Science of the Total Environment 992 (2025) 179945

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Noise pollution monitoring at pedestrian level by autonomous vehicles in urban areas

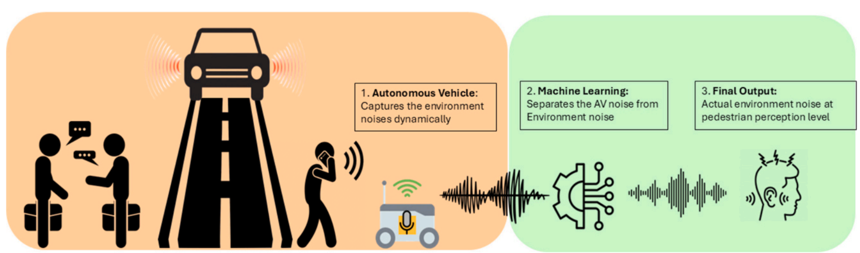
Behzad Ajdari [a,*], Nikoo Salimi [a], Lucanos Strambini [b,c], Elvezia Maria Cepolina [d]

[a] University of Genoa, DIEC, Genoa, Italy
[b] CNR - Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, Pisa, Italy
[c] CNR ICSC - Interdepartmental Centre for Cities Science, Italy
[d] University of Genoa, DIME, Genoa, Italy

HIGHLIGHTS

- Autonomous vehicles enable urban areas dynamic noise monitoring at pedestrian height.
- Machine learning helps predict and isolate vehicle noise for accurate assessment.
- The system outputs actual environmental noise at pedestrian perception level.
- Tested across three urban areas, achieving high prediction accuracy.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Pavlos Kassomenos

Keywords:
Environmental noise monitoring
Noise exposure
Mobile noise sensors
Small ground autonomous vehicle
Pedestrians
Mobility
Urban areas

ABSTRACT

This study presents a mobile noise monitoring system designed to assess pedestrian-level environmental noise using an autonomous ground vehicle (Yape) equipped with a calibrated sound level sensor. Unlike fixed monitoring stations or simulation-based models, our approach enables dynamic, high-resolution data collection along entire urban areas. Key challenges included ensuring that recorded noise accurately reflected pedestrian exposure and isolating the vehicle's own acoustic emissions from ambient noise. To address this, we collected training data by positioning a calibrated reference sensor at predetermined distances from the Yape and developed a machine learning model to estimate true environmental noise levels. The model achieved a high predictive accuracy ($R^2 = 0.94$, RMSE = 1.3 dBA), demonstrating reliable separation of Yape-generated noise from environmental signals. The system was tested across three distinct streets in Genoa, Italy, and captured a wide range of urban acoustic profiles. This method offers a scalable solution for urban noise mapping and has potential applications in smart city planning, pedestrian comfort analysis, and real-time environmental monitoring.

1. Introduction

Noise pollution is a major public health issue, yet it often goes unnoticed due to its delayed health effects. The World Health Organization (WHO) has ranked noise pollution as one of the main environmental problems in Europe (WHO Regional Office for Europe, 2011). Several clinical studies have shown that even low-level environmental noise can, over time, cause problems ranging from mild attention and

* Corresponding author.
E-mail address: behzad.ajdari@edu.unige.it (B. Ajdari).

https://doi.org/10.1016/j.scitotenv.2025.179945
Received 13 February 2025; Received in revised form 24 May 2025; Accepted 15 June 2025
Available online 26 June 2025

concentration issues to serious cardiac and neurological conditions (Babisch, 2005; Stanovská et al., 2024). The lack of attention by public administrations to the issue of noise pollution has coincided with rapid technological development in urban areas, such as the increase in motorized traffic, the rise of on-demand delivery services using combustion-engine scooters, and the expansion of construction activities. These changes have contributed to a worsening of exposure to environmental noise, both in terms of sources and levels. As an example, the European Commission's Environmental Noise Directive (2024) reports that we now have 20 % of the EU population living in areas where noise levels are considered harmful to health.

Although the topic of noise pollution seems to have attracted growing interest in recent years, forecasts today indicate a potential worsening of the problem with a prospect of a 3 % increase in the population that will suffer the effects of environmental noise by 2030, even if actions are adopted to mitigate and contain the sources of environmental noise (European Environment Agency (EEA), 2024).

The Outlooks suggest that, without further measures, it would be difficult to reduce the number of people disturbed by environmental noise, particularly transport noise, significantly by 2030 and would require regulatory or even legislative changes.

The threshold for excessive exposure, according to the EU, is 55 dB Lden (level day-evening-night) (Environmental Noise Directive, 2024). Lden refers to an A-weighted average sound pressure level over a whole year, accounting for days, evenings, and nights, with evenings including a weighting of 5 dB and nights including a weighting of 10 dB.

Authorities must compare this threshold to actual population exposure levels. In practice, noise levels (Lden) are traditionally mapped statically across different locations using noise maps, which are then overlaid with population density data. Noise mapping is especially important in urban areas where the population density, and therefore the number of people exposed to noise pollution and its effects on both physical and mental health, is higher (Clark et al., 2020; WHO Regional Office for Europe, 2011). Noise maps represent the spread and intensity of noise in different city areas; therefore, they can only pinpoint what city areas exceed allowable noise limits and what mitigation actions can be taken into consideration to reduce noise levels.

Noise maps show sound levels recorded by fixed-location monitoring sensors and, where sensor data are unavailable, estimated levels based on average traffic volumes, vehicle types, and road surface conditions (Khajehvand et al., 2021), since road traffic is the primary source of urban noise.

Dynamic noise maps, an advanced version of traditional noise maps, capture how noise levels vary over time, which improves their value for regulation. Estimating noise emissions on roads without sensors is a key challenge in creating dynamic traffic noise maps, particularly due to the limited availability of monitoring devices (Lan and Cai, 2021). The prevailing approach to this challenge uses total daily traffic flow as a non-acoustic parameter related to road traffic to estimate noise emissions on non-monitored roads (Benocci et al., 2019; Lan et al., 2020; Zambon et al., 2017). In addition to traffic flow, traffic speed significantly influences traffic noise. High-resolution road segment speed data, often available from digital maps, can also be used to dynamically update noise maps (Lan and Cai, 2021).

Many previous studies have focused primarily on modeling the physical intensity of traffic noise, often using metrics such as Leq or Lden, without directly considering the dynamics of human exposure or the variability in individual sensitivity to noise (e.g., Zambon et al., 2017; Benocci et al., 2019; Lan and Cai, 2021). Researchers have started to understand the distribution of the human population over time and space and how these factors could contribute to noise pollution impacts. A new method was also introduced to evaluate urban traffic noise exposure in this regard by considering the variations in exposure based on the population profile. The model accounts for the distribution of people with respect to their ages and specific kinds of acoustic needs, enabling a finer analysis of the effects of noise on those groups (Wang et al., 2022).

This research shifts the focus from average population exposure to the noise experienced by individuals. Noise thresholds should reflect the levels a person encounters during daily activities across different locations and times (day, evening, and night). As a result, noise monitoring is no longer associated with a fixed location but is linked to an urban mobility experience. While short-term exposure to high-intensity noise, such as in occupational or industrial environments, may cause temporary auditory effects like threshold shifts, pedestrian exposure to continuous traffic noise in urban settings is generally below such thresholds. However, chronic exposure to moderate environmental noise remains a public health concern.

Moreover, the noise to which a person is exposed in a day is the result of noise emitted by all sources, which includes, but is not limited to, road traffic. However, the number of people exposed to noise from road traffic significantly surpasses those affected by noise originating from railways, aviation, or industrial activities. This pattern is consistent across Europe, in both metropolitan and rural areas (EEA, 2024).

This study proposes tracking daily urban noise exposure—similar to occupational noise monitoring—using a small autonomous vehicle equipped with an acoustic sensor to measure ambient noise and estimate individual exposure, both indoors and outdoors. The dynamic sensor calculates individual-level exposure metrics, such as the number of noise events a person experiences—data critical for assessing health impacts (Lex Brown, 2014).

The proposed system can detect the total noise from various sources to which a person is exposed while moving through an urban environment, as well as the noise detected by the vehicle itself, which may differ from surrounding ambient noise.

Recent research has highlighted the growing importance of non-transportation noise sources in urban environments, such as construction, nightlife, and technical systems (Michaud et al., 2022; Tong et al., 2021; Zambon et al., 2020). These findings underscore the need for comprehensive noise monitoring approaches that capture diverse sources across space and time. This study introduces an autonomous ground vehicle equipped with acoustic sensing and machine learning to assess total environmental noise exposure along pedestrian routes. To the best of our knowledge, this is the first study to integrate such a mobile system with automated data capture and prediction tools to assess pedestrian-level environmental noise exposure across complete urban trajectories.

This study addresses two main research questions:

a. Noise levels can vary significantly depending on the sensor's height from the ground. Ideally, the sensor should be positioned at human ear level to best represent real-world exposure. However, this may conflict with certain design or functional requirements of the autonomous vehicle.

R&D Question 1: What is the optimal sensor height from the ground on a small autonomous vehicle to minimize interference from the vehicle's own noise while accurately capturing ambient sound experienced by pedestrians?

b. Second key challenge is distinguishing the noise generated by the autonomous vehicle itself from the surrounding environmental noise. Accurately isolating the vehicle's acoustic emissions is essential to avoid distortion in the recorded data.

R&D Question 2: How can environmental noise levels be predicted from onboard sensor data while accounting for the noise produced by the vehicle in motion?

The paper is organized as follows: Section 2 describes the materials and methods, including the YAPE ecosystem. It addresses the first R&D question by identifying the optimal sensor height. Moreover, Section 2 details the design and methodology of the prototype system, focusing on the second R&D question regarding the estimation of environmental

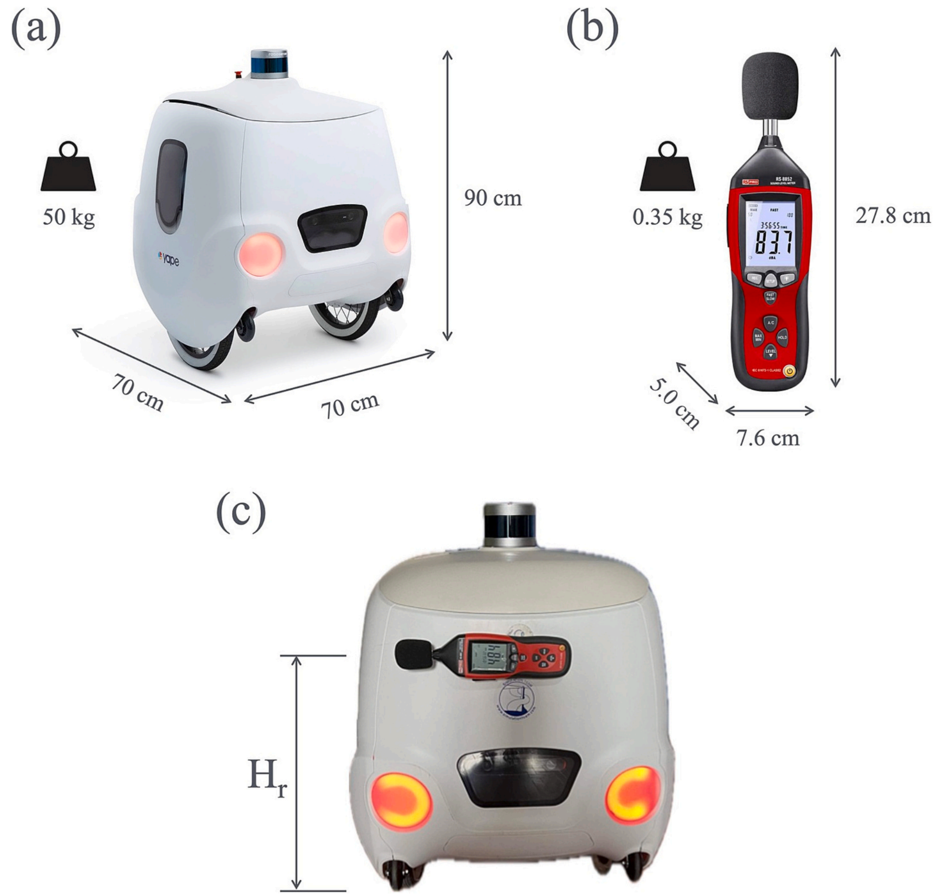


Fig. 1. Image and size characteristics of the: a) Yape droid, b) noise sensor (RS Pro DT-8852), and c) prototype of the innovative system developed for the outdoor monitoring of the noise pollution at pedestrian level.

noise pollution levels while accounting for the noise contribution of autonomous vehicles. Noise pollution in urban environments is analyzed in Section 3. Section 4 concludes the study by summarizing the key findings and implications for future research.

2. Material and methods

This section outlines the components and methodology of the system used to measure pedestrian-level noise exposure in urban environments. The core of the system is the YAPE autonomous vehicle equipped with an acoustic sensor to collect noise data along city routes. A combination of experimental and data-driven approaches was employed to calibrate the sensor placement, isolate the autonomous vehicle's noise emissions, and assess the environmental noise experienced by pedestrians. Technical specifications, sensor setup details, and machine learning implementation are provided in Appendices A to D. What follows is a focused summary of the system architecture, key measurement strategies, and the framework used to assess urban noise exposure.

2.1. YAPE system and sensor setup

The system is built on the Yape autonomous ground vehicle, a small self-driving platform capable of navigating urban pedestrian environments. Yape was adapted to carry a noise sensor for the purpose of monitoring acoustic exposure along walking routes. Its mobility allows for coverage across a wide range of public spaces, including streets, sidewalks, and open areas.

The noise sensor installed on the vehicle was selected for its suitability in urban environmental monitoring. It captures A-weighted sound levels relevant to human perception and health standards. The goal is to document the total noise pedestrians are exposed to, including contributions from traffic, human activity, construction, and other urban sources.

Fig. 1 shows the Yape and the RS Pro 8852 sound level meter used in the experiment. As illustrated in Fig. 1c, the sound level meter was externally mounted on the front side of the Yape, with the microphone fully exposed to the ambient environment. Full technical specifications of both the Yape and sensor are provided in Appendix A.

2.2. Sensor positioning methodology

The primary goal of this study was to measure the noise levels experienced by pedestrians in urban outdoor settings. Since the average height of a human ear is approximately 1.6 m from the ground, we needed to determine the optimal positioning of the noise sensor on the Yape to accurately capture this experience.

Two potential approaches were considered: (1) mounting the sensor at the human ear height of 1.6 m, which would present significant challenges for Yape stability and maneuverability; or (2) mounting the sensor at a more practical height on the Yape and based on previous studies (Montes González et al., 2020) using correction factors if necessary. We conducted controlled experiments to determine if the latter approach could provide sufficiently accurate measurements.

Experiments were designed to measure sound level variations between two sensor heights from the ground (50 cm and 160 cm) at varying distances from sound sources. Two common urban noise types were tested: human speech (calibrated to 60 dBA at 1 m) and traffic noise (calibrated to 70 dBA at 1 m). Measurements were taken at horizontal distances of 1, 2, and 3 m from each sound source in a controlled indoor environment (Detailed specifications of the experimental setup,

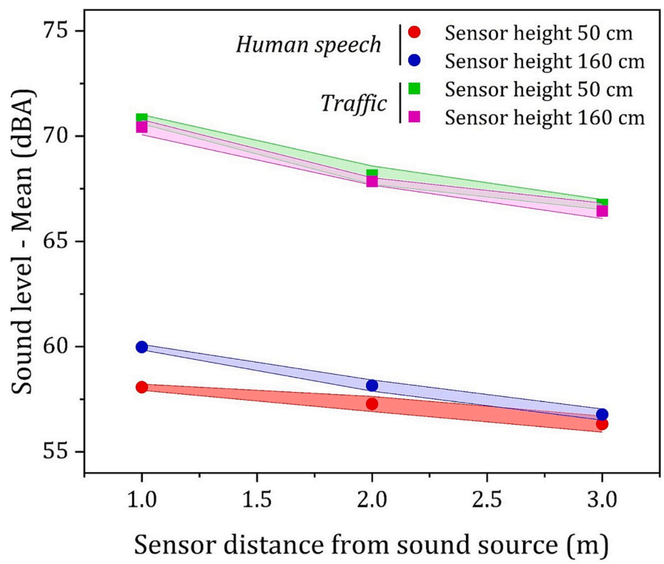


Fig. 2. Sound level variation by changing distance of sensor from the source: mean values are presented as dot/square while related colored bands refer to the standard deviation (STD) values.

including acoustic source characteristics, audio stimuli specifications, and environmental controls, are provided in Appendix B).

The results, shown in Fig. 2, demonstrate that sound level differences between the two sensor heights from the ground were minimal, particularly as distance from the source increased.

As shown in Fig. 2, mean sound levels (dBA) were recorded at sensor heights of 50 cm and 160 cm for both human speech and traffic noise, at distances of 1, 2, and 3 m. Red and blue circles represent human speech at 50 cm and 160 cm, respectively; green and magenta squares represent traffic noise at the same heights. Shaded areas indicate standard deviations. For human speech, the largest difference (1.9 dBA) between sensor heights occurs at a distance of 1 m and decreases with distance, reflecting ground reflection effects. For traffic noise, variations between heights are minimal due to the source height (100 cm) being close to the midpoint between sensors. All differences fall well below the 3 dBA threshold generally considered the minimum difference detectable by the average human listener in real-world settings (Bies et al., 2018; Fastl and Zwicker, 2007). To further evaluate whether ground surface type affects recorded noise levels, we conducted additional outdoor experiments across two pedestrian surfaces: mosaic tiles, modular plastic flooring. Across all conditions, the measured differences remained consistently below 3 dBA. These findings indicate that minor surface material variations within urban pedestrian environments have a limited effect on sound propagation. This result aligns with (Attenborough, 2002), who notes that acoustically hard surfaces such as concrete or asphalt exhibit similar impedance characteristics and produce minimal differences in ground-induced attenuation, especially at

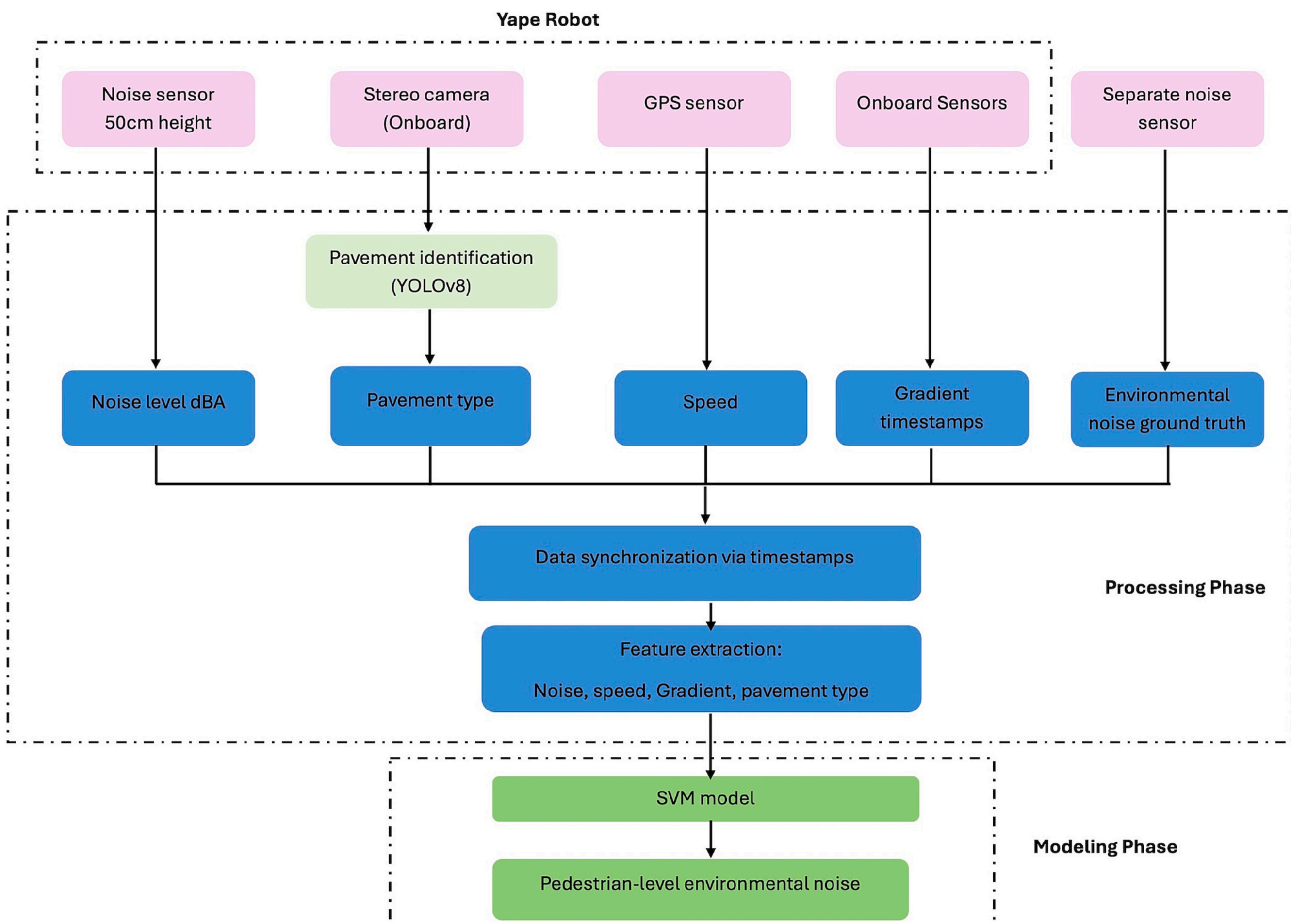


Fig. 3. Overview of the data processing pipeline for estimating environmental noise using the Yape robot. Inputs from noise, image, and motion sensors are synchronized to extract features such as pavement type, speed, and slope, which isolates the ambient environmental noise level from the robot's own contributions.

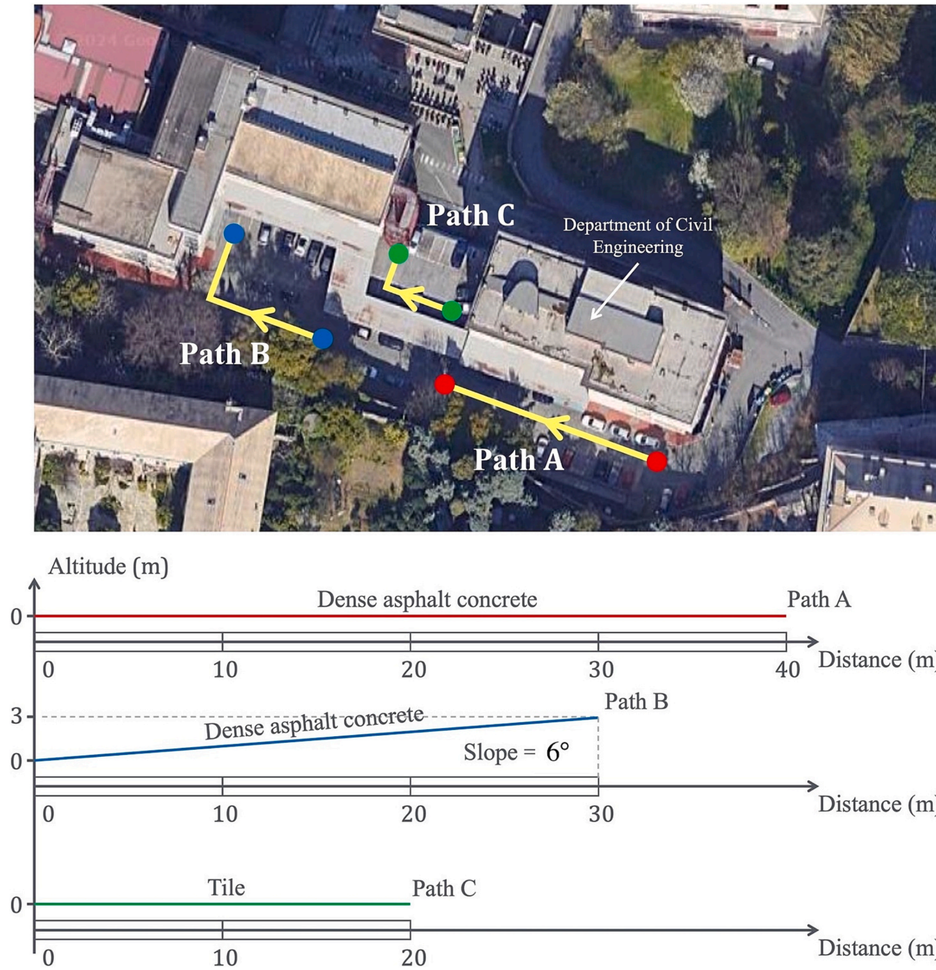


Fig. 4. Plan view and features of the three outdoor test path defined: two path on dense asphalt concrete, path A (red dot) flat road and path B (blue dot) road with a slope of 10 % (6°), and tiled path C (green dot) flat road.

frequencies relevant to urban environmental noise.

Based on these findings, we determined that mounting the sensor at 50 cm height from the ground on the Yape would provide measurements sufficiently representative of pedestrian-level noise exposure without requiring complex physical modifications to the Yape or mathematical correction factors. The sensor was therefore positioned on Yape as shown in Fig. 1c. The microphone was consistently front-mounted and exposed throughout all experiments, meaning the recorded noise levels may have been influenced by the vehicle's direction of travel relative to surrounding noise sources.

This approach provides a practical solution that balances measurement accuracy with operational feasibility. The slight overestimation of noise levels at the 50 cm height (compared to 160 cm) was deemed acceptable given its minimal magnitude and the advantages of maintaining the Yape's stability and maneuverability.

2.3. Environmental noise modeling strategy

To quantify the level of environmental noise that pedestrians are exposed to, it is first necessary to separate the noise generated by the Yape mobile sensing platform from the ambient urban soundscape. This section outlines the methodology developed to achieve this goal. The approach involves identifying the primary explanatory variables that influence noise emissions from the Yape and using these variables in a machine learning framework to isolate the surrounding environmental noise. The process integrates sensor data from GPS, onboard systems, and a stereo camera, and aligns all data streams through timestamp synchronization to form a structured dataset suitable for analysis. A schematic overview of the sensing and processing architecture is shown in Fig. 3.

2.4. Identification of the explanatory variables for the noise emitted by Yape

Traffic noises are caused by many sources, inclusive of:

1. Power train noise. For conventional vehicles, this includes engine, transmission, and exhaust system components. Exhaust noise, in particular, is a major contributor to power train noise in internal combustion engine vehicles, especially during acceleration or high load. In contrast, electric vehicles produce significantly less power train noise. Their electric motors are relatively quiet, and the dominant source of vehicle noise in urban settings becomes the tyre-road interaction. Yape, being an electric vehicle, emits negligible power train noise from its motor.
2. Gear train noise. Yape's motor is not only used to move Yape but also to keep it balanced on its two wheels, so noise due to it is generated even in standby conditions. As it concerns the gear train, since Yape has only one gear, it has no clutch or gearbox. Noises generated by gear trains are, therefore, not present in Yape.
3. Noise caused by aerodynamic turbulence. Given the low operating speed of Yape, wind noise is minimal.
4. Interaction between tyre and road surface. This is the most important origin of traffic noise. Such interaction produces two major

Table 1
Outdoor scenarios analyzed in the controlled environment.

Path	Road pavement surface material	Road gradient	Modes
A	dense asphalt concrete	0° flat road	(1)
A	dense asphalt concrete	0° flat road	(2)
A	dense asphalt concrete	0° flat road	(3)
B	dense asphalt concrete	6° slope	(1)
B	dense asphalt concrete	6° slope	(2)
B	dense asphalt concrete	6° slope	(3)
C	Tiles	0° flat road	(1)
C	Tiles	0° flat road	(2)
C	Tiles	0° flat road	(3)

mechanisms responsible for external noise emissions (Sandberg and Ejsmont, 2002):

- vibro-dynamic noise produced by vibrations of the tyres resulting from the contact of the tread with the road surface, often referred to as structure-borne noise;
- aerodynamic noise caused by compression and subsequent decompression of air trapped between the tyre tread and the road surface.

We, therefore, need to monitor, instant by instant, the environmental factors that determine the vibration of Yape's electric motor and the type of tyre/road pavement interaction.

Vibration from the engine is a function of Yape's speed and road gradient, while tyre/road pavement interaction is mainly due to road pavement characteristics. We have therefore considered the following values for these explanatory variables on which the sound emitted by Yape depends:

1. The speed of the Yape:
   a. Stand-by condition: 0 km/h
   b. Operational speed: an average of 5–6 km/h
2. Road gradient:
   a. 0° (flat road)
   b. 6° slope, which is the maximum incline Yape can handle smoothly
3. Road pavement surface material
   a. Tiles
   b. A mix of dense asphalt concrete (0/11) and stone mastic asphalt (0/11)

The noise emitted by Yape is considered to be uniquely attributable to the present value of these variables. Other parameters have been neglected. Acceleration and deceleration of the vehicle are very low since the maximum Yape speed is very low. The reference conditions are: dry road surface and an air temperature of 20 °C.

The target was to identify the contribution of Yape to the noise levels recorded by the sensor. To accomplish this, each experiment was conducted in three Modes:

1. sensor on Yape while it was moving;
2. sensor on Yape while it was in standby mode without moving (the Yape's engine is turned on but not moving);
3. sensor resting on a stationary stand at the same height from the ground as it had when mounted on Yape (the Yape's engine is turned off).

The difference between the noise levels recorded in the moving mode (1) and the standby mode (2) was assumed to be the noise contribution due to Yape's movement, while the difference between the standby mode (2) and mode (3) was attributed to Yape's presence and operation in a non-moving state.

2.5. Experiments to test the impact of the identified explanatory variables

The experiments have been run in off-peak hours so that there are no other active noise sources in the environment and only the noise generated by Yape is recorded. The experiments were conducted within the Opera Pia campus of the University of Genoa, which provided a controlled and representative urban environment to assess Yape's noise in different scenarios.

Three different paths were defined (A, B, and C), as shown in Fig. 4. The paths were typically 20–40 m long and were laid out on both asphalt and tile. The characteristics of the paths are shown in Table 1, together with the motion conditions of the Yape in the relevant experiments: mode (1) and mode (2). Six scenarios have been, therefore, identified to take into account different combinations of values of the explanatory variables identified above. For each scenario, the noise level was also measured in mode (3).

Four replications of the experiment were performed in each scenario. The number of replicates for the noise measurement experiment was determined based on guidelines from established standards and literature in the field of acoustics and noise measurement. The International Standard ISO 9613-2:2024 "Acoustics - Attenuation of sound during propagation outdoors - Part 2: General method of calculation" recommends taking at least three measurements at each position to account for temporal fluctuations in the noise source and environmental conditions. Similarly, the ANSI S12.9-1988/Part 2 "Quantities and Procedures for Description and Measurement of Environmental Sound, Part 2: Measurement of Long-term, Wide-area Sound" suggests taking multiple measurements, with a minimum of three replicates, to improve reliability and account for variations over time. In line with these recommendations, four replicates were conducted to ensure the accuracy.

The average value over the four replications of the noise levels recorded by the sensor was then noted for each scenario. In Table 2, the average noise level, measured in dBA, refers to Modes (1) and (2), while the average noise level of the Environment, measured again in dBA, refers to measurements recorded in Mode (3).

From the analysis of Fig. 5 and Table 2, it can be deduced the following:

In Fig. 5, the vertical lines above each bar represent 95 % confidence intervals, calculated from repeated noise measurements under each experimental condition. These intervals represent the 95 % confidence range for the true mean sound level.

1. Effect of the Yape's speed: the noise level differences were consistently higher between the Yape moving at an average speed of 5 km/

Table 2
Experimental results.

Route	Surface	Slope	Avg speed km/h	Avg noise level (dBA)	Avg noise of environment dBA	Difference
A (outdoor)	Asphalt	0°	5	61.4	51	10.4
A (outdoor)	Asphalt	0°	0	59.1		8.1
B (outdoor)	Asphalt	6° (ramp)	5	63.5	50.5	13
B (outdoor)	Asphalt	6° (ramp)	0	58.8		8.3
C (outdoor)	Tile	0°	5	59.5	50.1	9.4
C (outdoor)	Tile	0°	0	57.9		7.8

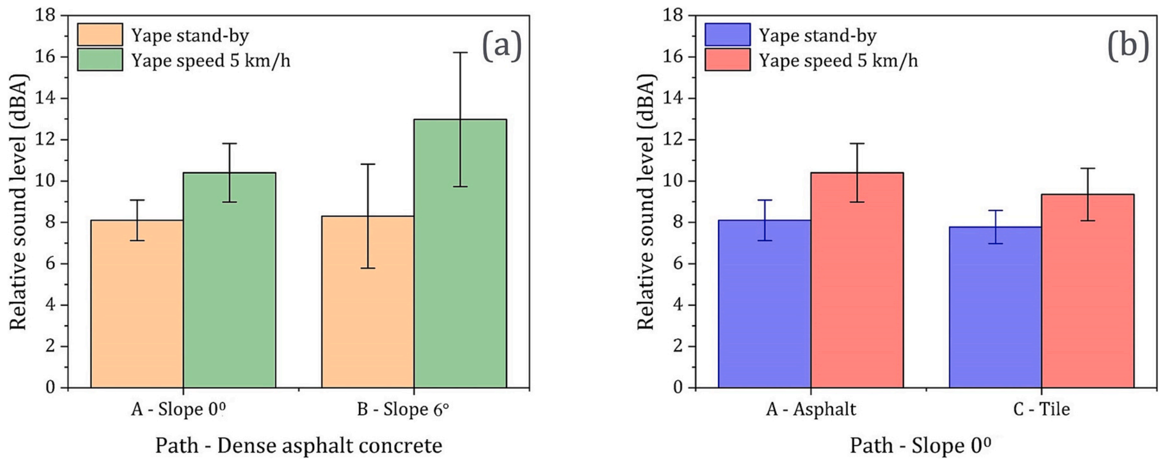


Fig. 5. Sound level values relative to the background noise level for different road slope (a) and road pavement material (b): comparison between the mean values and STD obatinad in Mode 1 (Yape speed 5 km/h) and Mode 2 (Yape stand-by).

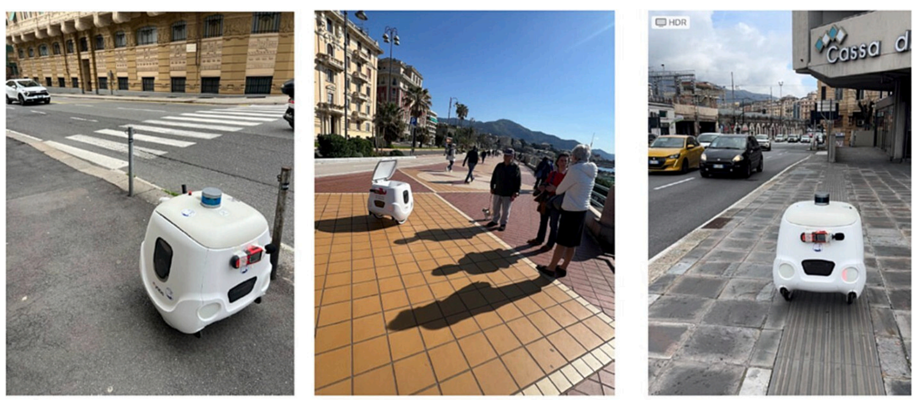


Fig. 6. images of the 3 selected streets within the urban center of Genoa.

h and its stationary condition. This indicates that the movement of the Yape contributes to the noise levels recorded by the sensor.

2. Effect of the road gradient: the maximum value of noise level (13 dBA) was recorded on outdoor asphalt 6° ramp at 5 km/h compared to stationary condition, for which it was found to be 8.3 dBA (Fig. 5a). The ramp probably forced the noise due to the supplement vibrations and mechanical noises from the Yape.
3. Effect of the road pavement surface material: outdoor tiles highlights noise level difference of 9.4 dBA at 5 km/h and 7.8 dBA when stationary (Fig. 5b). They have less impact on noise compared to asphalt, even as their sounds are increased when movement is initiated.
4. Stationary Mode: a stationary mode seems to generate noise impacts that can be pretty much similar on tiles and asphalt equivalent to 8 dBA in difference. This suggests that when the Yape is not moving, the type of surface has a negligible impact on the noise level recorded by the sensor.

We formulated a hypothesis regarding the explanatory variables that influence noise generated by Yape. In controlled laboratory experiments, where the values of each explanatory variable are known and measured, we examined how different combinations of these independent variables affect the Yape noise level, which is considered the dependent variable. However, it was not feasible to test all possible combinations of the independent variables and record the resulting Yape noise for each scenario.

2.6. *Machine learning approach for environmental noise assessment*

This study proposes a machine learning-based methodology for estimating the environmental noise pollution-related variables that impact the noise produced by Yape. The specific model does not make explicit the noise contribution of Yape; rather, it directly estimates environmental noise levels using recorded sensor data and explanatory variables about the noise produced by Yape during its activities.

Explanatory Variables and Target: The independent variables (features) considered for the machine learning model are as follows:

1. Noise level recorded by the sensor mounted on Yape.
2. Speed of Yape
3. Road gradient
4. Road pavement surface material

The dependent variable (target) under consideration is the external environmental noise pollution level indicative of the noise present in the environment outside.

Methodological Innovation: In contrast to the traditional way of assessing environmental noise after first separating the Yape vehicle's noise, this method eliminates that step, and machine-learned estimation instead calculates an environmental noise pollution parameter using recorded sensor data and the identified explanatory variables.

2.7. Conceptual framework

As shown in Fig. 3, the methodology establishes that the independent variables have some relation with the target variables. The model obtained is trained with data obtained under varying conditions of operation by Yape: speed, road gradient, and surface material, which leads to robustness and accuracy in field applicability.

This proposed methodology enables an easy analysis without isolating the contribution of Yape from an environmental point of view and directly evaluates the noise It reduces computational complexity since that would allow, in principle, real-time applicability in the case of external monitoring in dynamic urban situations. Finally, the introduction of the most relevant factors that affect noise propagation improves the accuracy of the sensitive noise level evaluation. Practical Implications: This methodology shows the potentiality of autonomous systems like Yape in the monitoring of noise in urban situations without significant interference with the data. This work represents a new application of machine learning to environmental noise assessment that can enable the development of more efficient and accurate monitoring techniques.

2.8. Data-set construction

Field experiments out of the laboratory were conducted in order to build the data set. This approach enables us to check if the identified explanatory variables are enough for the assessment of the environmental noise pollution level with a satisfactory level of reliability. A dataset consisting of 8139 samples has been built from the 3 routes in the urban area in Genoa, Italy, shown in Fig. 6: Corso Italia, Corso Gastaldi, and Corso Buenos Aires.

Two main challenges were faced:

1. Automatically identifying the value of the independent (explanatory) and dependent variables in real-time.
2. Predicting the environmental noise pollution.

Both challenges were addressed by using machine learning techniques.

a) Automatically identifying the value of the independent and dependent variables in real time

The speed of the Yape:

GPS data collected from Yape, along with a high-precision mobile GPS device (iPhone-based), served as the primary source of positional data. By analyzing the stability of coordinates and speed, the methodology aimed to distinguish between stationary and moving periods of Yape, as the data collection process is dynamic and requires the Yape to consistently move, though it may need to stop momentarily to avoid collisions or cross streets.

The implementation of machine learning algorithms classifies the positions of Yape based on GPS data. Following Support Vector Machines or SVM, a powerful type of supervised learning algorithm first proposed by Cortes and Vapnik in 1995-the researchers learned effectively from labeled GPS data and classified the periods as either stationary or moving with high accuracy. The capability of handling complexities in the pattern of data and adapting to various environments made it a useful tool for the exact determination of the Yape's position during the whole process of data gathering. The training of all the SVM models was done within MATLAB, along with its robust machine learning toolkit used during preprocessing, feature extraction, and model training of the data.

Road gradient:

Another important track characteristic is the gradient of the road. The Yape continuously measured the gradient of the road along its route. In fact, elevation changes were picked up from onboard sensors every second to monitor the terrain in high resolution. This sampling frequency allowed the determination of the slope of the road according to the path of the Yape. The methodology purports to generate a full overview of the topographic demands the Yape faced uphill and downhill on integrating this gradient data with the positional data. The Yape company provides access to this data through a dedicated platform at a very fine resolution, allowing detailed research on the performance of the Yape at various urban environments.

Road pavement surface material:

Images from Yape's stereo camera were used to classify road surfaces into two categories: asphalt and tile. A machine learning model was trained to assign a surface label to each image, enabling this variable to be included in the environmental noise assessment framework. YOLOv8 (a real-time object detection model) was selected for its accuracy and efficiency in image classification.

Details of the model architecture, training process, and performance evaluation are provided in Appendix C.

The noise recorded by the sensor installed on Yape:

These data are provided by the system defined in Section 2.2.

Environmental noise pollution level (pure, without the noise of Yape):

Environmental noise pollution levels, excluding the noise generated by Yape itself, were recorded using a dedicated noise sensor with the same characteristics as the one mounted on the Yape (RS Pro 8852). The sensor was manually moved at walking speed during data collection. To avoid contamination from the Yape's noise emissions, the sensor was positioned at a safe lateral distance from Yape while maintaining synchronization in time and location. Training data was manually recorded using a noise sensor positioned near Yape, ensuring an uncontaminated ground truth reference for model training. We verified the sensor's placement by comparing SPL measurements with the Yape turned on and off. The difference was consistently below 0.5 dBA, confirming that the Yape's noise did not affect the readings. This collection of parallel data was essential for training the machine learning algorithm to distinguish environmental noise from the Yape's own emissions. By capturing both signals concurrently, the methodology allowed accurate isolation of Yape's acoustic contribution and supported a more precise evaluation of its environmental impact across various urban settings.

Synchronization issue

The problem is to associate each time instant in which noise level data has been recorded with a road pavement surface material, road gradient and the speed of the Yape. To do this, it is necessary to synchronize the images from Yape's stereo camera, the data from GPS with the data recorded by the 2 noise sensors.

This synchronization is achieved through the fusion of multiple sensors on the Yape. A GPS device captures precise latitude, longitude, and speed at each second, timestamped to the accurate time. The noise sensors record sound levels with matching timestamps. An onboard camera continuously captures ground images to detect and classify road surface types like asphalt or tiles. By combining and synchronizing these data streams based on their common timestamps, each noise measurement can be associated with the road pavement surface material, road gradient, and the speed of the Yape.

b) Predicting environmental noise pollution level

The rich dataset, consisting of 8139 samples, provided a foundation for training a machine learning model using Support Vector Machines (SVM), known for its robustness in handling complex patterns in data.

The input variables for the algorithm consisted of several key features:

1. Noise level recorded by the sensor installed on Yape
2. Speed of the Yape
3. Road gradient
4. Road pavement surface material

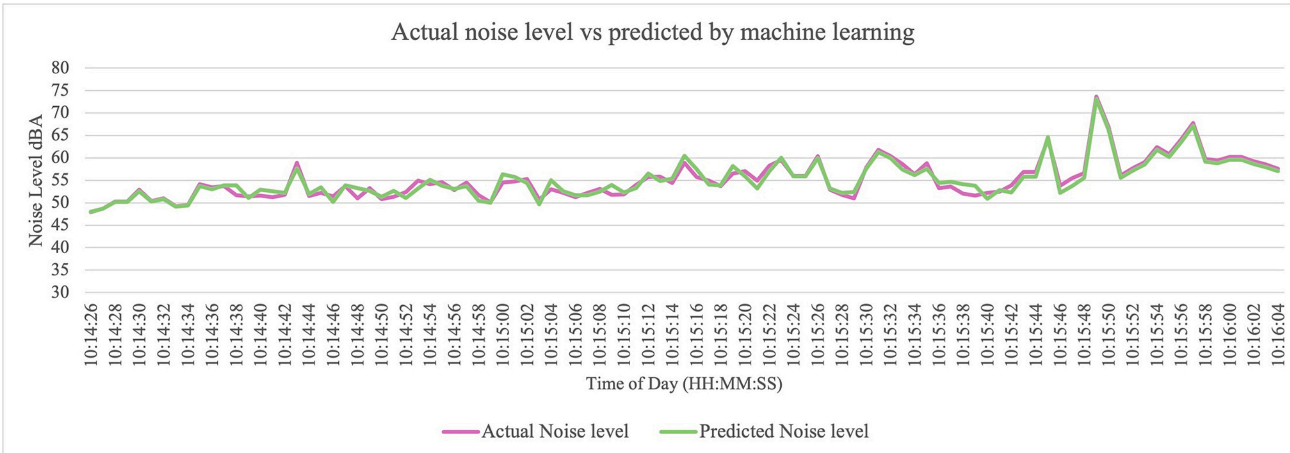


Fig. 7. Example of the deviation between the observed environmental noise (red) and the predicted environmental noise levels (dBA) by the SVM estimation method (green) in corso italia.

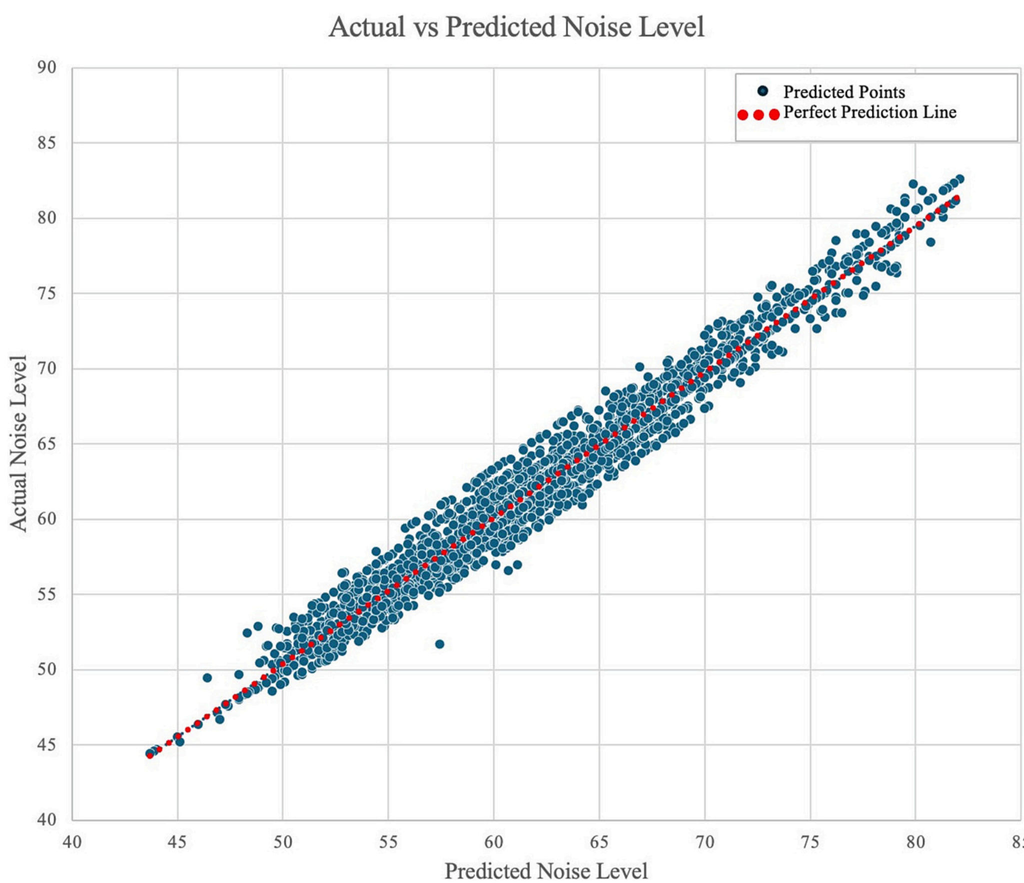


Fig. 8. Plot of the *actual* vs. *predicted* environmental noise pollution levels.

The output variable (target) of the algorithm was the true environmental noise pollution level. This output represents the actual ambient noise in the environment, carefully isolated from any noise generated by the Yape itself.

2.9. Training and testing: performance of SVM in predicting

To estimate the environmental noise levels unaffected by Yape's own acoustic emissions, a Support Vector Regression (SVR) model was trained using four explanatory variables: the noise level recorded by the sensor mounted on Yape, the Yape's speed, road gradient, and pavement surface material. This approach enables the model to infer true environmental noise levels without explicitly separating Yape's contribution during preprocessing.

We trained and evaluated the model on a dataset of 8139 samples, split into 70 % for training and 30 % for testing, following established practice (Géron, 2019). The SVR implementation used a polynomial kernel to capture nonlinear relationships, as recommended in prior urban noise modeling research (Torija and Ruiz, 2015).

To minimize overfitting and ensure generalization, we split our data

Table 3
Traffic volume and route specifications.

Route	Length (km)	Entire route width (m)	Private transport/ traffic flow (vehicle/h)	Private transport/ saturation index (%)	Public transport/ traffic flow (bus/h)
Corso Italia	1.1	30	1000–2400	0.25–0.5	8
Corso Gastaldi	1.0	15–25	2400–3000	0.75–1	24
Corso Buenos Aires	0.5	18	1000–2400	0.5–0.75	85

Fig. 9. Locations of the 3 streets in the Genoa urban area.

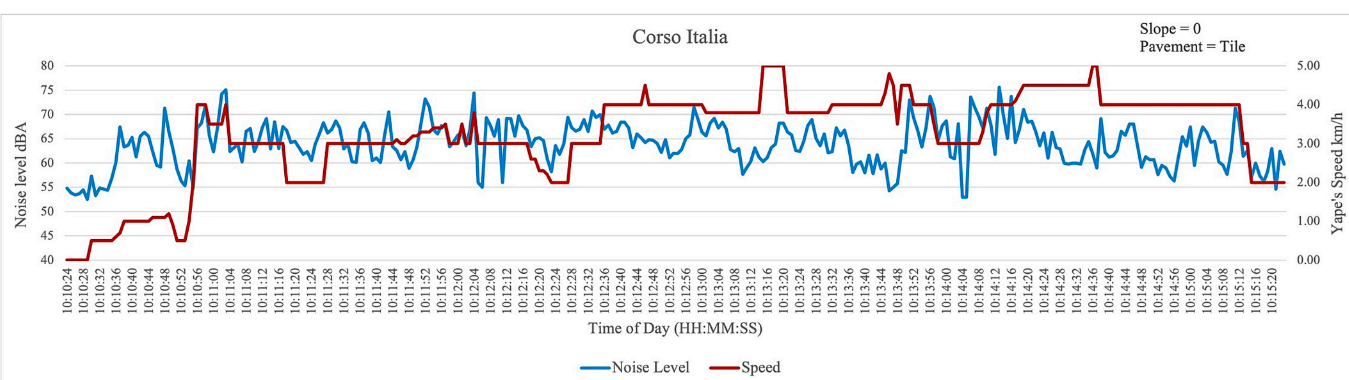


Fig. 10. Corso Italia: 5-minute noise level profile.

by route: Corso Italia and Corso Gastaldi were used for training (5697 samples), and Corso Buenos Aires for testing (2442 samples), avoiding temporal autocorrelation within each route. Additionally, 5-fold cross-validation was applied using non-overlapping time segments to evaluate the model's robustness across varied environmental conditions. Further details are provided in Appendix C.3.

Performance evaluation showed that the model achieved an R^2 value of 0.935 and a Root Mean Square Error (RMSE) of 1.27 dBA, meaning predictions deviated by just ±1.27 dBA on average from reference measurements. To assess robustness, 95 % confidence intervals were calculated across 5-fold cross-validation, yielding $R^2 = 0.935$ [0.927, 0.942] and RMSE = 1.27 dBA [1.21, 1.32]. As shown in Fig. 7, the predicted noise levels closely align with the actual values across varying conditions. Further validation in Fig. 8 illustrates that the majority of predicted environmental noise levels fell within ±2 dBA of the measured values.

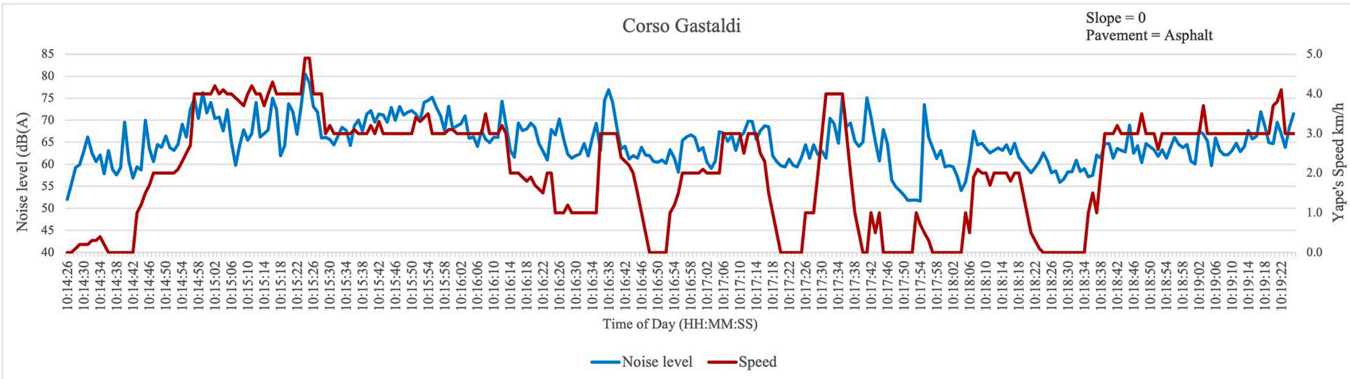


Fig. 11. Corso Gastaldi: 5-minute noise level profile.

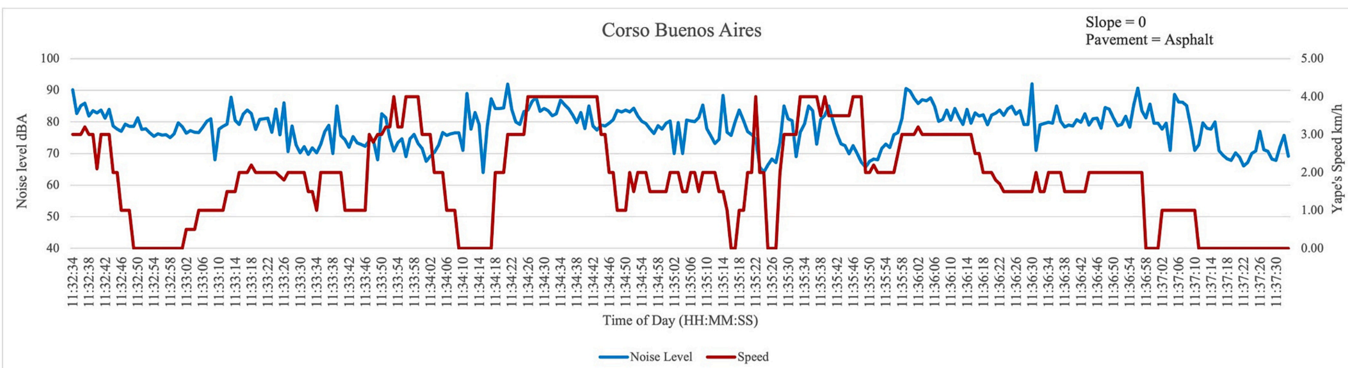


Fig. 12. Corso Buenos Aires: 5-minute noise level profile.

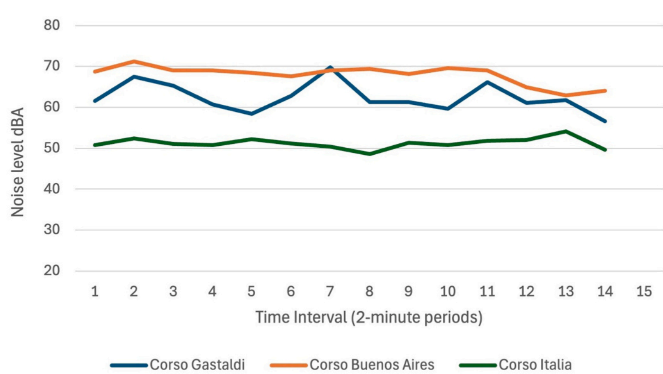


Fig. 13. Comparison of median of the predicted environmental noise levels in the 3 Urban Streets.

These results are consistent with or superior to those reported in previous studies. For instance, Torija & Ruiz ((2015)) achieved an R^2 of 0.82 and RMSE of 2.6 dBA using neural networks for urban noise modeling, while Can et al. (2014)) reported an RMSE of 2.2 dBA for mobile measurement-based noise mapping.

Although XGBoost outperformed SVR in raw accuracy ($R^2 = 0.957$, RMSE = 1.03 dBA), and Random Forest also performed well ($R^2 = 0.913$, RMSE = 1.47 dBA), SVR was chosen for its significantly lower computational demands. It achieved 2.5× faster inference speeds and required less memory, making it more suitable for integration into the resource-constrained environment of the Yape.

A detailed explanation of the kernel selection, RMSE calculation, cross-validation method, and performance benchmarking is provided in Appendix C. Limitations related to weather and future enhancements incorporating meteorological sensors are also discussed there.

3. Analysis of noise pollution in urban environments

The effectiveness of the model has been demonstrated, allowing for the analysis of noise pollution in three urban streets with different topological and morphological characteristics: Corso Gastaldi, Corso Buenos Aires, and Corso Italia.

Data were collected on Thursday, March 14 from 10 am to 12 pm (see Table 3).

The flow data and the saturation indices related to private transport have been reported in PUMS (Urban Plan for Sustainable Mobility, 2019). The data refers to the morning peak hour (07.30–08.30). The data related to public transport refers to the time window 10 am to 12 pm (Linee bus urbane, Genova, n.d.). While not used in modeling, they help contextualize the routes.

Blue Route (Corso Italia): this route spans a total length of 2.7 km, of which a 1.1 km segment was selected for the experiment. The entire width of Corso Italia is 30 m, with 20 m dedicated to the street and 10 m allocated for pedestrian walkways. The chosen segment covers an open, wide area with minimal traffic. Characterized by extensive views and a mostly tiled surface, the Blue Route offers a calm journey with limited vehicle interference. The wide, open layout of this part of Corso Italia, with one side facing the nearby sea without any obstacles, contributes to the relatively low traffic congestion along this section of the route.

Brown Route (Corso Buenos Aires): Corso Buenos Aires represents a canyon narrow street, for a length of 500 m. The tight passage between buildings on opposite sides has a width of 18 m, with 13 m allocated for the street and 5 m for pedestrian walkways. It should be noted that the flow that passes by this road is mainly composed of buses (6 lines in each direction; each line in the morning has an average frequency of 10bus/h) and this creates a dynamic and lively, noisy, atmosphere in the

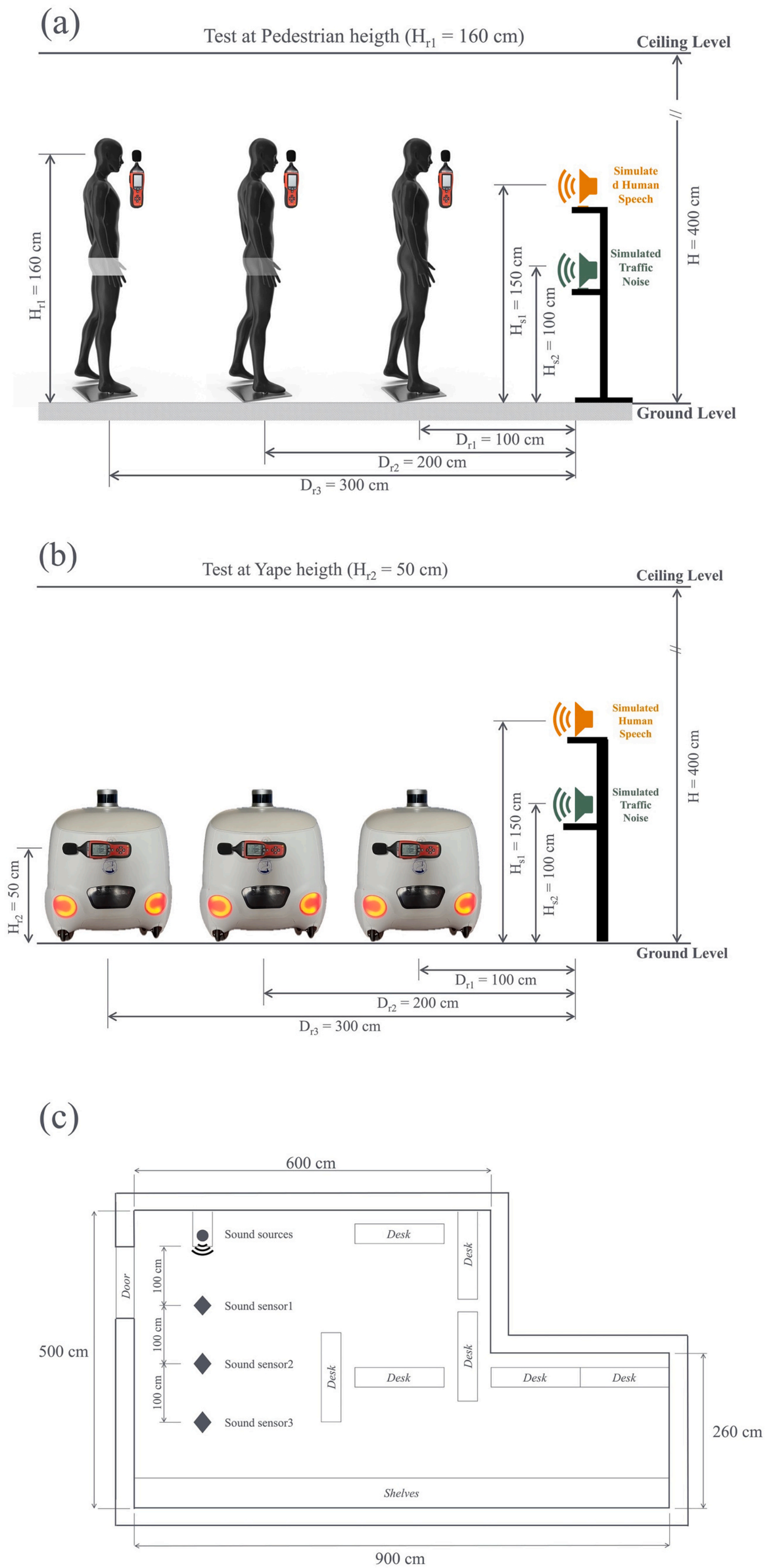


Fig. B.1. Sketches of the experimental set-ups in the laboratory: lateral views with the position of acoustic sources and the receiver for the test at the pedestrian height (a) and at the Yape height (b) and plan view (c) of the indoor environment with the relative position of acoustic sources and the receiver.

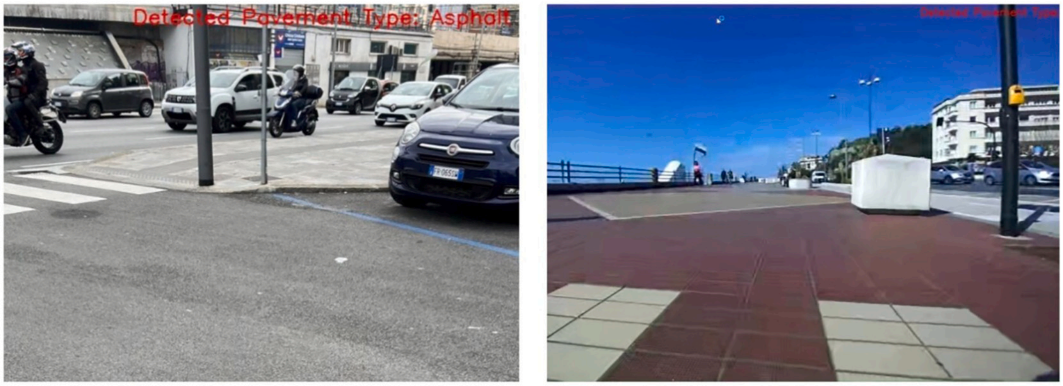


Fig. C.1. The two classes of the road pavement surface material: Asphalt on the left and Tile on the right.

packed surroundings. The street's narrow width and high density of nearby buildings amplify the noise levels from the constant flow of vehicles, while the surface type is asphalt.

Purple Route (Corso Gastaldi): exploring an open environment with a mix of road and railway traffic, the Purple Route follows Corso Gastaldi for a length of 1.1 km. The width of the entire street varies between 15 and 25 m, with the narrowest part being close to the Brignole side, where one side is directly beside the train rails. In this section, 12 to 23 m are dedicated to the street, while 3 to 7 m are allocated for pedestrian walkways. The surface is mostly covered by asphalt, with some sections having tiles. This route presents a balanced blend of ambient noise from both vehicle and rail sources (the contribution of bus flow here is limited: 2 lines with an average frequency of 12 bus/h for each direction). While the open layout provides space, the closeness to the railway line introduces an additional source of noise pollution along the Purple Route.

Fig. 9 shows the locations of the 3 routes on the Genoa map.

Fig. 10 shows 5 min of the predicted environmental noise level for Corso Italia, displaying a fairly consistent pattern with regular fluctuations. It reports predicted environmental noise levels every second but displays them with labeled intervals of approximately 4 s. There are occasional spikes reaching around 70 dBA, but the majority of the readings fall between 55 and 62 dBA. This pattern supports the description of a generally quiet environment with intermittent louder events, possibly due to passing vehicles or other urban activities. Corso Italia demonstrated the lowest average noise level (51.1 dBA). This result aligns with its characteristics: a wide, open space with one side facing the sea and low traffic volume. The sea likely acts as a natural sound buffer, while the open layout prevents noise amplification. The wide range between maximum (75.2 dBA) and minimum (55 dBA) noise levels, coupled with a standard deviation of 5.0 dBA, suggests a relatively consistent noise environment with occasional loud events.

Fig. 11 shows 5 min of the predicted environmental noise level for Corso Gastaldi. It reports predicted environmental noise levels every second with labeled intervals of approximately 4 s. Corso Gastaldi presented moderate results across most metrics. Its average noise level of 62.0 dBA, with the Yape moving at an average speed of 2.9 km/h, reflects its medium traffic volume. The highest predicted environmental noise level is 82.1 dBA, which can be attributed to the periodic passage of noisy trains. This route exhibited the highest standard deviation (6.6 dBA) among the three, meaning it's the most variable noise environment. This is likely due to intermittent high-noise events from passing trains, which cause sharp peaks in the otherwise quieter background noise.

The Gastaldi graph shows more pronounced fluctuations compared to Corso Italia. There are multiple instances where the noise level rapidly rises to a peak of around 80 dBA, likely corresponding to train passages. The baseline noise seems higher than Corso Italia, generally ranging between 50 and 70 dBA. This graph visually confirms the higher variability in noise levels described by the standard deviation.

While these fluctuations are attributed to rail passages based on timing and context, the current system does not include an automated method for noise source classification. As such, distinguishing between overlapping sources, such as road traffic and rail noise, remains a limitation of this approach. Future developments may incorporate spectral features or source-separation techniques to enhance the system's ability to differentiate noise types in complex urban environments.

Fig. 12 shows 5 min of predicted environmental noise levels for Corso Buenos Aires. It reports predicted environmental noise levels every second but displays them at a different ratio (with labeled intervals of approximately 4 s). The Corso Buenos Aires graph shows the highest overall noise levels among the three routes. It is around 70–80 dBA, averaging 77.4 dBA, with frequent peaks up to 90 dBA, and a Standard Deviation of 5.5. In contrast with Gastaldi, the number of quiet periods is much less, which aligns with the setting described as constant noise. There are also fewer extremely fluctuating peaks than in Gastaldi - again in agreement with its moderate standard deviation.

In analyzing 30-minute predicted environmental noise level data from three urban streets, the chosen method employs median values calculated over 2-minute intervals. This approach strikes a balance between data reduction and preservation of temporal trends, resulting in 15 data points per street for comparison. The median was selected over the mean due to its robustness against outliers, which are common in urban acoustic environments (Can et al., 2014). Using the median it is less influenced by very short extreme noise events, which often distort the results of averages, and hence provides a better indication of typical noise levels in each street. The 2-minute interval was chosen to capture short-term variations in the urban soundscape while still offering a manageable number of data points for analysis and visualization. This enables one to compare directly the general noise levels between the three streets, without any influence of short-term variations. Considering these considerations, the median of each 2-minute sample of noise level was to underline the general acoustic character for each street, while differences dealt with the characteristic noise environment of each. This will be of particular use in the cases when the general noise exposure levels between more and less structured urban areas are under consideration and momentary changes are of little concern.

Fig. 13 reveals distinct acoustic profiles for Corso Gastaldi, Corso Buenos Aires, and Corso Italia. Corso Buenos Aires has shown invariably higher values of noise, while Corso Gastaldi is more changeable, with medians varying from the mid-50s to the upper 60s dBA, showing more dynamics in its acoustic behavior. On the other side, Corso Italia represents the lowest noise levels among the three, with low values in the low 50s dBA range and never or hardly reaching 55 dBA. This behavior may correspond to the fact that Corso Buenos Aires is exposed to the most considerable traffic or urban activity, whereas Corso Italia should be characterized as a relatively quiet street. Corso Gastaldi exhibits noise levels somewhere between, with very high variability potential indication of the combination of busy and quiet periods.

Although predictive models like Cnossos-EU are widely used for environmental noise estimation, their road traffic, railway, aircraft, and industrial noise modules rely on detailed input data such as vehicle flow and standard source parameters, limiting their ability to capture real-time urban acoustic variability, such as transient noise events or micro-scale spatial effects (Kephalopoulos et al., 2012). In contrast, our measurement-based approach using the Yape platform captures the full acoustic environment at pedestrian level, including non-traffic sources such as construction, and crowd noise. Future work should include a direct comparison with CNOSSOS-EU predictions to quantitatively evaluate differences and demonstrate how real-time mobile sensing can complement traditional noise modeling.

While this study was conducted on three streets in Genoa, the sensing and prediction framework was intentionally designed using four physically interpretable features—noise level, Yape speed, road gradient, and pavement material—which are generally applicable across urban contexts. Nonetheless, we recognize that environmental and acoustic conditions may differ substantially in other cities. For example, tropical regions may introduce persistent biological sounds, snowy environments can alter sound propagation, and dense high-rise areas may increase reverberation. These factors could influence model performance and require retraining with localized data or hardware adaptations such as weather sealing, weatherproofing sensors or GPS alternatives. Future work will focus on validating the system in diverse urban morphologies and climatic conditions to assess its broader applicability.

4. Conclusions

This study demonstrates the potential of autonomous mobile sensing systems to advance urban noise monitoring through a human-centered approach. Using the Yape platform, we developed a method that captures pedestrian-level noise exposure across full urban trajectories, addressing two key challenges: identifying an effective sensor placement at 50 cm and training a machine learning model that separates environmental noise from the vehicle's own acoustic emissions. The model achieved high performance (R squared equals 0.935), confirming the system's reliability in complex urban conditions.

Unlike traditional fixed-sensor networks, which map noise at static points, our system captures the total acoustic environment encountered by individuals in motion. It reflects real-world exposure more accurately by accounting for diverse sources such as traffic, construction, and other transient urban activities. The sensor's low mounting height further enhances measurement relevance by aligning closely with human perception.

The mobile, autonomous nature of the system enables high-resolution spatial monitoring with minimal labor requirements, offering a scalable and consistent alternative to wearable dosimeters and fixed stations. While not intended to replace existing infrastructure, this method provides a complementary tool that enhances spatial coverage and measurement granularity.

However, there are practical limitations to deploying Autonomous Vehicle (AV)-based systems in all contexts. Poor internet connectivity may hinder real-time data transmission, rough or uneven terrain could affect AV mobility and sensor stability, and social challenges such as vandalism, dense crowds, or restricted pedestrian areas may pose operational constraints. These considerations highlight the importance of tailoring deployments to local conditions and identifying appropriate use cases for AV-based monitoring.

In practical terms, the methodology supports a range of urban applications. The data collected can update digital representations of city noise conditions, inform targeted mitigation strategies, and enable personalized acoustic comfort services based on individual needs. By focusing on the actual experience of environmental noise, this system contributes meaningfully to public health, urban planning, and environmental quality.

To date it is not easy, neither technically nor from a legislative point of view, to use autonomous vehicles on public urban contexts. However, great effort are being made in overcoming the difficulties. The European community has focused heavily on autonomous driving and large projects have been funded to test it in the field (for example the SHOW project). The Yape prototype was successfully used in an urban pedestrian area of a Greek city, continuously for several months (Cepolina et al., 2024; Malisetty et al., 2025; SHOW Project, 2025) in a pilot experimentation of the SHOW project. Compared to other pilots of the SHOW project, which involved autonomous road vehicles, the difficulties encountered with Yape were lower precisely due to the reduced speed of the vehicle and the environment for which it was designed (indoor and outdoor pedestrian spaces) which is certainly less regulated than the road.

In future work, we will incorporate meteorological factors such as wind, rain, and temperature through additional sensors, evaluate the system's adaptability across different climate zones and urban settings, develop a semi-automated labeling process based on frequency and amplitude thresholds to improve data reproducibility, and explore the integration of onboard vision systems to estimate real-time traffic flow and vehicle density for enhanced contextual analysis.

CRediT authorship contribution statement

Behzad Ajdari: Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – review & editing. **Nikoo Salimi:** Software, Data curation. **Lucanos Strambini:** Visualization, Validation, Methodology, Conceptualization, Writing – review & editing. **Elvezia Maria Cepolina:** Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research has been partially funded by European Union's Horizon 2020 research and innovation programme under grant agreement No. 875530 SHOW - SHared automation Operating models for Worldwide adoption.

Appendix A. System components and specifications

A.1. YAPE ecosystem

The Yape system (Yape S.r.l., 2022) is an innovative self-driving solution designed for low-contact services and last-mile delivery operations. Also, it is adaptable for operating in both indoor and outdoor environments. This ecosystem makes autonomous navigation by integrating several components:

- Admin Platform (software application): This platform allows users to manage the digital maps of the locations where the robots operate. It also provides the latest software updates for both the Control Room and the robots.
- Control Room (software application): Operators use this tool to monitor the fleet's movements in real-time on a digital map. It also enables manual intervention when necessary, ensuring smooth and secure operations.
- Delivery Platform (software application): This app allows the user to place a delivery request.
- Fleet of Droids (1 or more): The Yape droids (Fig. 1a) are designed for SAE (Society of Automotive Engineers) Level 4 automation, which means they are capable of fully autonomous driving within predefined areas and conditions, without requiring human intervention, although a human can take control if needed. This classification is defined by the SAE J3016 standard. The droids have reached a Technology Readiness Level (TRL) of 7. Docking stations also help the droids operate autonomously. Each droid measures 70 × 70 × 90 cm, weighs 50 kg, and can carry loads up to 10 kg with a maximum speed of 9.6 km/h.

Yape droids are equipped with advanced technology such as a lidar sensor for obstacle detection and pathfinding, three high-definition cameras, and a stereo camera to enable operators to remotely monitor delivery operations. Their lithium-ion batteries provide up to 8 h of uninterrupted operation, and the droids can handle inclines of up to 12°.

The complete region where the droids would operate must be covered with 4G/5G.

A.2. Noise sensor

The RS Pro DT-8852 sound level meter, manufactured by RS Components S.r.l., was selected for its balance of measurement accuracy, environmental durability, and suitability for dynamic urban noise monitoring. It complies with IEC 61672-1:2002 Class 2 standards and is capable of measuring sound pressure levels from 30 dB to 130 dB with a resolution of 0.1 dB and an accuracy of ±1.4 dB. The device supports both A-weighting and C-weighting filters and provides readings via an illuminated LCD display. It features fast and slow time constants for averaging and operates across an effective frequency range of 31.5 Hz to 8 kHz. Its fast response time of 125 ms allows for capturing transient acoustic events common in pedestrian-level urban environments. Designed for outdoor use, it functions reliably in temperatures from 0 °C to 40 °C and relative humidity up to 90 % (non-condensing). Before each measurement session, the sensor was calibrated using a certified 94 dB, 1 kHz acoustic calibrator to ensure consistent accuracy throughout the study. These specifications make the RS Pro DT-8852 a robust and reliable instrument for capturing real-world environmental noise levels during mobile monitoring deployments.

The frequency response limitation of this sensor is not expected to significantly affect the accuracy of the measured noise levels. According to Yang et al. (2020), the dominant acoustic components of road traffic noise are concentrated between 100 Hz and 5000 Hz, with most of the energy observed between 250 Hz and 4000 Hz depending on vehicle type and speed. Although some acoustic content exists above 5 kHz, its contribution to total sound pressure levels is limited. Moreover, A-weighting substantially attenuates high-frequency bands above 8 kHz, minimizing their influence on reported values. Based on this evidence, the sensor's range is considered sufficient for capturing the most perceptually and environmentally relevant components of urban noise exposure.

This is an adaptable sensor suitable for several applications in industrial, environmental, and office noise monitoring. Its selection was informed by the framework developed by Wessels and Basten (2016), which classifies environmental sensor networks into four categories based on criteria including hardware cost, versatility, scalability, dependability, and precision. The RS Pro 8852 fits within Category 2 of this framework, offering a practical balance between the performance of high-cost, dedicated monitoring equipment and the affordability of lower-grade sensors. This category is aligned with the needs of the present study, which require both reliability and deployment flexibility in complex urban environments.

Appendix B. Experimental design for sensor positioning

This appendix provides detailed information on the experimental procedures used to evaluate sensor placement on the Yape. It includes descriptions of the noise sources, sensor configurations, controlled testing environment, and complete measurement results used to assess the acoustic accuracy at different heights.

To find the best sensor placement, we conducted experiments by changing the height of the sensor and observing how it affects the noise data we collect. This was done to make sure our measurements closely match what a person walking in an urban area would hear.

B.1. Experimental setup

Acoustic Source: The acoustic source used in the experiments was a Beolit 20 loudspeaker (Bang & Olufsen) capable of generating sound energy enough to cover an area of up to 50 m^2. The loudspeaker, which has a working frequency range between 37 and 20,000 Hz and reaches a maximum SPL of 93 dB, was mounted on a fixed stand at a given height above the ground (Hs). To more closely approximate real-life urban conditions, two major contributors to urban noise pollution were chosen: road traffic and human speech.

1. Road Traffic Noise (Hs1 = 1.5 m):

The loudspeaker produced calibrated road traffic noise at an equivalent continuous sound level (Leq) of 70 dBA, measured at a distance of 1 m, which is within the lower range of the WHO Guidelines for Community Noise, 1999, which indicate Leq values between 70 and 90 dBA near busy roads. The acoustic signal was a pre-recorded sample of traffic noise with frequencies between 100 Hz and 2 kHz and temporal variations in order to simulate real-life traffic.

2. Human Speech Noise (Hs2 = 1 m):

The conversational speech noise was calibrated to an equivalent continuous sound level (Leq) of 60 dBA, measured 1 m from the loudspeaker, following the Noise and Hearing Conservation Technical Manual (Occupational Safety and Health Administration (OSHA), 2022). The speech audio included a frequency range from 100 Hz to 8 kHz and balanced male and female voice samples representative of typical conversational speech frequencies.

The audio stimuli were obtained from a specialized professional audio library that focuses on environmental soundscapes, referred to as "(*Sound of Italy - Italian Sound Effects Library, n.d.*)." Although specific recording specifications were not disclosed, standard industry equipment utilized probably included shotgun microphones (for instance, Sennheiser MKH-8060) and stereo arrays (such as AKG C480B). The audio files were supplied in an uncompressed, high-resolution WAV format, which guarantees precise reproduction during monophonic playback. This methodology ensured that the audio samples accurately represented genuine urban soundscapes found in Italy.

Receiver Configuration: The receiver (sensor) was placed on a stationary support and measured the sound pressure levels (SPL) at some points relative to the noise source. These levels depended on the distance of the receiver from the source and the acoustic properties of the surroundings.

Measurements were carried out with the receiver at two vertical heights:

- Hr = 50 cm: This value is the height of the sensor mounted on the Yape.
- Hr = 160 cm: This is the average ear height of a pedestrian and it is used as the reference target height.

The horizontal distances from the source, Dr. = 1 m, 2 m, and 3 m were considered resulting in six different sensor positions in a plane perpendicular to the ground and passing through the source and the receiver. The configurations are presented in Fig. B.1a, and b.

Experimental Environment: The experiments took place in a controlled indoor area, 5 × 6 m in size, with a 4-meter-high ceiling (Fig. B.1c). The walls and ceiling were covered by drywall, while the flooring was tiled, providing quite a neutral acoustical environment.

To ensure accurate sound measurements the ambient background noise was very minimal and constant, remaining negligible in comparison to the sound output. Also, the office space was equipped with standard office furnishings, including three desks and five chairs, to create a real yet controlled testing environment.

This setup allowed for a necessary balance between a realistic urban auditory environment and strict, replicable parameters for the experiments.

B.2. Experimental results

Table B.1
Sound level variation by changing the distance of the sensor from the source.

Noise type	Speaker's height (cm)	Sensor's height (cm)	Horizontal distance to source (m)	Average sound level	Noise variation with sensor height (dB)
Human speech	150	50	1	58.1	1.9
Human speech	150	160	1	60.0	
Human speech	150	50	2	57.3	0.9
Human speech	150	160	2	58.2	
Human speech	150	50	3	56.3	0.5
Human speech	150	160	3	56.8	
Traffic noise	100	50	1	70.8	0.4
Traffic noise	100	160	1	70.4	
Traffic noise	100	50	2	68.2	0.3
Traffic noise	100	160	2	67.9	
Traffic noise	100	50	3	66.8	0.3
Traffic noise	100	160	3	66.5	

Table B.1 summarizes the sound level measurements for human speech and traffic noise under varying experimental conditions, including changes in sensor height (50 cm and 160 cm) and horizontal distance from the source (1 m, 2 m, and 3 m). The table also highlights the noise variations resulting from sensor height differences.

The results show that such differences in sensor height actually result in variations in the measured sound levels by 0.3 dBA to 1.9 dBA when the sensor height is increased from 50 cm to 160 cm. The most varied measured values for human speech were 1.9 dBA at 1 m while that of traffic noise had lesser variations, reaching a maximum of 0.5 dBA at that distance. The increased horizontal distances created reductions in these variations for both noise types caused by a diminishing influence of the ground reflection and a larger dispersion of the sound waves.

Appendix C. Machine learning implementation details

C.1. YOLOv8 implementation details

To classify road surface materials along Yape's route, images captured by the stereo camera were processed using a convolutional neural network. Fig. C.1 shows two sample images collected from the Yape-mounted camera.

YOLOv8 (You Only Look Once, version 8) was selected for this classification task due to its strong performance in real-time object detection and image classification. The model was configured to assign a single label per image, distinguishing between two surface types: asphalt and tile.

The input images were resized to 64 × 64 pixels and divided into training and validation sets. We the model for 50 epochs, with performance tracked through training/validation loss and validation accuracy.

The model showed strong convergence behavior, with both loss curves stabilizing after initial epochs and no sign of overfitting. Top-1 accuracy exceeded 95 % early in training and stabilized close to 100 %, confirming the model's ability to reliably distinguish between the two pavement types.

This classification output was used to annotate each sensor data point with a corresponding surface label, supporting noise analysis as a function of road material.

C.2. SVM model implementation specifics

C.2.1. Kernel selection and model objective

A Support Vector Regression (SVR) model was implemented to estimate the true environmental noise level, independent of the noise produced by Yape itself. The model uses the following four explanatory variables:

- Measured noise level from Yape's onboard sensor
- Speed of the Yape
- Road gradient
- Road surface material (classified via YOLOv8)

A polynomial kernel of degree 3 was selected to capture nonlinear interactions between these features, consistent with prior studies in environmental acoustics (e.g., Torija and Ruiz, 2015).

C.2.2. Training strategy and cross-validation

The dataset consisted of 8139 samples, randomly split into:

- 70 % for training (5697 samples)
- 30 % for testing (2442 samples)

To ensure generalizability and prevent overfitting, we applied 5-fold cross-validation during training. Hyperparameters were tuned using grid search across folds, optimizing for both R^2 and RMSE on validation sets.

All input features were standardized prior to training. The SVR model was implemented using MATLAB's machine learning toolbox.

C.2.3. Model performance and robustness

Model accuracy was evaluated using:

- R^2 (coefficient of determination)
- RMSE (Root Mean Square Error)

Across 5-fold cross-validation, the SVR model achieved:

- Mean $R^2 = 0.935$, 95 % CI: [0.927, 0.942]
- Mean RMSE = 1.27 dBA, 95 % CI: [1.21, 1.32]

These results indicate both high predictive accuracy and robust generalization.

C.2.4. Model selection and efficiency

In addition to Support Vector Regression (SVR), we evaluated Random Forest and XGBoost (Boosted Trees) on the same dataset (8139 samples). XGBoost demonstrated the best predictive performance with an R^2 of 0.957 and RMSE of 1.03 dBA, followed by SVR ($R^2 = 0.935$, RMSE = 1.27 dBA) and Random Forest ($R^2 = 0.913$, RMSE = 1.47 dBA). However, SVM was chosen due to its significantly lower computational requirements, offering 2.5× faster inference speed and lower memory usage compared to XGBoost. This made it more suitable for deployment on our autonomous urban monitoring platform, where embedded efficiency and quick real-time predictions are crucial. Additionally, SVR achieved prediction errors well within acceptable urban acoustic tolerances (differences under 3 dBA are generally imperceptible). As the dataset expands in future implementations, we may revisit ensemble models like XGBoost to leverage their scalability. For this study, SVR offered the best trade-off between performance, simplicity, and real-world applicability.

C.2.5. Model validation and overfitting mitigation

To assess and mitigate overfitting risks, we implemented a 5-fold cross-validation strategy during training of the Support Vector Regression (SVR) model. The dataset of 8139 samples was randomly partitioned into five folds, ensuring that each sample contributed to both training and validation. The model achieved consistent performance across folds, with a mean R^2 of 0.935 (95 % CI: [0.927, 0.942]) and a mean RMSE of 1.27 dBA (95 % CI: [1.21, 1.32]), indicating reliable generalization to unseen data. Additionally, the model was tested on hold-out segments with environmental conditions not represented in the training folds, yielding $R^2 = 0.91$. These results suggest the model effectively captures underlying acoustic relationships without overfitting to specific routes. Model inputs were limited to four interpretable features—sensor-recorded noise, Yape speed, road gradient, and surface material classification—to further reduce the risk of learning spurious patterns. We recognize the need for future validation across more diverse urban conditions, and ongoing data collection is underway to expand and diversify the training set. To further minimize temporal autocorrelation in mobile sensor data, we avoided random sampling within the same route. Instead, training and testing sets were separated by route: data from Corso Italia and Corso Gastaldi were used for training, and Corso Buenos Aires was reserved for testing. This ensured that the model was evaluated on a different temporal and spatial context than it was trained on, reducing the chance of overlap due to similar traffic or environmental conditions. Additionally, cross-validation was performed using non-overlapping segments (i.e., distinct time intervals within each route) to promote generalization across time windows.

C.3. Limitations and future enhancements

This study was conducted under stable weather conditions, but environmental factors significantly impact outdoor acoustic measurements. Wind effects are substantial, with (Tanaka and Shiraishi, 2008) demonstrating that wind direction and speed can cause variations up to 16 dB in measured traffic noise levels. Precipitation generates impact noise and alters surface acoustics, while temperature gradients create refraction effects that can extend urban noise propagation distances, particularly during nighttime inversions (Van Renterghem and Botteldooren, 2010). These variables, though not addressed in our current methodology, represent critical considerations for comprehensive noise monitoring. Future implementations of our mobile sensing platform will incorporate meteorological sensors to record environmental parameters alongside acoustic data, enabling both weather-based data filtering and the development of correction algorithms to enhance measurement accuracy across diverse urban conditions.

Data availability

Data will be made available on request.

References

Attenborough, K., 2002. Sound propagation close to the ground. Annu. Rev. Fluid Mech. 34 (1), 51–82. https://doi.org/10.1146/annurev.fluid.34.081701.143541.

Babisch, W., 2005. Guest editorial: noise and health. Environ. Health Perspect. 113 (1). https://doi.org/10.1289/ehp.113-a14.

Benocci, R., Molteni, A., Cambiaghi, M., Angelini, F., Roman, H.E., Zambon, G., 2019. Reliability of Dynamap traffic noise prediction. Appl. Acoust. 156, 142–150. https://doi.org/10.1016/j.apacoust.2019.07.004.

Bies, D.A., Hansen, C.H., Howard, C.Q., 2018. Engineering Noise Control (Fifth Edition). CRC Press.

Brown, Lex, 2014. (PDF) An Overview of Concepts and Past Findings on Noise Events and Human Response to Surface Transport Noise. ResearchGates. https://www.researchgate.net/publication/287032221_An_overview_of_concepts_and_past_findings_on_noise_events_and_human_response_to_surface_transport_noise.

Can, A., Dekoninck, L., Botteldooren, D., 2014. Measurement network for urban noise assessment: comparison of mobile measurements and spatial interpolation approaches. Appl. Acoust. 83, 32–39. https://doi.org/10.1016/j.apacoust.2014.03.012.

Cepolina, E.M., Malisetty, V.A.B., Bruzzone, A., 2024. Small autonomous vehicles in pedestrian contexts: a first analysis of logistics performances in terms of commercial speed. Sustainability 16 (14), 5844. https://doi.org/10.3390/su16145844.

Città Metropolitana di Genova & Comune di Genova, 2019. Urban plan for sustainable mobility. https://pums.cittametropolitana.genova.it/.

Clark, C., Crumpler, C., Notley, H., 2020. Evidence for environmental noise effects on health for the United Kingdom policy context: a systematic review of the effects of environmental noise on mental health, wellbeing, quality of life, cancer, dementia, birth, reproductive outcomes, and cognition. Int. J. Environ. Res. Public Health 17 (2), 393. https://doi.org/10.3390/ijerph17020393.

Environmental Noise Directive—European Commission, 2024, December 11. https://environment.ec.europa.eu/topics/noise/environmental-noise-directive_en.

European Environment Agency (EEA), 2024. Sustainability of Europe's Mobility Systems: Transport Noise. Publications Office. https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems/transport-noise.

Fastl, H., Zwicker, E., 2007. Hearing area. In: Fastl, H., Zwicker, E. (Eds.), Psychoacoustics: Facts and Models. Springer, pp. 17–22. https://doi.org/10.1007/978-3-540-68888-4_2.

Géron, A., 2019. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, Second edition. O'Reilly.

Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F., European Commission, 2012. Common Noise Assessment Methods in Europe (CNOSSOS-EU): To Be Used by the EU Member States for Strategic Noise Mapping Following Adoption as Specified in the Environmental Noise Directive 2002/49/EC. OPEU. https://doi.org/10.2788/31776.

Khajehvand, M., Rassafi, A.A., Mirbaha, B., 2021. Modeling traffic noise level near at-grade junctions: roundabouts, T and cross intersections. Transp. Res. Part D: Transp. Environ. 93, 102752. https://doi.org/10.1016/j.trd.2021.102752.

Lan, Z., Cai, M., 2021. Dynamic traffic noise maps based on noise monitoring and traffic speed data. Transp. Res. Part D: Transp. Environ. 94, 102796. https://doi.org/10.1016/j.trd.2021.102796.

Lan, Z., He, C., Cai, M., 2020. Urban road traffic noise spatiotemporal distribution mapping using multisource data. Transp. Res. Part D: Transp. Environ. 82, 102323. https://doi.org/10.1016/j.trd.2020.102323.

Linee bus urbane | AZIENDA MOBILITA' E TRASPORTI SpA. (n.d.). Retrieved May 23, 2025, from https://www.amt.genova.it/amt/trasporto-multimodale/linee-bus-urbane/.

Malisetty, V.A.B., Patatouka, E., Raptis, O., Antonakopoulou, A., Amditis, A., Silani, E., Cepolina, E.M., 2025. Field-driven lessons learned for delivery robots logistics operation in urban environment. In: Cornet, H., Gkemou, M. (Eds.), Shared Mobility Revolution. Springer Nature Switzerland, pp. 97–117. https://doi.org/10.1007/978-3-031-71793-2_7.

Michaud, D.S., Marro, L., Denning, A., Shackleton, S., Toutant, N., McNamee, J.P., 2022. Annoyance toward transportation and construction noise in rural suburban and urban regions across Canada. Environ. Impact Assess. Rev. 97, 106881. https://doi.org/10.1016/j.eiar.2022.106881.

Montes González, D., Barrigón Morillas, J.M., Rey Gozalo, G., Godinho, L., 2020. Evaluation of exposure to road traffic noise: effects of microphone height and urban configuration. Environ. Res. 191, 110055. https://doi.org/10.1016/j.envres.2020.110055.

Occupational Safety and Health Administration (OSHA), 2022. Technical Manual (OTM): Section III: Chapter 5—Noise (OSHA Technical Manual (OTM)). https://www.osha.gov/otm/section-3-health-hazards/chapter-5#appendixb3.

Sandberg, U., Ejsmont, J.A., 2002. Tyre/Road Noise Reference Book. INFORMEX.

SHOW Project, 2025. CORDIS | European Commission. https://cordis.europa.eu/project/id/875530/results.

Sound of Italy—Italian sound effects library. (n.d.). Sound of Italy - Italian Sound Effects Library. Retrieved January 5, 2025, from https://www.soundofitaly.it/.

Stanovská, M., Tomášková, H., Šlachtová, H., Potužníková, D., Argalášová, L., 2024. Health impact of environmental and industrial noise – a narrative review. Med. Pr. 75 (5), 425–431. https://doi.org/10.13075/mp.5893.01491.

Tanaka, S., Shiraishi, B., 2008. Wind effects on noise propagation for complicated geographical and road configurations. Appl. Acoust. 69 (11), 1038–1043. https://doi.org/10.1016/j.apacoust.2007.07.007.

Tong, H., Aletta, F., Mitchell, A., Oberman, T., Kang, J., 2021. Increases in noise complaints during the COVID-19 lockdown in Spring 2020: a case study in Greater London, UK. Sci. Total Environ. 785, 147213. https://doi.org/10.1016/j.scitotenv.2021.147213.

Torija, A.J., Ruiz, D.P., 2015. A general procedure to generate models for urban environmental-noise pollution using feature selection and machine learning methods. Sci. Total Environ. 505, 680–693. https://doi.org/10.1016/j.scitotenv.2014.08.060.

Van Renterghem, T., Botteldooren, D., 2010. The importance of roof shape for road traffic noise shielding in the urban environment. J. Sound Vib. 329 (9), 1422–1434. https://doi.org/10.1016/j.jsv.2009.11.011.

Wang, H., Wu, Z., Chen, J., Chen, L., 2022. Evaluation of road traffic noise exposure considering differential crowd characteristics. Transp. Res. Part D: Transp. Environ. 105, 103250. https://doi.org/10.1016/j.trd.2022.103250.

Wessels, P.W., Basten, T.G.H., 2016. Design aspects of acoustic sensor networks for environmental noise monitoring. Appl. Acoust. 110, 227–234. https://doi.org/10.1016/j.apacoust.2016.03.029.

WHO Regional Office for Europe (with Who Regional Office for Europe), 2011. Burden of Disease from Environmental Noise: Quantification of Healthy Life Years Lost in Europe. WHO Regional Office for Europe.

Yang, W., Cai, M., Luo, P., 2020. The calculation of road traffic noise spectrum based on the noise spectral characteristics of single vehicles. Appl. Acoust. 160, 107128. https://doi.org/10.1016/j.apacoust.2019.107128.

Yape S.r.l., 2022, May 11. Yape. https://yapemobility.it/.

Zambon, G., Benocci, R., Bisceglie, A., Roman, H.E., Bellucci, P., 2017. The LIFE DYNAMAP project: towards a procedure for dynamic noise mapping in urban areas. Appl. Acoust. 124, 52–60. https://doi.org/10.1016/j.apacoust.2016.10.022.

Zambon, G., Muchetti, S.S., Salvi, D., Angelini, F., Brambilla, G., Benocci, R., 2020. Analysis of noise annoyance complaints in the city of Milan, Italy. J. Phys.: Conf. Ser. 1603 (1), 012029. https://doi.org/10.1088/1742-6596/1603/1/012029.