Available online at www.sciencedirect.com

ScienceDirect

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

Original Research Paper

Scaling model for a speed-dependent vehicle noise spectrum

Giovanni Zambon [a], H. Eduardo Roman [b], Roberto Benocci [a,*]

[a] Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milano 1-20126, Italy
[b] Department of Physics, University of Milano-Bicocca, Milano 1-20126, Italy

HIGHLIGHTS

- Cluster analysis of vehicle noise spectra.
- Derivation of speed-dependent analytical functions fitting the average spectrum profiles.
- Prediction of vehicle speed based on noise spectrum patterns.

ARTICLE INFO

Article history:
Available online 19 May 2017

Keywords:
Car noise spectrum
Cluster analysis
Analytical model of noise spectrum
Speed-dependent analytical function

ABSTRACT

Considering the well-known features of the noise emitted by moving sources, a number of vehicle characteristics such as speed, unladen mass, engine size, year of registration, power and fuel were recorded in a dedicated monitoring campaign performed in three different places, each characterized by different number of lanes and the presence of nearby reflective surfaces. A full database of 144 vehicles (cars) was used to identify statistically relevant features. In order to compare the vehicle transit noise in different environmental condition, all 1/3-octave band spectra were normalized and analysed. Unsupervised clustering algorithms were employed to group together spectrum levels with similar profiles. Our results corroborate the well-known fact that speed is the most relevant characteristic to discriminate between different vehicle noise spectrum. In keeping with this fact, we present a new approach to predict analytically noise spectra for a given vehicle speed. A set of speed-dependent analytical functions are suggested in order to fit the normalized average spectrum profile at different speeds. This approach can be useful for predicting vehicle speed based purely on its noise spectrum pattern. The present work is complementary to the accurate analysis of noise sources based on the beamforming technique.

* *Corresponding author*. Tel.: +39 02 6448 2108; fax: +39 02 6448 2794.
E-mail addresses: giovanni.zambon@unimib.it (G. Zambon), eduardo.roman@mib.infn.it (H. E. Roman), roberto.benocci@unimib.it (R. Benocci).

Peer review under responsibility of Periodical Offices of Chang'an University.
http://dx.doi.org/10.1016/j.jtte.2017.05.001

1. Introduction

A sound generating device is generally identified according to its acoustic features which are referred to as acoustic signatures and are usually employed to discriminate among vehicles. Such signatures in case of moving sources are mostly linked to the engine vibrations and tire-road friction (Sandberg and Ejsmont, 2002). Among the techniques used to extract features in time-frequency domain there are Short Time Fourier Transform and Wavelet Transform (Munich, 2004; Sun and Qi, 2008). A signal processing time-domain technique used in sensor arrays for directional signal transmission or reception (Van Veen and Buckley, 1988), denoted as Beamforming, has been recently employed in car by-pass noise identification (Ballesteros et al., 2015). This technique allows one not only to locate the main noise sources during the pass-by of a vehicle, but also their characterization in terms of source strength.

Pattern recognition has been the subject of many studies which included techniques of compressed sensing (Candes et al., 2006; Donoho, 2006; Donoho et al., 2006) to disclose important acoustic features of an unknown signal and principal component analysis to dimensionally reduce such features employed to describe the difference between signals (Averbuch et al., 2012; Wang and Qi, 2002; Wu et al., 1999). For a review of such techniques we refer to Kakar and Kandpal (2013). Other advanced methods are based on laser Doppler vibrometry (Ometron, 2013), source height measurement (Glegg and Yoon, 1990), sound intensity (GMNA, 2005), nearfield acoustic holography (Ruhala, 1999) and spatial transformation of sound fields (Hald, 1995).

Different projects have been developed to determine common noise assessment methods for road, railway, aircraft and industrial noise in order to improve the reliability and the comparability of results across the EU Member States (CNOSSOS-EU (Kephalopoulos et al., 2012)). Further projects include prediction algorithms in order to forecast noise from specific sources such as road, rail traffic, aircraft and industrial sites (HARMONOISE (Bullen, 2012; Salomons et al., 2011), IMAGINE project (CORDIS, 2012)), providing a procedure to be adopted for strategic mapping as defined by the environmental noise directive. As for traffic noise, in the Harmonoise/Imagine model, particular attention has been paid to rolling noise, split into the vehicle and track contribution as well as wheel-surface interaction. The algorithms also include noise transport due to the combined effects of air absorption, the ground effect, shielding by topography (including barriers or buildings), atmospheric refraction, and atmospheric scattering. It is known that in the case of traffic noise, the events associated with each single vehicle transit are mainly random and characterized by fast or occasional sequences according to high or low vehicle flow rates. Monitoring traffic noise in a mid-long period (days or weeks) is usually carried out with 1 s temporal resolution recording both the spectrum at 1/3-octave band and the equivalent A-weighted level. In addition, traffic noise usually displays high variability in the noise spectra. For this reason, in this paper we address the question of the measured spectra variability and perform a detailed analysis by considering their average behavior. From these considerations, we developed a completely "blind" approach based upon the building up of statistically relevant classes characterized by "similar" spectrum profiles. The content of each group was then cross-checked against all the available vehicle characteristics. It is found that speed alone is sufficient to discriminate between different clusters, and, correspondingly an analytical model is presented. The model should be useful as a predictive instrument in a number of environmental applications. A similar approach has been applied to recorded hourly noise level profiles in the Dynamap-Life project (Zambon et al., 2014, 2015, 2016a,b).

The paper is organized as follows: In Section 2 we discuss the statistical analysis of the noise spectra produced by 144 vehicles pass-by, regarding their clustering and composition. In Section 3, we present an analytical model which fits the mean noise spectra very well, allowing us to make predictions for different car speeds. Finally, in Section 4 we present the conclusions.

2. Statistical analysis of noise spectra

2.1. *Database description*

Considering the well-known features of the noise emitted by moving vehicles, a monitoring campaign was planned in order to obtain detailed information on the moving sources. The monitoring has been performed in three different places, each characterized by different numbers of lanes and the presence of nearby reflective surfaces. The experimental measurements are part of a project (Zambon and Radaelli, 2012) aimed at providing a control system for real time field monitoring of single vehicle emission. The recorded data have been obtained using a 1/3-octave band spectrum, and further analyzed (Peeters and Blokland, 2007), in order to make it in accordance with the accepted standards ISO 362-1 (International Organization for Standardization, 2007) and Regulation No. 51 (GRB Expert Group, 2007). As equipment, a sound level meter providing the acquisition of A-weighted equivalent level, $L_{Aeq,100ms}$, (for 1/3-octave band analysis in the frequency range 25 Hz to 10 kHz), a speed detector (radar) and a camera were employed. The camera was used for the identification of regular transits in double lane roads which might be influenced by overtaking and to read the vehicle plate number. The latter was used to obtain information on the year of registration, unladen mass, engine size, power and fuel. In all measuring sites the microphone was positioned 5.5 m from the axes of the road and 1.2 m above the road surface. All the measurements were synchronized by means of a photocell. More details on the experimental setup and procedure are given in Zambon and Radaelli (2012). In this work, we consider the original 1/3-octave band spectra recorded during the vehicle pass-by, which contain the whole vehicle information such as speed, fuel, age, etc.

2.2. *Clustering of vehicle spectra*

The database made of 144 vehicles (cars) was used to identify statistically relevant features. Because of the non-homogeneity of the equivalent level (A-weighted) dataset, L_{Aeq}, due to

different conditions of the road such as its geometry, the presence of reflecting surfaces and obstacles in sound propagation and types of paving, we proceeded normalizing each jth 1/3-octave band spectrum.

The normalized spectrum, S_j, can therefore be regarded as dependent on each vehicle feature and written as the ratio between the spectrum level and its mean value $\overline{L_{Aeq_j}}$ (Eq. (1)).

$$S_j(f,k_1,k_2,\cdots,k_n) = L_{Aeq_j}(f,k_1,k_2,\cdots,k_n)\Big/\overline{L_{Aeq_j}}(k_1,k_2,\cdots,k_n) \qquad j=1,2,\cdots,144 \tag{1}$$

where f and k_n represent the 1/3-octave band frequency and the generic vehicle feature, respectively, n is the number of features considered.

Unsupervised clustering algorithms were employed to group together normalized spectrum levels found to be "close" to one another. Various algorithms, such as hierarchical agglomeration (Ward, 1963), K-means algorithm (Hartigan and Wong, 1979), and partitioning around medoids (PAM) (Rousseeuw and Kaufman, 1990), were considered, and their results compared. The number of clusters was chosen as a compromise between satisfactory discrimination and the need to limit the number of groups. Euclidean distance was chosen as metric among observations.

The statistical software R (The R project), a free software environment for statistical computing and graphics, was employed for the analysis. The package "clValid" (Brock et al., 2008) was used for validating the results. All the clustering algorithms were ranked based on their performance as determined simultaneously by all the validation measures (Pihur et al., 2007). Thus, the optimal list, obtained through a score assigned by each validation index, gives a two-cluster hierarchical agglomeration at the first place followed by K-means and PAM methods, each one yielding also a two-cluster separation.

In addition, we compared the results obtained by the hierarchical algorithm with two clusters with those of the K-means with three clusters. The result with three clusters was considered in order to achieve better discrimination.

The values for S_j are not normally distributed over all frequencies. Therefore, to check the statistical independence of the two (three) average cluster profiles, a student's t-test was performed on data following a normal distribution, whereas a nonparametric Mann–Whitney–Wilcoxon's (MWW's) test was applied to non-normal populations. The two-tailed p-values for both tests were much lower than the $\alpha = 0.05$ significance level (or 95% confidence level) with the exception of the 400 Hz 1/3-octave band in case of two clusters (Fig. 1).

This means that, at 95% confidence level, the two-cluster average spectrum profiles can be considered independent with the exception of the 400 Hz 1/3-octave band. In fact, the latter represents the point where the two average profiles intersect, as we will see below. Both the student's t-test and the MWW's test confirm the same result, representing, therefore, a good discrimination between the two average spectrum profiles.

The same procedure has been performed for a three-cluster result. In this case, we found a much broader feature (Figs. 2 and 3). Criticality is found between cluster 1 vs cluster 3 where the two-tailed p-values exceed the threshold level in correspondence of the following 1/3-octave bands: 80, 200, 250, 400, 500, 630, 800, 1000, 1600 Hz. In case of cluster 1 vs cluster 2 it occurs at 400 Hz 1/3-octave band, and for cluster 2 vs cluster 3 at 315 and 400 Hz. Both the student's t-test and MWW's test are in agreement to each other.

As mentioned above, the distribution of S_j is not normal for some frequencies, therefore in order to obtain a meaningful and more accurate estimator of the population, the mean and its confidence level were calculated applying a resampling procedure: the bootstrap method (Efron and Tibshirani, 1993), that consists in considering 1000 samples with replication randomly extracted from our initial sample.

Figs. 4 and 5 illustrate the multi-dimensional scaling (MDS) results applied to the data to provide a visual representation of the pattern of proximities among the data. The distinction among clusters, marked by different colors, is rather good. In particular, we can observe that cluster 1, for the two-cluster solution, splits into two sub-clusters (cluster 1 and cluster 3) in case of the three–cluster result. However, as we can see from these results, the many over-threshold peaks of Fig. 3 (cluster 1 vs cluster 3) manifest in the close vicinity of the two clusters (blue and red) in Fig. 5.

According to these results, we proceeded with the study of the mean noise spectra associated with each cluster

$$C_i(f,k_1,k_2,\cdots,k_n) = \sum_{j\in \text{cluster i}} S_j(f,k_1,k_2,\cdots,k_n)\frac{1}{m_i} \tag{2}$$

where m_i is the number of cars in the ith cluster.

Fig. 6 shows the two normalized cluster average spectra, C, for the two-cluster result. As we can observe, each normalized spectrum can be divided into three regions: a low frequency interval up to 200 Hz, a medium interval of frequencies

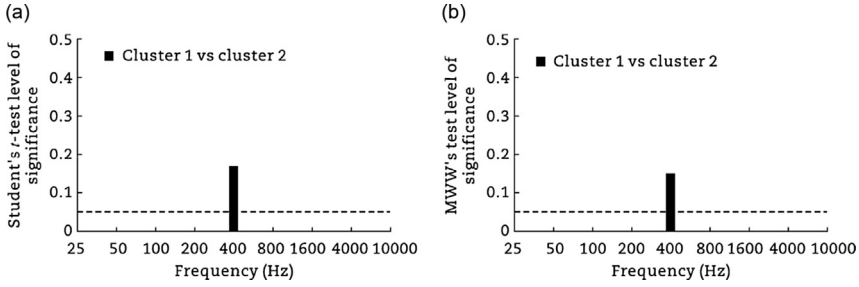

Fig. 1 – Results for two clusters. (a) Student's t-test. (b) Mann–Whitney–Wilcoxon's test.

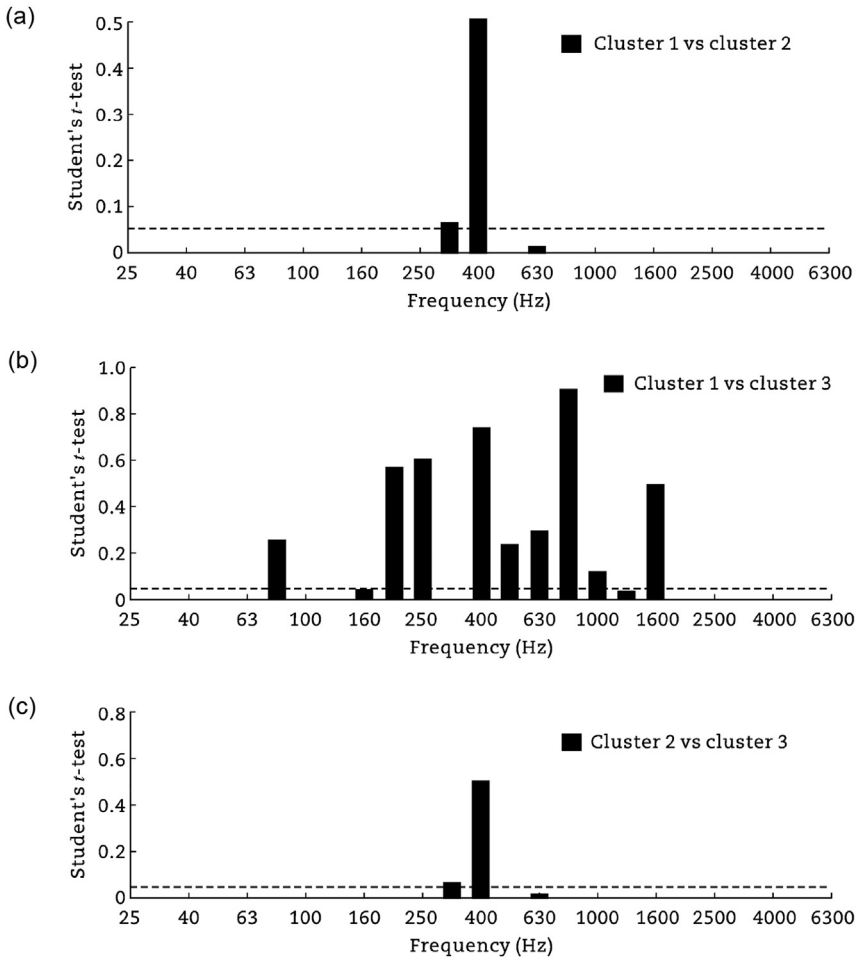

Fig. 2 – Student's t-test results for three clusters.

between 200 and 1000 Hz and a high frequency interval for values higher than 1000 Hz. In the low frequency region, the two curves are well-separated and the bumps which appear between 50 and 100 Hz, may be attributed to engine rotations. In the central frequency region, the curves tend to get closer and overlap in agreement with the results of independence tests. At higher frequencies, the spectra display bumps around 1000 Hz, followed by well-separated decreasing tails. The source of the bump at 1000 Hz may be attributed to tyre-pavement interaction.

As a comparison, we also considered a clustering obtained by the *K*-means algorithm with three clusters. This choice, as mentioned above, was made by the capability of such algorithm to perform an acceptable cluster separation and therefore potentially reveal more details on the cluster content. The results of such analysis are reported in Fig. 7.

As in the previous case for two-cluster results, we observe the typical features of the contribution due to engine rotations at low frequencies, a mid-frequency range where the normalized spectrum profiles tend to overlap assuming a flat trend, and a high frequency interval due to the tyre-road interaction manifested by the presence of bumps followed by well-separated spectrum tails.

2.3. *Clusters composition*

In order to determine the composition of each cluster and therefore to assess the parameters driving the cluster formation, each group was cross-checked with each parameter identifying the examined vehicles. Among these, the vehicle speed showed to prevail over the other characteristics (Benocci et al., 2014). In particular, Fig. 8 shows the histogram and the probability density of the content of each cluster with respect to the measured vehicle speed (27–70 km/h). Cluster 1 represents vehicles with a mean speed $v_1 = 46.8$ km/h (106 cases), whereas cluster 2 is characterized by vehicles with a mean speed $v_2 = 39.4$ km/h (38 cases), suggesting the possibility of naturally grouping spectra of vehicles according to their speed. In other words, clusters do not show any sensitive dependence on other parameters like engine power, fuel type, year of registration, vehicle weight, engine size and power. Therefore, the normalized cluster means spectrum, *C*, now becomes

$$C(f, k_1, k_2, \cdots, k_n) \equiv C(f, v) \tag{3}$$

explicitly showing the dependence on the mean cluster speed, v.

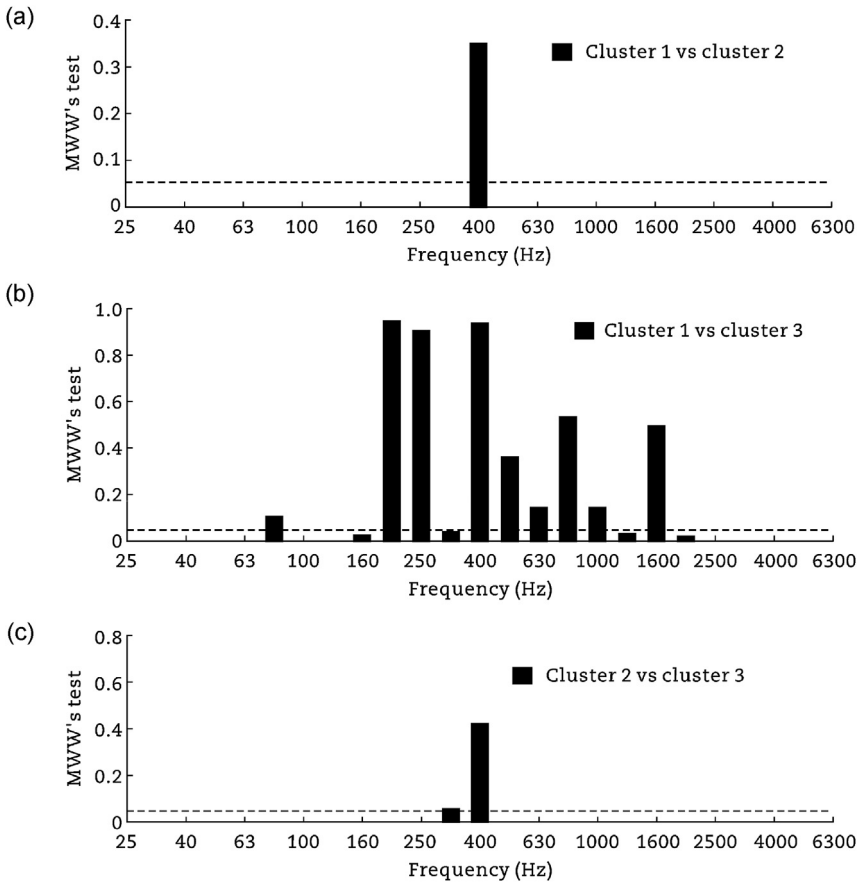

Fig. 3 – Mann–Whitney–Wilcoxon's test results for three clusters.

The same result is obtained for three clusters as shown in Fig. 9. In this case, speed range is 27–70 km/h, and the three clusters are characterized by the following mean vehicle speeds: $v_1 = 49$ km/h for cluster 1, $v_2 = 38.5$ km/h for cluster 2 and $v_3 = 44.5$ km/h for cluster 3. As is apparent, the previous cluster 1 has roughly split into two smaller sub-groups. Therefore, it seems that the approach aimed at describing the traffic in terms of car speed is reflected in the formation of clusters, each one characterized mainly by its distinct speed distribution which can be further implemented by using three clusters.

3. Analytical model

The peculiar spectrum profile, associated with each cluster and characterized by a different mean speed, can be extended to the speeds not included in the present set of data. For this reason, the speed-dependent spectrum was fitted by a set of functions each defined in a specific frequency range.

As suggested by the results shown in Fig. 6, the spectrum is described by a power-law and a Gaussian curve for

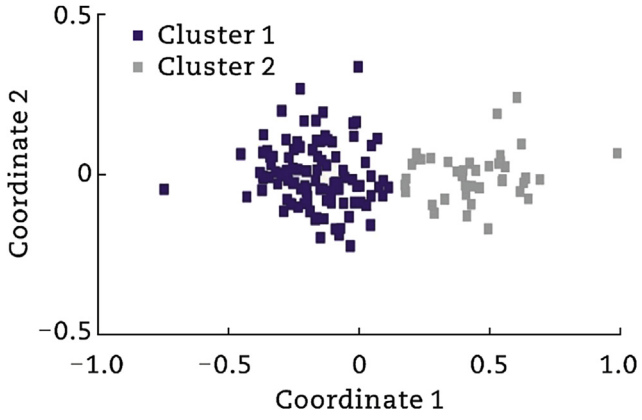

Fig. 4 – Multi-dimensional scaling of level profiles with the solution at two clusters.

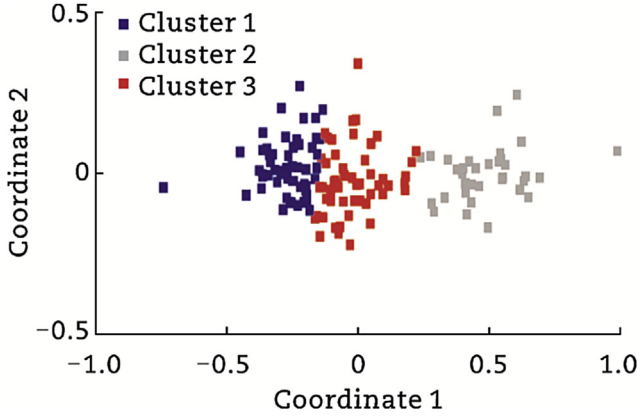

Fig. 5 – Multi-dimensional scaling of level profiles with the solution at three clusters.

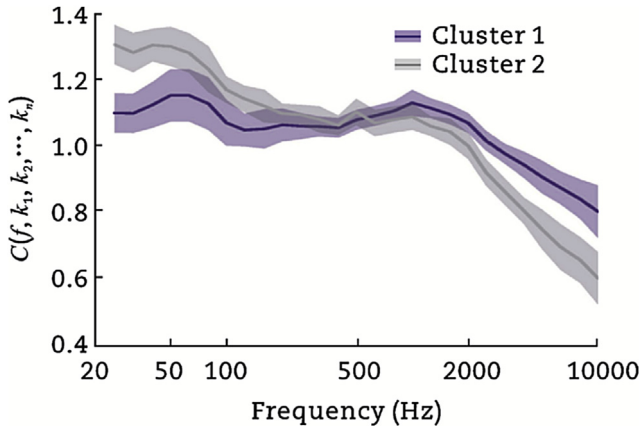

Fig. 6 – Average normalized spectrum profiles $C(f, k_1, k_2, \cdots, k_n)$ for each hierarchical cluster and their confidence levels (one standard deviation).

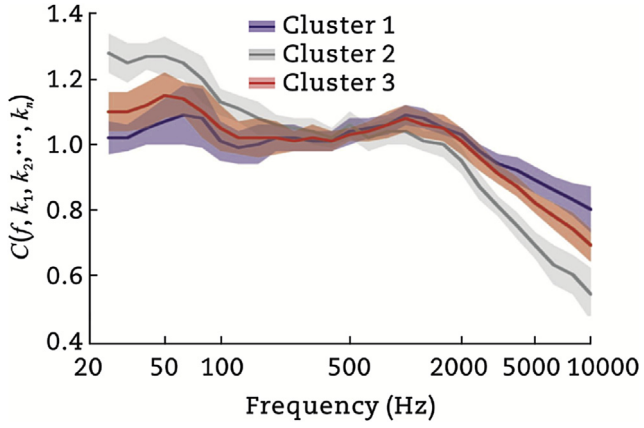

Fig. 7 – Average normalized spectrum profiles $C(f, k_1, k_2, \cdots, k_n)$ for each K-means cluster and their confidence levels (one standard deviation).

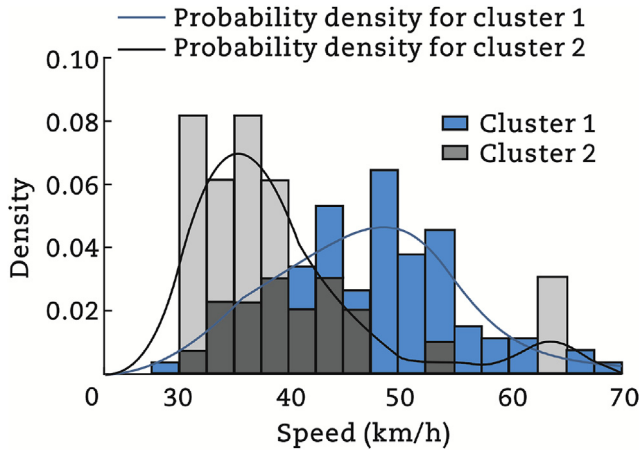

Fig. 8 – Histogram and probability density vs vehicle speed for the hierarchical two-cluster results.

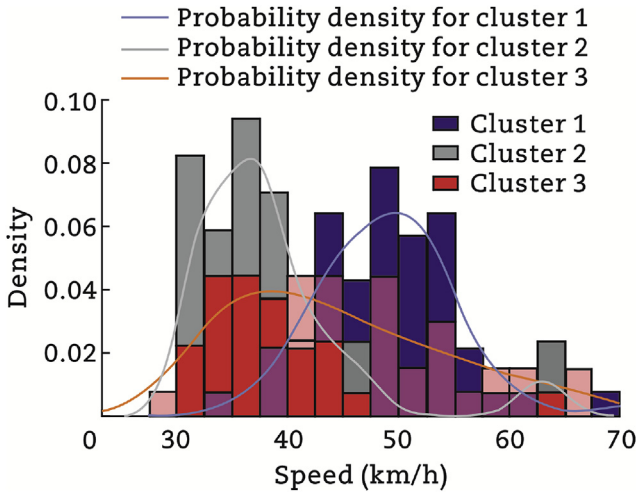

Fig. 9 – Histogram and probability density vs vehicle speed for the K-means three-cluster results.

frequencies below 500 Hz, it can be fitted by a Gaussian function in the range between 500 and 1600 Hz, whereas for higher frequencies a power-law represents the optimal fit. Eq. (4) describes this approach for the normalized mean cluster speed-dependent spectrum $C(f, v)$.

$$C(f,v) = \begin{cases} A(v)\left(\frac{f}{f_0}\right)^{-\alpha(v)} + \frac{E}{\sqrt{2\pi}\sigma} e^{-\frac{(f-f_G(v))^2}{2\sigma^2}} & f \le 500 \text{ Hz} \\ \frac{E}{\sqrt{2\pi}\sigma} e^{-\frac{(f-f_G(v))^2}{2\sigma^2}} & 500 \text{ Hz} < f < 1600 \text{ Hz} \\ B(v)\left(\frac{f}{f_0}\right)^{-\alpha(v)-\beta(v)} & f \ge 1600 \text{ Hz} \end{cases} \quad (4)$$

where α and β are the two speed-dependent power-law exponents, A and B are the two speed-dependent power-law constants, E is the amplitudes of the Gaussian distribution, f_G is the positions of the center of the Gaussian peaks, α is the full widths at half maximum, $f_0 = 1$ Hz.

Results of the fitting for the two and three cluster profiles are shown in Figs. 10 and 11.

In Fig. 10, two-cluster profiles are characterized by $v_1 = 46.8$ km/h and $v_2 = 39.4$ km/h. The corresponding fits

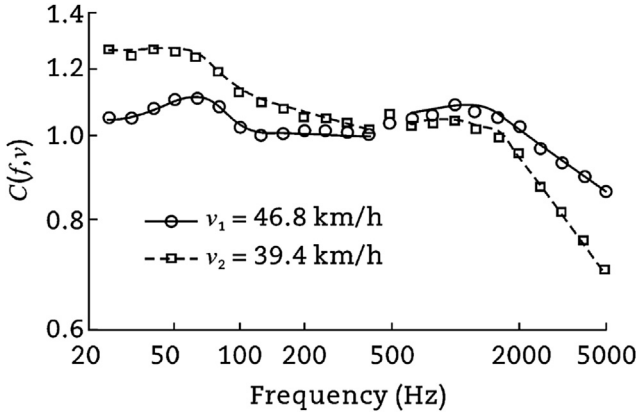

Fig. 10 – Results of the fitting for two-cluster profiles.

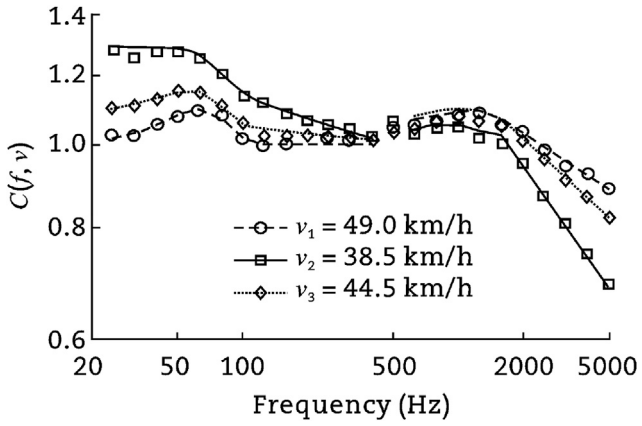

Fig. 11 – Results of the fitting for three-cluster profiles.

obtained using Eq. (4) are displayed by the lines. The values of the fitting parameters are reported in Table 1. In Fig. 11, three-cluster profiles are characterized by $v_1 = 49$ km/h, $v_2 = 38.5$ km/h and $v_3 = 44.5$ km/h. The corresponding fits obtained using Eq. (4) are displayed by the lines. The values of the fitting parameters are reported in Table 2.

Table 1 – Results of the fitting procedure of two-cluster profiles.

Two-cluster profile	α	β	A (1/Hz)	B (1/Hz)	E	σ (Hz)	f_G (Hz)
$v_1 = 46.8$ km/h							
$f \leq 500$ Hz	0.01	–	1.06	–	5	20	63
500 Hz $< f <$ 1600 Hz	–	–	–	–	20	250	1100
$f \geq 1600$ Hz	0.01	0.17	–	4.0	–	–	–
$v_2 = 39.4$ km/h							
$f \leq 500$ Hz	0.07	–	1.55	–	5	20	57
500 Hz $< f <$ 1600 Hz	–	–	–	–	20	250	900
$f \geq 1600$ Hz	0.07	0.27	–	12.5	–	–	–

Table 2 – Results of the fitting procedure of three-cluster profiles.

Three-cluster profile	α	β	A (1/Hz)	B (1/Hz)	E	σ (Hz)	f_G (Hz)
$v_1 = 49$ km/h							
$f \leq 500$ Hz	0.001	–	1.00	–	5	20	63
500 Hz $< f <$ 1600 Hz	–	–	–	–	20	250	1120
$f \geq 1600$ Hz	0.001	0.159	–	3.45	–	–	–
$v_2 = 38.5$ km/h							
$f \leq 500$ Hz	0.080	–	1.63	–	5	20	57
500 Hz $< f <$ 1600 Hz	–	–	–	–	20	250	900
$f \geq 1600$ Hz	0.080	0.270	–	13.50	–	–	–
$v_3 = 44.5$ km/h							
$f \leq 500$ Hz	0.020	–	1.14	–	5	20	58
500 Hz $< f <$ 1600 Hz	–	–	–	–	20	250	1000
$f \geq 1600$ Hz	0.020	0.205	–	5.60	–	–	–

In order to obtain a predicting model of the average spectrum profile for a set of vehicles moving at different speeds from those under investigation in the present set of data, the calculated fitting parameters, referring to the average cluster behavior, have been in turn fitted by a speed-dependent function characterized by the empirical parameters a, b, g, m, and q. In particular, we employed the following forms.

$$\{\alpha(v), \beta(v)\} \sim a e^{bv} \tag{5}$$

$$\{A(v), B(v)\} \sim y = y_0 + g e^{hv} \tag{6}$$

$$f_G(v) \sim mv + q \tag{7}$$

Therefore, an exponential speed-dependent function for the power-law coefficients, $\{\alpha, \beta\}$, and a linear function for the position of the center of the Gaussian peak, f_G, were adopted.

In Tables 3–5, the calculated coefficients are reported according to the fitting functions illustrated in Eqs. (5)–(7) and in Figs. 12–16 such coefficients are illustrated together with their fitting curve and the corresponding 95% confidence level.

An interesting feature which emerges from this analysis is related to the dependence of the bumps in the mean spectrum profiles of Figs. 10 and 11 on the vehicle speed (Eq. (7) and Figs. 15 and 16). This result, therefore, confirms the link between speed and the spectrum shift, due to engine rotations and tire-surface interaction, towards higher frequencies.

4. Conclusions

Contrary to beamforming technique whose main objective is the noise source identification and characterization, in our work we focused on the spectrum emitted by the pass-by of cars. Analyzing such spectra, we conclude that the only characteristic differentiating statistically obtained groups is the vehicle speed. In particular, a full database of 144 car noise

Table 3 – Fit for coefficients $\alpha(v)$, $\beta(v)$ according to $y = ae^{bv}$, with $y = \{\alpha, \beta\}$.

	a	b (h/km)
α	4086	−0.28
β	2.21	−0.05

Table 4 – Fit for coefficients $A(v)$, $B(v)$ according to $y = y_0 + ge^{hv}$, with $y = \{A, B\}$.

	y_0 (1/Hz)	g (1/Hz)	h (h/km)
A	0.96	14,601	−0.26
B	2.57	203,247	−0.25

Table 5 – Fit for the low and high frequency position of the center of the Gaussian peak according to $f_G = mv + q$.

f_G (Hz)	m (Hz·h/km)	q (Hz)
Low frequency Gaussian fit	0.63	32
High frequency Gaussian fit	22.20	38

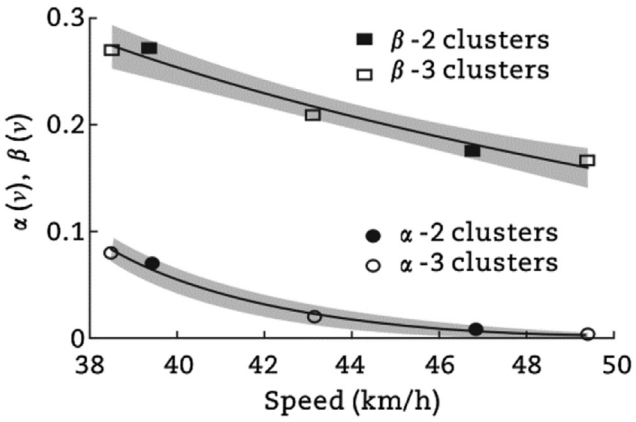

Fig. 12 – Fit for coefficients $\alpha(v)$, $\beta(v)$ and their 95% confidence levels.

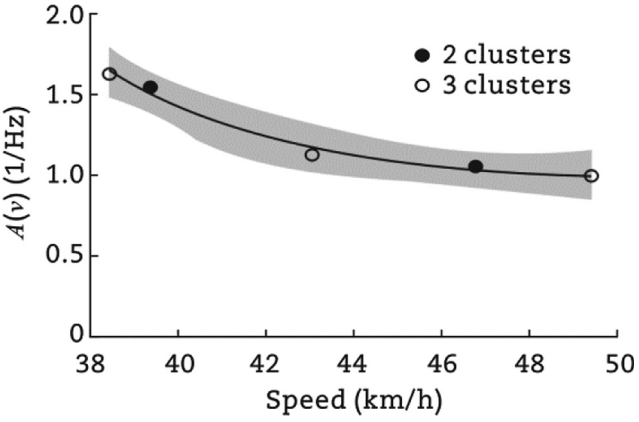

Fig. 13 – Fit for coefficient $A(v)$ and its 95% confidence level.

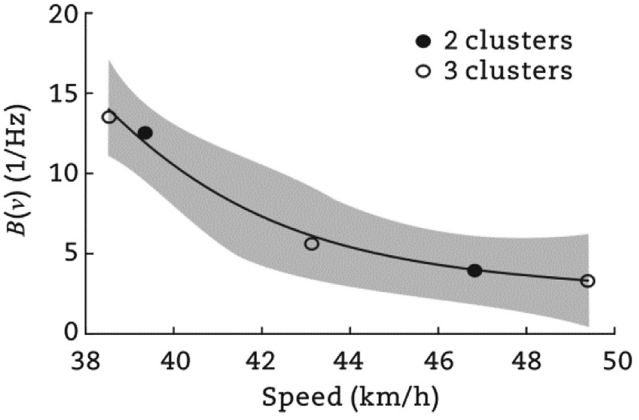

Fig. 14 – Fit for coefficient $B(v)$ and its 95% confidence level.

spectra was investigated to identify common features related to speed, engine power, engine size, weight, fuel type and year of registration. Spectra were statistically analyzed according to unsupervised cluster algorithms in order to group together "similar" spectrum profiles. The package "clValid" was employed to determine the optimal number of clusters and clustering algorithms. For comparison and to achieve better

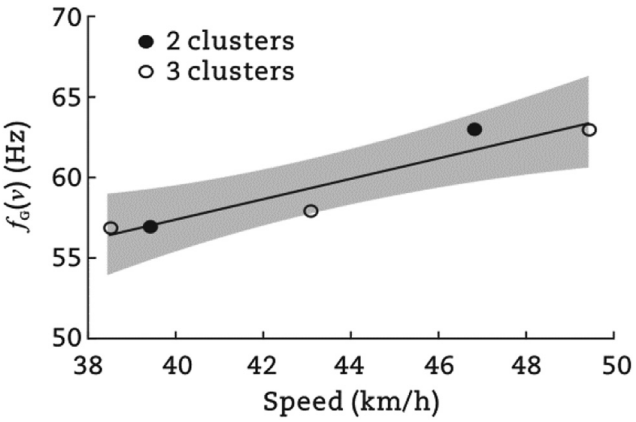

Fig. 15 – Fit for coefficient $f_G(v)$ and its 95% confidence level in the low frequency range.

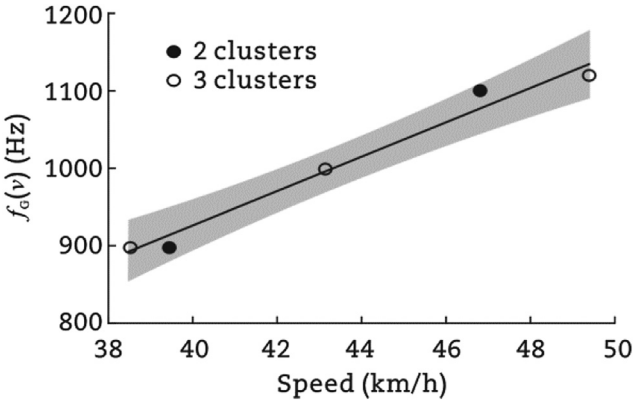

Fig. 16 – Fit for coefficient $f_G(v)$ and its 95% confidence level in the high frequency range.

resolution, out of this rank we considered the solution obtained by hierarchical algorithm for two clusters and the solution using *K*-means method for three clusters. The obtained groups resulted to be clearly composed by spectra belonging to cars with different mean speed: $v_1 = 46.8$ km/h and $v_2 = 39.4$ km/h for the two-cluster, and $v_1 = 49$ km/h, $v_2 = 38.5$ km/h and $v_3 = 44.5$ km/h for the three-cluster solution. The profiles for the three-cluster solution, characterized by $v_1 = 49$ km/h and $v_3 = 44.5$ km/h mean speed, come mainly from the splitting of the profile characterized by $v_1 = 46.8$ km/h mean speed of the two-cluster solution. The other parameters characterizing each car transit did not show significant relations with the statistically composed groups. Thus, the two-cluster solution turns out to be optimal, although the three-cluster one allows for a better speed resolution, still displaying distinct (i.e., independent) spectrum profiles. A set of speed-dependent analytical functions was also suggested in order to fit the normalized average spectrum profile at different speeds. This approach can be useful for predicting vehicle speed based purely on its noise spectrum pattern. The present work is complementary to the accurate analysis of noise sources based on the beamforming technique (Ballesteros et al., 2015).

REFERENCES

Averbuch, A., Rabin, N., Schclar, A., et al., 2012. Dimensionality reduction for detection of moving vehicles. Pattern Analysis and Applications 15 (1), 19–27.

Ballesteros, J.A., Sarradj, E., Fernández, M.D., et al., 2015. Noise source identification with Beamforming in the pass-by of a car. Applied Acoustics 93, 106–119.

Benocci, R., Roman, H.E., Zambon, G., et al., 2014. Traffic acoustical classes based upon vehicle characteristics. In: Forum Acusticum, Krakow, 2014.

Brock, G., Pihur, P., Datta, S., et al., 2008. clValid: an R package for cluster validation. Journal of Statistical Software 25, 1–28.

Bullen, R., 2012. The Harmonoise noise prediction algorithm: validation and use under Australian condition. In: Acoustics 2012, Fremantle, 2012.

Candes, E., Romberg, J., Tao, T., 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory 52 (2), 489–509.

CORDIS, 2012. Final Report Summary: IMAGINE Improved Methods for the Assessment of the Generic Impact of Noise in the Environment. Available at: http://cordis.europa.eu/result/rcn/47869_it.html (Accessed 3 October 2012).

Donoho, D.L., 2006. Compressed sensing. IEEE Transactions on Information Theory 52 (4), 1289–1306.

Donoho, D.L., Tsaig, Y., Drori, I., et al., 2006. Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit. 2006-2. Department of Statistics, Stanford University, Stanford.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall/CRC, London.

General Motors Company (GMNA), 2005. Road Tire Noise Evaluation Procedure. GMN7079TP. GMNA, Englewood.

Glegg, S., Yoon, J., 1990. Determination of noise source height, part I: the measurement of equivalent acoustic source height above a reflecting surface. Journal of Sound Vibration 143 (1), 19–37.

GRB Expert Group, 2007. Uniform Provisions Concerning the Approval of Motor Vehicles Having at Least Four Wheels with Regard to Their Noise Emission, Revision 1. Regulation No. 51. Regulation UN/ECE. Geneva.

Hald, J., 1995. Technical Review No. 1: Spatial Transformation of Sound Fields in the Automotive Industry Pass-by Noise Measurements. Available at: https://www.bksv.com/media/doc/bv0046.pdf (Accessed 19 April 2013).

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a K-means clustering algorithm. Journal of the Royal Statistical, Series C (Applied Statistics) 28 (1), 100–108.

International Organization for Standardization (ISO), 2007. Measurement of Noise Emitted by Accelerating Road Vehicles–Engineering Method – Part 1: M and N Categories. ISO 362-1. ISO, Geneva.

Kakar, V.K., Kandpal, M., 2013. Techniques of acoustic feature extraction for detection and classification of ground vehicles. International Journal of Emerging Technology and Advanced Engineering 3 (2), 419–426.

Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F., 2012. Common Noise Assessment Methods in Europe, (CNOSSOS-EU). Publication Office of the European Union, Luxembourg.

Munich, E.M., 2004. Bayesian subspace methods for acoustic signature recognition of vehicles. In: The European Signal Processing Conference (EUSIPCO-04), Vienna, 2004.

Ometron, 2013. Areas of Vibrometer Applications. Available at: http://www.ometron.com/applications/index.html (Accessed 19 April 2013).

Peeters, B., Blokland, G.V., 2007. The Noise Emission Model for European Road Traffic. Available at: https://www.mp.nl/sites/all/files/publicaties/IMA55TR-060821-MP10%20-%20IMAGINE%20Deliverable%20D11.pdf (Accessed 3 October 2012).

Pihur, V., Datta, S., Datta, S., 2007. Weighted rank aggregation of cluster validation measures: a monte carlo cross-entropy approach. Bioinformatics 23 (13), 1607–1615.

Rousseeuw, P.J., Kaufman, L., 1990. Finding Groups in Data. Wiley, Hoboken.

Ruhala, R., 1999. A Study of Tire/pavement Interaction Noise Using Nearfield Acoustical Holography (PhD thesis). The Pennsylvania State University, University Park.

Salomons, E., Van Maercke, D., Defrance, J., et al., 2011. The Harmonoise sound propagation model. Acta Acustica United with Acustica 97 (1), 62–74.

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

Sun, Y., Qi, H., 2008. Dynamic target classification in wireless sensor networks. In: 19th International Conference on Pattern Recognition, Tampa, 2008.

The R Project for Statistical Computing. Available at: http://www.r-project.org/(Accessed 3 October 2012).

Van Veen, B.D., Buckley, K.M., 1988. Beamforming: a versatile approach to spatial filtering. IEEE ASSP Magazine 5 (2), 4–24.

Wang, X., Qi, H., 2002. Acoustic target classification using distributed sensor arrays. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, 2002.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58 (301), 236–244.

Wu, H., Siegel, M., Khosla, P., 1999. Vehicle sound signature recognition by frequency vector principal component analysis. IEEE Transactions on Instrumentation and Measurement 48 (5), 1005–1009.

Zambon, G., Radaelli, S., 2012. Measurements of noise from single road vehicle in urban traffic. In: Ninth European Conference on Noise Control, Prague, 2012.

Zambon, G., Benocci, R., Bisceglie, A., 2015. Development of optimized algorithms for the classification of networks of road stretches into homogeneous clusters in urban areas. In: The 22nd International Congress on Sound and Vibration (ICSV), Florence, 2015.

Zambon, G., Benocci, R., Angelini, F., et al., 2014. Statistics-based functional classification of roads in the urban area of Milan. In: The 7th Forum Acusticum, Krakow, 2014.

Zambon, G., Benocci, R., Brambilla, G., 2016a. Cluster categorization of urban roads to optimize their noise monitoring. Environmental Monitoring and Assessment 188 (1), 1–11.

Zambon, G., Benocci, R., Brambilla, G., 2016b. Statistical road classification applied to stratified spatial sampling of road traffic noise in urban areas. International Journal of Environmental Research 10 (3), 411–420.

Giovanni Zambon has been working at Department of Earth and Environmental Sciences (DISAT) at University of Milano-Bicocca since 1993. In 1999 he graduated in physics. In 2003 he was named as teaching fellow. Research activities developed at DISAT focus on indoor acoustic and outdoor acoustic. He is technical adviser for University of Milan, for Milan Municipality, Environmental Lombardy Foundation (FLA), Management for the Airport of Linate and Malpensa (SEA). Since the academic year 2004/2005 at present he is the teacher of the course on "environmental acoustic" and of the "laboratory of environmental physics" for bachelor in science and technology for the environment. He is advisor of many bachelor and PhD thesis.

Dr. H. Eduardo Roman is currently working at the Plasma Research Group of Department of Physics and at Department of Earth and Environmental Sciences of University of Milano-Bicocca. He has been awarded a visiting professorship at Pohang University of Science and Technology (POSTECH) working at the Division of IT Convergence Engineering, Pohang, Korea. He has been guest scientist at the Max Planck Institute for the Physics of Complex Systems (Dresden) and as research fellow at Department of Physics of University of Milan. His interests cover different areas of complex systems such bio-molecules, nano-materials, fractals and finance. He is Fedeor Lynen Fellow of the Alexander von Humboldt Foundation. He is the author of seven books and over hundred publications in international journals.

Dr. Roberto Benocci is currently working at Department of Earth and Environmental Sciences of University of Milano-Bicocca, Italy. His interests cover the field of plasma physics with particular regards to laser plasma interaction for inertial confinement fusion and fast ignition research. He has been recently involved in researches in the field of environmental impact of electromagnetic radiation, acoustic noise statistical analysis and building acoustics, and been a member of the "Piero Caldirola" International Centre for the Promotion of Science. He is the author of over 40 papers in international scientific journals.