Model to predict the level and the temporal and spectral composition of the sound pressure in urban soundscapes based on neural networks

Torija Martinez, AJ, Ruiz, DP and Ramos-Ridao, A

10.13140/2.1.3238.1127

Model to predict the level and the temporal and spectral composition of the sound pressure in urban soundscapes based on neural networks

Conference Paper · January 2009

DOI: 10.13140/2.1.3238.1127

3 authors, including:

Antonio J. Torija
University of Salford

Diego P. Ruiz
University of Granada

Edinburgh, Scotland
EURONOISE 2009
October 26-28

Model to predict the level and the temporal and spectral composition of the sound pressure in urban soundscapes based on neural networks

Antonio J. Torija[a]
Diego P. Ruiz[b]
Department of Applied Physics Faculty of Sciences
Campus Fuentenueva s/n
University of Granada
18071 Granada (Spain)

Ángel Ramos-Ridao[c]
Department of Civil Engineering E.T.S. of I.C.C.P.
Campus Fuentenueva s/n
University of Granada
18071 Granada (Spain)

ABSTRACT

Environmental noise is a factor with a significant impact on personal well-being. One of the principal challenges of urban planning is the creation of soundscapes capable of providing inhabitants with a high quality of life. In the framework of the European legislation, the urban planners need tools that allow approaching the final goal of the design, planning and assessment of the soundscapes to adapt them to the needs required by the population. Nowadays, competent authorities have models of L_{Aeq} prediction to undertake this task, nevertheless, is not enough only with the analysis of the A-weighted equivalent sound pressure level (L_{Aeq}), but it is necessary to analyze the temporal and spectral composition of the sound pressure of the considered soundscape. For this reason, in this work a model for the prediction of the A-weighted (L_{Aeq}) and no weighted (L_{eq}) equivalent sound pressure level, the difference between both descriptors, the temporal sound level variability, the impulsiveness of the sound level, as well as the sound level for each of the 1/3-octave bands between 20-20000 Hz, has been developed. For that purpose, due to the great urban complexity, with a wide range of variability in the great amount of relevant acoustic variables, together with the great amount of input and output variables in the model we have based our model in a backpropagation neural network, with which a prediction with an average mistake of 1.62±1.63 dB and an average factor r^2 0.87 has been obtained.

1. INTRODUCTION

Classical environmental noise assessment basically focuses on reducing noise levels of unwanted sounds. Little or no attention was paid to detailed acoustical features of the sound, let alone to the meaning that the listener gave to this sound. More recently, the concept of (urban) soundscape design – an idea originating in the early seventies[1] – has

[a] Email address. ajtorija@ugr.es
[b] Email address. druiz@ugr.es
[c] Email address. ramosr@ugr.es

gained renewed interest[2-4]. Soundscape research takes a positive approach. The sonic environment is studied in a particular physical and social context taking into account its typical use. The description and evaluation of the sonic environment approaches the level of detail that music or landscape researchers have reached – or have been trying to reach – for many years. Appraisal by population of the soundscape is a key factor[5].

On the other hand, the great complexity of the urban agglomerations causes a challenge problem to solve from the point of view of the prediction of environmental acoustical variables. In the urban agglomerations it appears many noise sources and spaces, and diversity becomes common. So, locations with different composition of road traffic and urban configurations, sources difficult to characterize such as leisure/commercial activities, existence of green spaces, etc, cause different sound levels, as well as different temporal and spectral compositions of measured noise. The heterogeneous physiognomy of the urban environments together with the characteristics of the environmental noise, with a great spatial, temporal and spectral variability makes its modelization and prediction a very complex and non-linear problem, which forces to use a powerful tool of data mining, artificial neural networks[6,7]. The artificial neural networks are a paradigm of learning and automatic processing that seeks as ultimate goal emulate the biological brain, or at least some of its functions, such as learning. This method provides flexibility, accuracy and some amount of fault tolerance in noisy and changing environments[8].

For this reason, in this paper we have developed a model based on a backpropagation neural network to obtain a suitable prediction of both the temporal and spectral characteristics and sound noise levels in different urban locations. This model contributes to the final goal of providing a tool for the design, planning and assessment of the urban soundscapes with the final goal of integrating them in the needs of the population and obtaining acoustically sustainable urban agglomerations.

2. METHODOLOGY

A. Data Collection Process

For the study of urban soundscapes in the city of Granada, we raised the idea of creating a system that, through neural networks, will help us to obtain more precise results. A whole of 80 locations of the city of Granada was selected, trying to include the great heterogeneity of acoustic relevant situations present in an urban environment. Situations with different conditions of road traffic, in different periods of the day, with different geometry and physiognomy of the location, in parks and urban squares (without the presence of road traffic), with the presence of noisy sources different from the traffic (commercial/leisure, construction...), etc. were selected. Once realized all the measurements, every 5 minutes, the calculation of the different acoustic descriptors was realized, so that from the different selected input variables the prediction of each one of the used acoustic indicators was realized, with 5 minutes like integration period.

The acoustical measurements were taken with a sound level meter type 1. Specifically, it has been used the sound level meter 2260 Observer™ with the software BZ7219.

B. Input and Output Variables for the Development of the Model

For the development of the prediction model, as we can see in table 1, it have been selected a series of input variables (33 variables), included in 2 modules, a module of sound spread and a module of sound emission, which in turn consists of 3 sub-modules. These sub-modules contain variables for the characterization of the sound environment, variables related to the temporal evolution of the sound pressure and variables for the characterization of the sound emission generated by the road traffic respectively. In addition, we have 36 output variables, the A-weighted equivalent sound pressure level (L_{Aeq}) and no weighted equivalent sound pressure level (L_{eq}), the difference between both descriptors, the temporal sound level variance, the impulsiveness of the sound level, as well as the sound level for each of the 1/3-octave bands between 20-20000 Hz.

Table 1: *Input and output variables for the development of the prediction model.*

Input Variables			Output Variables
Module	Sub-module	Variable	
Sound Emission	Sound Environment	Type of Day	
		Day Period	
		Commercial or Leisure Environment	
		Appearance of Construction Works	
		Type of Localization	
		Appearance of Water Fountains	
		Appearance of Vegetation	
	Temporal Evolution of the Sound Pressure	Stabilization Time	L_{Aeq}
		Type of Traffic Flow	
		Anomalous Sound Events no related to Traffic	L_{eq}
		Anomalous Sound Events related to traffic	L_{eq} - L_{Aeq}
	Road Traffic Noise Emission	Ascendant/ Descendant Road Traffic Flow	TSLV
		Number of Vehicles with Siren	CF
		Average Speed	1/3-octave bands between 20-20000 Hz
		Traffic Slope	
		Number of Lanes Upward/Downward	
		Type of Pavement	
		Condition of Surface	
Sound Spread		Street Geometry	
		Street Width	
		Street Height	
		Roadway Width	
		Distance Source-Receptor	

With regard to the road traffic flow, 5 types of vehicles have been considered: light vehicles, heavy vehicles, buses, motorcycles-mopeds and urban cleaning vehicles.

For the assessment of the temporal variability of the sound pressure and the impulsiveness of the sound level, in this work the descriptors Temporal Sound Level Variance (TSLV) and Crest Factor (CF) have been used[9,10].

C. Structure of the Model based on Backpropagation Neural Network

For its property to generalize, and its possibility of supervised learning, we have used an artificial neural network, variant Levenberg-Marquardt with Bayesian regulation backpropagation (Trainbr like training function). The structure of this ANN (shown in figure 1) is 33 inputs variables, 36 neurons on the hidden layer and 36 output variables.

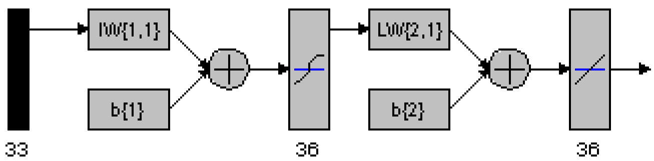

Figure 1: Structure of the proposed neural network.

Once selected the structure of the neural network, shown in figure 1, we built 5 different data sets. The 5 data sets are constructed randomly form the 533 input records. Therefore, we obtained five sets of data containing a training set and a test set, each one formed in turn by different records. The training sets contain 350 records and the test sets contain 183 records. The ANN was run five times, each one with different initial weights, with each of the five sets of data. This give us 25 trials to evaluate the accurate of the prediction of the different acoustics descriptors used in this work.

3. RESULTS

A. Acoustical Characteristics of Urban Soundscapes

Before analyzing the precision and accuracy of the neural network developed for the prediction of the sound level and temporal and spectral composition of the urban soundscapes a cluster analysis has been realized to identify the main types of soundscapes present in the city and, to observe the acoustical characteristics that have each of them. We have grouped in 5 main types to the different urban soundscapes. As we can observe in table 2, four of the five types of urban soundscapes are strongly dominated by the road traffic. The soundscapes 2 and 5 are characterized by the presence of a very high light vehicles flow (soundscape 2) and high heavy vehicles, buses and motorcycles-mopeds flow. The soundscapes 1 and 3 have similar road traffic flow, differing in what the soundscape 1 is characteristic of the downtown with a intermittent traffic flow and the appearance of a large number of anomalous sound events[11] and, the soundscape 3 is typical of the exterior edge of the city, with a high traffic speed and with wide avenues (free field).

The soundscapes 2 and 5 have the highest values for the acoustic descriptors L_{Aeq} and L_{eq} (table 3). In addition, these two soundscapes have the highest sound level for the totality of the 1/3-octave bands and, its composition in low-middle frequencies is very important (figure 2). The soundscapes 1 and 3 have similar values for the descriptors L_{Aeq} and L_{eq} (table 3), nevertheless, due to the characteristics of the soundscape 1 (intermittent traffic flow, anomalous sound events,...), this has a very high value in the level of temporal variability (TSLV) and impulsiveness (CF) of the sound pressure level. For what it concerns to the spectral composition (figure 2), the soundscape 1 has a greater composition in low frequencies, due to the high traffic slope, bad condition of road surface, geometry of the street, etc., whereas the soundscape 3 has a slight higher composition of middle frequencies.

Table 2: *Description of the analyzed urban soundscapes.*

Soundscape	Description
1	Appearance, in street of the downtown (very narrow street type "U") of intermittent traffic flow, with very high stabilization time of the sound level, principally, during the evening-night and weekends measurements periods. Appearance of a large amount of anomalous sound events, due, principally, to the high traffic slope or to the bad condition of the road surface.
2	Main routes of distribution of road traffic, with very high light vehicles flow circulating to high speed and, with commercial/leisure activities. Streets type "U" with a large amount of traffic lanes (high street width) and with great street height.
3	Appearance of a not very high road traffic flow, to high speed, in locations placed in the exterior edge of the city. Streets similar to the type "free field", characterized by its very large width and a not very high height.
4	Parks, squares and pedestrian streets, with the presence of vegetation (and water fountains, especially in squares and parks), with the absence of road traffic in its proximities.
5	Main routes of distribution of road traffic, with very high heavy vehicles, buses and motorcycles-mopeds flow circulating to high speed and, with commercial/leisure activities. Streets type "U" with a large amount of traffic lanes (high street width) and with great street height.

Finally, the soundscape 4 has the lowest value in all the acoustic descriptors (table 3). It has the lowest value of the sound pressure level, the temporal variability and the impulsiveness of all the urban soundscapes analyzed. Moreover, on having observed its spectral composition (figure 2) we verify that the content in low frequencies is very small, due to the absence of road traffic in its proximities.

Table 3: *Acoustic descriptors value of the analyzed urban soundscapes.*

Acoustic Descriptor	Soundscape 1	Soundscape 2	Soundscape 3	Soundscape 4	Soundscape 5
L_{Aeq} [dB(A)]	66.50	71.10	66.35	63.59	74.11
L_{eq} [dB]	76.15	80.61	74.91	72.61	83.21
L_{eq}-L_{Aeq}	9.65	9.51	8.57	9.02	9.10
TSLV [dB^2]	14.23	7.18	8.39	1.30	5.53
CF	1.07	0.95	0.83	0.49	1.03

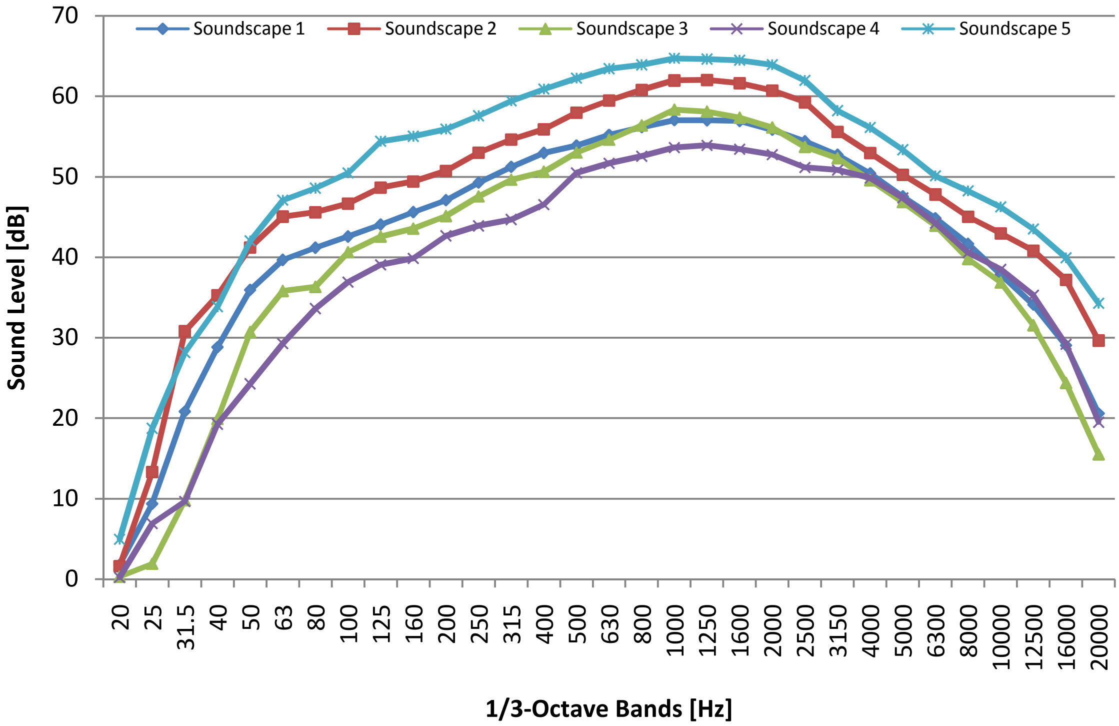

Figure 2: Spectral composition of the analyzed urban soundscapes.

B. Prediction Model based on Neural Network for Urban Soundscapes

Once selected the structure of the neural network, it was trained and tested 25 times, 5 times for each of the previously established 5 training-test sets. The obtained results appear in the table 4 (training sets) and table 5 (test sets). With regard to the descriptor L_{Aeq} (range value 0-100 dB(A)) the mean prediction error (MPE) range between 0.68±0.62 and 0.71±0.69, for training sets and between 0.82±0.76 and 0.90±0.81, for test sets. The correlation factor (R^2) range between 0.96-0.97 (training sets) and between 0.93-0.95 (test sets). The MPE and R^2 values for the descriptor L_{eq} (range value 0-100 dB) are 0.86±0.80 - 0.94±0.81 and 0.91-0.92, respectively (training) and, 0.90±0.73 – 1.08±0.84 and 0.90-0.91, respectively (test). For the descriptor L_{eq}-L_{Aeq} (range value 0-30), the MPE and R^2 values are 0.78±0.72 - 0.90±0.80 and 0.82-0.86, respectively (training) and, 0.86±0.70 – 1.05±0.95 and 0.78-0.82, respectively (test). The descriptor TSLV (range value 0-60 dB^2) has the MPE values, 0.89±0.88 - 0.92±0.90 (training) and 0.96±1.05 – 1.18±1.12 (test), and the R^2 values, 0.98 (training) and 0.96-0.98 (test).

Table 4: *Mean prediction error (MPE) and R^2 value of the training sets.*

Acoustic Descriptor	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
	MPE	R^2	MPE	R^2	MPE	R^2	MPE	R^2	MPE	R^2
L_{Aeq} [dB(A)]	0.71±0.64	0.96	0.70±0.67	0.96	0.71±0.69	0.96	0.68±0.62	0.97	0.70±0.63	0.97
L_{eq} [dB]	0.86±0.80	0.92	0.94±0.81	0.91	0.89±0.76	0.92	0.88±0.78	0.92	0.89±0.79	0.92
L_{eq}-L_{Aeq}	0.85±0.78	0.83	0.90±0.80	0.82	0.86±0.78	0.83	0.78±0.72	0.86	0.83±0.76	0.84
TSLV [dB^2]	0.91±0.84	0.98	0.92±0.87	0.98	0.89±0.88	0.98	0.92±0.90	0.98	0.91±0.92	0.98
CF	0.15±0.17	0.86	0.15±0.17	0.86	0.15±0.16	0.87	0.14±0.16	0.87	0.14±0.16	0.87
1/3-octave bands level	1.54±1.46	0.89	1.59±1.46	0.89	1.57±1.48	0.89	1.51±1.44	0.90	1.52±1.45	0.90

Table 5: *Mean prediction error (MPE) and R^2 value of the test sets.*

Acoustic Descriptor	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
	MPE	R^2	MPE	R^2	MPE	R^2	MPE	R^2	MPE	R^2
L_{Aeq} [dB(A)]	0.82±0.76	0.93	0.86±1.05	0.93	0.85±0.95	0.94	0.83±0.92	0.95	0.90±0.81	0.95
L_{eq} [dB]	0.90±0.73	0.91	1.08±0.84	0.90	0.97±0.85	0.91	1.05±0.89	0.91	1.04±0.86	0.91
L_{Aeq}-L_{eq}	0.86±0.70	0.82	1.05±0.95	0.78	0.92±0.83	0.82	1.02±0.93	0.82	0.97±0.91	0.80
TSLV [dB^2]	0.96±1.05	0.98	1.14±1.25	0.97	1.10±1.15	0.97	1.18±1.12	0.97	1.07±1.18	0.96
CF	0.15±0.15	0.85	0.16±0.18	0.87	0.15±0.16	0.89	0.16±0.16	0.89	0.14±0.15	0.89
1/3-octave bands level	1.70±1.69	0.83	1.79±1.86	0.85	1.76±1.78	0.86	1.72±1.73	0.87	1.76±1.72	0.86

For the descriptor CF (range value 0-10), the MPE and R^2 values are 0.14±0.16 - 0.15±0.17 and 0.86-0.87, respectively (training) and, 0.14±0.15 – 0.16±0.18 and 0.85-0.89, respectively (test). Finally, for the 1/3-octave bands (range value 0-90 dB) the average MPE range between 1.51±1.44 and 1.59±1.46, for training sets and between 1.70±1.69 and 1.79±1.86, for test sets. The correlation factor (R^2) range between 0.89-0.90 (training sets) and between 0.83-0.87 (test sets).

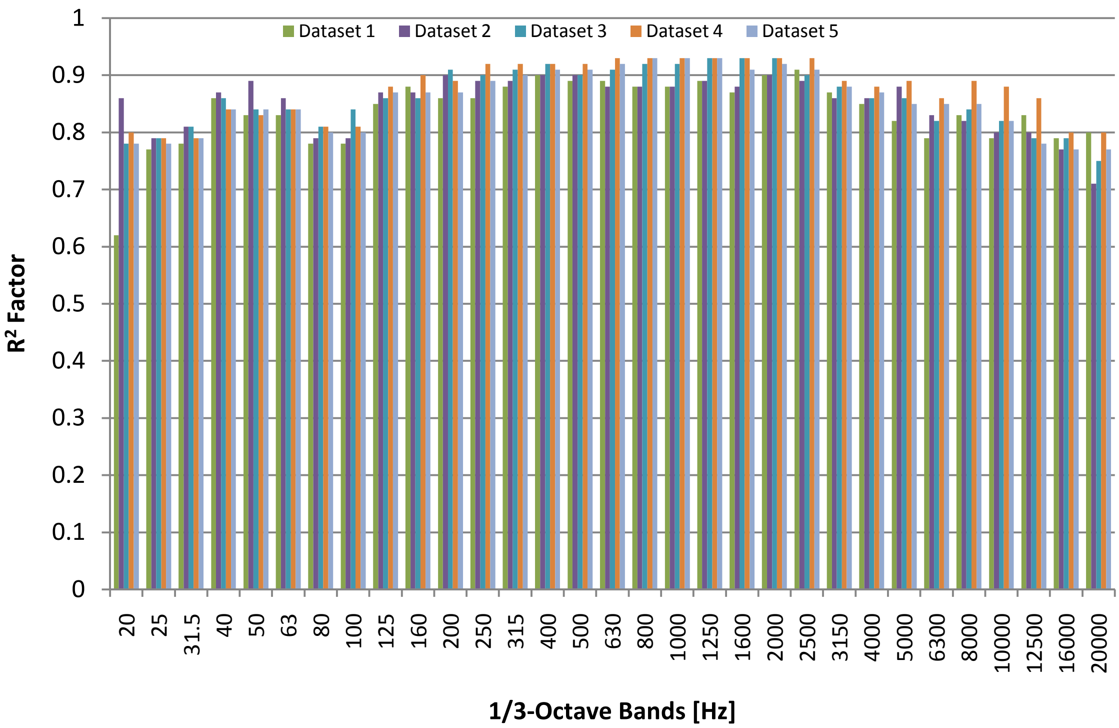

Figure 3: R^2 value for the 1/3-octave bands (20-20000 Hz) of the test sets.

As we can observe in the figure 3, the correlation factor of the prediction varies depending on considered 1/3-octave band. The bands with a greater value of the R^2 factor are located between 125-3150 Hz, whereas the bands with a correlation degree lower are located between 20-31.5 Hz, 80-100 Hz and 16-20 KHz. Nevertheless, observing the average value of the R^2 factor for the 5 datasets, we verify that the minimum value is 0.77, achieving values of 0.92.

In view of the obtained results, we can establish that the proposed neural network is capable of predicting with a considerable precision and accuracy both the sound pressure level (A-weighted and no weighted) and the temporal and spectral composition of the different types of situations presented to the network, situations that include a great heterogeneity and complexity (characteristics of the urban agglomerations), representing different soundscapes, each of which with totally different characteristics.

C. Analysis of the Tricky Cases

Once analyzed the capacity of prediction of the neuronal network for the set of 533 records, which include, as we have mentioned previously, a great heterogeneity of acoustic relevant situations, we have analyzed the response of the model to a series of tricky cases, situations where we find extreme values in some of the input variables. This kind of situations are very frequent in the urban agglomerations. The selected tricky cases can be observed in table 6.

Table 6: *Description of the analyzed tricky cases.*

Case	Description
1	Urban park.
2	Night period with very high stabilization time of the sound level.
3	Urban square with water fountain
4	Pedestrian walkway
5	Location with very low road traffic flow
6	Location with great traffic slope
7	Location with a large amount of vehicles with siren
8	Traffic congestion
9	Commercial location without road traffic in its proximity
10	Very narrow street

In the table 7, we can observe that for each of the acoustic descriptors, between 60 % and 90 % of the studied tricky cases have a smaller MPE than the average MPE of 5 test datasets for the same descriptor (table 5). For the descriptors L_{Aeq}, L_{eq} and L_{eq}-L_{Aeq} the percentage of tricky cases with MPE lower than the average MPE of 5 test sets is 90 %, whereas for the descriptors TSLV, CF and the spectral composition (1/3-octave bands) it is 60 %, 70 % and 80 %, respectively.

Table 7: *Mean prediction error (MPE) value of the analyzed tricky cases.*

Acoustic Descriptor	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10	Average
L_{Aeq} [dB(A)]	0.40	0.27	0.42	0.63	0.27	1.11	0.30	0.29	0.35	0.62	0.47±0.26
L_{eq} [dB]	0.52	0.19	0.25	0.46	0.71	1.03	0.65	1.15	0.64	0.98	0.66±0.32
L_{Aeq}-L_{eq}	0.38	0.23	0.24	0.02	1.03	0.02	0.47	0.93	0.49	1.68	0.55±0.52
TSLV [dB^2]	0.23	1.70	0.40	0.42	1.48	4.50	4.92	0.37	0.51	0.42	1.50±1.77
CF	0.13	0.32	0.05	0.12	0.02	0.18	0.62	0.07	0.11	0.02	0.16±0.18
1/3-octave bands level	1.55	2.10	0.72	1.51	1.95	1.68	1.07	1.32	1.78	1.55	1.52±1.34

The cases in which we can observe a higher prediction error are the case 2 (night period with very high stabilization time of the sound level), 5 (location with very low road traffic flow), 6 (location with great traffic slope) and 9 (commercial location without road traffic in its proximity) in which we can verify as the maximum prediction error appears in the descriptor TSLV and in the spectral composition (cases 2, 5 and 6) and in the descriptor L_{eq}-L_{Aeq} and in the spectral composition (case 9).

In the figure 4, it has been showed the mean prediction error of the analyzed tricky cases for the 1/3-octave bands between 20 Hz – 20 KHz.   In this case we verify as the greater values of mean prediction error appear between 40 Hz - 100 Hz and above 6300 Hz.

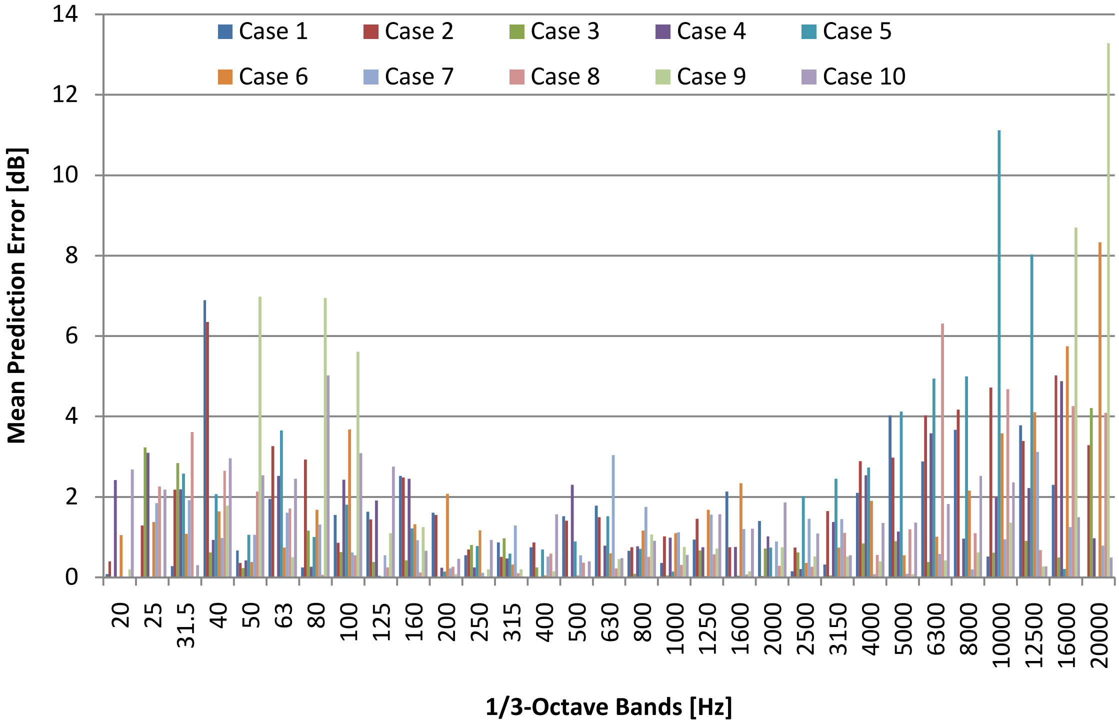

Figure 4: Mean prediction error (MPE) value for the 1/3-octave bands (20-20000 Hz) of the analyzed tricky cases.

4. CONCLUSIONS

The urban soundscapes are strongly affected by the road traffic, nevertheless, in an urban environment we can find situations in which the road traffic does not have a direct incidence (e.g. urban parks, pedestrian walkways, commercial places, etc.). In addition, the impact of the road traffic will be different depending on multiple factors, e.g. traffic intensity, type of traffic flow, type of vehicle, traffic speed and slope, etc. and, the sound spread will be different depending on the geometry of the street, type of road surface, etc. This generates the appearance of different kinds of urban soundscapes, which, as we have observed in this work (section 3.A), have totally different acoustical characteristics. All this gives the city a wide variety of situations, which must be modeled in a way that we can predict the temporal and spectral characteristics and the sound pressure level, with the final goal to assess and manage the different urban soundscapes, which are perceived and interpreted by the population, besides for non acoustic factors as information that contributes, value for the population, social context, etc., depending on its magnitude degree (sound pressure level) and its temporal and spectral composition.

In view of the obtained results (sections 3.B y 3.C), we can conclude that the proposed neural network achieves a good prediction of the temporal and spectral composition and of the sound level (A-weighted and no weighted) with great accuracy. This methodology and the obtained results allow us to be optimistic with the possibility of having a tool for the prediction of both the temporal, spectral composition and sound noise levels to issue the integration of the acoustic variables in the town planning for obtaining soundscapes adapted to the exposed population.

REFERENCES

1. R. M. Schafer, The Soundscape: Our Sonic Environment and the Tuning of the World. Destiny Books, Rochester, Vermont (USA) 1994.

2. D. Dubois, C. Guastavino and M. Raimbault, A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories, *Acta Acustica united with Acustica*, 92(6), pp. 865-874, (2006).

3. B. De Coensel and D. Botteldooren, The quiet rural soundscape and how to characterize it, *Acta Acustica united with Acustica*, 92(6), pp 887-897, (2006).

4. B. Berglund and M. E. Nilsson, On a tool for measuring soundscape quality in urban residential areas, *Acta Acustica united with Acustica*, 92(6), pp. 938-944, (2006).

5. See for example the special issue on soundscapes, *Acta Acustica united with Acustica,* 92(6), (2006).

6. Y. Avsar, A. Saral, M. T. Gönüllü, E. Arslankaya and U. Kurt, Neural network modelling of outdoor noise levels in a pilot area, *Turkish Journal of Engineering & Environmental Sciences,* 28, pp. 149-155, (2004).

7. G. Cammarata, S. Cavalieri and A. Fichera, A neural network architecture for noise prediction, *Neural Networks,* 8(6), pp. 963-973, (1995).

8. J. C. Patra and G. Panda, ANN-based intelligent pressure sensor in noisy environment, *Measurement,* 23, pp. 229-238, (1998).

9. A. J. Torija, D. P. Ruiz and A. Ramos, A method for prediction of the stabilization time in traffic noise measurements, in *Proceedings of 19th International Congress on Acoustics,* Madrid, Spain, 2007.

10. A. J. Torija, D. P. Ruiz, D. Botteldooren and B. De Coensel, Taking into account the temporary and spectral structure of the sound energy for the characterization of the annoyance generated by the road and railway traffic, in *Proceedings of the 5th Iberian Congress of Acoustics (Acústica '08)*, Coimbra, Portugal, 2008.

11. A. J. Torija, D. P. Ruiz and A. Ramos, Obtaining of a factor to describe the anomalous sound events in traffic noise measurements, in *Proceedings of 19th International Congress on Acoustics,* Madrid, Spain, 2007.