24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

SOUNDSCAPE CHARACTERIZATION AND CLASSIFICATION: A CASE STUDY

Giovanni Brambilla

CNR-IDASC, Institute of Acoustics and Sensors "O.M. Corbino" of the National Research Council of Italy, Via del Fosso del Cavaliere 100, 00133 Roma, Italy

Francesca Pedrielli

CNR-IMAMOTER, Institute for Agricultural and Earthmoving Machines of the National Research Council of Italy, Via Canal Bianco 28, 44124 Ferrara, Italy
email: f.pedrielli@imamoter.cnr.it

Massimiliano Masullo

Department of Architecture and Industrial Design, Università degli Studi della Campania "L.Vanvitelli", Via San Lorenzo, 81031 Aversa (CE), Italy

In the historic center of Alghero, a seaside town in Sardinia-Italy, a soundwalk with 18 participants was performed. Ten sites distributed along the soundwalk path have been selected to be representative of different environments, namely in pedestrian areas, in an urban garden, along the seafront and so forth. Binaural recordings were carried out with simultaneous subjective appraisals of the sonic environment and of other features, like perceived landscape quality and its influence on soundscape ratings. The subjective data were collected by a questionnaire filled in by the subjects at each site. Acoustic parameters have been determined from binaural recordings and together with the subjective data have been analyzed by statistical procedures of feature extraction and classification, as well as to develop a model to predict classification membership. The unsupervised algorithm of hierarchical clustering was applied. Among the solutions the one that groups the data into three categories looked the most appropriate considering the characteristics of the sites. In order to develop a model for predicting the cluster membership, multinomial logistic regression has been performed by the "caret" package available in the "R" software using k-fold cross validation (k=10) and 5 repetitions. The available data have been split into two sets, one used for training the model and the other used for testing it. The results in terms of classification performance indices are rather promising.
Keywords: soundscape, soundwalk, classification, modelling

1. Introduction

The way a place sounds may be a key factor in the development of a sense of the place itself [1]. In recent years several authors have attempted to identify the informational, aesthetic or affective qualities of sound which help to confer quality on a given landscape [2]. In this framework soundscape, which is different from the sonic environment as it is a perceptual construct whereas the latter is a physical phenomenon, can provide useful hints to improve the quality of the environment also in relation to the several uses and expectations of its users. Due to the large diversity of soundscapes, their classification [3] is fundamental to guide towards their design and improvement. In addition, the development of models based on the environment features and able to predict the soundscape category are also useful tools for managing and planning.

For the above purposes, soundwalk [4, 5] is a tool frequently used to collect acoustic data on the sonic environment under investigation and subjective appraisals on its features (i.e. acoustic, visual, aesthetic, etc.) given by a sample of population through questionnaires.

The present paper describes the case study undertaken in the historic center of Alghero, a seaside town in Sardinia-Italy, where a soundwalk with 18 participants was performed along a path where ten spot sites were selected to represent different environments. Statistical procedures of feature extraction and classification have been applied, as well as multinomial logistic regression in order to develop a model to predict the cluster membership. The results in terms of classification performance indices are rather promising.

2. Experimental procedure

The soundwalk was performed in May, during a sunny afternoon from 6 to 7 pm. The 18 participants were divided into two groups, each formed by 9 subjects walking along the planned path in two opposite directions (clockwise and counter-clockwise) to minimize two types of biases on ratings:

- sequential order bias, that may occur because the subject’s rating is conditional on her/his rating order;
- sequential history bias, that may occur because the evaluation of soundscape of a site may directly depend on the quality of the previous one(s).

The path and the selected 10 spot sites were available to each subject on her/his smartphone by means of “Google map” application (Fig. 1).

Figure 1: Soundwalk path and spot sites.

The groups were led by investigators who walked across the study area and stopped at the 10 selected locations. The sites were chosen to be representative of different environments (pedestrian areas, urban garden, along the seafront and so forth). At each site, participants were required to listen to the sonic environment for about two minutes and then fill in the structured questionnaire, available on her/his smartphone by means of “Google Form” application.

At the same time, the investigator in the group took a binaural recording (lasting 2 minutes) of the sonic environment by means of a calibrated binaural headset connected to a mobile digital sound recorder and measurement system (SQobold-Head Acoustics). From the binaural recordings the acoustic parameters listed in Table 1 have been determined for each site and channel (left and right ear.

The questionnaire used to collect subjective ratings was structured into 5 sections. The first one dealt with name of the subject, age, group membership, spot site and time of filling the questionnaire. In section 2 the perceived sounds, separated into traffic, technological, anthropic and natural sounds, were rated on a 5-point scale from 1 (not heard at all) to 5 (very often heard). In section 3, both the perceived quality of the soundscape and its congruence with the site were rated on a scale from 0 (very much negative) to 10 (very much positive). The next section dealt with the appraisal of eight attributes of the soundscape similar to those reported in [6] (eventful, exciting, pleasant, calm, uneventful, monotonous, annoying, chaotic), given on a 5-point Likert scale from 1 (completely disagree) to 5 (completely agree). In the last section, both the perceived quality of the landscape and its influence on the soundscape ratings were rated on a scale from 0 to 10 (from bad and not at all to very good and very much, respectively).

Table 1: Acoustic parameters determined for each channel and site

A-weighted continuous equivalent level L_{Aeq} [dB(A)]	*Roughness R* [asper]
Percentile levels L_{A10}, L_{A50}, L_{A90} [dB(A)]	*Fluctuation Strength F* [vacil]
Loudness N, N_5, N_{50} [sone]	*Centre of gravity logG of 1/3 octave band spectrum* (80÷8000 Hz) [7]
Sharpness S [acum]	

3. Statistical analysis procedures

The statistical analysis of the collected data was carried out by the software “R”, an open-source programming environment for data analysis, graphics and statistical computing [8], with the extension of some specific packages.

The dataset included 180 observations (9 subjects × 2 groups × 10 sites) for each of the 29 variables, namely 16 formed by subjective responses and 13 by the acoustic parameters, each scaled to get standardize scores.

In order to reduce the number of variables, multicollinearity was investigated by means of Spearman’s rank correlation matrix on which a cut-off value set at $\rho = \pm 0.4$ was applied. Subjective and acoustic variables with a correlation coefficient below this value have been considered poorly correlated and kept for further analyses.

Afterwards, to explore the classification of the sites on the basis of the kept variables cluster analysis was performed. On the kept variables, different clustering methods available in the “clValid” R package [9] were applied. In particular, six methods were considered, that is hierarchical, Partitioning Around Medoids (PAM), k-means, DIvisive ANAlysis clustering (DIANA), Model-base clustering and Self-Organizing Tree Algorithm (SOTA). The range of clusters was set between 2 and 10, corresponding the former to the minimum and the latter to the maximum of discrimination among the sites. In addition, the Euclidean distance among the observations was considered. The clustering performance of the methods was ranked according to 7 parameters, that is Connectivity, Silhouette Width and Dunn Index combining measures of compactness and separation of the clusters, the average proportion of non-overlap (APN), the average distance (AD), the average distance between means (ADM) and the figure of merit (FOM). The selected method was applied considering the chosen number of clusters to get the classification of the sites.

The next step was focused on developing models to predict the cluster membership on the basis of the selected features of the sites. For this purpose the “caret” R package, acronym for “Classification And REgression Training” [10], was used. The dataset needed to be divided into two subsets, one for training the model and the other to test it and evaluate its classification performance. In particular, the responses given by the subjects in group 1 were used as ‘training dataset’ while those given by the subjects in group 2 were used as ‘test dataset’. This separation was preferred to a random one because the binaural recordings were carried out along with the appraisals of group 1, while no recordings were taken for group 2. The multinomial logistic regression was applied to develop the model because the dependent variable (cluster membership) is categorical with more than two categories. The k-fold cross validation was used, with k = 10 and

5 repetitions. The importance of the predictors in the model (independent variables) was also computed, as well as the classification performance of the model determined comparing the predicted classification on the test data subset with that obtained by the cluster analysis.

4. Results and discussion

The sonic environment at the 10 sites was sufficiently different and ranged from 75 down to 53 dB(A), as shown by the measured L_{Aeq} values plotted in Fig. 2(a) in descending order. Based on this parameter, the sites may be roughly grouped into three categories, namely a group with L_{Aeq} levels greater than 65 dB(A), sites 4 and 6, another with L_{Aeq} between 55 and 65 dB(A), sites 2, 3, 5, 7 and 9, and the third group with L_{Aeq} below 55 dB(A), sites 1, 8, and 10.

In Fig. 2(b) the mean value ± 1 standard deviation of subjective responses on the perceived quality of the soundscape at each site are reported. Site 3 shows significant difference on ratings between the two groups, most likely due to sequential history bias because group 1 rated site 3 after site 2 (higher mean score), whereas group 2 made the appraisal at site 3 after site 4 (lower mean score). The scores tend to increase with decreasing of L_{Aeq}, except for site 9 where soundscape was rated good even if L_{Aeq} was rather higher than levels observed at sites 1, 8 and 10. This confirms, once again, the need of additional acoustic parameters and non-acoustic features to properly characterize the soundscape.

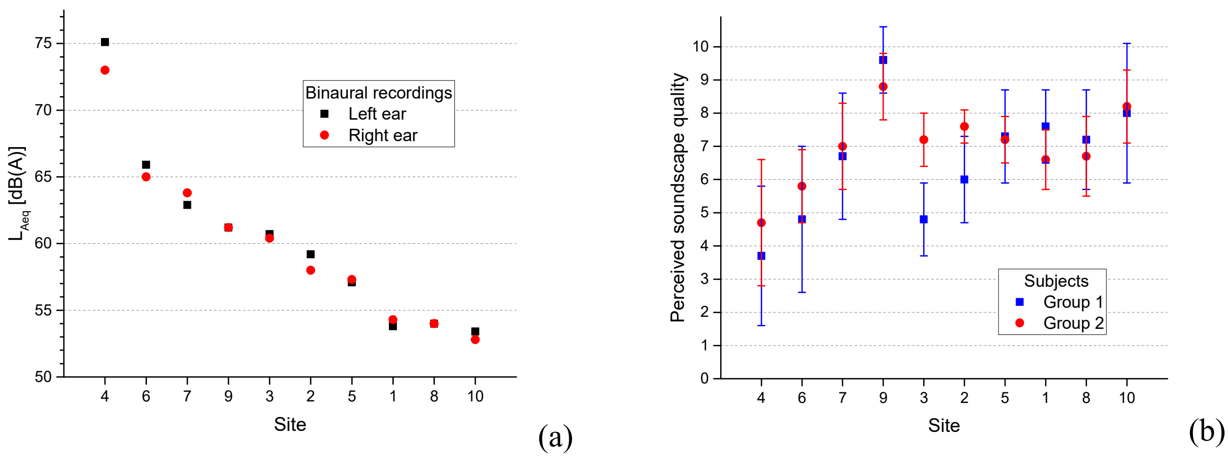

Figure 2: Differences among the 10 sites in terms of (a) L_{Aeq} equivalent levels and (b) perceived quality of the soundscape (mean value ± standard deviation).

The Spearman's rank correlation matrix and the cut-off value set at $\rho = \pm 0.4$ showed that only 10 variables had lower correlation (correlogram in Fig. 3), namely those relating to:

- the type of perceived sounds, that is *technological (TS)*, *anthropic (AS)*, or *natural (NS);*
- the subjective *expectation (EX)* of the soundscape;
- the soundscape features of being either *eventful (EV)* or *monotonous (MO)*;
- the self-rated *influence of the landscape on the perceived soundscape quality (AV)*;
- the *roughness (R)*;
- the *difference* $L_{A10} - L_{A90}$ *(SC)*;
- the *center of gravity of the sound spectrum* in terms of *logG*.

The output of the validation procedure on the selected six clustering methods indicated as best solution the hierarchical agglomeration using the Ward algorithm with discrimination into two groups. In particular, site 4 was in a group and all the other sites were grouped in the remaining cluster, as shown in Table 2. This table reports the distribution of the observations among the clusters for each site, either for two and three groups clustering. It is worth noting that the number of observations at site 4 was 17 because a few missing values were present in the questionnaire filled in by one subject and, therefore, the corresponding observation was not considered in the analysis. The dendrogram in Fig. 4 shows the agglomeration process: the higher the height, the greater the distance between clusters.

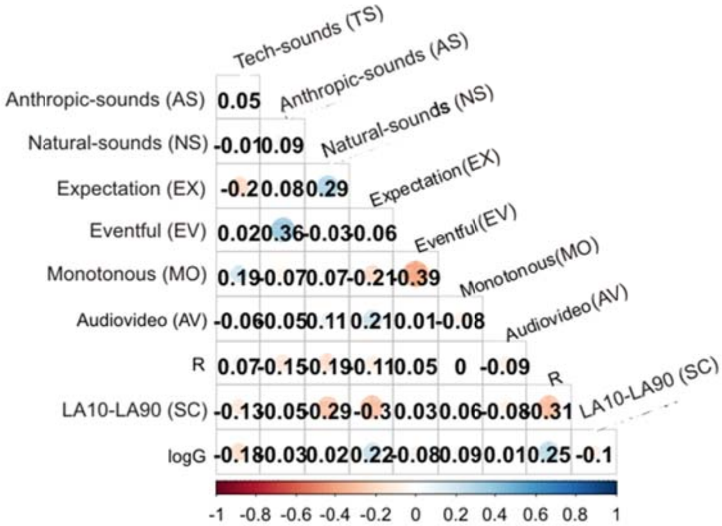

Figure 3: Spearman's correlogram of the 10 variables.

Table 2: Distribution of observations for each site among clusters

			Sites									
		Tot.	1	2	3	4	5	6	7	8	9	10
2 clusters	1	162	18	18	18	--	18	18	18	18	18	18
	2	17	--	--	--	17	--	--	--	--	--	--
3 clusters	1	108	18	17	14	--	16	--	16	9	17	1
	2	54	--	1	4	--	2	18	2	9	1	17
	3	17	--	--	--	17	--	--	--	--	--	--

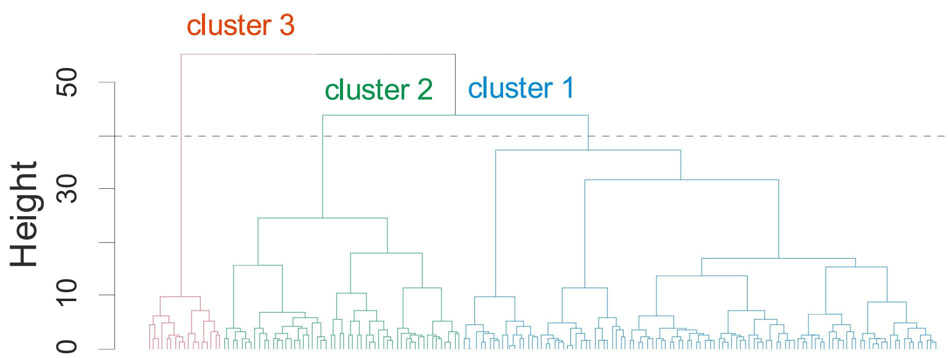

Figure 4: Dendrogram of hierarchical clustering.

The 2-cluster solution has the advantage that all the observations in a site belong to the same cluster. This behavior does not often occur in the 3-cluster solution where the observations in 8 sites are distributed between cluster 1 and 2. For these sites the cluster containing the majority of observations was taken as cluster membership. The observations at site 8 are equally split between cluster 1 and 2 and, therefore, this site was associated to both clusters. Notwithstanding a less clear distinction, it has been decided to keep the classification into three groups to get a higher discrimination among them.

To have insights and cues to associate meanings to the statistical classifications of the sites, the responses to specific questions in the questionnaire were analyzed: the 8 attributes of the soundscape, the perceived type of sounds, the perceived quality of the soundscape and its congruence with the site (expectation) and the perceived quality of the landscape. Because the observations at site 8 were equally split between cluster 1 and 2, half percentages of scores were considered for each cluster.

The spider plot in Fig. 5(a) shows the percentage of subjects giving a score from 4 to 5 on the Likert's scale for the 8 features of the soundscape. It can be seen that cluster 3 (site 4) is associated to a chaotic, monotonous and annoying soundscape, whereas soundscape features of sites in cluster 1 differ to those in cluster 2 mainly for being more pleasant and calm, and less chaotic.

Regarding the other appraisals, the spider plot in Fig. 5(b) shows the percentage of subjects giving a score from 4 to 5 on the Likert's scale for the perceived type of sounds (traffic, technological, anthropic, and natural) and the percentage of subjects giving a score from 6 to 10 on the rating scale for the soundscape and landscape quality, as well as the expectation. Fig.5(b) shows that in cluster 3 (site 4) the predominance of traffic sounds reduces the perceived quality of soundscape and landscape; the expectation is also not high. In cluster 1 the technological sounds are perceived less than in cluster 2, contrarily to the anthropic and natural sounds leading to an increase in soundscape quality (+ 12%). This confirms the important role played by the type of sound sources in the soundscape appraisal.

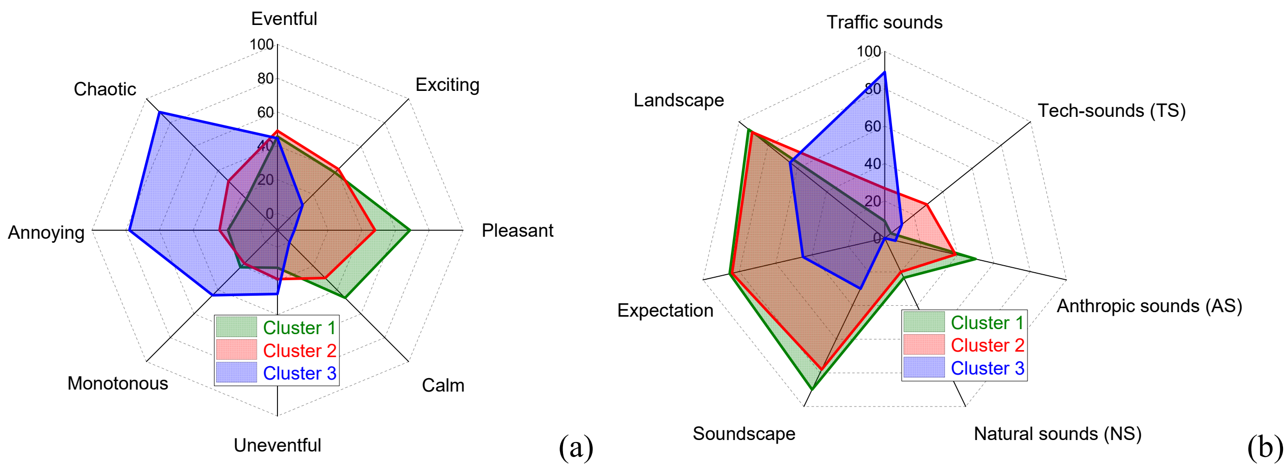

Figure 5: Features and statistical classification of the sites in terms of: (a) 8 different features of the soundscape; (b) other appraisals on the perceived type of sounds, soundscape and landscape quality and expectations.

Regarding the classification model, because the dependent variable (cluster membership) to be predicted had three levels (categories), multinomial logistic regression was applied considering all the 10 variables used for the cluster analysis. After the training process, taking cluster 1 as reference, the model equations in terms of probability P of an observation to belong to either cluster 2, or cluster 3, were those reported in Eq. (1) and Eq. (2):

$$\ln\left(\frac{P2}{P1}\right) = -1.71+1.61TS-0.31AS-0.50NS-1.38EX-0.47EV+0.44MO-0.42AV-1.25R-3.23SC-2.92\log G \quad (1)$$

$$\ln\left(\frac{P3}{P1}\right) = -2.79+0.27TS-0.82AS-0.43NS-0.40EX-0.17EV-0.14MO+0.13AV+2.10R-0.16SC-0.56\log G \quad (2)$$

where P1 is the probability of an observation to belong to cluster 1;
P2 is the probability of an observation to belong to cluster 2;
and so for.

The importance of the predictors in the model is given in Fig. 6 in relative terms. It can be seen that the three acoustic parameters have greater importance than the seven subjective variables.

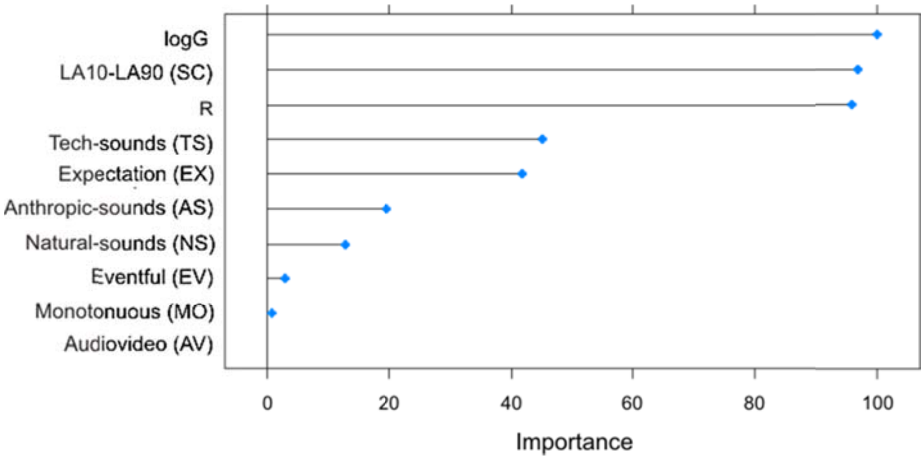

Figure 6: Relative importance of the predictors in the model.

The classification model was then applied to the test dataset in order to evaluate its classification performance and the obtained confusion matrix is reported in Fig. 7. The results are very good, being the model accuracy 0.97 and the Cohen's kappa κ = 0.93. Table 3 reports additional performance parameters for each category.

Figure 7: Confusion matrix of the model applied to the test dataset.

Table 3: Classification performance of the model for each category

	Category		
Parameter	1	2	3
Sensitivity	0.97	0.95	1
Specificity	0.97	0.97	1
Prevalence	0.67	0.23	0.10
Detection rate	0.64	0.22	0.10
Detection prevalence	0.66	0.24	0.10

5. Conclusions

In soundscape studies, binaural recordings and questionnaires are commonly used to describe sound environments to characterize different places within cities. The survey undertaken in the historic center of Alghero has confirmed the validity of soundwalk as a method to get acoustic parameters and subjective appraisal for soundscape characterization. Statistical procedures of feature extraction may assist to reduce the number of the subjective and objective variables to be used for categorization of environments, by means of clustering, and for developing classification models. In the described case study 10 out of 29 collected variables were selected, including 7 subjective variables and 3 acoustic parameters.

Even though the outcome of the case study can not be generalized, the methodologies used for both the field survey and the statistical analyses have given satisfactory results.

The output of hierarchical clustering with Ward algorithm showed a satisfactory distinction of the sites into three groups. This categorization was associated to specific features of soundscape to get insights and cues to identify factors influencing the perceived quality of the environment, such as the type of heard sound sources. The three main clusters identified by this analysis, seem to suggest the following interpretation of the datasets partitions. Cluster 1 identifies a soundscape calm and pleasant with the presence of anthropic or natural sounds and an overall good quality of landscape and soundscape. Cluster 2 is similar to Cluster 1 with a soundscape calm, pleasant but exciting and eventful, probably due to the presence of technological sounds. Cluster 3 identifies a soundscape chaotic, annoying but also monotonous and uneventful with the presence of almost all traffic sounds and with an overall medium quality of landscape and soundscape.

The classification model that was developed based on the 10 selected variables showed a good performance with high accuracy (0.97) and an optimal correspondence between predicted and obtained cluster membership (Cohen's kappa κ=0.93).

Further studies can investigate the classification performance of models developed with different algorithms, such as random forest, neural networks, and so on.

Acknowledgements

The Authors are grateful to Aniello Pascale and Francesco Sorrentino for their cooperation in the binaural recordings and determination of acoustic parameters.

REFERENCES

1 Schafer, R. M., *The Tuning of the World*, Knopf, New York (1977).

2 Carles, J. L., Barrio, I. L. and de Lucio, J. V., Sound influence on landscape values, *Landscape and Urban Planning*, 43, 191-200, (1999).

3 Rychtáriková, M. and Vermeir, G., Soundscape categorization on the basis of objective acoustical parameters, *Applied Acoustics*, 74, 240–247, (2013).

4 Semidor, C., Listening to a city with the soundwalk method, *Acta Acust. united Acust.*, 92, 959–964, (2006).

5 Jeon, J. Y., Hong, J. Y. and Lee, P. J., Soundwalk approach to identify urban soundscapes individually, *J. Acoust. Soc. Am.*, 134, 803–812, (2013).

6 Axelsson, Ö, Nilsson, M. E. and Berglund, B., A principal components model of soundscape perception, *J. Acous. Soc. Am.*, 128(5), 2836–2846, (2010).

7 Raimbault, M., Lavandier, C. and Bérengier, M., Ambient sound assessment of urban environments: field studies of two French cities, *Applied Acoustics*, 64, 1241-1256, (2003).

8 R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, (2013).

9 Brock, G., Pihur, V., Datta, S. and Datta, S., clValid: An R Package for Cluster Validation, *J. Stat. Softw.*, 25, 1-22, (2008).

10 Kuhn, M., Building predictive models in R using the caret package, *J. Stat. Softw.*, 28, 1–26, (2008).