Research Article

Open Access

M. Smiraglia, R. Benocci*, G. Zambon, and H.E. Roman

Predicting Hourly Traffic Noise from Traffic Flow Rate Model: Underlying Concepts for the DYNAMAP Project

Abstract: The DYNAMAP project aims at obtaining a dynamic noise map of a large residential area such as the City of Milan (Italy), by recording traffic noise from a limited number of noise sensors. To this end, we perform a statistical analysis of road stretches and group them into different clusters showing a similar measured hourly traffic noise behavior. In the same way, we group simulated hourly traffic flow rates and compare their compositions with those of the traffic noise groups. The best agreement with the traffic noise was found by using the so-called normal traffic flow rate, yielding overlaps between 68 and 97%. Finally, we derive a simple analytical model to predict the hourly traffic noise from the simulated normal traffic flow, in very good agreement with the measured values.

1 Introduction

The Environmental Noise Directive (END) requires regular updating of noise maps to be implemented every five years to check and report to the regulator about the changes occurred during the reference period in the urban area [1]. This updating process is usually achieved using a standardized approach, consisting in collating and processing information through acoustic models to produce the up-dated maps [2, 3]. This procedure is time consuming and costly. To make the updating of noise maps easier and more cost effective, there is a need of integrated systems that incorporate real-time measurements and processing to assess the acoustic impact of noise sources. To this end, a dedicated project, named DYNAMAP, has been proposed and co-financed in the framework of the LIFE 2013 program, to develop a dynamic approach of noise mapping, being able to update environmental noise levels through a direct link with a limited number of permanent noise monitoring terminals.

The DYNAMAP project has the aim to group road network stretches in homogeneous clusters representing a possible method to reduce the size of the monitoring terminals network. Roads sharing the same characteristics for some parameters such as vehicles' flow rate capacity, number of lanes, etc., are grouped together. Such parameters are usually included in the functional classification of roads and linked to the role played in the urban mobility. However, this classification generally does not reflect the actual use of roads and, therefore, the actual noise emission. For a better description of the real behavior of noise in complex scenarios such as the road network of the city of Milan, we approach the problem considering an aggregation method based upon similarities among the 24-h continuous acoustic monitoring of the hourly equivalent noise levels, L_{Aeqh}. For this purpose, we put together historical noise monitoring campaigns carried out in Milan in the recent past, supplemented by additional measurements. The measurements we are dealing with here refer to 24-h continuous data from 58 monitoring stations, homogeneously distributed over the entire urban zone. Once normalized, such trend profiles provide a tool to group together roads according to their vehicular dynamics, and therefore allowing for a more real description of such road networks.

In this work, we study in detail the question of predicting traffic noise from calculated values of mean hourly traffic flow rate. Although this problem has been addressed several times in the literature [4–10], our discussion ex-

M. Smiraglia, G. Zambon: Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra. Università degli Studi di Milano–Bicocca, Piazza della Scienza 1, 20126 Milan, Italy
*Corresponding Author: R. Benocci: Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra. Università degli Studi di Milano–Bicocca, Piazza della Scienza 1, 20126 Milan, Italy; Email: roberto.benocci@unimib.it
H.E. Roman: Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra. Università degli Studi di Milano–Bicocca, Piazza della Scienza 1, 20126 Milan, Italy; Dipartimento di Fisica "G. Occhialini", Università degli Studi di Milano–Bicocca, Piazza della Scienza 3, 20126 Milan, Italy

tends those results to hourly noise and traffic flow rate patterns. We implement two different models for dealing with the traffic flow rate, and compare the results with measured noise patterns. Specifically, we study both the measured noise and the simulated traffic flow rate clusters, their correlations and road stretches compositions. Secondly, starting from the traffic flow rates we aim at predicting the hourly behavior of the noise stretches of the whole urban map. This prediction is based entirely on the traffic flow rate information, and its validity is checked on the available noise measurements. To be noted is that here the measurements of traffic noise were performed only temporarily (one day). Eventual anomalous noise sources (ambulances, road maintenance perturbations, general noises of anthropomorphic origin, etc.) have been removed from the raw data by analyzing the corresponding frequency spectra. Therefore, the results obtained here are relevant and representative for the present purposes. The influence of possible seasonal effects is the subject of a future study. The present results constitute the basis for determining the choice of a non-acoustic parameter to be used within the DYNAMAP project, which is intended to be based on the recording of the traffic noise data continuously over a limited set of monitoring stations, thus taking any seasonal variations into account implicitly.

The paper is organized as follows. First, in Sect. 2, we review works related to our paper. In Sect. 3, we deal with the clustering of both noise and flow rate data, followed by a discussion of the behavior of different types of traffic flow rates. In Sect. 4, we introduce two types of fitting models of hourly road noise as a function of the corresponding flow rate. We present two typical results for the two types of fitting functions used. We conclude the section by showing the predictions of noise using mean values of the model parameters for each cluster considered. Finally, our conclusions are reported in Sect. 5.

2 Related Works

Road traffic is the main source of noise in residential areas and its assessment and management is, therefore, strictly linked to such issues. For this reason, monitoring traffic and noise in urban areas has been the object of many studies, and their results have been used to build up noise maps to determine the population exposure to environmental noise. In the following, we summarize some of the results which are related to the present work, and provide us with a 'state-of-the-art' in the field.

In Ref. [4], a review of 20 past survey procedures shows that the surveys can be categorized into four types: random sampling, sampling by land-use category, receptor-oriented sampling and source-oriented sampling. Various weaknesses in the different types are examined and it is suggested that several survey types and various survey objectives are incompatible. Receptor-oriented surveys would appear to offer the best opportunity for gathering noise level data which can be generalized from site-specific information to the exposure of a population. Disaggregation of noise by source type during a measurement program could make the collection of noise data in urban noise surveys more efficient.

In Ref. [5], it is shown that environmental noise levels can vary over a wide range of values as a result of the diversity in the site conditions and different activities which necessarily occur during the time field measurements are carried out. This variability often yields a lack of consensus about how to estimate and present it when applying standards and regulations. In order to estimate the statistical variability, a large measurement database has been acquired under field conditions, consisting of noise recordings over two weeks, at 50 separate locations in residential areas which mainly affected by road traffic noise. The authors show that increased variability occur at the lower values of both logarithmic and arithmetic means of $L_{Aeq.}$ It is concluded that the observed relationships may be of help when estimating the noise level variability and the uncertainty associated with a noise measurement affected by road traffic or other environmental noise sources.

In Ref. [6], the authors explore the temporal and spatial variability of traffic noise in the City of Toronto. They collected real-time measurements of traffic noise at 554 locations across Toronto between June 2012 and January 2013. At each site, the measurements extended for a period of 30 min during daytime, and further measurements were made at 62 locations randomly selected from the first set of places, which exhibited high correlation (Pearson's correlation coefficient (r): 0.79). In addition, continuous measurements of noise were recorded for seven days at ten sites. It was observed that noise variability was predominantly spatial in nature, rather than temporal: spatial variability accounted for 60% of the total observed variations in traffic noise. Traffic volume, length of arterial road, and industrial area were the three most important variables explaining the majority of the spatial variability of noise (R^2 = 0.68 to 0.74). It is found that 80% of the sampled locations exceeded the guideline (*i.e.* 55 dBA, 16 h) of the Ministry of the Environment of the Province of Ontario. These findings suggested ubiquitous traffic noise ex-

posure across Toronto and that noise variability was explained mostly by spatial characteristics.

In Ref. [7], an overview of epidemiological studies is presented in the field of community noise and cardiovascular risk. Risk estimates are derived from individual studies and given for 5 dB(A) categories of the average A-weighted sound pressure level during the day. The noise sources considered in the studies are road and aircraft noise, while the health parameters considered are mean blood pressure, hypertension and ischaemic heart disease, including myocardial infarction, both from children and adults. Interestingly, the authors show evidence that correlations between transportation noise and cardiovascular risk have increased since the previous review published in Noise and Health in the year 2000.

In Ref. [8], authors discuss efficient decision-making in noise control actions, and classify a given location in a sensitive area according to the different prevailing traffic conditions. The paper outlines an expert system aimed to help urban planners to classify urban locations based on their traffic composition. Several machine learning type of algorithms are used based on multi-layer Perceptron and support vector machines with sequential minimal optimization. A combination of environmental variables are used as input variables. The procedure was tested on a full database collected from the city of Granada (Spain), including urban locations with road-traffic as dominant noise source. Among all the possibilities tested, support vector machines based models achieves the better results in classifying the considered urban locations into the four categories observed, with average values of weighted *F*-measure and Kappa statistics up to 0.9 and 0.8, respectively. Regarding the feature selection techniques, attribute evaluation algorithms (such as ReliefF and mRMR) achieve better classification results than subset evaluation algorithms in reducing the model complexity, and so relevant environmental variables are chosen for the proposed procedure. Results show that these tools can be used for addressing a prompt assessment of potential road-traffic-noise related problems, as well as for gathering information in order to take more well-founded actions against urban road-traffic noise.

In Ref. [9], the authors show that categorization is a powerful method for describing urban sound environments. The procedure is based on mobile measurements, followed by a statistical clustering analysis selecting relevant noise indicators for a better classification of sound environments. Analysis consists of a 3 days + 1 night survey where geo-referenced noise measurements were collected over 19 1-h sound-walk periods in a district of Marseille, France. The clustering analysis shows that a limited subset of indicators is sufficient to discriminate sound environments. Three indicators emerge from the clustering, that is, the L_{Aeq}, the standard deviation σL_{Aeq}, and the sound gravity spectrum SGC [50 Hz–10 kHz], are consistent with previous studies on sound environment classification. Interestingly, the procedure enables the description of the sound environment, which can be classified into homogenous sound environment classes by means of the selected indicators. The procedure can be adapted to any urban environment, and can, for example, favorably enhance perceptive studies by delimiting precisely the spatial extent of each typical sound environment.

In Ref. [10], the authors present an analysis of urban traffic within the H.U.S.H. (Harmonization of Urban Noise reduction Strategies for Homogeneous action plans), a project co-funded by European Community's Life+Program and it focuses on the harmonization of national and European legislations, regarding noise management tools. The paper is concerned with different vehicular traffic scenarios. Tools regarding design and traffic management are used for analysis and evaluation, as well as a computational model for traffic management validated with measured data. The results of simulations carried out on traffic flows, related to the different scenarios considered, have been used as input data for the acoustic model, leading to the definition of relationships existing between changing traffic flows and the reducing environmental noise.

3 Clustering

In our analysis we consider the hourly equivalent level L_{Aeqh} [dB] which has been measured in 58 sites in the city of Milan over one entire day and corresponding to 8 functional road classifications (from A to F and sub-groups), according to the official Italian classification of roads. For simplicity, we have not distinguished between sub-groups and have just considered the main four classes of interest (A, D, E, F). Data were recorded on weekdays and in absence of rain as prescribed by the current Italian legislation [11]. Due to the different monitoring conditions, such as different distances from the road and also to the condition of the street itself (*e.g.* its geometry, the presence of reflecting surfaces and obstacles along sound propagation and types of paving, etc.), we need to normalize the recorded data for each site appropriately. We define the normalized equivalent noise level, δ_{ij}, according to:

Table 1: Composition of clusters for different number of groups compared to the road functional classes. The composition is represented in the following way. From the total roads considered (58), there are 3 from Class A, 9 from D, 19 from E and 27 from F. Within each class, roads belong either to Cluster 1 or to Cluster 2. For instance, Class A (3) has roads only on Cluster 1, Class D (9) has 7 roads on Cluster 1 and 2 on Cluster 2. The total composition in our data-base made of 58 measurements has 62% or roads on Cluster 1 and 38% on Cluster 2.

Cluster	Road Class (functional classification)				Total
	A	D	E	F	
1	3 (100%)	7 (78%)	16 (84%)	10 (37%)	36 (62%)
2	0 (0%)	2 (22%)	3 (16%)	17 (63%)	22 (38%)

$$\delta_{ij} = L_{Aeqhj}/L_{AeqMaxj} \quad (i = 1, \dots\dots\dots, 24\,h; \quad (1)$$
$$j = 1, \dots\dots\dots, 58)$$

where the index i^{th} refers to the hour of the day and j^{th} to the corresponding site. Here, we have taken the normalization factor $L_{AeqMaxj}$ [dB] corresponding to the peak (maximum) noise value within the hourly series.

The traffic flow rate, F (number of vehicles per hour), for each street is taken from a standard and well tested model. The traffic data were provided by the AMAT agency, in charge of the traffic mobility management of the city of Milan [12]. The simulation model of the road network is defined as a "macro-scale traffic static allocation equilibrium model". The origin/destination matrix is based on the search for the path that minimizes the "generalized cost" between each zone pair. The parameters defining such "cost" rely on the travel time, the kilometric cost and eventually the toll charge. The time is converted into monetary value through a parameter denoted as "time value", which depends on the user category considered. The assigning model is of multi-class type, that is three different matrices are defined corresponding to cars, commercial vehicles and motorbikes, in order to take into account the different user behaviors. Such origin/destination matrices for each category have been obtained through dedicated mobility survey on the entire city of Milan, and its outskirts.

If one disregards the type of vehicle, we obtain what is called the normal flow rate, denoted as F_n. By considering the fact that large vehicles, like trucks for instance, have a larger contribution to traffic noise, one can also define the so called equivalent flow rate, F_{eq}, which weights differently the vehicle type (see below). Similarly, as for the noise, we normalize the traffic flow rate according to its peak value, F_{nMax}, as follows,

$$\lambda_{ij} = F_{nij}/F_{nMaxj} \quad (i = 1, \dots\dots\dots, 24\,h; \quad (2)$$
$$j = 1, \dots\dots\dots, 58)$$

and similarly, F_{eqMax} for the equivalent flow rate.

Unsupervised clustering algorithms are employed to separate normalized noise levels into groups which be-

have similarly. Several algorithms (hierarchical agglomeration [13], K-means algorithm [14], and partitioning around medoids (PAM) [15]) are considered. In general, we choose the number of clusters in such a way as to obtain a reasonable compromise between satisfactory discrimination between the elements, and the need to keep the number of groups to a minimum. We employ the Euclidean distance as the underlying structural metric.

We use the statistical software R [16] for computing and analysing the data. The validation of the results is performed using the package "clValid" [17]. The clustering results are ranked using an index based on both the performance and the validation measures for each algorithm [18]. In this way, the optimal list is obtained yielding a two-cluster hierarchical agglomeration at the first place, followed also by a two-cluster groups by the K-means and PAM methods.

The obtained noise clusters are composed of roads belonging to different classes, as reported in Table 1. For the four-cluster solution (not shown here), which is commensurate with the number of road functional classification, a poor match is found. The F class is found over all the four clusters, whereas the remaining classes are distributed in the first two groups. This confirms that the road traffic is primarily linked to the effective urban mobility rather than its functional classification.

Regarding the clustering process results, the two-cluster solution represents, therefore, a satisfying balance between an adequate differentiation among time patterns and the need to get a simple practical solution. The two clusters appeared to be formed primarily by the contributions from different temporal profiles belonging to roads of class A, D and E for cluster 1 (made up of 36 temporal profiles corresponding to a 62%) and of roads of class F for cluster 2 (made up of 22 temporal profiles corresponding to a 38%). This result confirms that the noise time patterns are not directly linked to the standard road classification. The same procedure is applied to cluster the traffic flow rate.
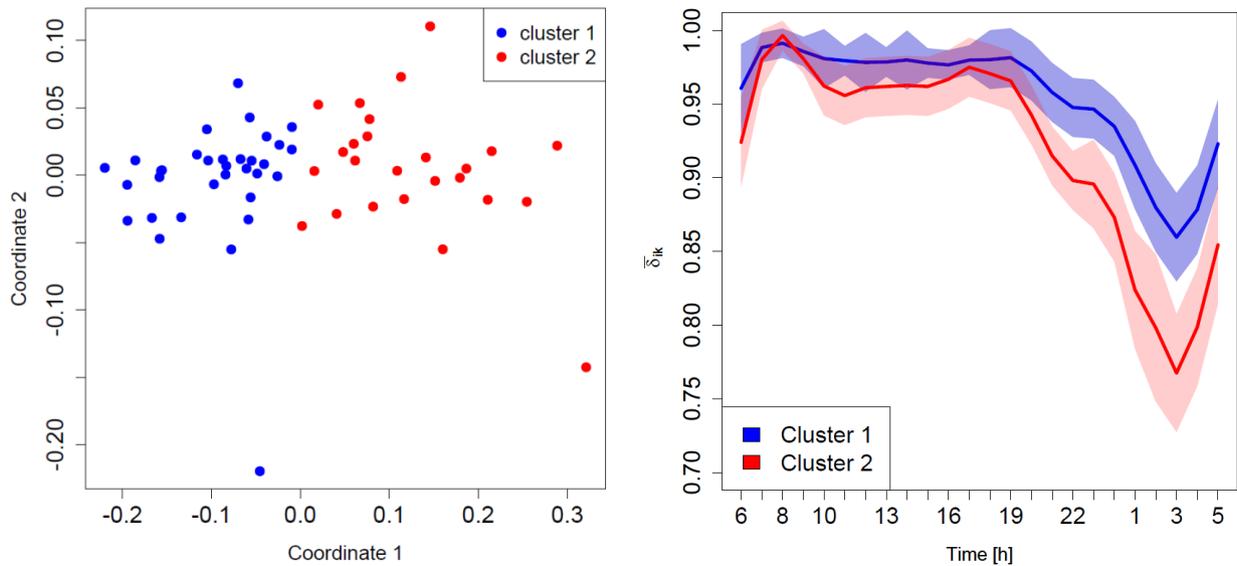
Figure 1 shows the results of hierarchical clustering

Figure 1: Statistical analysis of measured hourly traffic noise. (Left panel) Clustering results from MDS, where the two clusters are marked in different colors. Cluster 1 has 33 elements (hereafter denoted also as Large cluster, L) and Cluster 2 has 25 elements (Small cluster, S). (Right panel) Average values of traffic noise, $\overline{\delta}_{ik}$ ($k = 1, 2$), as a function of the intraday hours, i, for each cluster. The standard deviation band (1σ) is included for convenience.
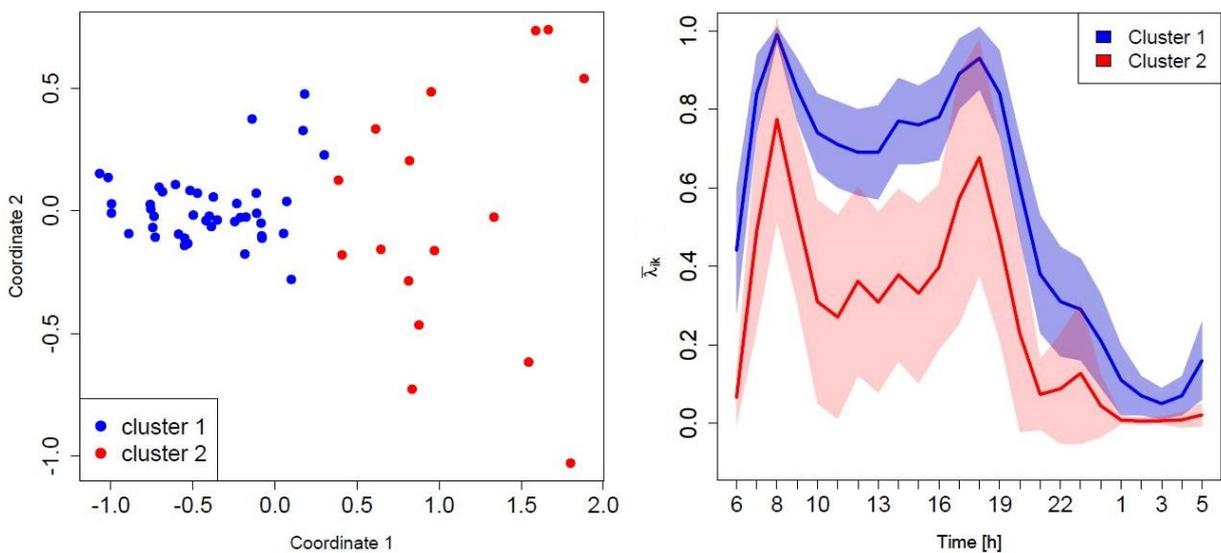

Figure 2: Statistical analysis of simulated hourly traffic flow rate. (Left panel) Clustering results from MDS, where the two clusters are marked in different colors. Large cluster with 42 elements (blue color), and Small cluster with 16 elements (red color). (Right panel) Average values of traffic flow rate, $\overline{\lambda}_{ik}$, as a function of the intraday hours, i, for each cluster, $k = 1, 2$. The standard deviation band (1σ) is included for convenience.

of noise data, for the 58 site measurements. In Fig. 1 (left panel), we show the scatter plot obtained from the Multi-Dimensional Scaling (MDS) results, providing a visual representation of the pattern of proximities among the data. MDS takes a set of dissimilarities and returns a set of points such that the distances between the points are approximately equal to the dissimilarities, in other words it displays the structure of (complex) distance-like data (a dissimilarity matrix) from a high dimensional space into a lower dimensional space without too much loss of information. The goal of MDS is to faithfully represent these distances within the lowest possible dimensional space [19]. The two coordinates shown in Figs 1 and 2, represent the best lower-dimensional space obtained from the MDS al-

Table 2: Comparison between the percentage number of elements in common for each cluster obtained from both the noise and traffic flow rates, by taking the noise clusters as reference. The best agreement is obtained for the normal flow rate, in which the large cluster has N_L = 42 elements, while the small one has N_S = 16 elements. This to be compared with the numbers N_L = 33 and N_S = 25 for the noise clusters, respectively.

	Noise S	Noise L		Noise S	Noise L
Normal Flow rate L	32	97	Equivalent Flow rate L	41	89
Normal Flow rate S	68	3	Equivalent Flow rate S	59	14
Log Norm. Flow rate L	55	97	Log Equiv. Flow rate L	55	97
Log Norm. Flow rate S	45	3	Log Equiv. Flow rate S	45	3

gorithm, where each point corresponds to a single road stretch including the whole hourly values of noise and traffic flow rate. In Fig. 1 (right panel), the corresponding mean values $\bar{\delta}_{ik}$ are reported for each cluster (Eq. 1), together with the hourly standard deviations. Figure 2 shows the corresponding results for the clustering and mean cluster values, $\bar{\lambda}_{ik}$ with (k = 1, 2), of the hourly traffic flow rate (Eq. 2).

By considering the results shown in Figs. 1 and 2, one can see that noise variations are limited to a rather narrow band, between 1 and 0.75, while traffic flow rate changes from 1 to near 0. The fact that noise does not decrease during night hours as much as the flow rate does, is due essentially to two facts: Locally, noise from neighboring streets have a non-vanishing contribution; and, secondly, vanishing traffic flow rates are actually a result of the traffic model employed, which has been designed to describe more accurately high traffic situations. This condition might not be fulfilled during night hours in secondary streets. Despite these particular differences, the global hourly behavior of both quantities is still quite strongly correlated. The different widths of the two bands, for noise and flow rate, can be taken into account by the linear fits we discuss later in Sect. 4.

The strong correlation between noise and flow rate is quite well documented in the literature (see *e.g.* [20]), and we will present further results supporting this view within the present context. We consider next the issue of comparing the cluster composition of both noise and flow rate, as obtained by using different definitions of the latter such as: The normal flow rate (as defined above), and the equivalent flow rate, obtained by weighting heavy vehicles, like trucks, differently from lighter ones such as cars. In our definition, we assume that a truck contributes 8 times the flow rate from a single car, reflecting the fact that it produces a higher noise. The number 8 used here has been determined empirically from direct measurements of both trucks and cars noises. Finally, we extend this study to the

corresponding logarithmic counterparts. This is due to the known fact that the logarithm of flow rate is well correlated with the associated traffic noise [20].

The results of clustering are reported in Table 2. As one can see, the normal flow rate clusters present a higher overlap with the corresponding noise clusters, yielding a 97% of superposition for the large clusters (33 and 42, for the noise and normal flow rate clusters, respectively), and 68% for the small ones (25 and 16, respectively).

As one can see from Fig. 3, the road stretch F30 belongs to Cluster 2 (S) for the noise levels, while it falls inside Cluster 1 (L) for the flow rate. This discrepancy, which occurs in a rather small number of cases, can be due to the single one-day measurement of the noise which is compared with a model calculation for the corresponding annual mean values. Despite this mismatch, the absolute errors associated with the present predictions are the following: ε_F = 0.04 (2.62 dB) to be compared with the mean cluster value ε_F = 0.022 (1.44 dB), while for the Log F we find, ε_{LogF} = 0.03 (1.9 dB) to be compared with the mean value ε_{LogF} = 0.02 (1.3 dB).

4 Model and Results

In this section, in view of the previous results suggesting a strong correlation between flow rate and noise, we wish to extend this concept towards a direct comparison between the two quantities. This has been done in the past (see *e.g.* [20]), but here we examine this relationship quantitatively in detail. We use two approaches, one is just a linear relation between flow rate and noise,

$$L_{Aeqh_{ij}}/L_{AeqMaxj} = \alpha \cdot F_{nij}/F_{nMaxj} + \beta, \qquad (3)$$

and a similar one for the logarithms of the flow rate,

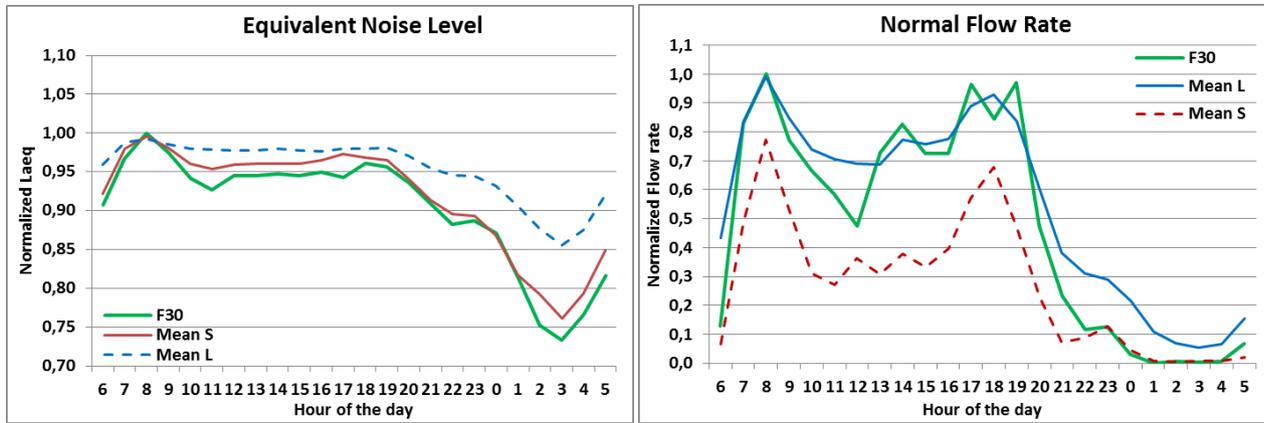$$L_{Aeqh_{ij}}/L_{AeqMaxj} = A \cdot \mathrm{Log}F_{nij}/\mathrm{Log}F_{nMaxj} + B \qquad (4)$$

Figure 3: Example of stretch (F30) which belongs to the 'small' cluster for noise while included in the 'large' cluster from the flow rate analysis. (Left panel) Equivalent noise level results for road F30. As shown, F30 belongs to the small cluster (Cluster 2). (Right panel) Traffic flow rate for F30 showing that it belongs to the large cluster (Cluster 1). In both cases, the continuous lines for the mean values correspond to the cluster at which F30 belongs to, while the dashed lines display the result for the opposite cluster.
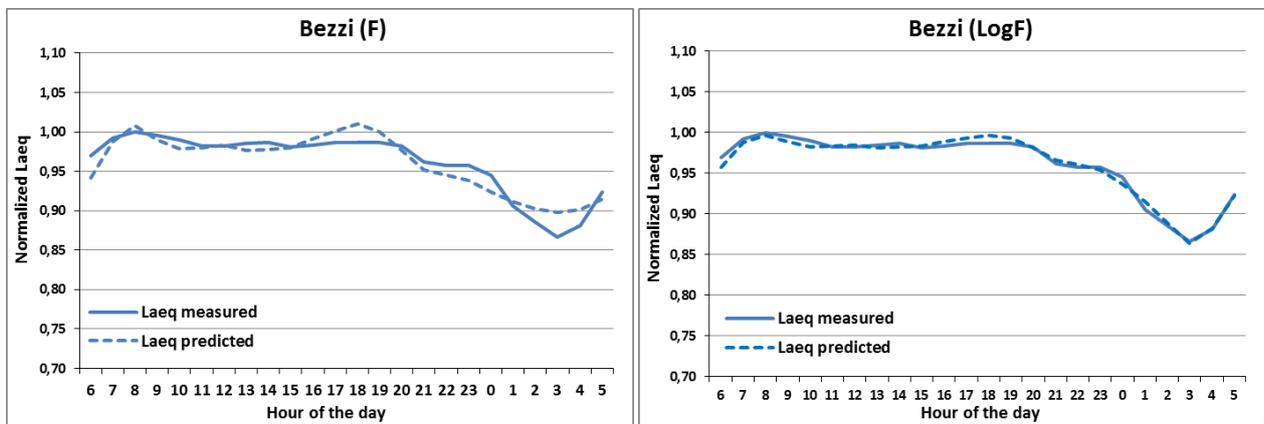

Figure 4: Fitting of the normalized equivalent level, LAeq, vs intraday hour, for the two models, Eq. (3) for F and Eq. (4) for Log$F$, for the street Bezzi. Least square fit parameters for: (F) α = 0.118, β = 0.892, and mean square fit error ε = 0.015 (1.079 dB); (Log$F$) A = 0.381, B = 0.616, and ε_{Log} = 0.006 (0.420 dB).

The unknown constants (α, β; A, B) are determined by a least square fitting. Despite the fact that the linear relation Eq. (3) might perform less accurately than the most accepted logarithmic dependence, we still keep it here to have a counterpart for a better assessment of the range of validity of the analytical results. To be noted is that Eq. (4) reduces to the linear behavior in Eq. (3) in the case of small variations of F_n. Two typical examples of the application of Eqs. (3) and (4) are shown in Fig. 4, for Bezzi street, belonging to the large cluster (L, cluster 2), and Fig. 5 for Garegnano street, belonging to the small cluster (S, cluster 1).

In keeping with a strategy aimed at predicting the traffic noise of an arbitrary street for which the parameters (α, β; A, B) are not known, we reconsider the examples discussed above, Bezzi and Garegnano streets, by using the corresponding mean cluster values ($\alpha_{L,S}$, $\beta_{L,S}$; $A_{L,S}$, $B_{L,S}$) for predicting the normalized hourly traffic noise.

The mean values are obtained for each cluster separately as, $\alpha_{L,S} = \left(\sum_{k=1}^{N_{L,S}} \alpha_k \right) / N_{L,S}$, and similar relations for the remaining parameters. Here, $N_{L,S}$ are the number of elements in each cluster (L,S). The results are shown in Figs. 6 and 7, respectively.

In Table 3 we report the relative and absolute errors (in dB) resulting from the fits using Eqs. (3) and (4) for the normal traffic flow rate and its logarithm. The cluster used are those obtained from the normal traffic low rates and logarithms. To obtain the absolute values of the error we evaluate the mean values within each normal flow rate cluster, yielding the values (65.5, 69.4) dB for F_n and (63.9, 69.4) dB for Log F_n, for Cluster 1 (Large) and 2 (Small), respectively.

As is apparent from the above results, the logarithmic fit is superior, suggesting an accurate way for predicting the hourly noise in a given street from the mean cluster
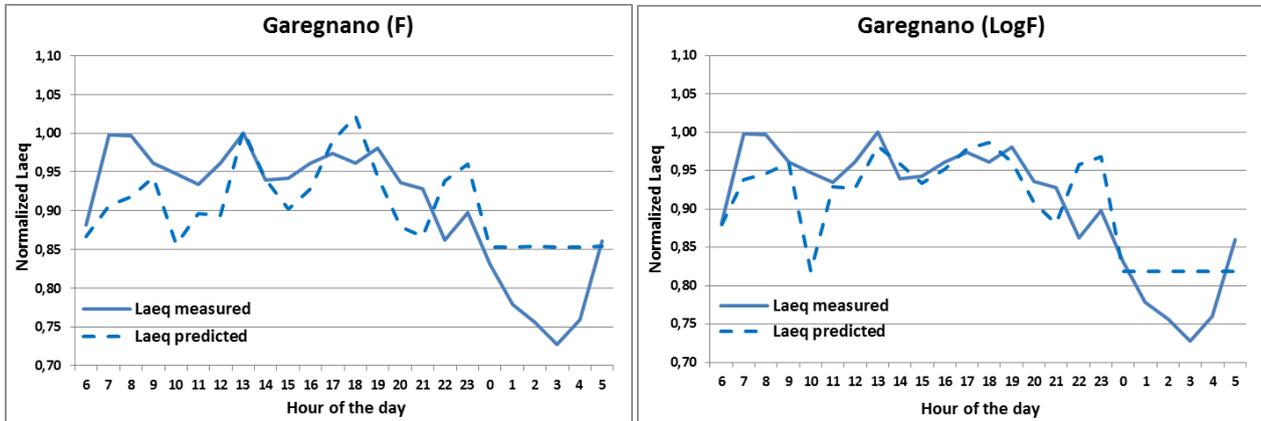
Figure 5: Fitting of the normalized equivalent level, LAeq, vs intraday hour, for the two models, Eq. (3) and Eq. (4), for a second example, street Garegnano. Least square fit parameters for: (F) $\alpha = 0.168$, $\beta = 0.853$, and mean square fit error $\varepsilon = 0.063$ (3.919 dB); (Log$F$) $A = 0.168$, $B = 0.818$, and $\varepsilon_{Log} = 0.051$ (3.175 dB).
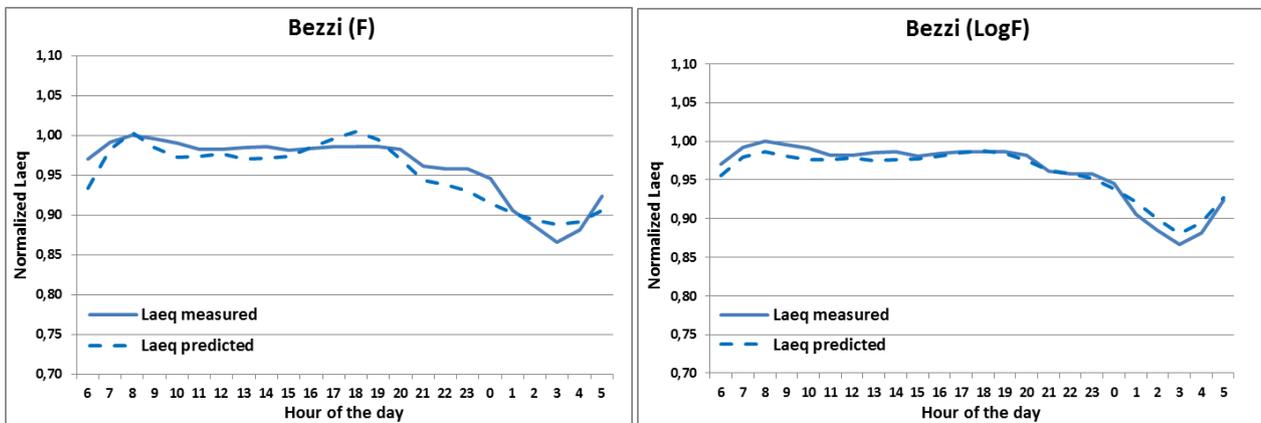

Figure 6: Fitting of the normalized equivalent level, L_{Aeq}, vs intraday hour, for the two models, Eq. (3) and Eq. (4), using the mean cluster values for the street Bezzi ($\alpha_L = 0.123$, $\beta_L = 0.881$) and ($A_L = 0.309$, $B_L = 0.678$), yielding mean square fits errors: $\varepsilon = 0.016$ (1.209 dB) (F) and $\varepsilon = 0.010$ (0.701 dB) (Log$F$).
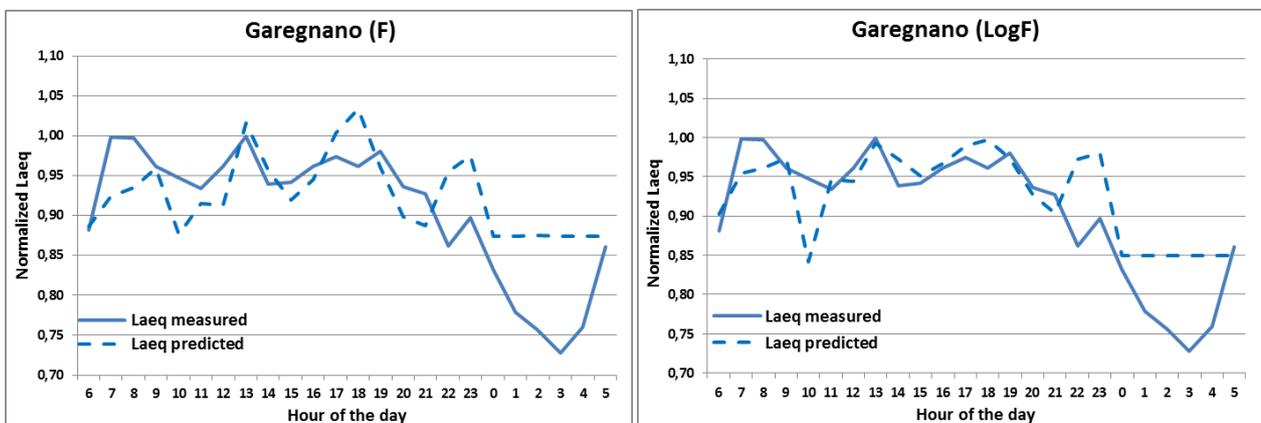

Figure 7: Fitting of the normalized equivalent level, L_{Aeq}, vs intraday hour, for the two models, Eq. (3) and Eq. (4), using the mean cluster values for the street Garegnano ($\alpha_S = 0.159$, $\beta_S = 0.874$) and ($A_S = 0.148$, $B_S = 0.849$), yielding mean square fits errors: $\varepsilon = 0.065$ (4.081 dB) (F) and $\varepsilon = 0.056$ (3.472 dB) (Log$F$).

Table 3: Mean errors from predicted noises with respect to measured ones, for the normal traffic flow rates and for their logarithms calculated within their corresponding clusters (cf. Table 1). The street parameters refer to the single fitting parameters for each street, averaged over the whole cluster (S, L). On the right side of the table we report the errors when the mean cluster values are used for obtaining the predicted noise. In parenthesis we display the absolute errors in dB.

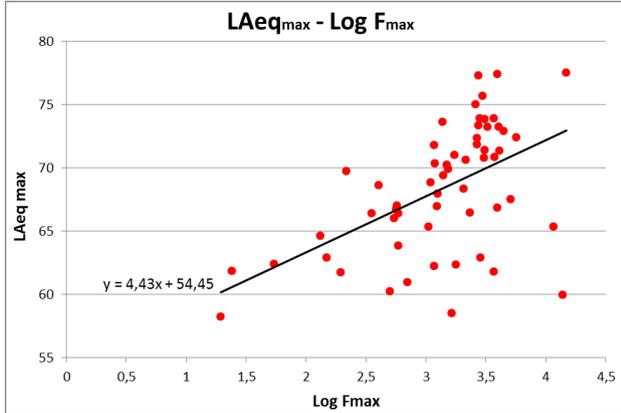| Cluster	Street parameters		Mean cluster parameters	
	$\overline{\varepsilon}_{\mathbf{F}}$	$\overline{\varepsilon}_{\mathbf{logF}}$	$\overline{\varepsilon}_{\mathbf{F}}$	$\overline{\varepsilon}_{\mathbf{logF}}$
L (1)	0.022 (1.531 dB)	0.017 (1.214 dB)	0.030 (2.063 dB)	0,026 (1.827 dB)
S (2)	0.051 (3.353 dB)	0.043 (2.761 dB)	0.056 (3.664 dB)	0,049 (3.162 dB)

Figure 8: Plot of the maximum equivalent level, L_{Aeqmax}, vs the logarithm of the maximum flow rate, Log F_{nmax} for all the stretches considered here. The straight line a least-square-fit with equation: $y = A_{max}x + B_{max}$, with A_{max} = 4.43 and B_{max} = 54.45, yielding a mean error (standard deviation of the data from the fit) of $\varepsilon \approx 6\%$.

values (A, B) to which it belongs. In order to determine the cluster membership, one can compare the hourly flow rate record of the street under consideration with the mean clusters results, and taking the one for which the sum of the squares of hourly differences is smaller. To be noted is that the exclusion of the single street from the database for the purpose of predicting its noise behavior from the flow rate, using either its own fit parameters or the mean cluster ones, does not change the result in a significant way. We have tested this for all road stretches considered and the differences fall much below the errors reported in Table 3.

In the applications, Eq. (4) can be used to predict the relative equivalent level of a road stretch *j*. In order to get the absolute value of noise, one needs to multiply it by $L_{Aeqmaxj}$. The latter can be estimated from the corresponding maximum value of the flow rate, Log F_{nmaxj}, according to

$$L_{Aeqmax} = A_{max} \cdot \text{Log} F_{nMax} + B_{max} \qquad (5)$$

where A_{max} and B_{max} are least-square fitting parameters. Results of such a fit for our 58 noise stations are reported in Fig. 8. The mean prediction error of the fit is given by the square-root of the mean square differences between the straight line and each measured value L_{Aeqmax}, yielding an error of about 6%. Some conspicuous deviations are observed for large values of F_{nmax} which can be attributed to the traffic model used to evaluate the normal flow rate. In what follows we briefly discuss how the present results could be of relevance to the DynaMap project.

5 Application to DynaMap

The original idea of DynaMap is that the cluster discriminant relies on a single hour, the rush-hour (typically 8:00–9:00 am), using the corresponding peak vehicular flow for separating roads into two sets. Roads having a rush hour peak flow above the threshold belong say, to Cluster 1, while those roads having smaller traffic flow rates, fall into Cluster 2. The clustering approach discussed here can be seen as an improvement on this hard threshold method, since we do not use just one single value of the traffic flow rate (rush hour), but the whole 24 hours profile. Regarding the real-time implementation of the acoustic map with an update time interval of (5, 10, 20, etc.) minutes, each stretch will just follow the corresponding Cluster (1 or 2) of noise sensors, by considering in addition the normalization noise value obtained from Eq. (5) if absolute noise predictions are required.

DynaMap requires the choice of an optimal non-acoustic parameter for each road stretch within the urban area of interest. Our present study suggests that a useful quantity can be obtained by considering the whole hourly traffic flow rate. In particular, the logarithm of the normal flow rate, which yields the smallest prediction error (see Table 3), can be suggested as the appropriate quantity for determining the actual non-acoustic discriminant parameter.

6 Conclusions

The present calculations suggest that the concept of clustering, for both the traffic noise and the normal traffic flow rate, is a very promising approach to deal with the difficult problem of predicting the traffic noise in a complex urban network. We can safely say that despite the complexity of such large human built-up area a rather simple scheme seems to emerge suggesting an accurate way to predict the time behavior of a single component in the network. The present results provide a basis for building up a quantitative relation between noise and flow rate temporal dependences, constituting the core structure within the DYNAMAP project. The latter is intended to be based on the recording of the traffic noise data continuously but over a limited set of monitoring stations. The traffic noise for the remaining (non-monitored) road stretches can be predicted using information of the corresponding traffic flow rates (non-acoustic parameter), which are available for the interested urban zone. This implies that seasonal variations are implicitly taken into account. The resulting dynamic maps are expected to be updated on a small time window. More generally, DYNAMAP, being a real time map of traffic noise, could be extended to inform about the volumetric concentration of pollutant agents in the atmosphere, thus providing a visual and constantly updated map of air quality, displaying also the value and evolution of meteorological parameters, such as air humidity or wind speed. This same idea could be expanded to create dynamic maps of human-caused environmental parameters, such as traffic density.

References

[1] Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental noise. Official Journal of the European Communities, L 189/12, 2002.

[2] W. Wei, T. V. Van Renterghem, B. De Coensel, D. Botteldooren, Dynamic noise mapping: A map-based interpolation between noise measurements with high temporal resolution, Applied Acoustics 101, 127–140 (2016). DOI: 10.1016/j.apacoust.2015.08.005

[3] B. De Coensel, K. Sun, W. Wei, T. Van Renterghem, Dynamic Noise Mapping based on Fixed and Mobile Sound Measurements, EuroNoise, May 31 – June 3, Maastricht (2015).

[4] A. L. Brown, K. C. Lam, Urban noise surveys. Appl. Acoust. 20, 23–39 (1987).

[5] J. Alberola, I. H. Flindell, A. J. Bullmore, Variability in road traffic noise levels. Appl. Acoust. 66, 1180–1195 (2005).

[6] F. Zuo, Y. Li, S. Johnson, J. Johnson, S. Varughese, R. Copes, F. Liu, H. Jiang Wu, R. Hou, H. Chen, Temporal and spatial variability of traffic-related noise in the City of Toronto, Canada. Science of the Total Environment 472, 1100–1107 (2014).

[7] W. Babisch, Transportation noise and cardiovascular risk: updated review and synthesis of epidemiological studies indicate that the evidence has increased. Noise Health 8, 1–29 (2006).

[8] A. J. Torija, D. P. Ruiz, Automated classification of urban locations for environmental noise impact assessment on the basis of road-traffic content, Expert Systems with Applications, 53, 1–13 (2016).

[9] A. Can, B. Gauvreau, Describing and classifying urban sound environments with a relevant set of physical indicators, Journal of the Acoustical Society of America 137, 208–218 (2015).

[10] G. Baldinelli, R. Bellomini, F. Borchi, M. Carfagni, S. Curcuruto, S. Luzzi, R. Silvaggio, M. Stortini, Correlation between traffic flows and noise reduction in HUSH project strategic actions, in Proceedings of Forum Acusticum 2011, Aalborg, Denmark (2011).

[11] Decree Ministry of Environment 16 March 1998, Tecniche di rilevamento e di misurazione dell'inquinamento acustico (Measurement techniques of noise pollution), Gazzetta Ufficiale serie generale (Official Journal of Italian Republic) n. 76, 1/4/1998.

[12] http://www.amat-mi.it/it/mobilita/dati-strumenti-tecnologie/

[13] J. H. Ward, Hierarchical Grouping to Optimize an Objective. Journal of the American Statistical Association, 58, 236–244 (1963).

[14] J. A. Hartigan, M. A. Wong, A K-means clustering, Applied Statistics, 28, 100–108 (1979).

[15] L. Kaufman, P. J. Rousseeuw, Finding Groups in Data. Wiley Series in Probability and Mathematical Statistics, Hoboken (NJ) (1990).

[16] http://www.r-project.org/

[17] G. Brock, P. Pihur, S. Datta, S. Datta, clValid: An R Package for Cluster. Journal of Statistical Software, 25, 1–22 (2008).

[18] V. Pihur, S. Datta, S. Datta, Weighted Rank Aggregation of Cluster Validation Measures: a Monte Carlo Cross-entropy Approach. Bioinformatics 23, 1607–1615 (2007).

[19] A. F. Zuur, E. N. Leno, G. M. Smith, Statistics for Biology and Health – Analysing Ecological Data, Springer, New York (2007). ISBN 978-0-387-45967-7 (Print), 978-0-387-45972-1 (Online).

[20] G. B. Cannelli, K. Glueck, S. Santoboni, A Mathematical Model for Evaluation and Prediction of Mean Energy level of Traffic Noise in Italian Cities, Acustica 53, 31 (1983).