Research Article

Francesc Alías*, Rosa Ma Alsina-Pagès, Ferran Orga, and Joan Claudi Socoró

Detection of Anomalous Noise Events for Real-Time Road-Traffic Noise Mapping: The DYNAMAP's project case study

https://doi.org/10.1515/noise-2018-0006
Received Dec 28, 2017; accepted Jun 26, 2018

Abstract: Environmental noise is increasing year after year, especially in urban and suburban areas. Besides annoyance, environmental noise also causes harmful health effects on people. The Environmental Noise Directive 2002/49/EC (END) is the main instrument of the European Union to identify and combat noise pollution, followed by the CNOSSOS-EU methodological framework. In compliance with the END legislation, the European Member States are required to publish noise maps and action plans every five years. The emergence of Wireless Acoustic Sensor Networks (WASNs) have changed the paradigm to address the END regulatory requirements, allowing the dynamic ubiquitous measurement of environmental noise pollution. Following the END, the LIFE DYNAMAP project aims to develop a WASN-based low-cost noise mapping system to monitor the acoustic impact of road infrastructures in real time. Those acoustic events unrelated to regular traffic noise should be removed from the equivalent noise level calculations to avoid biasing the noise map generation. This work describes the different approaches developed within the DYNAMAP project to implement an Anomalous Noise Event Detector on the low-cost sensors of the network, considering both synthetic and real-life acoustic data. Moreover, the paper reflects on several open challenges, discussing how to tackle them for the future deployment of WASN-based noise monitoring systems in real-life operating conditions.

Keywords: Environmental noise pollution, dynamic noise maps, anomalous noise event, road-traffic noise monitoring, wireless acoustic sensor network, feature extraction, machine learning, acoustic event detection, equivalent noise level

Corresponding Author: Francesc Alías: GTM - Grup de recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull, C/Quatre Camins, 30, 08022 Barcelona, Spain;
Email: francesc.alias@salle.url.edu
Rosa Ma Alsina-Pagès: GTM - Grup de recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull, C/Quatre Camins, 30, 08022 Barcelona, Spain; Email: rosamaria.alsina@salle.url.edu
Ferran Orga: GTM - Grup de recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull, C/Quatre Camins, 30, 08022 Barcelona, Spain; Email: ferran.orga@salle.url.edu
Joan Claudi Socoró: GTM - Grup de recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull, C/Quatre Camins, 30, 08022 Barcelona, Spain; Email: joanclaudi.socoro@salle.url.edu

1 Introduction

Environmental noise is increasing year after year and becoming a growing concern in urban and suburban areas, especially in large cities, since it does not only cause annoyance to citizens, but also harmful effects on people. Most of them focus on health-related problems [1], being of particular worry the impact of noise on children [2], whose population group is especially vulnerable. Other investigations have also shown the effects of noise pollution in concentration, sleep and stress [3]. Finally, it is worth mentioning that noise exposure does not only affect health, but can also affect social and economic aspects [4].

Among noise sources, road-traffic noise is one of the main noise pollutants in cities. According to the World Health Organization (WHO), at least one million healthy life years are lost every year from traffic-related noise in western Europe [5]. For instance, it was recently stated that transportation noise alone accounts for 36% of the total burden of disease attributable to urban planning, an even higher percentage than the one caused by air pollution in Barcelona [6].

The European Union (EU) has reacted to this alarming increase of environmental noise pollution, especially in large agglomerations, by approving the Environmental Noise Directive 2002/49/EC (END) [7]. In accordance with the END, the Common Noise Assessment Methods in Europe (CNOSSOS-EU) has been defined to improve the consistency and comparability of noise assessment results across the EU Member States [8]. The main pillars of the END are the following: *i)* determining the noise exposure, *ii)* making the updated information related to noise available to citizens, and *iii)* preventing and reducing the environmental noise where necessary. Moreover, the END requires the European Member States to publish noise maps and action plans every five years for agglomerations with more than 100,000 inhabitants, major roads, railways and airports, besides introducing the need to discern between different sound sources [7]. Working towards the same goals, new mitigation regulations and strategies have been proposed in order to evaluate and combat specific noise sources, such as: *i)* road-traffic [9, 10], *ii)* railways [11, 12], *iii)* airports [13, 14], *iv)* industries [15, 16] and *v)* wind turbines, which have recently been recognized as highly annoying [17, 18].

The classic approach of building noise maps from costly expert-based measurements has recently undergone a change of paradigm to address the END regulatory requirements thanks to the emergence of Wireless Acoustic Sensor Networks (WASNs), which allow the ubiquitous measurement of environmental noise in real time. WASNs have been successfully applied in several cities, such as Barcelona (Spain) [19], Pisa (Italy) [20], Monza (Italy) [21], Halifax (Canada) [22] or the National Highway of Burdwan (India) [23], to name a few. However, most of these WASN-based environmental noise monitoring projects address the problem in a holistic way by representing the global noise levels of the area of interest without differentiating the contribution of noise sources. As an exception, in [24], sound recognition is applied together with a subjective survey in order to cross both the acoustic and the subjective perception of noise components. Nevertheless, since the identification of sound sources is conducted after computing the noise levels, this information cannot be used to modify the noise map calculation dynamically. In order to identify specific sound sources automatically, several acoustic event detection algorithms have been proposed to address different urban environment applications, mainly focused on surveillance. Some of them are focused on noise source identification [25–27], while other projects are centered on the separation between target and interfering signals with the final goal of noise monitoring in cities [28, 29].

In this context, the LIFE DYNAMAP project[1] aims to deploy a low-cost hybrid WASN to tailor noise maps that represent the acoustic impact of road infrastructures in real time, using a Geographic Information System (GIS) platform. The project includes two pilot areas in Italy: the A90 motorway surrounding Rome (suburban area) and the district 9 of Milan (urban area) [30].

In an attempt to monitor the impact of road infrastructures solely, those events unrelated to Road-Traffic Noise (RTN), denoted as Anomalous Noise Events (ANEs) (*e.g.*, birds, people talking, sirens, etc.), should be removed from the noise map generation due to the non-negligible impact [31] of both individual and aggregated ANEs on the A-weighted equivalent noise levels (L_{Aeq}) computation [32]. Therefore, ANEs should be detected automatically in order to obtain a reliable picture of citizens' exposure to RTN by means of the generated noise maps. Within the DYNAMAP project, this goal has been pursued by the design and development of an Anomalous Noise Event Detector (ANED) [28, 33]. The ANED algorithm is asked to identify ANEs in real time (every second) and designed to run in the low-cost acoustic sensors of the WASN (see Figure 1).

This work describes the different ANED approaches developed incrementally to fulfill the DYNAMAP project specifications hitherto, considering both synthetic and real-life acoustic data. The paper also discusses and suggests some potential solutions for the main open challenges of deploying WASN-based noise monitoring systems in real-life operating conditions.

The paper is structured as follows. Section 2 reviews the state of the art about acoustic sensing in urban environments. Section 3 describes the main elements of the ANED algorithm, and the process followed to characterize the acoustic environment of operation. Next, Section 4 details the main results obtained considering both synthetic and real-life acoustic datasets, together with the analysis of the impact of ANEs on the noise maps generation. Section 5 discusses the main open challenges of the problem at hand. The paper ends with the conclusions in Section 6.

2 State of the art about acoustic sensing in urban environments

In this section, firstly, we review several representative WASN-based projects focused on measuring the quality of life of citizens due to noise pollution. Secondly, we briefly describe the acoustic event detection literature.

1 http://www.life-dynamap.eu/

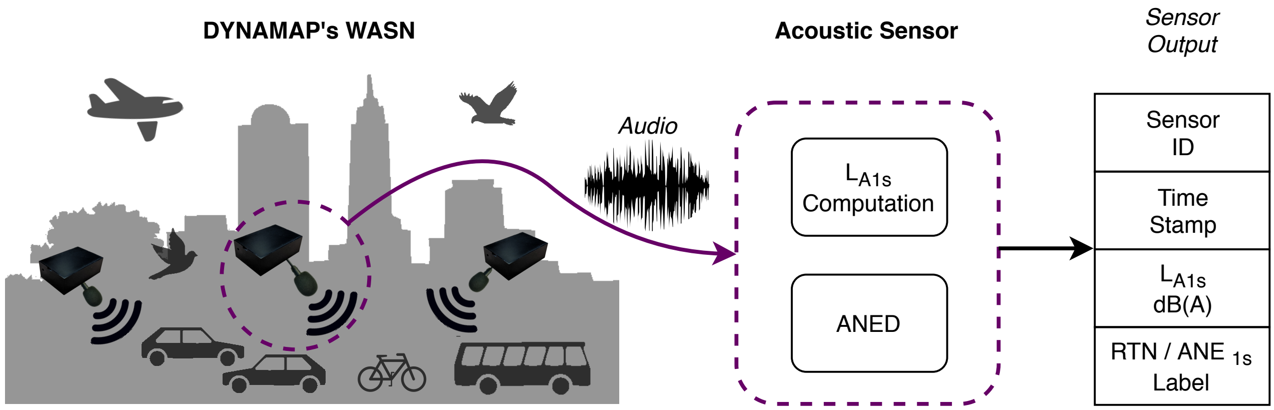

Figure 1: Block diagram of the DYNAMAP's WASN, describing the main processes run in the low-cost acoustic sensors every second, including the computation of the A-weighted equivalent noise level (L_{A1s}) and the Anomalous Noise Event Detector (ANED), which labels the input acoustic data as Road Traffic Noise (RTN) or Anomalous Noise Event (ANE) every second. Moreover, the sensor sends its identifier (ID) and the time stamp to the central server.

2.1 WASN-based acoustic urban sensing projects

In order to satisfy the increasing demand for automatic monitoring of the noise levels in urban areas, several WASN-based projects are being developed in different countries; some of these projects include other environmental measurements.

The SENSEable project in Pisa is based on the smart city concept to measure the noise level in several points of the city in real time [20]. A noise monitoring network is being deployed in Barcelona in order to manage the resources efficiently and to reduce the impact of urban infrastructures on the environment [19] and also, recently in Monza, by a LIFE project that implements also a low-cost system [21]. The Fi-Sonic project, which is based on the FIWARE platform, is mainly focused on continuous environmental noise monitoring and surveillance; several sound events can be identified for surveillance purposes, with the signal processing conducted on a centralized server [26]. The CENSE (Characterization of urban sound environments) project in France [34], aims to develop a new methodology for the production of realistic noise maps, based on an assimilation of simulated and measured data through a dense network of low-cost sensors. The RUMEUR (Urban Network of Measurement of the sound Environment of Regional Use) wireless network [35] developed in the Paris region by BruitParif, is a project that includes both high-accuracy equipment for critical places (*e.g.*, airports) together with less precise measuring equipment placed in other locations where the goal is only to evaluate the equivalent noise level. The IDEA (Intelligent Distributed Environmental Assessment) project [36] measures noise and air quality pollution levels in urban areas in Belgium. The MESSAGE (Mobile Environmental Sensing System Across Grid Environments) project [37] monitors noise, carbon monoxide, nitrogen dioxide, temperature, humidity and traffic occupancy/flow, providing real-time noise data levels in the United Kingdom, being the case-study applications conducted in London.

Other projects aim to monitor the urban noise in real time, such as the UrbanSense project in Canada [22], which also aims to monitor other pollutants as carbon dioxide (CO_2) and carbon monoxide (CO). Finally, some projects focus in certain areas, as for example highways. In [23], five points along the National Highway of Burdwan have been monitored with an audiometer in order to register the acoustic equivalent level, besides conducting the corresponding statistical analyses. In [38], the urban sound environment of New York City is monitored using a low-cost static acoustic sensing network named SONYC; the goal of this project is to monitor the noise pollution in the city providing an accurate description of its acoustic environment. In [39], the LIFE MONZA project addresses the issue of the definition, the criteria for the analysis and the management of the Noise Low Emission Zones by means of the use of a low-cost WASN.

The aforementioned projects aim to monitor the noise in determined areas using low-cost WASNs. However, as far as we know, none of them intend to remove any anomalous events which bias the traffic noise map measurement, both in urban and suburban scenarios, since they are not designed to monitor this specific noise pollutant. In the Smart Sound Monitoring project, De Coensel *et al.* conducted a study in [24] that crosses acoustic information with subjective perception surveys, enabling the typology of the acoustic information to be considered. Besides a sound recognition system is applied in order to give information about the detected sounds and establish a relation between the identified events and the perception surveys. However, the system identifies events only to give information but not to remove these sounds from the noise map. In [29], the authors introduce a supervised noise source classifier, learned from a small amount of manually classified recordings, to automatically detect the activity of anomalous noise sources; nevertheless, this pilot has been implemented only in one environment, on a rock crushing site. It has not been yet exported to usual urban and suburban environments with wider types of interfering signals. With the purpose of generalizing the solution to the interfering problem, the DYNAMAP project aims to monitor road traffic noise in suburban and urban areas reliably, after removing the anomalous noise events from the road traffic noise map computation [30].

2.2 Detection of acoustic events in urban environments

In the literature, we can find several works showing that the detection of events non-related to road traffic noise is a very challenging task [40, 41]. On the one hand, when focusing on environmental sound detection, acoustic events tend to be disconnected one from another (see [42] and references therein), unlike speech or music, besides presenting a great variability, which depends on the acoustic environment [43]. On the other hand, the occasional, unpredictable and diverse nature of ANEs makes their identification particularly complex in real-life scenarios [25, 44].

The problem of environmental Acoustic Event Detection (AED) has attracted the scientific interest of the research community recently, *e.g.*, as can be observed in the Workshops on Detection and Classification of Acoustic Scenes and Events (DCASE) 2016 [45] and 2017 [46]. In general terms, two general approaches have been applied to environmental AED: detection-and-classification and detection-by-classification [47]. The former allows detecting sudden changes of the background noise (*i.e.*, acoustic salience) at the cost of discerning between different kinds of sounds. The latter classifies all the audio segments by means of the parametrization and classification of the input acoustic events within a predefined set of classes, after the classifier has been trained with enough representative data per class (*i.e.*, multi-class AED). Given that the determination of the acoustic salience of an acoustic event is not enough to discriminate between road traffic noise and anomalous events, the recognition of ANEs should follow the detection-by-classification approach.

In this context, classic multi-class AED can not be directly applied to the problem at hand due to its unbounded nature (*i.e.*, the number of classes composing the anomalous noise events can not be known beforehand). Alternatively, the AED-based novelty detection focuses on the major acoustic class (*i.e.*, road traffic noise, in this case) by means of the One-Class Classification (OCC) machine learning approach (see [48] and references therein). This approach has been applied in urban environments for surveillance applications oriented to the identification of gun-shots, broken glasses, and screams, among others [25, 49]. Another urban event detection proposal [50] includes also normal and abnormal (or anomalous) audio events such as screams, shouting or asking for help, which are collected in real-life outdoor public surveillance scenarios. The acoustic data is parametrized every 30 ms using a multi-domain feature vector including different audio descriptors, such as Mel Frequency Cepstral Coefficients (MFCC), MPEG-7 Low-Level Descriptors (LLD) and Perceptual Wavelet Packets (PWP). The parametrized audio frames are fed into different probabilistic classifiers based on Gaussian Mixture Models (GMM) [25], Hidden Markov Models (HMM) and Universal HMM [50], following a class-specific modelling and universal background modelling.

However, this approach omits considering valuable information about ANE. To this aim, in [28], a two-class classification scheme has been introduced to detect ANE mixed with RTN in real-life environments, outperforming the OCC-based approach when having enough representative data to model the minority class (*i.e.*, the anomalous noise events). Further details of this approach are described in the following section.

3 Developing an anomalous noise event detector in real-life environments

The DYNAMAP project goal is to generate noise maps using only road traffic noise contribution; so, it requires the detection of all non-traffic events in order to discard them. According to the project requirements, ANEs are considered to be those acoustic events that do not come from the engines of the vehicles or are not derived from the normal contact of their tires with the pavement, *i.e.* not re-

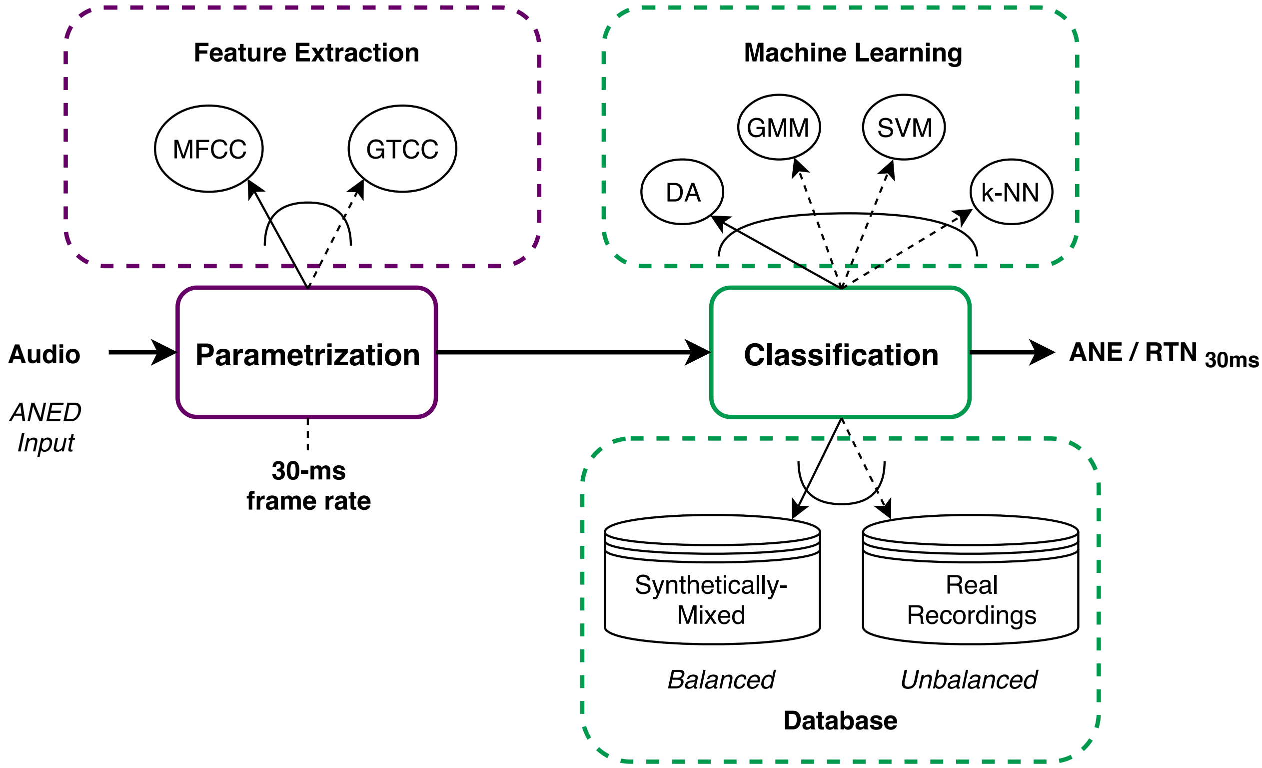

Figure 2: Block diagram of the frame-level decision phase of the ANED algorithm, composed of two main stages: parametrization of the input audio extracting the MFCC or GTCC at a 30-ms frame rate, and its binary classification as RTN or ANE. Four different classifiers (DA, GMM, SVM and k-NN) have been considered hitherto, which have been trained using both real-life or synthetically-mixed acoustic databases.

flecting regular road traffic noise [7]. Moreover, the Anomalous Noise Event Detector envisioned by the DYNAMAP project has to be designed and integrated in the low-cost sensors of the WASN to discard ANEs before integrating the information to the GIS-based road traffic noise map. As far as the ANED operating specifications are concerned, the project consortium agreed to set the output decision of the ANED at every second, focusing the discrimination between ANE and RTN based on their spectral differences [30].

In the pursuit of this goal, the ANED development has been addressed incrementally throughout the course of the project. In the following sections, the key elements considered up to date to tackle the detection of ANEs in real time have been defined.

3.1 Detection of ANEs in real time

In order to identify and subsequently remove the impact of ANEs on the road traffic noise computation in real time, an ANED has been designed, developed and implemented to run on the low-cost acoustic sensors of the WASN deployed in the two pilot areas of the DYNAMAP project: urban (Milan) and suburban (Rome) environments. As aforementioned, the ANED is asked to provide a binary decision between ANE and RTN every second (see Figure 1).

The ANED has been designed following the classic architecture of machine hearing systems [42]: audio parametrization and classification (see Figure 2). Audio parameterization by means of feature extraction aims to provide a meaningful and compact description of an input audio frame. Among the different possibilities implemented to parametrize the acoustic events, up to now, we have opted for computing spectral-based features, being either MFCC [51] or Gammatone Cepstral Coefficients (GTCC) [52]. While the former have been widely used in speech recognition and for the identification of other type of audio signals, the latter has shown better performance in environmental sound classification tasks [52].

The ANED algorithm has been designed to output a binary label every second (RTN/ANE$_{1s}$) based on a two-stage decision scheme (see Figures 2 and 3). First, the ANED categorizes the audio input as RTN or ANE following a two-class detection-by-classification AED approach (see Figure 2). Due to the short nature of some ANEs (*e.g.*, impulse-

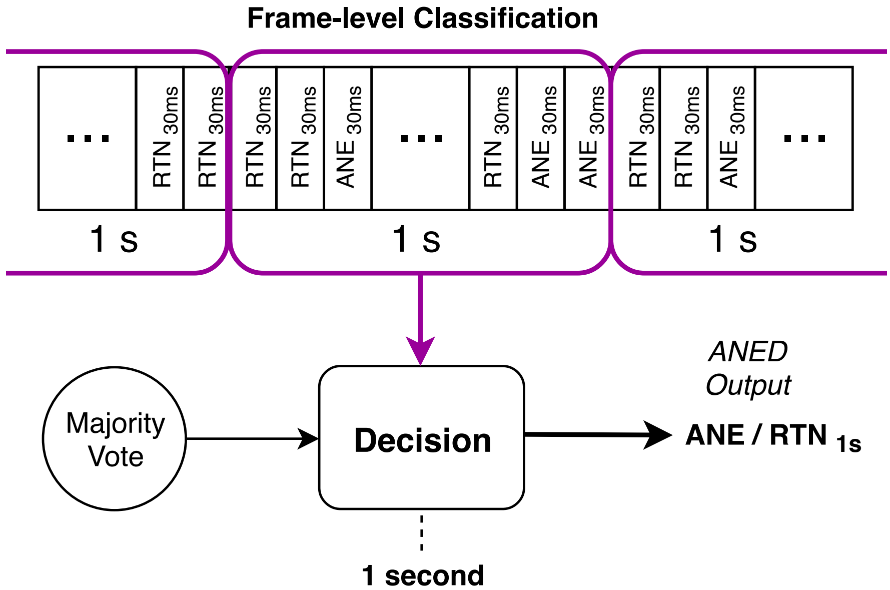

Figure 3: Block diagram of the majority vote stage applied after the frame-level classification, which outputs the ANED's binary decision every second (RTN/ANE$_{1s}$).

like events such as dog barks or door closings), the input audio data is analyzed every 30 ms in order to detect them at the frame-level. It should be noted that different core machine learning techniques have been considered to train the frame-level classifier, such as: Discriminant Analysis (DA); GMM; Support Vector Machines (SVM); and k-Nearest Neighbors (k-NN), as depicted in Figure 2. Second, the frame-level decisions obtained every 30 ms are integrated every second as a binary output label to fulfill the project specifications (see Figure 3). Until now, this decision has been made through a simple yet effective majority vote. As a result, the next steps of the map generation are informed whether the corresponding L_{A1s} belongs to RTN or it should be discarded since it is contaminated with ANEs.

Due to the supervised nature of the frame-level classification of the ANED, it becomes essential to have enough representative data to build the acoustic models of both classes: RTN (also including background city noise) and ANEs. In the following section, we describe the main steps followed to characterize the acoustic environment, which has addressed considering both synthetic and real-life data throughout the project.

3.2 Characterization of the acoustic environment

The acoustic database employed to train the ANED is one of the key elements to tackle the problem at hand. The classifier should be trained with enough representative RTN and ANE examples if both acoustic classes have to be modelled properly. Those examples can be either obtained from online repositories or from real-life recordings, where ANE and RTN should be mixed synthetically or appear naturally merged with city background and traffic noise, respectively. During the course of the project, we have observed the complexity of both designing artificially-mixed acoustic data that represent real-life ANEs in urban and suburban environments reliably, and collecting representative events in the real field.

A preliminary version of the acoustic dataset was synthetically generated, following similar previous works [40, 44]. The database was designed in a balanced manner (250 seconds of RTN and 300 seconds of ANEs). In order to simulate real-life conditions, we mixed on-site recordings from a ring road of the city of Barcelona (*i.e.*, similar to the Rome suburban environment) together with several ANE obtained from the Freesound[2] open database. The acoustic salience of the mixed ANEs with respect to RTN was set to be either +6 dB or +12 dB, being computed their Signal-

2 https://freesound.org

to-Noise Ratio (SNR) as follows:

$$SNR = 10\,log_{10}\left(\frac{P_{ANE}}{P_{RTN}}\right) \quad (1)$$

where P_{ANE} and P_{RTN} are estimations of the acoustic power of ANEs and RTN, respectively.

In the following sections, we describe the main steps followed to obtain a real-life acoustic database to characterize both urban and suburban pilot areas of the DYNAMAP project.

3.2.1 Collecting ANE in real-life urban and suburban areas

The DYNAMAP's WASN is currently being deployed in two pilot areas to validate the performance of the designed road traffic noise monitoring system in both urban and suburban scenarios [30]. The key characteristics of the urban and the suburban pilot areas are described below.

Regarding the urban area, Milan has 1.3 M citizens [53], increasing to 3.2M inhabitants in its metropolitan area. The selection of the pilot area took into account the following attributes [54]: extension, number of citizens, noise exposure levels and linear road length. As a result, district 9 was selected as the most suitable area of the city to deploy the pilot of the WASN. It is to note that this district contains two sensitive sites that should be protected from noise: the largest hospital of Milan and the University of Milano-Bicocca.

Subsequent studies [55] have concluded that the recording locations within Milan's district 9 can be acoustically clustered in two groups, according to the distribution of the A-weighted traffic-noise hourly equivalent noise levels (L_{Aeqh}). Between 08:00 and 09:00, the first cluster presents a higher peak than the second group, while during the night period (20:00 and 05:00), the first cluster shows lower averaged values than the second one.

As far as the suburban scenario is concerned, Rome is the largest and most populated city of Italy, with 2.8 M inhabitants [53], reaching 4.3 M residents if all the metropolitan area is considered. The A90 ring-road surrounding Rome has been found to be critical in terms of noise pollution since it is a six-lane 68 km-long motorway with presence of multiple noise sources, *e.g.* railways and crossing roads. The sensors of the WASN deployed in this pilot area are being installed in the message panels and road signs at critical locations, usually intersections and crossing roads [56].

In order to obtain an accurate picture of the real-life acoustic urban and suburban environments where the

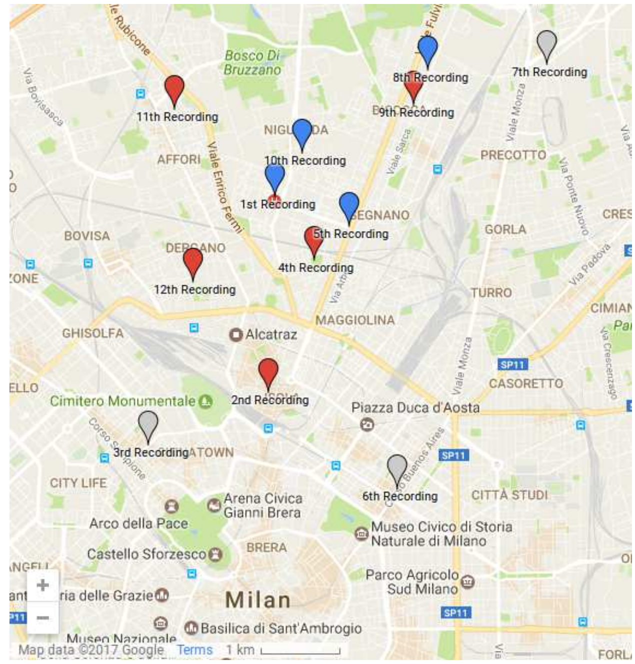

Figure 4: Recording campaign sites of the Milan urban area.

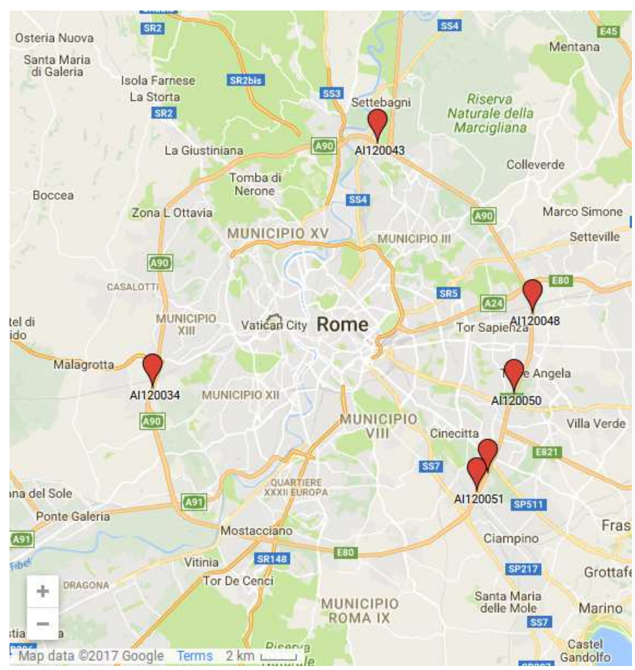

Figure 5: Recording campaign sites of the Rome suburban area.

ANED is asked to operate, a recording campaign was conducted between the 18^{th} and 21^{st} May 2015, covering both Milan (urban) and Rome (suburban) pilot areas. Several recording sites were selected taking into consideration diverse representative traffic conditions and acoustic characteristics within the pilot areas defined in the previous section. The specific recording sites of Milan are depicted in

Figure 4, while the recording spots along the A90 ring-road portals of Rome are shown in Figure 5. The next section describes the main characteristics of the gathered acoustic data.

3.2.2 Characteristics of the real-life urban and suburban acoustic database

After the manual annotation of all the gathered acoustic data during the recording campaign, an audio database of 9h and 8 min of real-life data was obtained, after removing some synchronization passages. The ANEs only represent the 7.5% of the labelled data [43]. In particular, 19 different ANE categories were identified perceptually, such as airplanes, bird songs, barking dogs, thunders, sirens, or people talking, among others. ANEs lasted from 30-40 ms for impulse-like sounds like door closing to 30 seconds for longer events like a siren pass-by, besides presenting SNR from -10 dB to +15 dB with a relevant diversity of intermediate SNR values. When comparing both environments, ANEs represent about 3.2% of the 4 h and 44 min database of the suburban environment, being 12.2% in the urban context (4 h and 45 min). Among them, ANEs with significant acoustic salience (*i.e.*, those with SNR ≥ 6 dB) only represent the 29% and the 1.85% of the total ANEs for the urban and suburban environments, respectively.

These values together with the in-depth analysis of the distribution of the collected acoustic data show the high diversity of ANEs in terms of their occurrence, duration and SNR with respect to the synthetically generated database, besides confirming the higher complexity and unbalanced nature of the classification problem in real-life environments (the reader is referred to [43] for further details). Furthermore, although both urban and suburban environments share some general characteristics, it is worth mentioning the significant differences found between them due to their specific acoustic nature. Thus, this entails, at least, a specific training of the ANED algorithm for each environment to address the problem appropriately.

4 Experiments and results

In this section, the main results obtained from the process of design and implementation of the ANED algorithm for real-time road traffic noise mapping are presented. Firstly, Section 4.1 presents an analysis of the impact of the real-life ANE audio passages on the computation of the averaged noise equivalent levels, showing some specific examples for illustrative purposes. Secondly, the results obtained from the different design phases of the ANED algorithm are described in Section 4.2, which encompasses our previous works ranging from those which used the synthetic audio database to the research conducted using audio datasets obtained from real-life urban and suburban environments.

4.1 Impact of ANEs on the L_{Aeq} computation

The removal of ANEs from the RTN map generation is based on the hypothesis that they may bias the L_{Aeq} value significantly. In this context, it seems reasonable to assume that individual ANEs with high acoustic salience and long duration (*e.g.*, a siren pass-by) could affect significantly the L_{Aeq} computation. Nevertheless, other combinations such as a large number of medium-impact ANEs or the presence of frequent impulse-like ANEs within the integration period considered to compute the L_{Aeq} could also have a similar effect.

In [32], the impact of both individual and aggregated ANEs on the A-weighted equivalent noise level computation is evaluated, proving their potential bias on the road traffic noise map computation. That work computed the deviation caused by ANEs on the L_{Aeq} value with respect to a predefined ground truth, obtained after the manual identification of all the events that should be removed from the equivalent RTN level computation [32]. In that research, two parameters were used to parametrize the anomalous noise events: the duration, which is obtained as the difference between the time stamp of the first and last sample of the ANE; and the SNR of the ANE under study.

That work concludes that both duration and SNR of the ANE are key variables to identify the most relevant events to take into account; however, there is a need to quantify the contribution of each ANE to the L_{Aeq}. This contribution is calculated as the difference between the L_{Aeq} considering the individual ANE and the L_{Aeq} replacing the event with the linear interpolation from the last and the following RTN sample of the original raw data.

In the DYNAMAP project the dynamic map is updated every five minutes during the day, because the L_{Aeq} values present higher variability depending on time; evening and night show more stable L_{Aeq} values [55]. To take into account all circumstances, the equivalent noise level and the impact of ANEs should be calculated over the minimum 5-min integration time, *i.e.* L_{A300s}. In Figures 6 and 7, a couple of examples collected in two different recording sites of the Milan pilot area are included for illustrative purposes. In those figures, the impact is plotted together along the

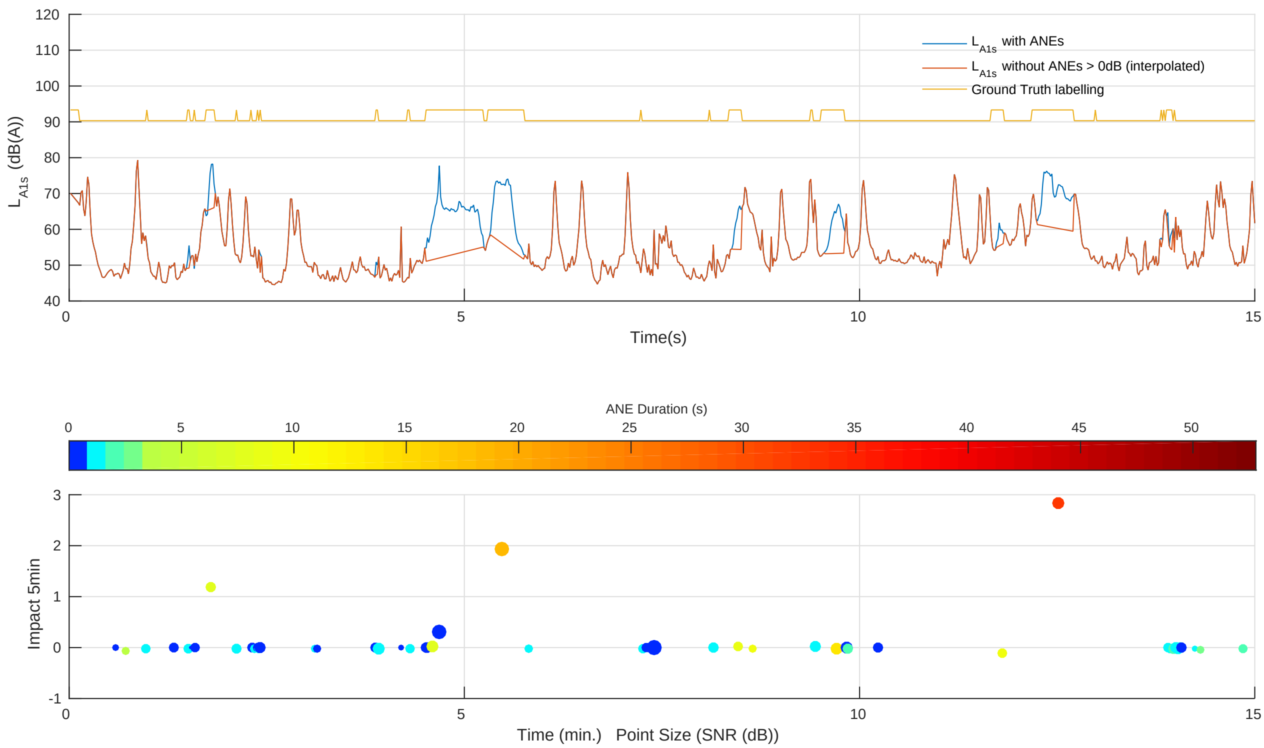

Figure 6: A detailed example of the Site 4 of the Milan pilot area recording [43], showing the bias of the ANEs on the L_{A300s} computation, including the evolution of the L_{A1s} curve with and without ANEs obtained from the ground truth manual labelling reference.

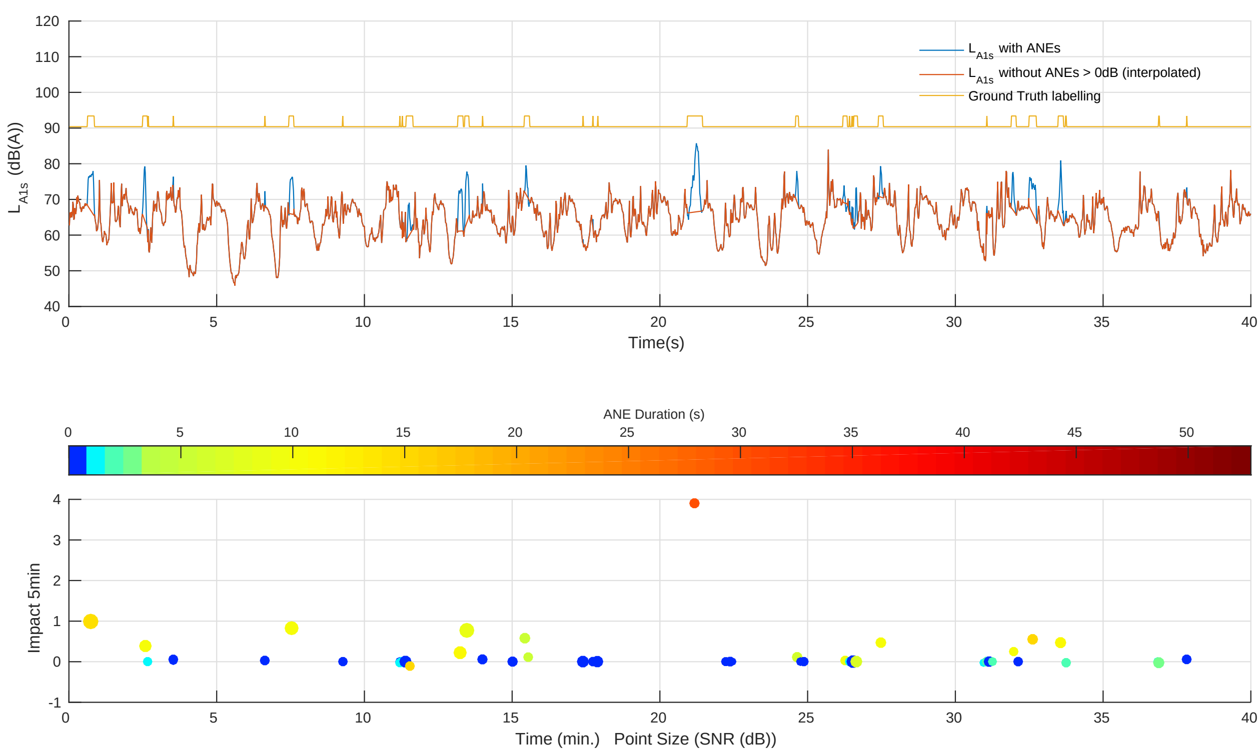

Figure 7: A detailed example of the Site 5 of the Milan pilot area recording [43], showing the bias of the ANEs on the L_{A300s} computation, including the evolution of the L_{A1s} curve with and without ANEs obtained from the ground truth manual labelling reference.

same temporal axis with the duration (color legend) and the SNR (size of the dot). The A-weighted equivalent noise level computed every second (L_{A1s}) of the raw audio and the resulting L_{A1s} of the same audio without ANEs is also shown for comparative purposes. This way, the reader may analyze in detail the impact of each and every individual ANE within the 5-min period it belongs to.

Both examples show the frequent occurrence of ANEs, with diverse duration and SNR. In particular, the most salient ANE in Figure 6 is a train pass-by in minute 12, which lasts 32 s, it has an SNR of 6.9 dB and it adds 2.8 dB to the corresponding L_{A300s}. Next, in minute 6, an 18-s train pass-by with an SNR of 13 dB contributes with 2 dB to the final L_{A300s} computation. And finally, another train pass-by is captured in the 2^{nd} minute that has a duration of 7 s, with an SNR of 2.6 dB and affecting the L_{A300s} in 1.2 dB. As the reader may observe, the recording of Site 4 contains basically train pass-bys in terms of ANEs occurrences.

As can be observed in Figure 7, the most salient ANE, a 30-s siren with an SNR of 5.9 dB, can be found in the 21^{st} minute. According to the graph, it contributes 3.9 dB to the final L_{A300s} computation. In the same recording, there are other ANEs biasing the L_{A300s}, but with a contribution of 1 dB or less to the L_{A300s} computation. However, as observed in [32], all ANEs should be taken into consideration in the computation, as many low-impact ANEs detected in the same period could have a significant aggregated impact.

4.2 Classification accuracy

In this section, we describe the main results obtained from the experiments conducted to evaluate the accuracy of the different implementations of the anomalous noise events detector within the DYNAMAP project.

In [44], a preliminary version of the ANED algorithm trained with the synthetically generated environmental acoustic database was tested using a 4-fold cross-validation scheme. As described in Section 3.2, two SNR values were set artificially by changing a gain factor for both ANEs and RTN audio passages (+6 dB and +12 dB) to generate the artificial mixtures. Moreover, two classifiers (k-NN and linear DA) and two audio parametrization techniques (MFCC and GTCC) were evaluated under supervised and semi-supervised two-class classification schemes. The results showed the superiority of the semi-supervised approach in most of the simulated scenarios in terms of the F1 measure of the ANE class, yielding around 80% of classification accuracy, and presenting better results for the +12dB than the +6dB ANEs, as expected a priori.

Unfortunately, these results proved to be non-generalizable to real-life conditions during the course of the project [57]. Two main conclusions were derived from the obtained results when tested with real-life data. First, the ANED performance fell around 25% on average with respect to the classification results observed when working on the synthetic balanced database. Second, the high complexity of real-life acoustic environments due to the larger diversity of ANEs, SNRs, etc. makes it difficult (or almost unfeasible) to model them artificially [43]. Thus, subsequent research was focused on training and evaluating the ANED with real-life acoustic data, discarding any strategy based on using non-realistic mixtures between ANEs and RTN, even though it was an approach widely considered in the literature related to audio event detection. In the following studies, we used the real-life database built from the recording campaign, and considered a larger diversity of core machine learning techniques for the two-class classification approach (see Figure 2).

The subsequent experiments using the real-life database considered 4-fold and leave-one-out cross-validation schemes to study the ANED performance at the frame-level (every 30 ms) and at the high-level (every 1 s) - see Figure 3 -, respectively [28]. The 4-fold scheme was employed to select the best core classifier of the ANED when making a binary RTN/$\mathrm{ANE}_{30\mathrm{ms}}$ decision at the frame-level (see Figure 2). The following machine learning approaches were compared: DA with *quadratic* discriminant function; k-NN with $K = 1$; SVM with *Radial Basis Funtion* kernel; and GMM with 256 components for both acoustic models (RTN and ANE). The obtained macro-averaged F1 values ranged from 58,01% to 74,43% for the suburban scenario, and from 68,96% to 81,44% for the urban environment, respectively. SVM and k-NN were the two classifiers that obtained the best F1 scores in both environments, followed by GMM and DA.

Moreover, since the ANED algorithm is asked to perform its classification in real time on a low-cost acoustic sensor, the computational cost of these core classifiers was also evaluated. As the training stage can be performed off-line, the study only focused on the computational load of the testing phase. The results showed that k-NN and SVM were the two core classifiers yielding the highest computational cost, while GMM and DA obtained the lowest scores (being the computational complexity of the GMM, around 1% and 2% of the computational complexities of k-NN and SVM, respectively). Therefore, from the 4-fold cross validation experiments, we selected the GMM approach as the ANED core classifier since it entailed the best trade-off between classification performance and computational load (*i.e.*, it can be run on the low-cost sensors of the WASN in real time).

Next, the leave-one-out strategy was used to evaluate the classification accuracy of the GMM-based ANED at the high-level one second rate (see Figure 3), using entire audio clips to assess the ANED decisions under simulated real-life operating conditions. When compared to the alternative Universal GMM-based OCC, the two-class GMM-based classifier presented significant improvements for both urban and suburban environments, yielding a relative increase of 18.7% and 31.8% in the frame-level macro-averaged F1 measure, respectively [28]. Moreover, the results also showed that the the classification accuracy increased in 3.8% and 4.4% for the suburban and urban environments, respectively, after the majority vote.

These results confirm that the two-class classification approach outperforms the OCC alternative when enough representative data can be found to model the minority class, in our case, the ANE category. Moreover, the experiments also demonstrate the major complexity of the suburban acoustic environment ahead the urban one in terms of anomalous noise events detection in real-life environments.

5 Discussion

In this work, we have described the key milestones of the DYNAMAP project approach to develop an automatic process to detect and remove those acoustic events unrelated to road traffic noise from the WASN-based noise map computation. In this section, we discuss some open challenges that should be addressed to improve the future versions of the ANED algorithm running on a WASN in real-life operating conditions.

5.1 Acoustic Databases

We have found that the performance of the ANED highly depends on the nature of the audio database used for the training stage. The preliminary ANED algorithm was trained and tested on a class-balanced synthetic database, obtaining very satisfactory results at the frame-level. However, when the same classifier was applied to real-life audio recorded data, *i.e.*, with unbalanced proportion of RTN and ANEs containing a high diversity of SNR values, the performance of the classifier fell dramatically. These results proved that the simulation of real-life operating environment by mixing RTN recorded from a real suburban soundscape with ANEs collected from online repositories is very complicated, becoming almost unfeasible to address the real-life problem properly [43].

Moreover, the significant acoustic differences observed between the urban and suburban scenarios (both in terms of ANE typology and SNR variations) suggest that the ANED implementation should discern, at least, between these two different acoustic environments [28, 43]. Nevertheless, and according to [55], we do not discard introducing another level of adaptation in the Milan pilot area to take into account the two observed acoustic clusters. The typology of the streets, of the traffic flow and the different types of noise propagation and ANEs is leading us to consider the possibility of designing more than one ANED for the urban scenario [58]. Therefore, the potential local adaptations should be accounted for when analyzing the degree of generalization of the ANED approach.

5.2 Classification

Another key characteristics of the acoustic data collected in real-life environments is their inherent unbalanced nature in terms of the occurrence of ANEs with respect to RTN. Obviously, during our research we have observed that RTN is the majority class while the ANE class contains a reduced number of heterogeneous samples. Moreover, by its definition, the ANE class is unbounded since it accounts for *any* sound event different from road traffic noise. These particular characteristics of the problem at hand makes it specially complex to address effectively, as already discussed in the literature. Up to now, we have observed that designing the ANED by means of a two-class classifier at the frame-level outperforms tackling the problem by only focusing on the majority class, *i.e.*, by means of the so-called OCC novelty detection approach, when enough representative data is obtained to model the minority class correctly.

Nevertheless, those techniques designed to minimize the unbalanced nature of the problem, such as reducing the RTN class samples or increasing the ANE set of the database [59], could be also considered in future works so as to improve the ANED classification performance. However, it is worth mentioning that oversampling the ANE subset synthetically has proven unrealistic to follow patterns in real-life environments (*e.g.*, considering the same type of acoustic sensors, recording conditions, events SNRs diversity, etc.). Whatever the solution adopted, it should include improvements in ANED reliability and performance in real-life operating conditions.

As for the machine learning approach, the current implementation of the ANED has considered the detection of ANEs as a two-class classification problem based on spectral parametrization and Gaussian Mixture Models [28]. However, other classic machine learning techniques could also have been considered. Among them, Deep Neural Networks (DNN) have recently gained relevance in acoustic event detection problems (*e.g.*, [40, 60] and those recently presented in DCASE competitions, see [45] and [46]). Nevertheless, a large amount of ANE examples should be gathered from real-life conditions to train DNN-based approaches appropriately, making this alternative a daunting task due to the nature of this kind of anomalous noise events. Finally, other approaches which consider not only the spectral but also the temporal evolution of the audio input, *e.g.* through spectro-temporal parameterization and/or classification schemes such as HMM could be considered, only if we can guarantee the real-time performance required by the problem at hand. Nevertheless, this research will be addressed for future work.

5.3 Real-world WASN operating challenges

Any WASN deployed to cover a dynamic acoustic mapping project presents a set of challenges associated with the hardware design and system maintenance. One of the key elements of a WASN is the inclusion of low-cost sensors, so as to ensure an affordable sensor network deployment, with a good trade-off between the cost of each sensor and the features the sensors offer (*e.g.*, being able to run the ANED and compute L_{Aeq} every second as in the DYNAMAP project).

The potential problems of low-cost technology are accuracy and reliability, as well as robustness over time. This last point is crucial both from an economic point of view and when considering the monitoring accuracy. On the one hand, the deployed WASN will have to cover maintenance costs, which should be minimized as much as possible. On the other hand, the noise map computation (and, thus, the ANED results) together with the subsequent action plans are based on the measurements of the low-cost acoustic sensor. If they stop working properly, any derived conclusions becomes inaccurate.

Moreover, the power supply of the devices is also an issue to take into account. Nowadays, most of the sensors deployed in the city are supplied by the public electrical network, but this limits the places where they can be located. One of the goals of the sensor devices design should be to minimize energy consumption and invest in affordable alternatives which can generate energy autonomously.

Furthermore, many cities already have a sensor network to capture other types of data, *e.g.*, CO_2, CO or air particle sensors, aggregating and transmitting the collected data through specific software platforms, as SENTILO [61], which is being used in Barcelona. Any new WASN design should take into account the compatibility issues with all those software platforms already processing data in smart cities, while no specific solution such as CityOS has been agreed as a general standard yet.

Data management is also a key issue in terms of design of a WASN. An acoustic network running in real time captures a large amount of raw data that should be processed properly to obtain the maximum information about the events and noise pollution occurring in urban and suburban areas. A deep study of the goals of the noise monitoring projects should be conducted in order to evaluate the impact of this question. From the DYNAMAP point of view, we have opted to process the data collected in each and every one of the sensors locally, and send the result to the central server, which is asked to subsequently tailor the GIS-based noise map. This way, the amount of data sent through the communication network is reduced dramatically with respect to centralized server-based WASN approaches (*e.g.*, Fi-Sonic project [26]), since the node only transmits the SensorID, the time stamp of the measure, the L_{Aeq} in dBA and the binary RTN/ANE label every second (see Figure 1). The map is then provided to both city management authorities and citizens through an intuitive web interface. However, other approaches may also prefer to register the acoustic environmental data in the cloud where it could be integrated with other city environmental sensors.

Finally, the deployment of WASN in urban environments can lead to a wide range of possible methods to analyze the acoustic events in the city. The acoustic sensors, further than evaluating the L_{Aeq} levels, can be used to run AED algorithms in order to detail the contents of the acoustic events in the street, such as: the typology of sounds identified (*e.g.*, man-made or coming from nature), the impact of these on the environmental noise, the subjective perceptual effects over the people living in some areas, etc. In any case, these algorithms should be designed to run in real-time within the sensors and to address the particularities of the real-life acoustic environments where they should operate.

6 Conclusions

In this work, we have described and reviewed the incremental development of the Anomalous Noise Event Detector envisioned by the DYNAMAP project to monitor the acoustic impact of road infrastructures in urban and suburban areas solely. From our research, we have proven the relevance of including an ANED to detect and avoid the impact of ANE on the road traffic noise map computation. Since both individual and aggregated ANEs can have a significant impact on the A-weighted equivalent noise level computation, their removal from the noise map generation provides the competent authorities with a reliable picture of the RTN affecting citizens.

The ANED has been designed and implemented as a two-stage binary classifier based on a two-class classifier at the frame-level followed by a majority voting scheme providing the required output decision every second. The frame-level classifier has been tested under simulated and real-life operating conditions in urban and suburban scenarios. The results obtained on real-life data have demonstrated the complexity (or even the unaffordability) of modelling the problem synthetically by means of building an artificially-mixed database composed of road traffic noise with superimposed anomalous noise events. Moreover, we have observed that the ANED is sensitive to the characteristics of the acoustic environment, entailing the suburban and the urban areas similar but different classification scenarios. Therefore, at least the ANED has to be trained with specific data for each acoustic environment where it is deployed. Nevertheless, we do not discard studying the potential improvements for adapting the ANED to the specific characteristics of the two clusters of Milan's district 9 in the near future.

Moreover, during the course of the project we have concluded that addressing the problem through a two-class classification approach outperforms the results obtained by the one-class classification alternative, when enough representative data are collected and labelled to train the minority class properly. Furthermore, we have observed that the implementation of the majority voting scheme acts as a post-processing stage that filters some spurious classification decisions of the ANED algorithm at the frame-level, thus, improving the performance of the detection system.

Nevertheless, as discussed, there are still several open questions that should be analyzed in detail. Among others, since the recording campaign only covered specific periods, the acoustic database is currently being completed by collecting new acoustic data from the sensors already deployed in the Rome and Milan pilot areas, mainly from night and weekend periods. Moreovoer, other problems derived from real-life operation have to be taken into account, as the huge amount of data to be processed when the WASN is working in real-time, and the accurate definition of the acoustic event detection algorithms to obtain valuable information. The accuracy of the data obtained using the low-cost acoustic sensors of the WASN and their robustness along time deserve special attention so as to obtain reliable conclusions from the road traffic noise monitoring network. Finally, we will continue to investigate on the ANED proposal by considering, for instance, other parametrization and/or classification schemes in order to improve its overall performance in real-life operating conditions.

Acknowledgement: The research presented in this work has been partially supported by the LIFE DYNAMAP project (LIFE13 ENV/IT/001254). Ferran Orga thanks the support of the European Social Fund and the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement of the Catalan Government for the predoctoral FI grant No. 2018FI_B100175. This work has been conducted in the framework of recognized as a Consolidated Research Group by the Government of Catalonia (ref 2017 SGR 966).

References

[1] Stephen A Stansfeld and Mark P Matheson. Noise pollution: non-auditory effects on health. British medical bulletin, 68(1):243–257, 2003.

[2] Valéria Regecová and Eva Kellerová. Effects of urban noise pollution on blood pressure and heart rate in preschool children. Journal of Hypertension, 13(4):405–412, 1995.

[3] Wolfgang Babisch, Bernd Beule, Marianne Schust, Norbert Kersten, and Hartmut Ising. Traffic noise and risk of myocardial infarction. Epidemiology, 16(1):33–40, 2005.

[4] Lisa Goines and Louis Hagler. Noise pollution: a modern plague. Southern Medical Journal, 100(3):287–294, 2007.

[5] World Health Organization and the Joint Research Centre of the European Commission. Burden of disease from environmental noise: Quantification of healthy life years lost in Europe. pages 1–126. World Health Organization. Regional Office for Europe, 2011.

[6] Natalie Mueller, David Rojas-Rueda, Xavier Basagaña, Marta Cirach, Tom Cole-Hunter, Payam Dadvand, David Donaire-Gonzalez, Maria Foraster, Mireia Gascon, David Martínez, Cathryn C. Tonne, Margarita Triguero-Mas, Antonia Valentín-Crespo, and Mark Nieuwenhuijsen. Health impacts related to urban and transport planning: A burden of disease assessment. Environment international, 107:243–257, 2017.

[7] E.U. 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, L 189/12, 2002.

[8] Stylianos Kephalopoulos, Marco Paviotti, and Fabienne Anfosso-Lédée. Common noise assessment methods in Europe (CNOSSOS-EU). Publications Office of the European Union, Report EUR 25379 EN:1–180, 2002.

[9] D.W. Morley, K. de Hoogh, D. Fecht, F. Fabbri, M. Bell, P.S. Goodman, P. Elliott, S. Hodgson, A.L. Hansell, and J. Gulliver. International scale implementation of the CNOSSOS-EU road traffic noise prediction model for epidemiological studies. Environmental pollution, 206:332–341, 2015.

[10] Alejandro Ruiz-Padillo, Diego P. Ruiz, Antonio J. Torija, and Ángel Ramos-Ridao. Selection of suitable alternatives to reduce the environmental impact of road traffic noise using a fuzzy multi-criteria decision model. Environmental Impact Assessment Review, 61:8–18, 2016.

[11] Gaetano Licitra, Luca Fredianelli, Davide Petri, and Maria Angela Vigotti. Annoyance evaluation due to overall railway noise and vibration in pisa urban areas. Science of the total environment, 568:1315–1325, 2016.

[12] Fernando Bunn and Paulo Henrique Trombetta Zannin. Assessment of railway noise in an urban setting. Applied Acoustics, 104:16–23, 2016.

[13] Paolo Gagliardi, Luca Fredianelli, Duccio Simonetti, and Gaetano Licitra. ADS-B System as a Useful Tool for Testing and Redrawing Noise Management Strategies at Pisa Airport. Acta Acustica united with Acustica, 103(4):543–551, 2017.

[14] Carlos Iglesias-Merchan, Luis Diaz-Balteiro, and Mario Soliño. Transportation planning and quiet natural areas preservation: Aircraft overflights noise assessment in a national park. Transportation Research Part D: Transport and Environment, 41:1–12, 2015.

[15] Stylianos Kephalopoulos, Marco Paviotti, Fabienne Anfosso-Lédée, Dirk Van Maercke, Simon Shilton, and Nigel Jones. Advances in the development of common noise assessment methods in Europe: The CNOSSOS-EU framework for strategic environmental noise mapping. Science of the Total Environment, 482:400–410, 2014.

[16] Julien Morel, Catherine Marquis-Favre, and Laure-Anne Gille. Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise: A laboratory study. Applied Acoustics, 101:47–57, 2016.

[17] P. Gallo, L. Fredianelli, D Palazzuoli, G. Licitra, and F. Fidecaro. A procedure for the assessment of wind turbine noise. Applied Acoustics, 114:213–217, 2016.

[18] David S. Michaud, Katya Feder, Stephen E. Keith, Sonia A. Voicescu, Leonora Marro, John Than, Mireille Guay, Allison Denning, D'Arcy McGuire, Tara Bower, Eric Lavigne, Brian J. Murray, Shelly K. Weiss, and Frits van den Berg. Exposure to wind turbine noise: Perceptual responses and reported health effects. The Journal of the Acoustical Society of America, 139(3):1443–1454, 2016.

[19] Júlia Camps. Barcelona noise monitoring network. In Proceedings of EuroNoise 2015, pages 2315 – 2320, Maastrich, Netherlands, 31 May - 3 June 2015. EAA-NAG-ABAV.

[20] Luca Nencini, Paolo De Rosa, Elena Ascari, Bruna Vinci, and Natalia Alexeeva. SENSEable Pisa: A wireless sensor network for real-time noise mapping. In Proceedings of the EuroNoise, Prague, Czech Republic, pages 10–13, 2012.

[21] Chiara Bartalucci, Francesco Borchi, Monica Carfagni, Rocco Furferi, and Lapo Governi. Design of a prototype of a smart noise monitoring system. In Proc. of the 24th International Congress on Sound and Vibration (ICSV24), July 2017.

[22] D Rainham. A wireless sensor network for urban environmental health monitoring: Urbansense. In IOP Conference Series: Earth and Environmental Science, volume 34(1), page 012028. IOP Publishing, 2016.

[23] Srimanta Gupta and Chitralekha Ghatak. Environmental noise assessment and its effect on human health in an urban area. Int. J. Environ. Sci., 1:1954–64, 2011.

[24] Bert De Coensel and Dick Botteldooren. Smart sound monitoring for sound event detection and characterization. In 43rd International Congress on Noise Control Engineering (Inter-Noise 2014), 2014.

[25] Stavros Ntalampiras, Ilyas Potamitis, and Nikos Fakotakis. Probabilistic novelty detection for acoustic surveillance under real-world conditions. IEEE Transactions on Multimedia, 13(4):713–719, 2011.

[26] J. Paulo, Pedro Fazenda, Tiago Oliveira, and João Casaleiro. Continuos sound analysis in urban environments supported by fiware platform. Proceedings of the EuroRegio2016/TecniAcústica'16, pages 1–10.

[27] Jiaxing Ye, Takumi Kobayashi, and Masahiro Murakawa. Urban sound event classification based on local and global features aggregation. Applied Acoustics, 117, Part B:246 – 256, 2017. Acoustics in Smart Cities.

[28] Joan Claudi Socoró, Francesc Alías, and Rosa Ma Alsina-Pagès. An anomalous noise events detector for dynamic road traffic noise mapping in real-life urban and suburban environments. Sensors, 17(10):2323, 2017.

[29] Panu Maijala, Zhao Shuyang, Toni Heittola, and Tuomas Virtanen. Environmental noise monitoring using source classification in sensors. Applied Acoustics, 129:258–267, 2018.

[30] Xavier Sevillano, Joan Claudi Socoró, Francesc Alías, Patrizia Bellucci, Laura Peruzzi, Simone Radaelli, Paola Coppi, Luca Nencini, Andrea Cerniglia, Alessandro Bisceglie, Roberto Benocci, and Giovanni Zambon. DYNAMAP - Development of low cost sensors networks for real time noise mapping. Noise Mapping, 3:172–189, May 2016.

[31] European Commission Working Group of the Exposure to Noise (WG-AEN). Good practice guide for strategic noise mapping and the production of associated data on noise exposure. Position Paper, Ver.2, pages 1–129, 2007.

[32] Ferran Orga, Francesc Alías, and Rosa Ma Alsina-Pagès. On the impact of anomalous noise events on road traffic noise mapping in urban and suburban environments. International Journal of Environmental Research and Public Health, 15(1):13, 2018.

[33] Rosa Ma Alsina-Pagès, Joan Claudi Socoró, and Francesc Alías. Detecting Anomalous Noise Events on Low-Capacity Acoustic Sensor in Dynamic Road Traffic Noise Mapping. In Multidisciplinary Digital Publishing Institute Proceedings, volume 2, page 136, 2018.

[34] Cense - characterization of urban sound environments. http://cense.ifsttar.fr/. Accessed: 2017-09-14.

[35] Fanny Mietlicki, Christophe Mietlicki, and Matthieu Sineau. An innovative approach for long-term environmental noise measurement: Rumeur network. In Proceedings of EuroNoise 2015, pages 2309 – 2314, Maastrich, Netherlands, 31 May - 3 June 2015. EAA-NAG-ABAV.

[36] Dick Botteldooren, Bert De Coensel, Damiano Oldoni, Timothy Van Renterghem, and Samuel Dauwe. Sound monitoring networks new style. In David J. Mee and Ian D.M. Hillock, editors, Acoustics 2011: Breaking New Ground: Proceedings of the Annual Conference of the Australian Acoustical Society, pages 93:1–93:5, Queensland, Australia, 2-4 November 2011. Australian Acoustical Society.
[37] Margaret Carol Bell and Fabio Galatioto. Novel wireless pervasive sensor network to improve the understanding of noise in street canyons. Applied Acoustics, 74(1):169–180, 2013.
[38] Charlie Mydlarz, Justin Salamon, and Juan Pablo Bello. The implementation of low-cost urban acoustic monitoring devices. Applied Acoustics, 117:207–218, 2017.
[39] Chiara Bartalucci, Francesco Borchi, Monica Carfagni, Rocco Furferi, Lapo Governi, Alessandro Lapini, Yary Volpe, Salvatore Curcuruto, Enrico Mazzocchi, Giuseppe Marsico, et al. Life monza: project description and actions' updating. Noise Mapping, 5(1):60–70.
[40] Yasutaka Nakajima, Masahiro Sunohara, Taisuke Naito, Norihito Sunago, Toshiya Ohshima, and Nobutaka Ono. DNN-based Environmental Sound Recognition with Real-recorded and Artificially-mixed Training Data. In Proc. 45th International Congress and Exposition on Noise Control Engineering (InterNoise 2016), pages 3164–3173, Hamburg, Germany, August 2016. German Acoustical Society (DEGA).
[41] P. Foggia, N. Petkov, A. Saggese, N. Strisciuglio, and M. Vento. Audio surveillance of roads: A system for detecting anomalous sounds. IEEE Transactions on Intelligent Transportation Systems, 17(1):279–288, January 2016.
[42] Francesc Alías, Joan Claudi Socoró, and Xavier Sevillano. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. Applied Sciences, 6(5):143, 2016.
[43] Francesc Alías and Joan Claudi Socoró. Description of anomalous noise events for reliable dynamic traffic noise mapping in real-life urban and suburban soundscapes. Applied Sciences, 7(2):146, 2017.
[44] Joan Claudi Socoró, Gerard Ribera, Xavier Sevillano, and Francesc Alías. Development of an anomalous noise event detection algorithm for dynamic road traffic noise mapping. In Proc. of the 22nd International Congress on Sound and Vibration (ICSV22), pages 1–8, Florence, Italy, July 2015. The International Institute of Acoustics and Vibration (IIAV).
[45] Annamaria Mesaros, Toni Kristian Heittola, Emmanouil Benetos, Peter Foster, Mathieu Lagrange, Tuomas Virtanen, and Mark D Plumbley. Detection and Classification of Acoustic Scenes and Events: Outcome of the DCASE 2016 Challenge. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017.
[46] Annamaria Mesaros, Toni Heittola, Aleksandr Diment, Benjamin Elizalde, Ankit Shah, Emmanuel Vincent, Bhiksha Raj, and Tuomas Virtanen. DCASE 2017 challenge setup: Tasks, datasets and baseline system. In DCASE 2017-Workshop on Detection and Classification of Acoustic Scenes and Events, 2017.
[47] Andrey Temko, Robert Malkin, Christian Zieger, Dušan Macho, Climent Nadeu, and Maurizio Omologo. CLEAR Evaluation of Acoustic Event Detection and Classification Systems, pages 311–322. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
[48] Marco A.F. Pimentel, David A. Clifton, Lei Clifton, and Lionel Tarassenko. A review of novelty detection. Signal Processing, 99:215–249, 2014.
[49] Francesco Aurino, Mariano Folla, Francesco Gargiulo, Vincenzo Moscato, Antonio Picariello, and Carlo Sansone. One-class SVM based approach for detecting anomalous audio events. In Intelligent Networking and Collaborative Systems (INCoS), 2014 International Conference on, pages 145–151. IEEE, 2014.
[50] Stavros Ntalampiras. Universal background modeling for acoustic surveillance of urban traffic. Digital Signal Processing, 31:69–78, 2014.
[51] Paul Mermelstein. Distance measures for speech recognition, psychological and instrumental. Pattern recognition and artificial intelligence, 116:374–388, 1976.
[52] X. Valero and F. Alías. Gammatone cepstral coefficients: Biologically inspired features for non-speech audio classification. IEEE Transactions on Multimedia, 14(6):1684–1689, 2012.
[53] Italian National Institute of Statistics. Demographic balance for the year 2016, 2017.
[54] Simone Radaelli. Basic noise maps calculation in milan pilot area. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings, volume 253, pages 6401–6409. Institute of Noise Control Engineering, 2016.
[55] Giovanni Zambon, Roberto Benocci, Alessandro Bisceglie, H Eduardo Roman, and Patrizia Bellucci. The life dynamap project: Towards a procedure for dynamic noise mapping in urban areas. Applied Acoustics, 124:52–60, 2017.
[56] Patrizia Bellucci and Francesca Romana Cruciani. Implementing the dynamap system in the suburban area of rome. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings, volume 253, pages 5518–5529. Institute of Noise Control Engineering, 2016.
[57] Joan Claudi Socoró, Xavier Albiol, Xavier Sevillano, and Francesc Alías. Analysis and automatic detection of anomalous noise events in real recordings of road traffic noise for the life dynamap project. In Proc. 45th International Congress and Exposition on Noise Control Engineering (InterNoise 2016), pages 6370–6379, Hamburg,Germany, August 2016. German Acoustical Society (DEGA).
[58] Giovanni Zambon, Hector Eduardo Roman, Maura Smiraglia, and Roberto Benocci. Monitoring and prediction of traffic noise in large urban areas. Applied Sciences, 8(2):251, 2018.
[59] Yanmin Sun, Andrew KC Wong, and Mohamed S Kamel. Classification of imbalanced data: A review. International Journal of Pattern Recognition and Artificial Intelligence, 23(04):687–719, 2009.
[60] Oguzhan Gencoglu, Tuomas Virtanen, and Heikki Huttunen. Recognition of acoustic events using deep neural networks. In Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European, pages 506–510. IEEE, 2014.
[61] M Bain. Sentilo—Sensor and Actuator Platform for Smart Cities, 2014.