STREET CLASSIFICATION BY MEANS OF CLUSTER CATEGORIZATION FOR THE NIGHT PERIOD

PACS: 43.50

Quintero Perez, Guillermo[1]; Romeu Garbi, Jordi[2]; Balastegui Manso, Andreu[3].
LEAM - Polytechnic University of Catalonia,
Colom 11
08222 Terrassa, Spain
Email: [1]guillermo.quintero.p@gmail.com; [2] jordi.romeu@upc.edu; [3] andreu.balastegui@upc.edu;

Keywords: Cluster classification; night period; land use

ABSTRACT

For spatial stratification, the street classification by typology has been widely used. Its main principle is that streets can be grouped according to their use for accessing or connecting an urban area, which improves the spatial sampling when the noise is due to vehicular transit. Therefore, when there is influence non traffic noise, especially during the night period, the use of this temporal stratification is not recommended. In order to obtain a classification of streets for the night period that takes into account the influence of noise sources related to land use, such as leisure noise, a cluster classification of streets is performed by using two different input parameters and a comparison between both is performed.

RESUMEN

Para la estratificación espacial, la clasificación de calles por tipología ha sido ampliamente utilizada. Su principio fundamental es que las calles se pueden agrupar según su uso para acceder o conectar un área urbana, lo que mejora el muestreo espacial cuando el ruido se debe al tránsito vehicular. Por lo tanto, cuando hay influencia de ruido diferente al tránsito vehicular, especialmente durante el período nocturno, no se recomienda el uso de esta estratificación temporal. Con el fin de obtener una clasificación de calles para el período nocturno que tenga en cuenta la influencia de fuentes de ruido relacionadas con el uso del suelo, tal como el ruido de ocio, se realiza una clasificación por clúster de calles utilizando dos parámetros de entrada diferentes y se realiza una comparación entre ellos.

1. Introduction

Exposure to high or constant noise pollution can propitiate negative health effects whether they are auditory, e.g. hearing loss or impairment, or non-auditory e.g. cardiovascular, stress, sleeping. It has been shown that, specifically for the night time exposure, the noise can propitiate hypertension and a reduction in the sleep quality[1,2]. According to the European Noise Directive 2002/49/EC, for strategic noise mapping, which is the main tool for assessment of peoples noise exposure, at least L_{night} and L_{DEN} should be presented for the minimum time of one year.

There are many studies that aim to improve the long-term estimation of traffic noise by means of sampling strategies. For the case of spatial strategies, it has been shown that a spatial classification based on street functionality improves the estimation of long term values[3–5]. The categorization by street functionality proposes an aggrupation of streets based on their traffic characteristics[3,6], which aims to reduce the number of sampling points for long term traffic noise assessment.

For the specific case of the night period, there could be other noise sources rather than the traffic noise itself, which could be related to land use such as leisure or recreational noise[6–11]. Therefore, the categorization of streets based on traffic characteristics is not recommended and a different scheme for its assessment should be used.

The present research proposes a spatial categorization of streets for the assessment of noise during the night period, which would take into account the influence of noise sources related to land use. The categorization is automatically performed by means of a clustering procedure. Two different input parameters are then evaluated, the night equivalent level (L_{night}) and the night hourly equivalent level ($L_{h,night}$).

2. Material and methods

2.1. Data under study

Between 2010 and 2015, noise measurements at an interval of 1 second using Class 1 soundmeters were taken in the city of Barcelona, in 14 different sampling points. The data was analyzed and the abnormal values which affect the annual level were removed. One year of measurement of each sampling point was used for the present study (the one with less days removed). In order to compare the street classifications, the sampling points were categorized according to the street functionality approach:

- Category 1: Urban ring roads or access roads.
- Category 2: Main streets.
- Category 3: Ordinary streets.
- Category 4: Pedestrian streets.

2.2. Methodology

Once having the each second measurements, the two input parameters for the clustering procedure should be computed. Then, $L_{hr,night}$ was computed according to

$$L_{h,night} = 10\log\left\{\frac{1}{60}\sum_{i=1}^{60} 10^{\frac{L_{eq,1s}(i)}{10}}\right\} - L_{week} \quad (1)$$

where h corresponds to each hour comprised from 11pm to 7am, and L_{week} is the week equivalent level computed as

$$L_{week} = 10\log\left\{\frac{1}{7}\sum_{i=1}^{7} 10^{\frac{L_{night}(i)}{10}}\right\} \quad (2)$$

L_{night} in Equation 2 is the night equivalent level, which is also the second input data, and it is computed according to

$$L_{night} = 10\log\left\{\frac{1}{8}\sum_{i=1}^{8} 10^{\frac{L_{h}(i)}{10}}\right\} - < L_{night} > \quad (3)$$

where L_h is the hour equivalent level between 11pm to 7am and $< L_{night} >$ is the long term equivalent level (i.e. using all the days of the year). The reason to subtract in Equation 1 the week equivalent level and in Equation 2 the year equivalent level is to perform the classification by the noise patterns rather that the noise level itself as stated in [12], where a clustering method is used to group streets based on the noise pattern instead of the traffic characteristics. After the calculations the data was stored in a local database to be processed afterwards.

The k-means clustering is the selected algorithm to perform the automatic classification. The optimal number of cluster is computed prior the cluster classification based on the silhouette algorithm.

The spatial and temporal characteristics (noise patterns) of both classification are then compared, as well as the noise level variability of the sampling points of each cluster.

3. Results

3.1. Sampling points classification

Figure 1 shows the silhouette value for a k-number of cluster from 4 to 10. The minimum value was set to 4 as the street functionality categorization resulted in 4 different categories and it was limited up to ten clusters as according to the number of sampling points, more clusters would not bring any important information. As can be seen in Figure 1 (a), by using the hourly value, the optimal number of clusters is 9, although the silhouette mean does not have high variation for the different values of k. For the case of the daily value (Figure 1 (a)), the optimal number of k is 9 as well, but it is seen that 4, 7, 8 and 10 clusters have almost the same silhouette mean.

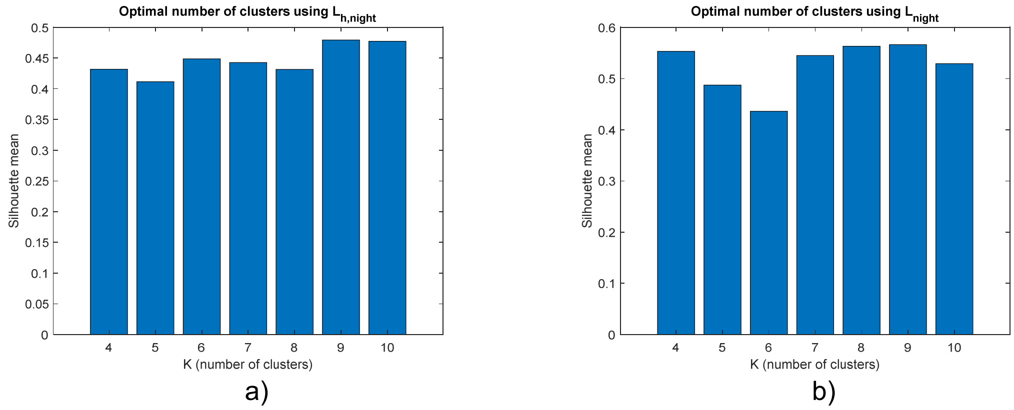

Figure 1. Silhouette mean for k=4-10 number of clusters using both input parameters.

In order to compare with the street functionality classification, the selected number of cluster was set to 4 and also to the optimal, k=9. Table 1 shows the cluster classification by using the two different inputs and k=4, 9. As can be observed, there is not a clear match between all categorizations. For the case of using k=9, it is observed that there are categories that are made up of only one street. As the spatial categorization is performed with the aim of grouping streets and reduce the number of sampling points, this value of k does not agree with the categorization approach. Taking as reference the street functionality classification (S.F.), for the case of k=4, it is observed that for L_{night} there are some matches for pedestrian and main streets. For the case of $L_{h,night}$, it is observed that pedestrian streets of the street classification match to categories 4 and 2 of the cluster categories, and it is important to remark that the corresponding sampling points of category 2, points 13 and 14, correspond to public squares. So it can be said that the using $L_{h,night}$ separates pedestrian streets into two categories, streets for pedestrian flow and public squares.

Table 1. Sampling point (S.P.) and its corresponding classification by employing the street functionality approach (S.F.) and by using the clustering procedure using the two different inputs setting k=4 and k=9.

S.P.	S.F.	k=4		k=9	
		L_{night}	$L_{h,night}$	L_{night}	$L_{h,night}$
1	4	4	4	5	1
2	4	1	4	6	1
3	2	3	1	2	4
4	1	3	1	3	2
5	2	1	1	4	2
6	2	1	1	1	4
7	4	1	4	6	1
8	2	3	1	2	2
9	2	3	1	2	2
10	2	3	3	3	3
11	2	3	1	3	7
12	4	4	4	5	9
13	4	4	2	9	6
14	4	4	2	8	5
15	2	1	1	1	4
16	3	2	1	1	4
17	3	3	1	3	4

Focusing on *k=4* clusters, Figure 2 shows the temporal evolution of each cluster for L_{night} and Figure 3 shows the evolution for $L_{h,night}$. It is observed that classifying street by $L_{h,night}$, highlights the similarity of the noise evolution of each day within the week, as it is observed that the hourly noise level of each day is similar in the whole week for all days. Then, the classification is performed for sampling points that has similar weekly noise evolution and similar daily noise evolution. For the case of L_{night} classification, it is observed that the grouping is performed only by the weekly noise evolution; Cluster 4 includes streets that has high noise levels during Friday and Saturday nights with a difference of about 6dB between the less noisy and the noisiest days. Cluster 2 has high noise levels from Thursday to Saturday, being Friday the day with highest levels. Cluster 3 is formed by streets where the noise level variation during the week is very low. Finally, Cluster 1 is comprised by streets that, similar to Cluster 1, have an increase in the noise levels during Friday and Saturday but in this case, the difference between lowest and highest noise level is around 2dB.

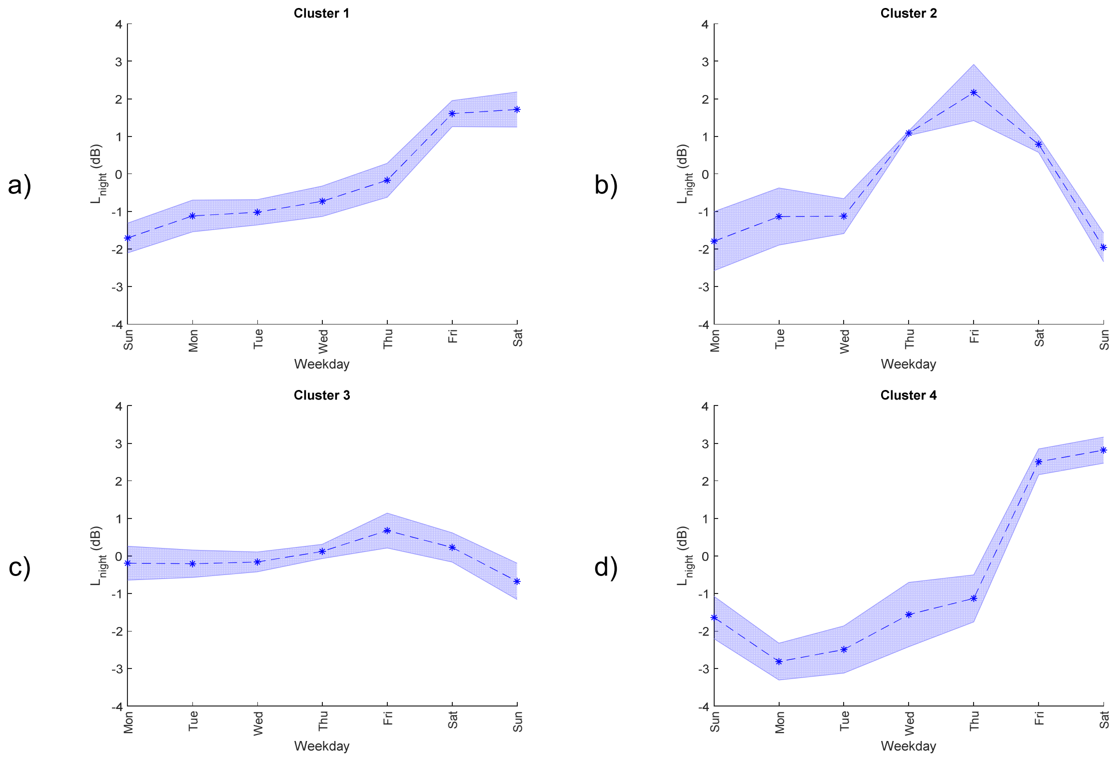

Figure 2. Temporal evolution of L_{night} for the 4 clusters and its corresponding standard deviation (shade).

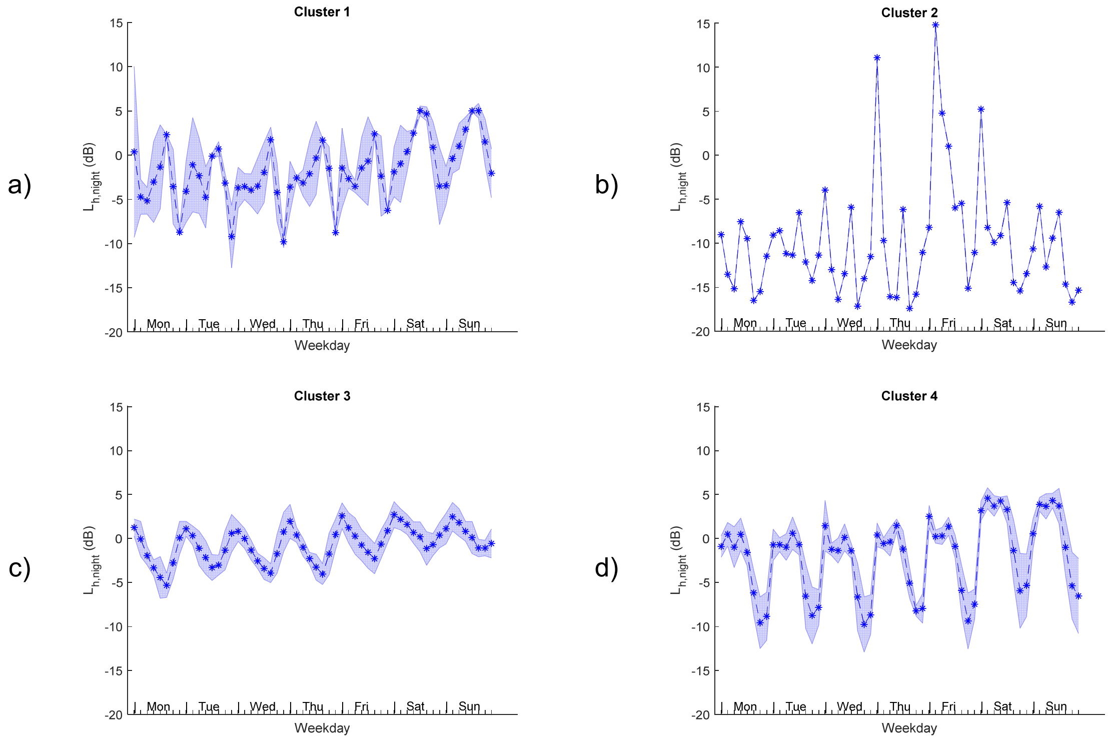

Figure 3. Temporal evolution of $L_{h,night}$ for the 4 clusters and its corresponding standard deviation (shade).

Based on previous analysis it can be interpreted that, for the night period, the influence of leisure activity could make higher noise levels during certain days (mostly on weekends) and that the classification by using L_{night} highlights this influence. Then, using L_{night} as the input, a principal component analysis (PCA) is performed in order to find more information regarding how the noise level of each day of the week affects the classification. The PCA brings evidence on how each variable is reflected in the clustering process and also let observe the configuration of each cluster itself. As can be seen in Table 2, 87% of the variability is explained by the 2 principal components. Figure 4 shows the plot of each sampling point by using the new coordinates defined by the two first principal components. As can be seen, Cluster 4 is totally in the right side of the plot, where the variables Friday and Saturday are located, this explains the high noise levels during this two days. Cluster 1 lays in the center of the axis, with half of the streets closer to Friday and Saturday. Cluster 3 is in the left side of the plot, having more influence of the weekdays rather than the weekend, this is the reason why the noise levels remains almost constant during the week. Finally, Cluster 2 is in the upper part of the plot, where Thursday is located. As observed in Figure 2 b), this cluster includes streets with high noise levels during Thursday.

Table 2. Coefficients for each day of the week and each principal components shown. It is also shown the percentage of variability explained by each one of the components.

Component	%	Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	78.51	-0.23	-0.49	0.36	-0.49	-0.30	0.18	0.45
2	9.22	-0.49	-0.29	-0.58	-0.11	0.56	0.10	0.11
3	4.88	-0.43	0.11	-0.20	0.60	-0.48	-0.03	0.42
4	3.63	-0.28	0.00	0.62	0.26	0.50	-0.44	0.16
5	2.58	-0.23	0.74	0.15	-0.24	0.16	0.49	0.22

6	0.74	0.38	0.23	-0.29	-0.27	0.07	-0.51	0.62
7	0.45	0.50	-0.24	0.07	0.42	0.30	0.51	0.39

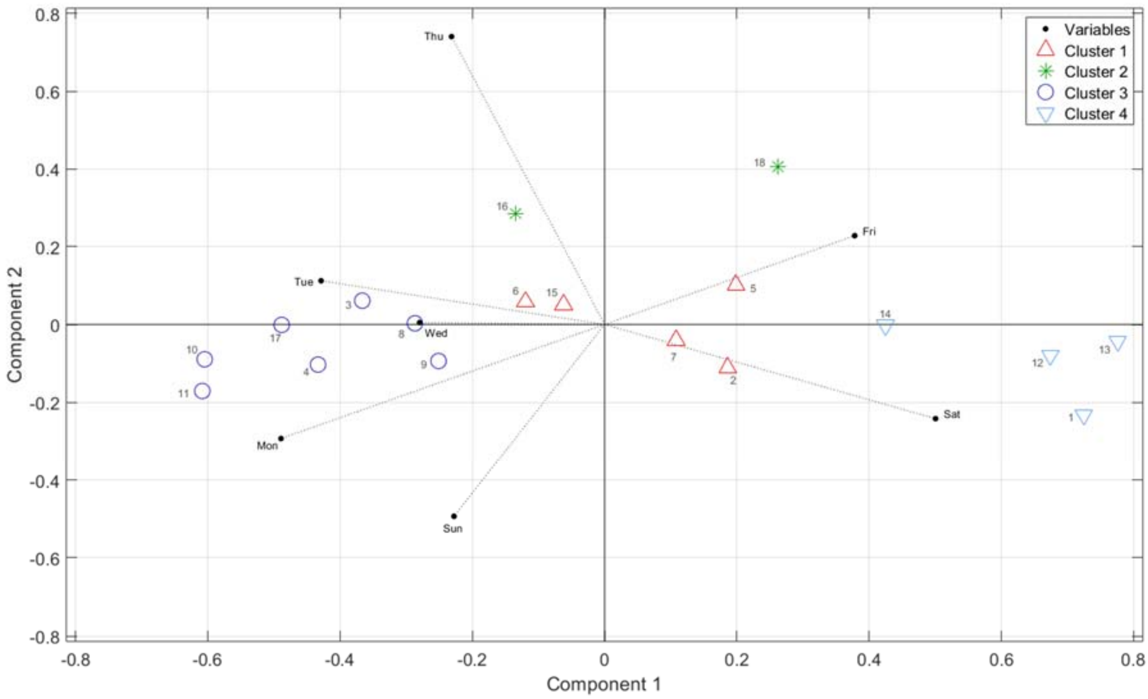

Figure 4. Sampling points scatter plot by using the corresponding coefficients for the first two principal components. The street number is shown, as well as their belonging cluster (symbol). The coefficients of the components 1 and 2 of each day of the week is also shown (black dots).

4. Conclusions

In order to perform a spatial classification of sampling points that takes into account the influence that other sources apart from traffic noise have in the noise levels evolution, a spatial classification based on cluster analysis is proposed. Two input parameters were tested, the each day equivalent night level L_{night}, and the hourly equivalent noise level $L_{h,night}$.

It was observed that the classification by $L_{h,night}$ highlights the similarity of the noise evolution of each day within the week, as it is observed that the hourly noise level of each day is similar for all days all along the week. For the case of using the daily L_{night}, it was observed that the sampling points are classified according to the increase in noise levels during weekends, which could be related to the leisure activities or land use in the zone.

A principal component analysis was performed to get more information about the variables and their relation to the clustering process. The analysis confirmed that the classification was done according to the weekly noise evolution. It could be observed that there is an increase of the noise levels during weekends (Friday and Saturday) in Cluster 1 and 4, which for the case of Cluster 2 includes Thursday, and different to the Cluster 1, where the noise level remains almost constant during the whole week.

References

1. Foraster M, Künzli N, Aguilera I, et al. High blood pressure and long-term exposure to indoor noise and air pollution from road traffic. *Environ Health Perspect*. 2014;122(11):1193-1200. doi:10.1289/ehp.1307156.
2. Basner M, McGuire S. WHO environmental noise guidelines for the european region: A systematic review on environmental noise and effects on sleep. *Int J Environ Res Public Health*. 2018;15(3). doi:10.3390/ijerph15030519.
3. Morillas JMB, Escobar VG, Sierra JAM, Vílchez-Gómez R, Vaquero JM, Carmona JT. A categorization method applied to the study of urban road traffic noise. *J Acoust Soc Am*. 2005;117(5):2844-2852. doi:10.1121/1.1889437.
4. Gajardo CP, Miguel J, Gozalo GR. Can weekly noise levels of urban road traffic , as predominant noise source , estimate annual ones ? 2016;140(5):3702-3709. doi:10.1121/1.4966678.
5. Quintero G, Balastegui A, Romeu J. Annual traffic noise levels estimation based on temporal stratification. *J Environ Manage*. 2018;206:1-9. doi:10.1016/j.jenvman.2017.10.008.
6. Romeu J, Jiménez S, Genescà M, Pàmies T, Capdevila R. Spatial sampling for night levels estimation in urban environments. *J Acoust Soc Am*. 2006;120(2):791-800. doi:10.1121/1.2215219.
7. Mioduszewski P, Ejsmont JA, Grabowski J, Karpi??ski D. Noise map validation by continuous noise monitoring. *Appl Acoust*. 2011;72(8):582-589. doi:10.1016/j.apacoust.2011.01.012.
8. Rey Gozalo G, Barrigón Morillas JM, Gómez Escobar V. Analyzing nocturnal noise stratification. *Sci Total Environ*. 2014;479-480(1):39-47. doi:10.1016/j.scitotenv.2014.01.130.
9. Morillas JMB, Escobar VG, Sierra JAM, Vílchez-Gómez R, Vaquero JM. Effects of Leisure Activity Related Noise in Residential Zones. *Build Acoust*. 2005;12(4):265-276. doi:10.1260/135101005775219085.
10. Ballesteros MJ, Fernández MD, Ballesteros JA. Acoustic evaluation of leisure events in two mediterranean cities. *Appl Acoust*. 2015;89:288-296. doi:10.1016/j.apacoust.2014.09.022.
11. Garrido MJB, Torija AJ, Fernandez MD, Ballesteros JA. Differences between road traffic and leisure noise in urban areas. Developing a model for automatic identification. *Acta Acust united with Acust*. 2016;102(1):35-44. doi:10.3813/AAA.918922.
12. Zambon G, Benocci R, Brambilla G. Cluster categorization of urban roads to optimize their noise monitoring. *Environ Monit Assess*. 2016;188(1):1-11. doi:10.1007/s10661-015-4994-4.