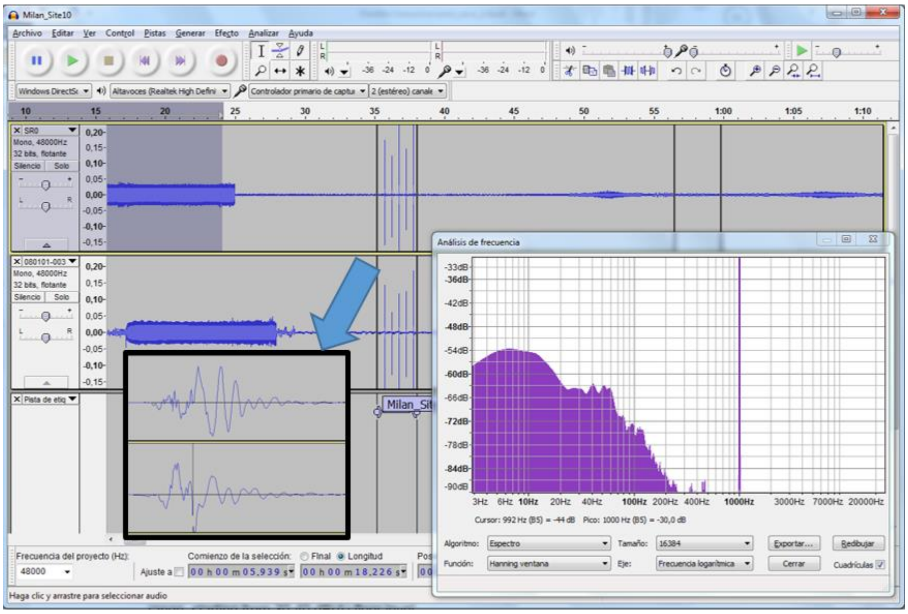


Figure 6 – Example of an Audacity audio project during the normalization process.

Audio labeling and clip exporting step. The step of labeling the audio files was toilsome and time consuming, as it entailed listening to nearly 10 hours of recordings to label audio passages according to a predefined annotation taxonomy that distinguishes among the following environmental noise events: road traffic noise (RTN), background noise (BCK) and anomalous events (ANE). RTN labels were assigned to all audio regions containing the pass-by of road vehicles, while BCK labels were reserved to those passages where it was difficult to identify the noise coming from vehicles since they contain the background noise of the city (e.g., quiet noise in a one-way street when no vehicles are present, but some distant traffic noise is perceived). In order to ensure that the ANED algorithm detects anomalous events appropriately, it is necessary to train it using samples of both RTN and BCK classes.

In turn, anomalous events (ANE) were labeled by using different subcategories, taking into account the diversity of acoustic phenomena gathered during the environmental recording campaign. These subcategories were defined in order to enrich the description of the occurred acoustic events, and were defined using the following labels (in descendent order of occurrence during the recording sessions): *peop* (people talking), *musi* (music in car or in the street), *tram* (stop, start and pass-by of tramways or trains), *sire* (sirens of ambulances, police, etc.), *stru* (noise of portals structure derived from its vibration, typically caused by the passing-by of very large trucks), *horn* (horn vehicles noise), *brak* (noise of brake or cars' trimming belt), *thun* (thunder storm), *bird* (birdsong), *trck* (noise when trucks or vehicles with heavy load passed over a bump), *door* (noise of house or vehicle doors, or other object blows), *airp* (airplanes), *wind* (noise of wind, or movement of the leaves of trees), *bike* (noise of bikes), *mega* (noise of people reporting by the public address station), *busd* (opening bus or tramway door noise), *chai* (noise of chains), and *dog* (barking of dogs).

All the ANE labels were associated to time intervals of the audio recordings only if they were easily identified subjectively. However, when an acoustic event was perceived but with high difficulty, that time region was marked with *cmplx* label (i.e., it was hard to distinguish the event from the background road traffic noise or from other acoustic events that simultaneously occur). Furthermore, noise of tyres when street bumps were considered within the RTN category.

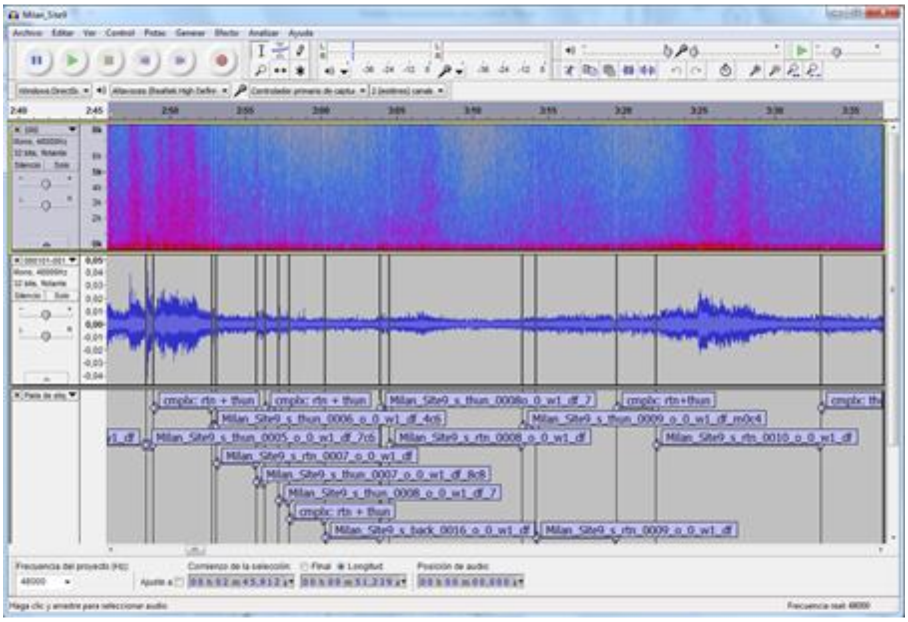


Figure 7 – Example of an Audacity audio project during the labeling process.

Figure 7 illustrates how the labelling process was conducted. To that effect, we listened to the two available audio tracks (one for each recording device), which were also visualized using spectrogram and time waveforms (see the two top tracks of Figure 7). The third track included the corresponding labels for the analyzed audio passage, where basically road traffic noise (*rtn*), background noise (*bck*) and thunders (*thun*) labels were used. Also, some regions labeled as complex scenes containing a mix of road traffic noise and thunders (*cmplx*: *rtn* + *thun*) are shown. Each one of the audio clips that was not labeled as a complex passage was exported as an independent .wav file (48 KHz and 16 bits/sample).

Next, the labeled audio clips were exported as independent *.wav* audio files using a sampling rate of 48 KHz and 16 bits/sample. Each filename contained the following parts: type of sensor (*s*: Bruel & Kjaer 2250 Sonometer, or *z*: Zoom H4n recorder plus Bluewave sensor), type of event (*rtn*, *bck*, *peop*, *musi*, *tram*, *sire*, *stru*, *horn*, *brak*, *thun*, *bird*, *trck*, *door*, *airp*, *wind*, *bike*, *mega*, *busd*, *chai*, or *dog*), order of appearance of this type of event in the same audio project (from 0 to 100), direction of measurement in relation with the traffic direction (*f* : forward, *b*: backward, *fb*: in both directions, or *o*: orthogonal), elevation angle of the measurements (0º, 45º or -45º), type of road (*h*: highway, *r1*: two-way wide road, *r2*: one-way wide road, *r3*: two way narrow road, *r4*: one way narrow road, *w1*: two-way wide road with wet pavement, *w3*: two-way narrow road with wet pavement), type of traffic (*df*: dense and fluid, *dr*: dense with retentions, *l*: low, or *vl*: very low). Additionally, ANE audio clips were also tagged with a computation of the relative amount of ANE amplitude with respect to BCK noise in dBs manually. This computation was performed individually for each audio clip, taking into account the signal to background noise ratio along the ANE with respect to a portion of audio of at least 30 immediately before or after the occurrence of the anomalous event. This extra information was added to include valuable information for the training step of the ANED algorithm (i.e., for excluding from training the anomalous events that have a very low amplitude with respect the background noise).

CONCLUSIONS AND FUTURE WORK

In this work, we have described the environmental noise recording campaign performed in May 2015 in the two pilot areas of the LIFE+ DYNAMAP project: Rome and Milan. The main goal of the campaign has been collecting enough representative acoustic data to train, validate and test the ANED algorithm included in the project to avoid including noise sources different from traffic when computing noise maps dynamically. After obtaining nearly 10 hours of audio, subsequent labelling and post-processing has led to 7 hours, 48 minutes and 38 seconds of RTN, 38 minutes and 37 seconds of BCK, and 25 minutes and 54 seconds of ANE. The rest of the recorded audio was labeled as complex audio passages. During this work, we have realized that the latter passages will need further analyses. Future work will be focused on training the ANED with the obtained acoustic database and validating its performance with respect to the results obtained previously based on synthetic databases.

ACKNOWLEDGMENTS

The research presented in this work has been partially supported by the LIFE DYNAMAP project (LIFE13-ENV/IT/001254). We would like to thank our colleagues in ANAS S.p.A. and Università de Milano-Biccoca for their support during the recording campaign.

BIBLIOGRAPHY

[1] W. Babisch; “Transportation noise and cardiovascular risk”, Noise&Health, 10, pp. 27–33, 2008.

[2] EU Directive 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise, Official Journal of the European Communities, L189/12, July 2002.

[3] S. Radaelli, P. Coppi, A. Giovanetti, R. Grecco; “The LIFE DYNAMAP project: automating the process for pilot areas location”, Proc. 22nd International Congress on Sound and Vibration, 2015.

[4] G. Zambon, R. Benocci, A. Bisceglie; “Development of optimized algorithms for the classification of networks of road stretches into homogeneous clusters in urban areas”, Proc. 22nd International Congress on Sound and Vibration, 2015.

[5] J. Socoró, G. Ribera, X. Sevillano, F. Alías; “Development of an anomalous noise event detection algorithm for dynamic road traffic noise mapping”, Proc. 22nd International Congress on Sound and Vibration, 2015.

[6] L. Nencini; “DYNAMAP monitoring network hardware development”, Proc. 22nd International Congress on Sound and Vibration, 2015.