DOI: 10.61782/fa.2025.0943

USING MICROPHONES IN OUTDOOR PASS-BY NOISE TESTING FOR VEHICLE SPEED MEASUREMENT

Fabio Lo Castro1* **Maria Luisa Ariza Alvarez[1]** **Massimiliano De Luca1** **Sergio Iarossi[1]**

[1] CNR-INM Sezione Acustica e Sensoristica O.M. Corbino, Rome Italy

ABSTRACT

In environmental noise pollution assessments, the Statistical Pass-By method (SPB, ISO 11819-1) is employed to evaluate different road surfaces based on their influence on traffic noise. It requires measuring the speed of heavy and light vehicles, with an error of less than 2.5 %, as they pass in front of and perpendicular to a Class I microphone, positioned 7.5 meters from the center of the road lane, along with the speedometer readings and the corresponding sound pressure levels (SPL). In this paper, we focus on evaluating the feasibility of using one or more microphones, instead of a speedometer and a video camera, to enhance performance in speed measurement, vehicle classification, vehicle counting, and sound pressure level (SPL) detection. The soundtracks used were recorded during the SPB measurements conducted along Highway A91 in Rome, Italy.

Keywords: *traffic noise, vehicle speed, road monitoring, sound measurement, microphone detection.*

1. INTRODUCTION

The accurate estimation of vehicle speed is a fundamental component of traffic management, urban monitoring, and road safety technologies. Traditionally, speed measurement has relied on radar, laser sensors, and video-based systems. However, these technologies often involve high costs [1]. In this context, passive acoustic sensing using microphones has emerged as a cost-effective and scalable alternative for vehicle speed estimation.

According to the Statistical Pass-By (SPB) method, which requires the use of microphones to measure sound pressure levels, a relevant question arises: can these microphones also be used to estimate vehicle speed through acoustic signal analysis? Specifically, two methodological approaches are investigated: a frequency-domain analysis, which exploits variations in the spectral components of recorded signals, and a time-domain analysis, which relies on signal amplitude and the time delay of signal arrival at multiple microphones to infer speed. The accuracy and error margins of both methods are evaluated and compared

Recent studies have demonstrated the potential of acoustic signal processing for vehicle speed estimation. For instance, a Doppler-based method employing spectrogram seam tracking was proposed to estimate vehicle speed from engine harmonics, achieving an average error between 0 and 10 km/h, with reduced accuracy at lower speeds (10–30 km/h) due to weaker Doppler shifts [2]. Similarly, multi-microphone arrays were utilized to enhance accuracy in dense traffic conditions [3].

Other studies, explored variations in noise levels as an indirect speed indicator, reporting errors of up to 17% at 50 km/h, improving to 8% at 120 km/h; however, it was concluded that the method does not meet the precision requirements of ISO 11819-1 [4].

[1]**Corresponding author*: fabio.locastro@cnr.it.*

In parallel, Doppler-shift techniques and various correlation-based methods within microphone arrays were reviewed to estimate traffic parameters [Borkar and Malik], while a high-resolution spectral method (AR-z spectrum) was proposed to detect harmonic signals in noisy environments, enhancing the identification of vehicular acoustic features [5].

The primary objective of this study is to determine which of the two proposed methodologies, frequency-based or time-domain analysis, yields the lowest error rate in speed estimation. To validate the accuracy of the proposed approach, simultaneous speed measurements were obtained using a laser-based sensor, providing a reliable reference for direct comparison with the results derived from acoustic signal analysis.

2. MATERIALS AND METHODS

2.1 Sound source description and models

The object of investigation is the speed of a vehicle as it passes in front of one or more microphones, as illustrated in Figure 1. The vehicle is assumed to have multiple acoustic sources. In the case of an internal combustion engine (ICE) vehicle, the primary sources of noise include the engine, typically located at the front of the vehicle, the exhaust system at the rear, tire and wheel noise from four or more contact points, and aerodynamic noise

The ICE vehicles dominate at lower speeds, but as speed increases beyond 40 km/h [6], road and tire noise (caused by rolling resistance, road surface texture, and tread pattern) becomes more prominent. Aerodynamic noise is considered significant when speeds exceed 100 km/h [7-8].
The spectrum of the ICE noise exhibits a fundamental frequency, f_0 (Eqn. 1) in the frequency range of 50 to 150 Hz [9], which is proportional to the number of cylinders, N, the firing frequency and its harmonics. These components are influenced by the engine's revolutions per minute (RPM) and the transmission gear ratios.

$$f_0 = \frac{N}{2}\left(\frac{RPM}{60}\right) \tag{1}$$

The Exhaust System Noise, due to the turbulent exhaust gas flow and muffler design, has a frequency range of 50 Hz to 3 kHz. The road/tire/wheel noise generates broadband noise starting from 125 Hz to 4 kHz, but is predominant at frequencies ranging from 500 to 2 kHz [10]. In electric vehicles, the electric motor noise is in the range of 1 kHz to 10 kHz [11]. It is quieter than an IC engine, but at speeds higher than 50 km/h the whole noise is comparable, if they mount the same tire size and model, due mainly to road/tire noise. Wind and Aerodynamic Noise Sources due to the Airflow around the vehicle, mirrors, and body gaps are in the frequency range of 500 Hz to 5 kHz. [12].
The model used to simulate the vehicle noise source is, as a first approximation, a moving monopole source with broadband noise attributed to tire–road interaction. This modeling choice is justified by the fact that not all recorded vehicle sounds in this study originate from ICE vehicles. Moreover, in some ICE vehicles, engine noise is significantly reduced due to improved acoustic insulation and vibration-damping engine mounts [13].

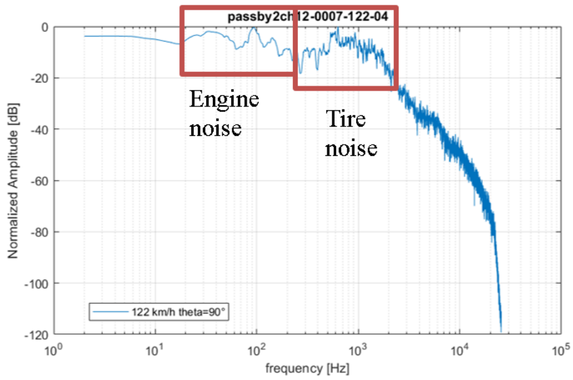

Figure 1. Example of a pass-by passenger vehicle normalized noise spectrum versus time.

2.2 Measurement set-up

Acoustic and speed measurements were conducted on the A91 highway in Rome, Italy, following the ISO 11819-1:2023 standard requirements. The highway consists of two lanes in each direction. For this study, two additional microphones were used for multisensor, located as illustrated in Figure 2.
Microphone M3 was placed 7.5 m from the centerline of the slow lane at a height of 3.0 m, while M1 and M2 were positioned 6.45 m from the centerline, also at 3.0 m in height. The distance between M3 and M1 (d_{13}) was 2.2 m, and the distance between M1 and M2 (d_1) was 1.9 m.
The microphones used were ½” Class I omnidirectional condenser microphones. Data acquisition was performed with a Sinus Apollo system, featuring four simultaneous acquisition channels, set to a sampling rate of 51,200 ksps.

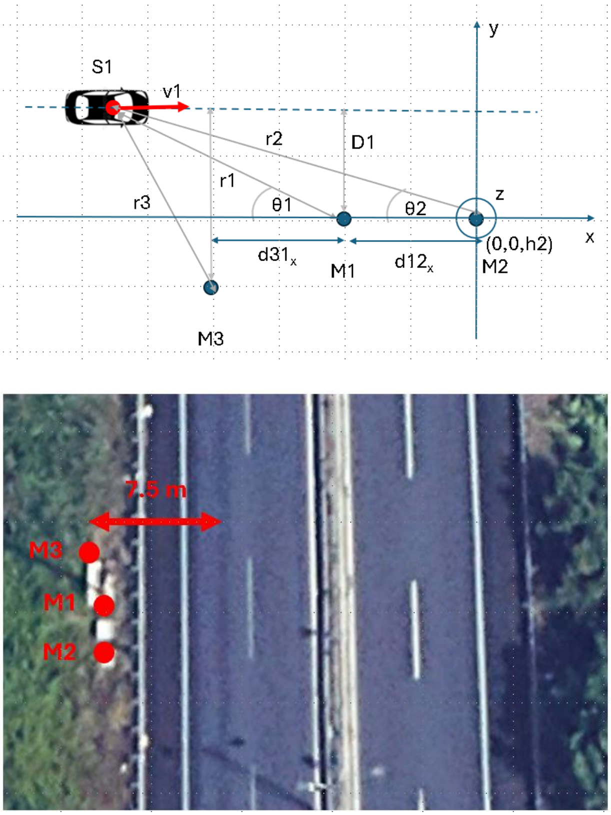

Figure 2. Measurement site and microphone positions; geometry of the problem, top view.

Vehicle speed was recorded using a SODI Scientifica KV Laser Statistic Analyzer, which provides an accuracy of ±1 km/h. This ensured precise speed measurements, allowing a reliable synchronization between noise levels and vehicle velocity, as required by the Statistical Pass-By (SPB) method.

The measurement site, located near the Smart Pole of the ANAS Smart Road project (PON-IR 2014-2020), was selected based on an inspection that ensured compliance with ISO 11819-1:2023 requirements. Particular attention was given to minimizing interferences from obstacles such as guardrails and road barriers, as their presence can alter sound measurements.

Tests were conducted at night when traffic density was lower, reducing the overlap of multiple sound sources and allowing a single vehicle pass-by. This also helped in ensuring that vehicles passed at a constant speed, preventing unwanted variations in noise levels and frequencies.

2.3 **Analysis methods**

In this study, both frequency-domain and time-domain analyses are employed for vehicle speed estimation. A single microphone (M3) is used for time- and frequency-domain analysis, leveraging sound pressure level (SPL) and the Doppler effect to estimate vehicle speed. Additionally, two microphones (M1 and M2) are utilized for time difference of arrival (TDOA) analysis to determine speed based on acoustic wave propagation delay. The inclusion of a third microphone (M3) enhances accuracy by precisely calculating the vehicle's distance from the microphones rather than relying on assumptions.

2.2.1 *Frequency-based analysis*

This approach utilizes variations in the spectral components of the acquired acoustic signal to infer vehicle speed, employing signal processing techniques.

Doppler shift and harmonic extraction

When a vehicle is moving relative to a fixed microphone position, the sound captured by the microphone presents changes in frequency content due to a Doppler effect. As the vehicle approaches, the sound is compressed, increasing its frequency; conversely, as it moves away, the frequency decreases. The frequency shift, $\Delta f = f(t) - f_0$, is measured in the spectrogram, generated by Short-Time Fourier Transform (STFT) or by the instantaneous phase (IF), obtained as the time derivative of the phase of the complex spectrum of the signal [14-15]. The source speed is estimated, as shown in Eqn. 2, based on the difference between the maximum and minimum detected harmonic's frequency.

$$|v_v| \cos(\theta) = \left(\frac{f_a - f_d}{f_a + f_d}\right) \quad (2)$$

Where: f_a and f_d are the frequencies during vehicle's approaching and departure the CPA; θ is the vehicle's approach angle; v_v is the vehicle's constant speed.

The frequency variation respect to f_o usually is less the 11% at the maximum highway speed.

.

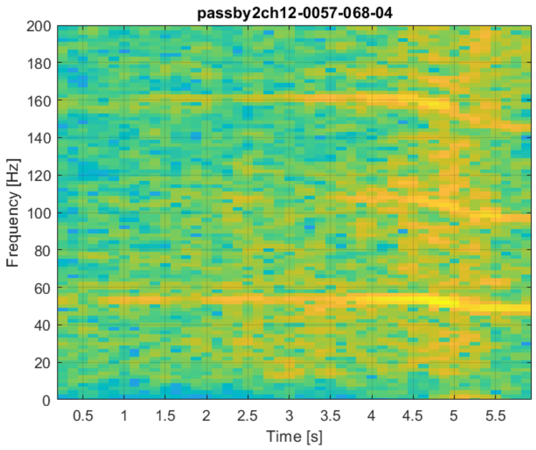

Figure 3. Example of a vehicle sonogram calculated using the STFT. The frequency variation of the engine harmonics due to the Doppler effect is observed.

To improve accuracy, engine harmonics are tracked, and spectral interpolation techniques are applied to refine signal variations [2].

2.2.2 *Time-Domain analysis*

In the time-domain analysis, two approaches were employed for vehicle speed estimation: one based on the measurement of sound pressure level over time, and the other on the time delay between signal arrivals at multiple microphones. In the first approach, the narrowing of the sound pressure level curve correlated with vehicle speed is used to estimate the speed. In the second method, correlation techniques combined with maximum likelihood estimation are applied to determine the time delay, which is then used to calculate the speed.

One microphone narrow envelope technique

This method relies on the observation that the sound pressure envelope of a passing vehicle exhibits a narrowing effect as vehicle speed increases [16]. This approach, which uses a single microphone, offers a low-cost and easy-to-deploy solution for traffic speed monitoring using minimal equipment, without requiring inter-sensor synchronization.

During the pass-by, the recorded acoustic signal shows a bell-shaped envelope centered around the vehicle’s closest point of approach (CPA) (Figure 4). As the vehicle moves faster, the duration of this envelope shortens, resulting in a steeper and narrower profile.

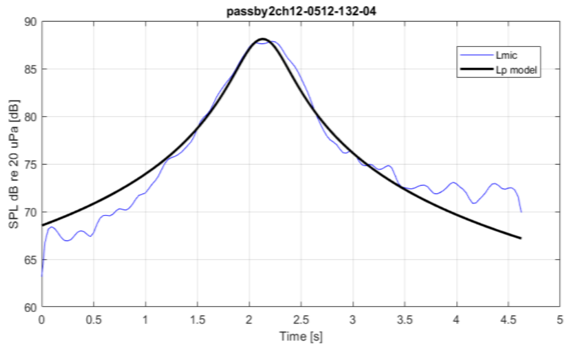

Figure 4. Sound pressure level recorded, Lmic, recorded at the microphone and curve modelled.

The sound pressure level, L_p, at the microphone can be written as:

$$L_p(t, v_0) = L_w - 8 - 20log_{10}(r(t, v_0)) + ID - A + n \quad (3)$$

Where: L_w is the sound power level of the vehicle; *r(t)* is the distance between the vehicle and the microphone changing in time; *ID* is the directivity index; *A* is the attenuation along the path; and *n* is the unwanted noise.

The distance between the microphone and the vehicle, which depends on time t and the vehicle's speed, v0, assumed constant during the measurement, can be expressed as:

$$r(t; v_0) = \sqrt{D^2 + \left(v_0(t - t_{CPA})\right)^2} \quad (4)$$

Where *D* is the perpendicular distance from the microphone when the vehicle is at the CPA at the time, t_{CPA}.

To mitigate the impact of variations in the distance between the vehicle and the microphone, amplitude normalization techniques and band-pass filtering are employed to eliminate low and high-frequency noise.

The signal has been analyzed in three frequency bands: (50–200) Hz, (500–3000) Hz, and (3000–8000) Hz, each corresponding to different sound sources as described in Paragraph 2.1. For each band, one or two sources were considered, located either at the center or at the front and rear of the vehicle, with corresponding modifications applied to Equation 4.

To determine the speed, v0, the vehicle speed v that minimizes the mean square error between the processed recorded sound pressure level, *y(t)*, and the model $L_p(t,v)$ is calculated as follows:

$$\underset{v}{\text{Arg min}} \left(\frac{1}{T}\sum_{t=0}^{T}\left(y(t) - L_p(t,v)\right)^2\right) \qquad (5)$$

The results obtained for each band have been averaged.

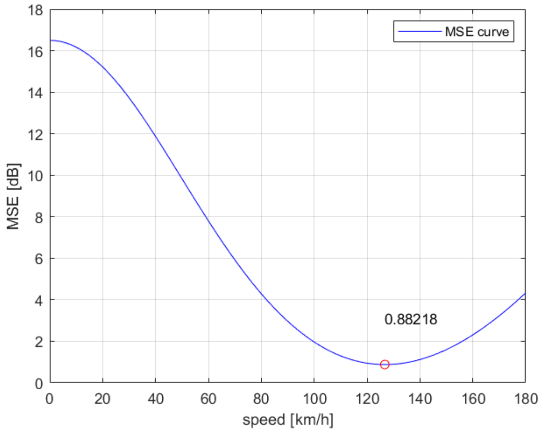

Figure 5. Curve of the mean square error of the Time Difference of Arrival (TDOA) estimation

Two-microphone technique and the Time Difference of Arrival (TDOA) estimation

In this method, at least two microphones are required to measure the TDOA, Δτ, of the acoustic wave generated by the vehicle, positioned at a known distance, *d*. It is given by:

$$\Delta\tau(t) = \frac{r_2(t) - r_1(t)}{c} \qquad (6)$$

Where r_1 and r_2 are the unknown vehicle distances from each of the two microphones, and *c* is the speed of sound. Δτ also represents the time delay between the acoustic signals received at each microphone. The received signals, $y_1(t)$ and $y_2(t)$, can be modelled as:

$$y_1(t) = h_1(t) * s(t) + n_1(t) \qquad (7)$$
$$y_2(t) = h_2(t) * s(t-\tau) + n_2(t)$$

Were *s(t)* is the source acoustic signal; $h_1(t)$ and $h_2(t)$ are the impulse responses from the instantaneous position of the moving vehicle and each microphone; $n_1(t)$ and $n_2(t)$ represent the unwanted signals, which also contain uncorrelated noise; *τ* is the delay of the second signal relative to the first signal, which is assumed to be the reference and not delayed. If the positions of the microphones are not sufficiently close to each other, the impulse responses cannot be simplified to an amplitude difference between the received signals.

Signal processing techniques are employed to identify similarities between signals captured by microphones, detecting the temporal offset between them. The maximum likelihood technique is applied, constructing a correlation matrix between microphone signals to identify the delay that maximizes correlation.

Due to the motion of the vehicle, the sound source is non-stationary, and thus the signal is also non-stationary. In this condition, the Least Mean Square (LMS) algorithm can be used to find the delay. However, if the process is considered quasi-stationary over a short time interval, this allows for the application of the Fourier transform, thereby enabling the use of the cross-correlation function, *R(t)*, defined as the product of the transformed signals [17], to estimate the delay

$$\underset{t}{\arg\max}\ (R_{ss}(t-\tau)) \qquad (8)$$

By substituting Eqn. 4 into Eqn. 6 for each distance, r_1 and r_2, the vehicle speed can be determined. [3]. As the vehicle approaches the CPA, the values of τ change rapidly, and the FFT-based correlation may fail. In this condition, it is useful to apply a fitting function. We use the logistic tanh function (Figure 6) to model the delay behavior:

$$\tau(t) = a \tanh\left(b(t-c)\right) \qquad (9)$$

Where a represents the delay when the vehicle is far from the microphone, *c* is the time corresponding to the CPA, and *b* is the slope of the curve, which is proportional to the vehicle speed.

$$v = b * D \quad \left[\frac{m}{s}\right] \qquad (10)$$

Where, *D* is the perpendicular distance from the vehicle to the microphones at the CPA.

This method has proven particularly useful in dense traffic environments, where tracking individual vehicles may be unfeasible. Passive sound detection using a microphone array has been shown to capture the overall acoustic pattern of traffic across multiple reference points [18].

By applying cross-correlation algorithms to the collected signal, the average vehicle speed within specific lanes can be estimated. Unlike Doppler-based methods, passive detection does not rely on a single vehicle's emitted signal but rather on the overall acoustic pattern of traffic, increasing its applicability in complex urban conditions.

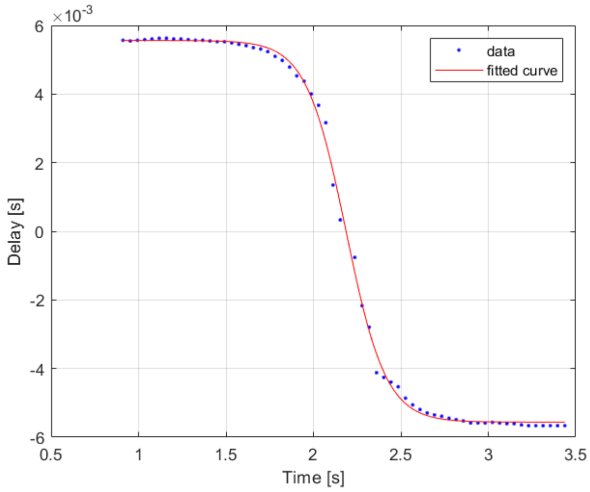

Figure 6. Signal time difference of wave arrival (blue dotted), and the fitting function.

3. RESULTS AND DISCUSSION

For speed calculation and accuracy evaluation, over 400 valid events were recorded, and the measured speeds were associated with these events. The distribution of the vehicle speed is shown in Figure 7. In each pass-by, only one vehicle was present, no vehicles in the other 3 lanes, and the difference between the maximum level during pass-by and the minimum before another pass-by was greater than 6 dB.

The classification was based only on vehicle type (passenger or heavy vehicles), without distinguishing between EVs and ICE vehicles.

Applying the different signal processing techniques, explained in the previous paragraph, to the recorded audio signals to estimate the vehicle speed, the results are summarized in Table 1 and Table 2.

Considering the increasing number of microphones from one to three, there is little improvement in the spread of the measured values from 10.0 km/h to 8.6 km/h. If it is considered the same number of two microphones, the method based on TDOA exhibits a better standard deviation from 8.9 km/h to 8 km/h.

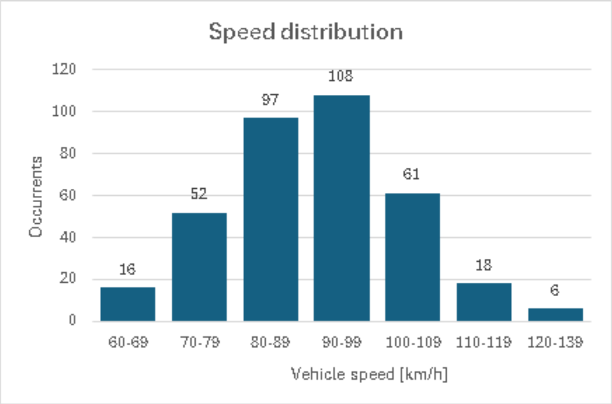

Figure 7. Vehicle speed distribution, measured with laser instrumentation.

The values at the extremes of the tables are not considered reliable due to the limited number of measured velocity cases in those ranges.

Table 1. Speed error estimation with the three methods based on the SPL envelope.

	1-Mic		2-Mic		3-Mic	
Speed [km/h]	**Mean [km/h]**	**Std [km/h]**	**Mean [km/h]**	**Std [km/h]**	**Mean [km/h]**	**Std [km/h]**
50-59	17.0	n.d.	14.1	n.d.	10.4	n.d.
60-69	5.7	9.0	7.7	11.1	9.6	11.7
70-79	1.4	8.9	5.0	9.2	7.6	11.0
80-89	0.1	9.9	4.7	11.0	5.9	10.3
90-99	-2.9	10.7	3.2	10.1	5.0	10.1
100-109	-1.4	9.9	0.3	8.8	3.0	9.6
110-119	-2.7	10.0	-4.7	10.3	-3.9	9.5
120-129	-14.2	12.8	6.0	7.3	0.5	8.1
130-139	-16.1	0.0	0.0	0.0	-2.7	2.7
Global	-0.9	10.0	3.1	8.9	4.5	8.6

Only 10 samples have been tested with the Doppler effect method, and without an autonomous algorithm. The algorithm is under development.

Table 2. Speed error estimation with the methods based on the Doppler effect (1-Mic) and TDOA (2-Mic)

	1-Mic		2-Mic	
Speed	**Mean**	**Std**	**Mean**	**Std**
[km/h]	**[km/h]**	**[km/h]**	**[km/h]**	**[km/h]**
50-59	n.d	n.d.	n.d.	n.d.
60-69	-1.6	n.d.	6.7	9.0
70-79	-12	n.d.	7.3	9.6
80-89	5.9	n.d.	5.6	9.4
90-99	3.3	n.d.	3.0	7.5
100-109	-3.3	n.d.	2.5	7.0
110-119	8.8	n.d.	-0.8	7.3
120-129	14.3	n.d.	-6.6	13.2
130-139	-3.7	n.d.	-4.9	4.9
Global	0.5	7.7	4.1	8.0

The high standard deviation observed in the speed evaluation is primarily attributed to limitations in the automatic processing. In the case of the SPL envelope-based methods, the model often fails to align accurately with the measured signal when the envelope's peak is flat rather than sharply defined, as assumed by the model. For the TDOA method, the error arises near the point of inversion of the time difference of arrival, where the fitting procedure may fail due to a limited number of data points or inaccuracies in their evaluation.

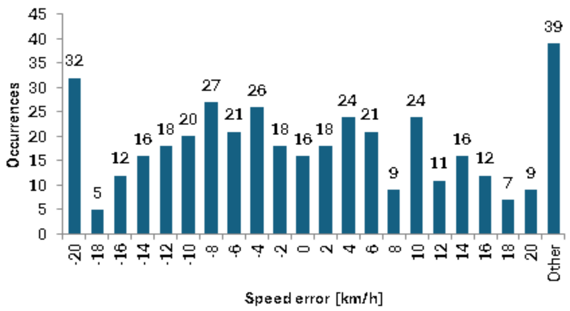

Figure 8. 1-Mic SPL envelope-based speed estimation method distribution error.

The distributions of the speed estimation error for the 1-mic SPL envelope-based and 2-mic TDOA are shown in Figure 8 and Figure 9. Although the standard deviation is higher for the first method, the failure rate, labeled as 'Other' in the figures, is lower than that of the second method, suggesting greater robustness.

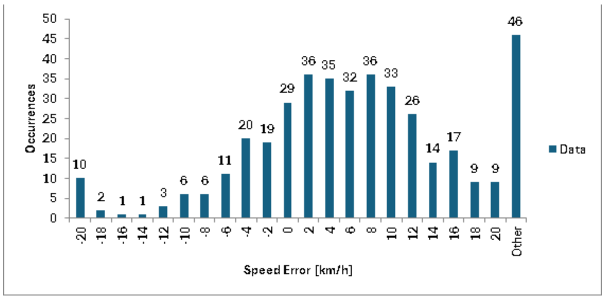

Figure 9. 2-Mic TDOA speed estimation method distribution error.

The other 2-mic method shows the same robustness, while the 3-mic method is the least robust method, with 61/401 fails.

4. CONCLUSIONS

In this study, a database of acoustic events, recorded during measurement sessions conducted in accordance with the ISO 11819-1 standard, was first established. Each event, corresponding to a single vehicle pass-by, was classified by vehicle type, either as a passenger vehicle or a heavy truck. This database was subsequently used to evaluate five elementary methods for estimating vehicle speed based on the pass-by sound generated by the vehicle itself. Four of these methods operate in the time domain, while one is based on frequency-domain analysis.

The uncertainty associated with the estimated values remains relatively high compared to other methods reported in the literature. This is primarily due to occasional failures in automatic processing, where the model fitting does not adequately match the measured data. Further studies are currently underway to address these issues and enhance the performance of the proposed algorithms

5. ACKNOWLEDGMENTS

This research was funded by the European Commission Executive Agency for Small and Medium-Sized Enterprises LIFE program (LIFE22-ENV-IT-LIFE SILENT - 101114310).

The authors declare no conflict of interest. The founders had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, and in the decision to publish the results.

6. REFERENCES

[1] F. Lo Castro, M.L Ariza Alvarez, M. De Luca, S. IAROSSI, P*roc INTER-NOISE 2024, Nantes, France*, Pages 7999 - 8999, pp. 8386-8397(12), DOI: https://doi.org/10.3397/IN_2024_4086

[2] S. Barnwal, R. Barnwal, R. Hegde, R. Singh, & B. Raj. (n.d.). Doppler-based speed estimation of vehicles using a passive sensor. *Proc. of an Unspecified Conf. or Journal.*

[3] F. Pérez-González, R. López-Valcarce, & C. Mosquera. (1998). Road vehicle speed estimation from a two-microphone array. *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP*), 2, 1321–1324.

[4] V. Rosão, E. Conceição, & L. Házyóvá. (2010). Method to determine the speed of vehicles by means of noise levels variation. *Proc. INTERNOISE 2010 Conf., Lisbon, Portugal.*

[5] H. Ding & B. F. Chao. (2015). Detecting harmonic signals in a noisy time-series: The z-domain autoregressive (AR-z) spectrum. *Geophys. J. Int., 201(3)*, 1287–1296. https://doi.org/10.1093/gji/ggv077

[6] Kephalopoulos S, Paviotti M, Anfosso-Lédée F. Common Noise Assessment Methods in Europe (CNOSSOS-EU). *EUR 25379 EN. Luxembourg (Luxembourg): Publications Office of the European Union;* 2012. JRC72550

[7] Beckenbauer, TRoad Traffic Noise. In: Müller, G., Möser, M. (eds) *Handbook of Engineering Acoustics. Springer, Berlin, Heidelberg 2013*. https://doi.org/10.1007/978-3-540-69460-1_15

[8] A. Lorea, V. Castelluccio, A. Costelli and M. Masoero, A Wind-Tunnel Method for Evaluating the Aerodynamic Noise of Cars, *SAE Transactions Vol. 95, Section 1 (1986*), pp. 1096-1105.

[9] T. Priede, Origins of automotive vehicle noise, *Journal of Sound and Vibration,V. 15, I. 1,1971*, pp. 61-73, ISSN 0022-460X, doi.org/10.1016/0022-460X(71)90360-9

[10] K. Iwao, I. Yamazaki A study on the mechanism of tire/road noise, *JSAE Review Volume 17, Issue 2*, April 1996, Pages 139-144

[11] Bao Meng, Chen Enwei, ,Lu Yimin, ,Liu Zhengshi, Liu Shuai, Vibration and noise analysis for a motor of pure electric vehicle, *Advanced Materials Research, ISSN: 1662-8985, Vols. 915-916,* pp 98-102.https://doi.org/10.4028/www.scientific.net/AMR.915-916.98

[12] George, A. and Callister, J., "Aerodynamic Noise of Ground Vehicles," *SAE Technical Paper 911027*, 1991, https://doi.org/10.4271/911027

[13] Somashekarappa M, A noise reduction technique in internal combustion engines: A multidisciplinary approach, *World Journal of Advanced Research and Reviews, 2020, 05(03*), 170–179

[14] Toshihiko Abe, Harmonics tracking and pitch extraction based on instantaneous frequency *Conference Paper in Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on · June 1995* DOI: 10.1109/ICASSP.1995.479804 · Source: IEEE Xplore

[15] Toshihiko Abe, Harmonics Estimation Based on Instantaneous Frequency and Its Application to Pitch Determination of Speech. *IEICE Transactions on Information and Systems · September 1995*

[16] Cevher, V., Chellappa, R., & McClellan, J. H. (2009). Vehicle speed estimation using acoustic wave patterns. *IEEE Transactions on Signal Processing, 57(1),* 30–47. https://doi.org/10.1109/TSP.2008.2005752

[17] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 24, no. 4, pp*. 320-327, August 1976, https://doi.org/10.1109/TASSP.1976.1162830

[18] P. Borkar & L. G. Malik. (2013). Review on vehicular speed, density estimation, and classification using acoustic signals. *Int. J. for Traffic and Transport Eng., 3(3)*, 331–343. https://doi.org/10.7708/ijtte.2013.3(3).08