JID: ESWA [m5G;January 27, 2016;15:46]

Expert Systems With Applications xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Automated classification of urban locations for environmental noise impact assessment on the basis of road-traffic content

Q1 Antonio J. Torija[a,*], Diego P. Ruiz[b]

[a] Institute of Sound and Vibration Research, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom
[b] Department of Applied Physics, University of Granada, Avda. Fuentenueva s/n, 18071 Granada, Spain

ARTICLE INFO

Keywords:
Machine-learning
Feature selection
Classification
Road traffic noise
Noise impact
Urban environments

ABSTRACT

Urban and road planners must take right decisions related to urban traffic management and controlling noise pollution. Their assessments and resolutions have important consequences on the annoyance of population exposed to road-traffic-noise and controlling other environmental pollutants (e.g. NOx or ultrafine particles emitted by heavy vehicles). One of the key decisions is the selection of which noise control actions should be taken in sensitive areas (residential or hospital areas, school areas etc), that could include costly measures such as reducing the overall traffic, banning or reducing traffic of heavy vehicles, inspection of motorbikes sound emission, etc. For an efficient decision-making in noise control actions, it is critical to classify a given location in a sensitive area according to the different prevailing traffic conditions.

This paper outlines an expert system aimed to help urban planners to classify urban locations based on their traffic composition. To induce knowledge into the system, several machine learning algorithms are used, based on multi-layer Perceptron and support vector machines with sequential minimal optimization. As input variables for these algorithms, a combination of environment variables was used. For the development of the classification models, four feature selection techniques, i.e., two subset evaluation (correlation-based feature-subset selection and consistency-based subset evaluation) and two attribute evaluation (ReliefF and minimum redundancy maximum relevance) were implemented to reduce the models' complexity. The overall procedure was tested on a full database collected in the city of Granada (Spain), which includes 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 4 categories observed, with values of average weighted *F*-measure and Kappa statistics (used as indicators) up to 0.9 and 0.8. Regarding the feature selection techniques, attribute evaluation algorithms (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.

1. Introduction

1.1. Urban road-traffic and noise

Road-traffic is known to be one of the main sources of pollution in urban environments (Nedic, Despotovic, Cvetanovic, Despotovic, & Babic, 2014). In many European urban areas, the road-traffic has been found as the predominant source of noise and most airborne pollutants (Can et al., 2011b). Both noise and air pollution are major environmental stressors that may lead to important psychological or physiological effects (Foraster et al., 2011). In terms of environmental noise, the influence of road-traffic-noise on human health has been analyzed by several studies (Babisch, 2006; Babisch et al., 2013; Brink, 2011; Caciari et al., 2013; Fyhri & Klaboe, 2009; Ising & Krupa, 2004; Muzet, 2007; Pirrera, De Valck, & Cluydts, 2010), which pointed out the road-traffic-noise not only as the most annoying noise source in urban environments (Calixto, Diniz, & Zannin, 2003), but also as a concern for public

Q2 * Corresponding author. Tel.: +44 2380532276.
E-mail addresses: ajtorija@ugr.es, ajtm19@gmail.com (A.J. Torija), druiz@ugr.es (D.P. Ruiz).

http://dx.doi.org/10.1016/j.eswa.2016.01.011
0957-4174/© 2016 Published by Elsevier Ltd.

health and environmental welfare (Kassomenos, Vogiatzis, & Bento Coelho, 2014). Furthermore, road-traffic-noise influences property prices in urban areas (Blanco & Flindell, 2011).

An important aspect to be considered is the composition of the road-traffic. The appearance of heavy vehicles and powered two wheelers (motorbikes and mopeds) in traffic lead to higher noise levels and reported annoyance (Braun, Walsh, Homer, & Chuter, 2013; Paviotti & Vogiatzis, 2012). Moreover, these road vehicles have been found as the most prevalent noticed-sound-events (NSE) in urban environments (Torija, Ruiz, Alba-Fernandez, & Ramos-Ridao, 2012). Under the assumption that sound has to be noticeable in order for it to contribute to an overall impression of annoyance, the NSE is a crucial factor to be considered for the evaluation of road-traffic-noise annoyance (De Coensel et al., 2009). Therefore, a tool for the identification of NSE might be used for the elaboration of action plans against environmental noise in urban environments.

Due to the good correlations found between noise levels and traffic intensity, some authors have approached the estimation of traffic parameters from recorded sound levels (Can et al., 2011a; Torija & Ruiz, 2012). Thus, for instance, Torija and Ruiz (2012) developed a series of classifiers to detect the urban scenarios where the percentage of heavy vehicles or motorcycles/mopeds is greater than a given threshold.

1.2. Applications of machine learning in environmental noise modeling

Machine learning algorithms have been widely applied to real-world environmental applications. As two of the most applied machine learning methods, artificial neural network (ANN) and support vector machine (SVM) are powerful algorithms for classification and regression problems. Thus, ANN- and SVM-based models have been developed in research fields such as, air pollution (Hájek & Olej, 2012), geology (Anifowose, Labadin, & Abdulraheem, 2015; Feng, Zhang, Zhang, & Wen, 2015), hydrology (Cho et al., 2014; Lafdani, Nia, & Ahmadi, 2013; Tan, Yan, Gao, & Yang, 2012; Xu & Liu, 2013), meteorology (Mercer, Dyer, & Zhang, 2013; Wu, Long, & Liu, 2015), renewable energy (Ekici, 2014; Gnana Sheela & Deepa, 2013; Mena, Rodríguez, Castilla & Arahal, 2014; Yadav & Chandel, 2014; Yadav, Malik, & Chandel, 2014; Yaïci & Entchev, 2014; Zheng & Qiao, 2013), or transportation (Jiang, Zhang, & Chen, 2014; Li et al., 2014; Ma, Tao, Wang, Yu, & Wang, 2015; Zhu, Cao, & Zhu, 2014).

Regarding noise related applications, several authors have used ANN algorithms to develop prediction models. Thus, Givargis and Karimi (2010) presented a multi-layer Perceptron (MLP) model which uses 5 input variables (hourly traffic flow, percentage of heavy vehicles, hourly mean traffic speed, gradient and angle of view) for the estimation of hourly A-weighted sound pressure level ($L_{Aeq,\,1\,h}$) in roads in Tehran at distances under 4 m from the near-side carriageway edge. In this work no significant difference was detected between the performance of the developed neural network and a calibrated version of the CORTN model (UK Calculation of Road Traffic Noise). Kumar, Nigam, and Kumar (2014) applied a multi-layer feed forward back propagation (BP) neural network, trained by Levenberg–Marquardt (L–M) algorithm, to develop an ANN model for predicting highway traffic noise. This model accurately estimated the 10 percentile exceeded sound level (L_{A10}) and the L_{Aeq} descriptor by accounting the input parameters found as more relevant to Indian highway traffic conditions (traffic volume, heavy vehicle percentage and average vehicle speed). Nedic et al. (2014) used 5 input variables (number of light motor vehicles, number of medium trucks, number of heavy trucks, number of buses and the average traffic flow speed) for the development of an ANN model for L_{Aeq} prediction in Serbian roads, which outperformed some classical noise prediction models. In order to assess road-traffic-noise in urban environments, Cammarata, Cavalieri, and Fichera (1995), using data collected with typical features of commercial, residential and industrial area, and with number of cars, number of motorcycles, number of trucks, average height of the buildings and width of the road as input variables, proposed a two cascading level neural architecture, where at the first level a learning vector quantification (LVQ) network filters the data discarding all the wrong measurements, while at the second level the BP algorithm predicts the sound pressure level (L_{Aeq}) in urban environments. Genaro et al. (2010) included 25 input variables, which were found as the whole variable set used by all the traditional noise prediction models evaluated. In this work, a MLP model was implemented to predict L_{Aeq} descriptor using data samples from the city of Granada (Spain). Also, a principal component analysis (PCA) was used to simplify the model (up to 11 input variables). This model outperformed the traditional noise prediction models. Torija, Ruiz, and Ramos-Ridao (2012), using a set of variables for the characterization of sound emission and propagation (20 input variables) and 821 samples collected in urban environments (Granada, Spain), developed an ANN model (trained by Levenberg–Marquardt variant with Bayesian regulation back-propagation algorithm) for the estimation of the L_{Aeq} descriptor, but also the estimation of parameters related to the temporal structure and spectral composition of urban sound environments ($L_{31.5-125\ Hz}$, $L_{160-1600\ Hz}$, $L_{2-10\ kHz}$, TSLV and CF). Moreover, a reduction of the input variables (up to 14) based on the analysis of the correlation coefficients and the distribution of the test residuals were performed.

Other applications of ANN in the acoustics field have been related to classification issues. Sánchez-Pérez, Sánchez-Fernández, Suárez-Guerra, and Carbajal-Hernández (2013) developed a model for aircraft classification with an identification performance above 85%. This model was based on the take-off noise signal segmentation (four segments) in time. Once extracted the different aircraft noise patterns, by using Linear Predictive Coding (LPC), the classification was addressed with the implementation of four parallel MLP (one for each segment). Moreover, a wrapper feature selection method was used for reducing the computational cost. Márquez-Molina, Sánchez-Fernández, Suárez-Guerra, and Sánchez-Pérez (2014) developed an aircraft take-off noises classification model. For the obtaining of the input variables, a feature extraction process of aircraft take-off signals was conducted through a 1/24 octave analysis and Mel frequency cepstral coefficients (MFCC), and the classification model was made by using two parallel feed forward neural networks (FFNN), achieving a total effectiveness of 83%. Torija and Ruiz (2012) performed an analysis to identify the 1/3-octave bands most influential on road-traffic intensity. Based on the gathered information, a series of MLP-based model were developed for the estimation of the overall road-traffic intensity and for the detection of conditions with percentage of heavy vehicles or motorbikes/mopeds larger than the usual values.

Although SVM algorithms have not been as extensively used in noise-related issues as ANN, some interesting applications could be highlighted. Barkana and Uzkent (2011) presented two stages classification method for the automatic recognition of environmental noises, where first, a feature extraction based on the pitch range was conducted, and second, SVM and k-means algorithms as classification techniques were trained on the extracted features. SVM classifier outperformed k-means by about 7%. Based on a previous study (Torija, Ruiz, & Ramos-Ridao, 2013) on the differentiation of urban soundscapes as a function of 14 acoustical descriptors and 15 semantic differential scales, Torija, Ruiz, and Ramos-Ridao (2014) implemented two techniques, SVM and SVM trained using sequential minimal optimization algorithm (SMO), for the development of a model for the classification of urban soundscapes (using the same 14 acoustical descriptors as input variables). According to

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

the results showed, SMO model (91.3% of instances correctly classified) outperformed SVM. Finally, in Torija and Ruiz (2015) is presented a comparative analysis of the performance of multiple linear regression (MLR), MLP, SMO and Gaussian processes for regression (GPR) algorithms in the estimation of L_{Aeq} in urban environments. Also, the performance of two feature-selection techniques, correlation-based feature-subset selection (CFS) and wrapper for feature-subset selection (WFS), and the data reduction technique PCA is evaluated. The use of WFS along with either SMO or GPR provided the best L_{Aeq} estimation. On the other hand, approaches based on fuzzy logic have been widely used for developing expert systems for assessing noise pollution (Zaheeruddin, 2006, 2008).

1.3. Objective and interest of this work

Information and communications technologies (ITCs) are now available to local authorities for addressing an effective management of urban environments aimed at improving the quality of life of the population. Sensor platforms allow the continuous monitoring of urban noise via inexpensive and highly accurate devices.

The objective of this paper is to exploit these data recorded by noise monitoring systems already available to city planners, to search for an automated procedure for the classification of urban locations according to their content in heavy vehicles and motorbikes/mopeds, which can be extended to a more comprehensive expert system for selecting noise control actions. With this automated classification, it can be suggested a whole procedure to environmental impact assessment of urban noise, as is proposed in this paper.

Given this objective, rules and decisions (knowledge base) are implemented in an automated procedure for classifying urban locations based on their traffic composition using machine learning algorithms to induce knowledge to an expert system. As a result of this research, along with a set of environmental variables selected for the characterization of the urban location, the proposed procedure uses a number of noise recorded metrics as input variables. Thus, once integrated in an expert noise monitoring system, it will allow an automatic classification of urban locations on the basis of their traffic content. For this procedure, several possibilities are suggested and tested in this work, such as several different machine learning classification algorithms or features selection techniques to select the most efficient ones and the selection of input variables.

Unlike the previous approaches briefly summarised in the preceding subsection, the developed classifier was aimed at classifying urban locations based on the content in heavy vehicles and motorbikes/mopeds using a set of environment variables (temporal period, road conditions, speed, and geometry of the locations) along with energy-equivalent sound level and 1/3-octave bands sound levels as input variables. Thus, from data gathered by noise monitoring systems the developed classifier will identify urban scenarios with high number of loud events in traffic. Information on number and source of loud events is a helpful knowledge in order to assess environmental noise impact and to define corrective measures. The number of loud events has been found to play an important role in explaining road-traffic-noise annoyance (Bartels, Márki, & Müller, 2015; Guski, 1999), so that the classification model presented in this paper might be used for addressing more effective actions in order to reduce noise impact in urban environments.

With the suggested procedure, city-planners can effectively know if a given location in a sensitive area (e.g. residential or hospital area) can be classified as dominated by motorbikes, heavy traffic, light traffic or several mixed traffic conditions. From this classification, they can adopt actions for timely and efficient controlling noise pollution and urban traffic management (Uzkent, Barkana, & Yang, 2011).

This paper is organized as follows. In Section 2 the methodology of this research is shown. In this section the fundamentals of a set of machine learning algorithms models suggested for the classification of urban locations according to their percentage of heavy vehicles (HV) and motorbikes/mopeds (MM) in urban traffic are described and adapted for the context of the problem issued here. Four feature selection techniques are also implemented for the development of the classification models. For the evaluation of the classification performance of the developed models, two statistic indicators were used, *F*-measure and Cohen's Kappa, and their practical implementation and interpretation is commented. Next, in Section 3, the different algorithms and methods were evaluated to suggest a final suggested procedure for classification. To accomplish this, we use a wide noise database measured in the city of Granada (Spain) previously tested in many studies. Thus, in this section it is firstly tested the classification performance of machine-learning algorithms based on the indicators defined in the previous section. In a second stage different combinations of classification algorithms and feature extraction techniques were implemented and tested. Finally in this section several statistical tests were used to evaluate the appearance of statistically significant differences among the developed models, based on the *F*-measure and the Kappa statistics. Taking into account these results, in Section 4 is given a discussion on the results obtained and it is suggested the 'optimal' structure (high accuracy and minimum computational/operational cost) for the developed classification model, as well as the suggested set of input variables to be used. From this discussion, a whole procedure is suggested for environmental noise impact assessment to help urban planners in this task. Finally some conclusions are driven in Section 5 to show the potential uses of the outlined procedure.

2. Methodology

As an application of machine learning algorithms to environmental modeling, in this paper a series of models were developed for the classification of urban locations according to their percentage of heavy vehicles (HV) and motorbikes/mopeds (MM) in traffic. These models were built on the basis of a series of recorded sound parameters and environment variables for the characterization of the temporal period and both the sound emission and propagation (Torija et al., 2010). A hierarchical cluster analysis (HCA) was conducted in order to group the urban locations considered in classes as a function of the HV and MM intensity. For the development of the classification models, four feature selection techniques – CFS, Consistency-based subset evaluation (CSE), ReliefF attribute evaluator (ReliefF), and minimum Redundancy Maximum Relevance (mRMR) – and two classification algorithms – MPL and SMO – were implemented.

2.1. Database

For the development of the classification models, a database of 508 instances was used. This database, which includes a series of urban locations with road-traffic as dominant noise source, was collected in the city of Granada (Spain). In each location, sound-level recordings and data for the road-traffic, temporal and geometrical characterization were taken at the same time. Table 1 shows the set of variables considered in this research. The time interval for the integration of the different sound parameters and other dynamic variables was 5 min, so that this research is framed in short-term modeling. The sound measurements were made with a type-1 sound-level meter (2260 Observer model with sound basic analysis programme BZ7219), using a tripod and wind shield, following international reference procedures, with the microphone mounted away from reflecting facades at a height of 4 m above

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

Table 1
Set of variables used for the development of classification models.

	Key	Variable	Type of variable	Units
Input variables	TD	Type of day (working day, Saturday or Sunday)	Discrete	–
	DP	Day period (day, evening or night)	Discrete	–
	AS	Average speed	Continuous	–
	GR	Gradient	Continuous	%
	NL	Number of lanes (1 to 4)	Discrete	–
	TP	Type of pavement (porous asphalt, smooth asphalt or paved)	Discrete	–
	CS	Condition surface (good, neither good nor bad or bad)	Discrete	–
	SG	Street geometry (“U” type, “L” type or free field)	Discrete	–
	SW	Street width	Continuous	m
	BH	Buildings height	Continuous	m
	RW	Roadway width	Continuous	m
	SRD	Source–receptor distance	Continuous	m
		$L_{eq,\ 5\ min}$	Continuous	dB
		$L_{f,\ 5\ min}$: f [31.5–10,000] Hz	Continuous	dB
Variables used for obtaining target categories	LV	Light vehicles intensity	Continuous	veh/5-min
	HV	Heavy vehicles intensity	Continuous	veh/5-min
	MM	Motorbike/Mopeds intensity	Continuous	veh/5-min

the local ground level. From these measurements, the 5-min energy equivalent sound level ($L_{eq,\ 5\ min}$) and the 5-min integrated 1/3-octave band sound levels ($L_{f,\ 5\ min}$) from 31.5 Hz to 10,000 Hz, were calculated to be included in the input variables set.

2.2. Clustering of selected locations

A HCA was applied to the set of selected urban locations (508 instances). The clustering was made by the Ward method and the squared Euclidean distance was the measurement unit. The input variables were light vehicles (LV), heavy vehicles (HV) and motorcycles/mopeds (MM) intensities, so the clustering was performed on the basis of the road-traffic characteristics. The determination of the number of clusters was based on the L method (Salvador & Chan, 2004), that finds the “knee” in a ‘number of clusters vs. clustering evaluation metric’ graph.

2.3. Machine learning classification algorithms

2.3.1. Multi-layer perceptron

MLP is an ANN architecture widely used in classification problems. A MLP consist of three layers: the input layer (whose nodes take input variables), the hidden layer (could have one or more hidden layers) and the output layer. The hidden nodes compute its output by:

$$O_j = h\left(\sum_{i=1}^{n} w_{ji}^1 x_i - \theta_j^1\right) \tag{1}$$

where O_j is the output of jth node in the hidden layer; $h(\cdot)$ is the transfer function from the input to the hidden layer; w_{ji}^1 is the weight between the ith node of the input layer and the jth node of the hidden layer; θ_j^1 is the bias value of the jth node in the hidden layer.

The outputs of the output layer are computed as follows:

$$z_k = g\left(\sum_{j=1}^{l} w_{kj}^2 O_j - \theta_k^2\right) \tag{2}$$

where z_k is the output of the kth node in the output layer; $g(\cdot)$ is the transfer function from the hidden to the output layer; w_{kj}^2 is the weight between the jth node of the hidden layer and the k-h node of the output layer; θ_k^2 is the bias value of the kth node in the output layer (Feng et al., 2015). In this work, all the transfer functions were sigmoid.

Given a training set $\mathfrak{D} = \{x_i, y_i\}_{i=1}^N, y_i \in \mathbb{R}$, the training error is minimized during the training process using the mean squared error the output (z) calculated by the network and the real one (y) as error function (E):

$$E(w) = \frac{1}{N}\sum_{k=1}^{N}(y_k - z_k)^2 \tag{3}$$

where N is the number of data points. For the optimal derivation of the weights for the MLP a BP algorithm was used, which updated the weights iteratively to minimize the error function (Kang & Cho, 2014). For a comprehensive description of MLP see Haykin (1999).

2.3.2. Sequential minimal optimization for support vector machine training

SVM algorithms are based on the structural risk minimization principle (Kang & Cho, 2014), which allow them to achieve superior generalization performance for classification problems (Burges, 1998; Vapnik, 1995, 1998).

Let a set of N training datapoints $\mathfrak{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^N$, where $\boldsymbol{x}_i$ is the ith input feature vector and $y_i \in \{-1, 1\}$ is the corresponding output class. The implementation of SVM searches the maximum margin hyperplane $\boldsymbol{w}^T\varphi(\boldsymbol{x}) + b = 0$ that separates the positive datapoints from the negative datapoints. Formulating the problem as a primal optimization, the following minimization is sought:

$$\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_i \xi_i \tag{4}$$

subject to

$$y_i(\boldsymbol{w}^T\varphi(\boldsymbol{x}_i) + b) \geq 1 - \xi_i,$$
$$\xi_i \geq 0, i = 1, \ldots, N,$$

where C is a penalty parameter that determines the trade-off between the training errors and the model complexity; φ is a non-linear mapping from an input space into a feature space; and ξ_i are the slack variables. This optimization problem is usually converted to the dual form through the following quadratic programming (QP) problem, which aims to maximize:

$$-\frac{1}{2}\sum_{ij}\alpha_i\alpha_j y_i y_j k(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sum_i \alpha_i \tag{5}$$

subject to

$$\sum_i \alpha_i y_i = 0,$$
$$0 \leq \alpha_i \leq C, i = 1, \ldots, N,$$

where α_i are Lagrange multipliers and $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is a kernel function. The resulting decision function, after the dual QP problem is solved, can be expressed as:

$$f(\boldsymbol{x}) = \boldsymbol{w}^T\varphi(\boldsymbol{x}) + b = \sum_{i=1}^{N}\alpha_i y_i k(\boldsymbol{x}_i, \boldsymbol{x}) + b$$
$$= \sum_{i \in SV}\alpha_i y_i k(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{6}$$

This decision function is expressed by only training datapoints with only nonzero $\propto_i$, which are called support vectors (SVs) (Feng et al., 2015; Kang & Cho, 2014). A comprehensive description of SVM can be found in Burges (1998) and Vapnik (1995, 1998).

In order to improve the efficiency of QP, Platt (1998) proposed SMO. SMO is a simple algorithm that decomposes the overall QP problem into QP sub-problems similar to Osuna's method. During the solution of the SVM QP problem, at every step, SMO chooses the Lagrange multipliers to jointly optimize, finds the optimal values for these multipliers, and updates the SVM to reflect the new optimal values (see Platt, 1998 for further details).

In this work, three kernel functions were used:

Polynomial (PN):

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i * \boldsymbol{x}_j + 1)^p \quad \text{with } p = 2 \textit{ in this work} \tag{7}$$

Radial basis function (RBF):

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp(\gamma \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2) \tag{8}$$

Pearson VII kernel function (PuK):

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = 1/[1 + (2\sqrt{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}\sqrt{2^{(1/\omega)} - 1}/\sigma)^2]^{\omega} \tag{9}$$

where γ, σ and ω are kernel parameters in the SVM feature space (Anifowose et al., 2015).

2.4. Feature extraction algorithms

For the selection of input variables two different approaches have been used, (i) subset evaluation (CFS and CSE), and (ii) attribute evaluation (ReliefF and mRMR).

2.4.1. Correlation-based feature-subset selection

CFS algorithm computes a metric based on the correlation between each feature and the output (relevancy) and on the correlation among the features in the subset (redundancy). This metric, which evaluates the merit of a given subset of features, is calculated as follows:

$$G_S = \frac{k\overline{r_{ci}}}{\sqrt{k + k(k-1)\overline{r_{ii'}}}} \tag{10}$$

where k is the number of features in the subset S; $\overline{r_{ci}}$ is the average correlation between the features in S and the target class; and $\overline{r_{ii'}}$ is the average correlation among the features in S (Hall & Smith, 1997).

Best-first algorithm (BFS) (with backward, forward and bi-directional direction) (Dechter & Pearl, 1985) and Linear forward selection algorithm (LFS) (Guetlein, Frank, Hall, & Karwath, 2009), which is an extension of BFS, were used as search methods.

2.4.2. Consistency-based subset evaluation

CFE evaluates the merit of a given feature subset on the basis of the inconsistency criterion (consistency is interpreted as zero inconsistency). The inconsistency rate is computed as follows (Liu & Setiono, 1996):

(i) Two instances are considered inconsistent if they match except for their class labels.
(ii) For all the matching instances (without considering their class labels), the inconsistency count is the number of the instances minus the largest number of instances among different class labels.
(iii) The inconsistency rate is the sum of all the inconsistency counts divided by the total number of instances.

This criterion, along with BFS and LFS as search strategies, was implemented to find the smallest subset of features with consistency equal to that of the full set of features.

2.4.3. Relief F attribute selection

ReliefF (Kononenko, 1994) is a feature selection strategy that chooses instances randomly, and changes the weights of the feature relevance based on the nearest neighbor. Thus, for a given attribute, ReliefF considers the value for the nearest instance of the same and different class. This algorithm estimates the ability of attributes to separate each pair of classes regardless of which two classes are closest to each other.

In this work, this feature selection strategy is implemented with a number of nearest neighbours to be considered for the attribute estimation equal to 10, 15 and 20. For each of the three conditions, this method was implemented with and without weighting nearest neighbours by their distance.

2.4.4. Minimum redundancy maximum relevance attribute selection

mRMR (Ding & Peng, 2005) is a feature selection strategy that seeks the selection of attributes under two assumptions, minimum redundancy (minRed) and maximum relevance (MaxRel). Minimum redundancy implies the selection of attributes that are mutually maximally dissimilar, so that the extracted feature subset gives a better representation of the entire dataset. The minRed condition can be expressed as follows:

$$\min W_I,\ W_I = \frac{1}{|S|^2}\sum_{i,j\in S} I(i,j), \tag{11}$$

where $I(i, j)$ represents the mutual information of two features; and $|S|$ is the number of features. The level of discrimination between classes is measured by the relevance. The MaxRel condition is to maximize the total relevance of all classes in S:

$$\max V_I,\ V_I = \frac{1}{|S|}\sum_{i,j\in S} I(h,i), \tag{12}$$

where h is the targeted class. Both conditions are combined in order to optimize a single criterion function.

Mutual information difference criterion (MID):

$$max(V_I - W_I) \tag{13}$$

Mutual information quotient criterion (MIQ):

$$max(V_I, /W_I) \tag{14}$$

In this work, the subsets of features derived from MaxRel, MID and MIQ criteria were evaluated.

It should be noted that neither ReliefF nor mRMR strategies lead to a selection of a feature subset (with a given number of features), but they rank attributes according to a given metric. For this reason, the number of features to be selected using these strategies was fixed to the number of features selected by CFS and CSE techniques. Table 2 summarizes all the feature extraction algorithms implemented in each machine-learning classification technique.

2.5. Model evaluation

For the evaluation of the classification performance of the models developed, two statistic indicators were used, *F*-measure and Cohen's Kappa. The *F*-measure provides a way of combining recall and prediction to get a single measure which falls between recall and precision. Thus, the *F*-measure is calculated as the harmonic mean of precision and recall and tends towards the lower of the two (Chinchor, 1992):

$$F - measure = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{1'}$$

Note that precision can be expressed as the ratio between the true positives (TP) and all the cases classified as positive, and recall represents the ratio between the TP and all the positive cases.

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

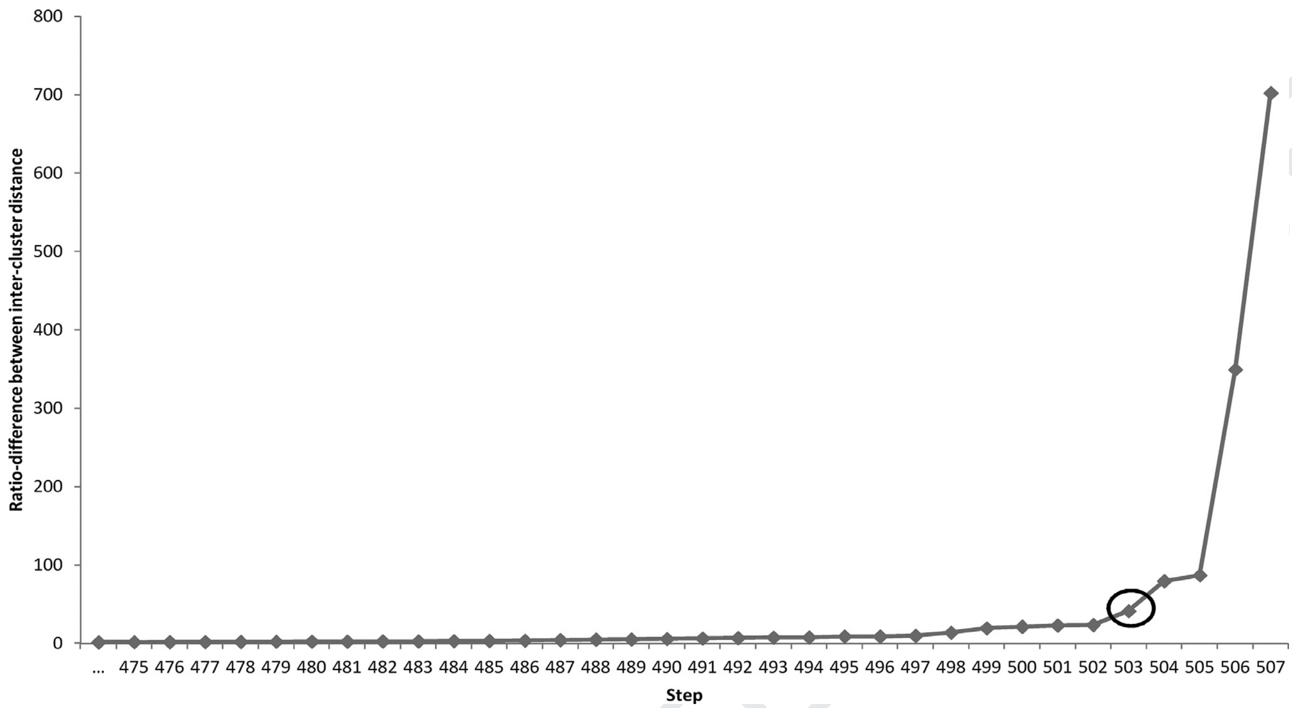

Fig. 1. Ratio difference between inter-clusters distance in the hierarchical cluster analysis performed with light vehicles, heavy vehicles and motorbikes/mopeds as input variables.

Table 2
Description of the feature extraction algorithms implemented.

Feature extraction algorithm	Parameters	Key
CFS	Best-first with backward search direction	CFS_Bw
CFS	Best-first with forward search direction	CFS_Fw
CFS	Best-first with bi-directional search direction	CFS_Bi
CFS	Linear forward selection algorithm	CFS_LFS
CFE	Best-first with backward search direction	CFE_Bw
CFE	Best-first with forward search direction	CFE_Fw
CFE	Best-first with bi-directional search direction	CFE_Bi
CFE	Linear forward selection algorithm	CFE_LFS
ReliefF	10 nearest neighbors	ReliefF_k10
ReliefF	10 nearest neighbours. Nearest neighbors weighted by distance.	ReliefF_k10w
ReliefF	15 nearest neighbors	ReliefF_k15
ReliefF	15 nearest neighbors. Nearest neighbours weighted by distance.	ReliefF_k15w
ReliefF	20 nearest neighbors	ReliefF_k20
ReliefF	20 nearest neighbors. Nearest neighbors weighted by distance.	ReliefF_k20w
mRMR	Maximum relevance criterion	mRMR_MaxRel
mRMR	Mutual information difference criterion	mRMR_MID
mRMR	Mutual information quotient criterion	mRMR_MIQ

Table 3
Overall and road-traffic intensity for each considered type of road vehicle. Moreover, the average L_{Aeq} value (integration time = 5 min) for each as measured for each cluster is shown.

Overall road-traffic intensity (veh/5-min)	LV	HV	MM	$L_{Aeq,\ 5\ min}$	Category
27.79	22.29 (80.21)	0.32 (1.16)	5.18 (18.64)	65.62	1
75.34	61.72 (81.92)	1.87 (2.48)	11.76 (15.60)	68.49	2
200.30	160.68 (80.22)	3.73 (1.86)	35.89 (17.92)	71.81	3
88.66	34.69 (39.13)	11.77 (13.27)	42.20 (47.60)	74.74	4

*In brackets it is shown the percentage of LV, HV and MM vehicles relative to the overall intensity.

The kappa statistic is directly interpretable as the proportion of agreement after chance agreement is excluded, and it is calculated as follows (Cohen, 1960):

$$kappa = \frac{P_0 - P_e}{1 - P_e} \tag{2'}$$

where P_0 is the observed proportion of agreement and P_e is the proportion of agreement expected by chance.

In the development of the classification models, a training-validation process was executed to minimize the estimation error in the training subsets and to maximize the generalization ability (in the test subsets). In order to avoid overfitting, the parameters related to the complexity of MLP (number of neurons in the hidden layer, learning rate, momentum) and SMO (C parameter) algorithms were carefully selected. Furthermore, the training process was carried out by using a 10-fold cross-validation standard scheme, where 10 training and 10 validation subsets were built. In each subset, 90% of samples were used in the training phase and the 10% of samples were used for validation. The values of the *F*-measure and kappa indicators were calculated as the arithmetic mean of the 10 validation subsets.

3. Results

3.1. Number of categories based on road-traffic content

Based on the results of the HCA, 4 categories have been identified according to the percentage of HV and MM in the evaluated urban locations. In Fig. 1, it is seen that in the step 503 there is a sudden increment in the value of the ratio-difference between the inter-cluster distance, which indicates that 4 categories can be determined. In Table 3, the value of road-traffic intensities and the corresponding percentage can be observed. Also, as seen in Table 3, no matter the overall road-traffic intensity, the increase

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

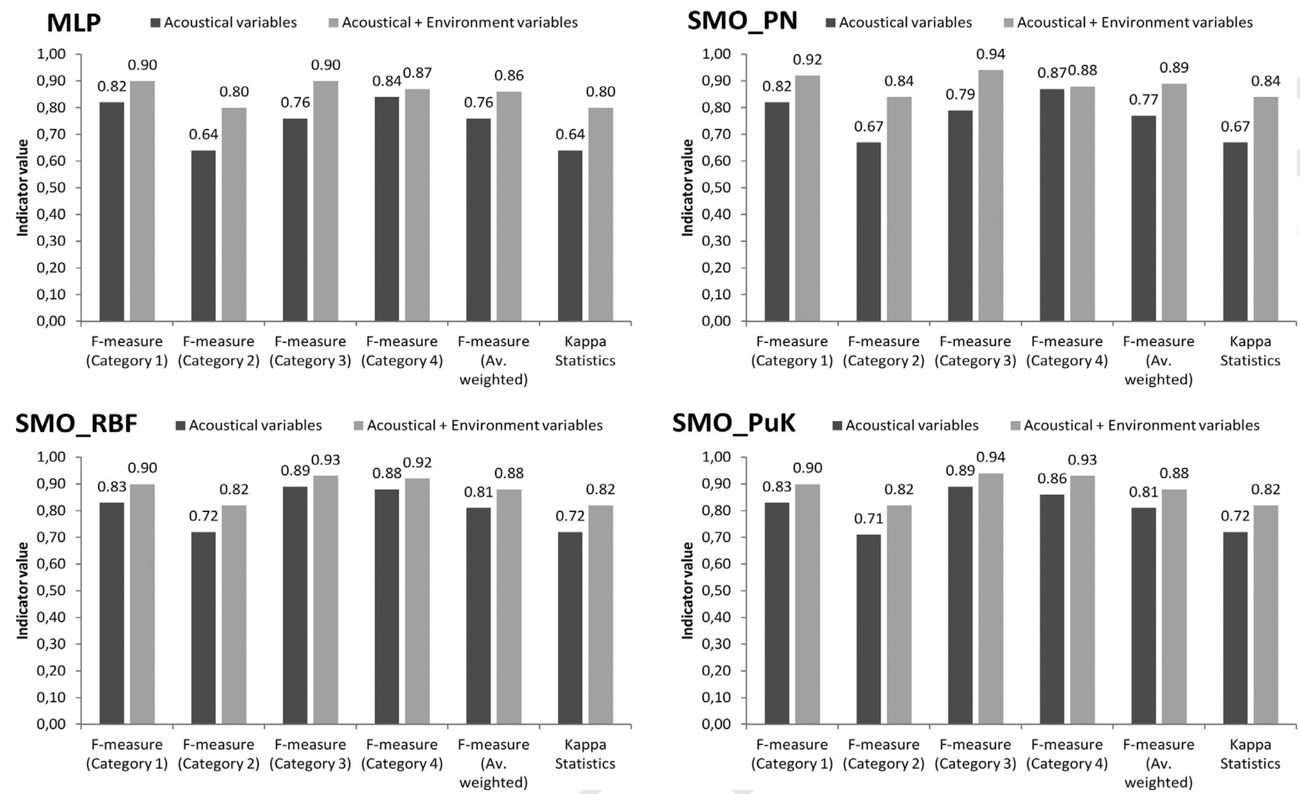

Fig. 2. Value of the *F*-measure indicator (weighted average and for categories 1, 2, 3 and 4) and Kappa statistics for each machine-learning classification algorithm with the whole set of input variables.

in the $L_{Aeq,\,5\,min}$ is driven by the increment of the HV and MM content.

3.2. Classification performance of machine-learning algorithms evaluated

Fig. 2 shows the performance (*F*-measure and Kappa statistics) of each machine-learning algorithm implemented (MLP, SMO_PN, SMO_RBF and SMO_PuK) in classifying the considered urban locations into the 4 categories observed. As seen in Fig. 2, using the measured spectra ($L_{f,\,5\,min}$) as input variables, values of average weighted *F*-measure and Kappa statistics ranging from 0.76–0.81 and 0.64–0.72 are obtained. The algorithms with the best performance are SMO_RBF and SMO_PuK, while MLP achieves the lowest classification values.

On the other hand, with the inclusion of the environment variables mentioned in Table 1 (temporal period, road surface, locations' geometry, circulation speed and gradient) the average value of the *F*-measure indicator increases between 8% (SMO_RBF) and 15% (SMO_PN). The average value of the Kappa indicator grows between 14% (SMO_RBF and SMO_PuK) and 25% (MLP and SMO_PN).

3.3. Evaluation of the feature extraction algorithms implemented for classification

In Table 4, the set of input variables selected by each feature selection algorithm is presented. The input variables are ordered by the value of the merit function in each feature selection technique. CFS algorithm selected the same subset of 13 input variables, with the 4 search methods considered. Regarding CFE algorithm, with Backward and Forward direction search methods 13 input variables were selected, while Bi-directional and LFS search methods selected subsets composed of 12 and 11 input variables, respectively. It should be noted that, for attribute evaluation algorithms (ReliefF and mRMR) a number of 13 input variables to be selected was fixed in order to ensure comparability among feature selection techniques.

Fig. 3 indicates that SMO algorithms outperformed MLP in classifying urban locations, regardless the feature selection technique implemented. On average, the feature selection techniques with the best performance are ReliefF_k10, ReliefF_k15, ReliefF_k20 and mRMR_MIQ. The highest values for both *F*-measure and Kappa statistics are achieved by SMO_PN with ReliefF_k10, ReliefF_k15, ReliefF_k20 as feature selection techniques. Moreover, as observed in Fig. 3, the reduction of input variables from the initial set of 40 to subsets of 11, 12 and 13 variables achieves by the feature selection algorithms implemented does not lead to a decrease in the classification performance. Thus, with the subset of input variables selected by the used feature selection techniques, the classification algorithms obtain values of *F*-measure and Kappa indicators in the same order of magnitude as with the total set of input variables, with the sole exception of CFE_Bw. Although there is not a clear tendency, it seems that attribute evaluation algorithms allows better classification performance. Table 5 shows the value of *F*-measure obtained by each classification algorithm with each category.

3.4. Statistical tests

A series of Mann–Whitney *U* tests were conducted to evaluate the appearance of statistically significant differences among the developed models, based on the *F*-measure (weighted average and categories 1, 2, 3 and 4), and the Kappa statistics. These statistical tests were performed to assess statistically significant differences ($p \leq 0.05$) among classification algorithms (Figs. 4 and 5), and among feature selection algorithms (Figs. 6 and 7). It should

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

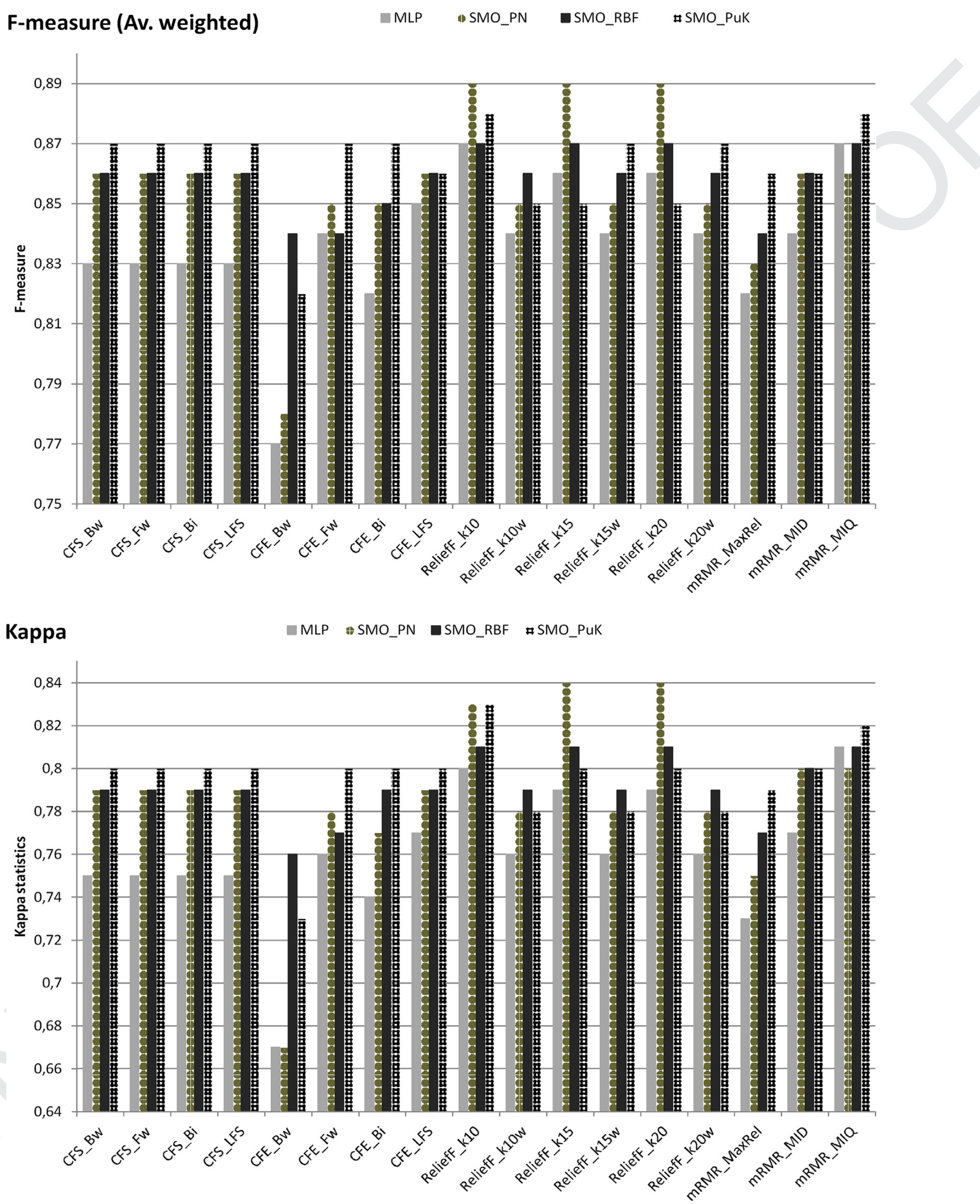

Fig. 3. Value of the weighted average *F*-measure indicator and Kappa statistics for each machine-learning classification algorithm and each feature selection technique.

be noted that the models were ordered on the basis of their median value.

As shown in Fig. 4 (*F*-measure) and Fig. 5 (Kappa statistics), SMO classification models significantly outperformed MLP models. In Figs. 6 and 7, it is observed that the feature algorithms with the best performance are ReliefF_k15, ReliefF_k20, mRMR_MIQ and ReliefF_k10. As for *F*-measure (Fig. 6), CFE_Bw algorithm statistically obtains the worst values, while mRMR_MIQ and ReliefF_k10 algorithms achieves similar values, and significantly outperformed CFE_Bw, mRMR_MaxRel and CFS algorithms. Regarding Kappa statistics (Fig. 7), mRMR_MIQ algorithm significantly improves CFE_Bw, mRMR_MaxRel, ReliefF_k10w, ReliefF_k15w and ReliefF_k20w, while ReliefF_k10 significantly outperformed all the feature selection algorithms, with the exception of CFE_Bi, ReliefF_k15, ReliefF_k20 and mRMR_MIQ.

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

	Median	MLP	SMO_PN	SMO_RBF	SMO_PuK
MLP	0.85		1	1	1
SMO_PN	0.87	1		0	0
SMO_RBF	0.88	1	0		0
SMO_PuK	0.88	1	0	0	

Fig. 4. Results of the Mann–Whitney *U* test for classification algorithms (*F*-measure). 1 = statistically significant difference and 0 = not statistically significant difference. ($p \leq 0.05$).

	Median	MLP	SMO_PN	SMO_RBF	SMO_PuK
MLP	0.76		1	1	1
SMO_PN	0.79	1		0	0
SMO_RBF	0.79	1	0		0
SMO_PuK	0.80	1	0	0	

Fig. 5. Results of the Mann–Whitney *U* test for classification algorithms (Kappa statistics). 1 = statistically significant difference and 0 = not statistically significant difference. ($p \leq 0.05$).

4. Discussion

The results obtained in this work, demonstrate that SMO models outperform MLP model in classifying the set of urban locations sampled into the corresponding category, which is based on the composition in HV and MM. Similar findings have been reached by several authors (Kang & Cho, 2014; Tan et al., 2012; Torija & Ruiz, 2015; Zeng & Qiao, 2013), pointing out SVM as the machine-learning method with the highest classification performance. SVM algorithm is based on structure risk minimization principle whereas ANN is based on empirical risk minimization principle. Thus, while SVM seeks to minimize the upper bound of a generalization error, ANN aims to minimize false classification error. Due to this principle, SVM is able to fix the overfitting problem

	Median	CFE_Bw	mRMR_MaxRel	CFS_Bw	CFS_Bi	CFS_Fw	CFS_LFS	ReliefF_k10w	ReliefF_k15w	ReliefF_k20w	CFE_Bi	CFE_Fw	mRMR_MID	CFE_LFS	ReliefF_k15	ReliefF_k20	mRMR_MIQ	ReliefF_k10
CFE_Bw	0.81		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
mRMR_MaxRel	0.86	1		0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
CFS_Bw	0.86	1	0		0	0	0	0	0	0	0	0	0	0	0	0	1	1
CFS_Bi	0.86	1	0	0		0	0	0	0	0	0	0	0	0	0	0	1	1
CFS_Fw	0.86	1	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1
CFS_LFS	0.86	1	0	0	0	0		0	0	0	0	0	0	0	0	0	1	1
ReliefF_k10w	0.87	1	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0
ReliefF_k15w	0.87	1	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
ReliefF_k20w	0.87	1	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
CFE_Bi	0.87	1	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0
CFE_Fw	0.87	1	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0
mRMR_MID	0.88	1	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0
CFE_LFS	0.88	1	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0
ReliefF_k15	0.88	1	1	0	0	0	0	0	0	0	0	0	0	0		0	0	0
ReliefF_k20	0.89	1	1	0	0	0	0	0	0	0	0	0	0	0	0		0	0
mRMR_MIQ	0.89	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0		0
ReliefF_k10	0.90	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	

Fig. 6. Results of the Mann–Whitney *U* test for feature selection algorithms (*F*-measure). 1 = statistically significant difference and 0 = not statistically significant difference. ($p \leq 0.05$).

	Median	CFE_Bw	mRMR_MaxRel	CFE_Bi	ReliefF_k10w	ReliefF_k15w	ReliefF_k12w	CFE_Fw	CFS_Bw	CFS_Bi	CFS_Fw	CFS_LFS	CFE_LFS	mRMR_MID	ReliefF_k15	ReliefF_k20	mRMR_MIQ	ReliefF_k10
CFE_Bw	0.70		0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
mRMR_MaxRel	0.76	0		0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
CFE_Bi	0.77	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0
ReliefF_k10w	0.78	0	0	0		0	0	0	0	0	0	0	0	0	0	0	1	1
ReliefF_k15w	0.78	0	0	0	0		0	0	0	0	0	0	0	0	0	0	1	1
ReliefF_k12w	0.78	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	1
CFE_Fw	0.78	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1
CFS_Bw	0.79	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	1
CFS_Bi	0.79	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	1
CFS_Fw	0.79	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	1
CFS_LFS	0.79	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	1
CFE_LFS	0.79	1	0	0	0	0	0	0	0	0	0	0		0	0	0	0	1
mRMR_MID	0.80	1	0	0	0	0	0	0	0	0	0	0	0		0	0	0	1
ReliefF_k15	0.81	1	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0
ReliefF_k20	0.81	1	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0
mRMR_MIQ	0.81	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0		0
ReliefF_k10	0.82	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0	

Fig. 7. Results of the Mann–Whitney *U* test for feature selection algorithms (Kappa statistics). 1 = statistically significant difference and 0 = not statistically significant difference. ($p \leq 0.05$).

Table 4
Attributes selected by each feature selection algorithm implemented.

CFS_ Bw	CFS_ Fw	CFS_ Bi	CFS_ LFS	CFE_ Bw	CFE_ Fw	CFE_ Bi	CFE_ LFS	ReliefF_ k10	ReliefF_k 10w	ReliefF_ k15	ReliefF_k 15w	ReliefF_ k20	ReliefF_k 20w	mRMR_Ma xRel	mRMR_ MID	mRMR_ MIQ
AS	AS	AS	AS	SRD	AS	AS	GR	NL	NL	NL	NL	NL	NL	NL	NL	DP
NL	NL	NL	NL	$L_{\text{40 Hz, 5 min}}$	RW	RW	RW	RW	AS	RW	AS	RW	AS	RW	TP	GR
SW	SW	SW	SW	$L_{\text{63 Hz, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{eq, 5 min}}$	AS	CS	SW	CS	SW	CS	$L_{\text{eq, 5 min}}$	RW	NL
RW	RW	RW	RW	$L_{\text{160 Hz, 5 min}}$	$L_{\text{40 Hz, 5 min}}$	$L_{\text{40 Hz, 5 min}}$	$L_{\text{40 Hz, 5 min}}$	SW	SG	SRD	SG	SRD	SG	$L_{\text{80 Hz, 5 min}}$	SRD	TP
SRD	SRD	SRD	SRD	$L_{\text{200 Hz, 5 min}}$	$L_{\text{50 Hz, 5 min}}$	$L_{\text{50 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	SG	BH	SG	BH	SG	BH	$L_{\text{100 Hz, 5 min}}$	$L_{\text{eq, 5 min}}$	SW
$L_{\text{eq, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{160 Hz, 5 min}}$	BH	RW	AS	RW	AS	RW	$L_{\text{125 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	RW
$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{1.25 kHz, 5 min}}$	$L_{\text{160 Hz, 5 min}}$	$L_{\text{160 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	SRD	DP	BH	DP	CS	DP	$L_{\text{160 Hz, 5 min}}$	$L_{\text{40 Hz, 5 min}}$	$L_{\text{eq, 5 min}}$
$L_{\text{63 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	$L_{\text{2kHz, 5min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{500 Hz, 5 min}}$	CS	SW	CS	SW	BH	SW	$L_{\text{200 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$
$L_{\text{80 Hz, 5 min}}$	$L_{\text{80 Hz, 5 min}}$	$L_{\text{80 Hz, 5 min}}$	$L_{\text{80 Hz, 5 min}}$	$L_{\text{3.15 kHz, 5 min}}$	$L_{\text{315 Hz, 5 min}}$	$L_{\text{800 Hz, 5 min}}$	$L_{\text{1.6 kHz, 5 min}}$	TP	SRD	TP	SRD	TP	SRD	$L_{\text{250 Hz, 5 min}}$	$L_{\text{80 Hz, 5 min}}$	$L_{\text{40 Hz, 5 min}}$
$L_{\text{125 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{5 kHz, 5 min}}$	$L_{\text{800 Hz, 5 min}}$	$L_{\text{2.5 kHz, 5 min}}$	$L_{\text{2.5 kHz, 5 min}}$	DP	TP	DP	TP	$L_{\text{315 Hz, 5 min}}$	TP	$L_{\text{315 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$	$L_{\text{80 Hz, 5 min}}$
$L_{\text{200 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{6.3 kHz, 5 min}}$	$L_{\text{2.5 kHz, 5 min}}$	$L_{\text{4 kHz, 5 min}}$	$L_{\text{10 kHz, 5 min}}$	GR	GR	$L_{\text{315 Hz, 5 min}}$	GR	$L_{\text{eq, 5 min}}$	GR	$L_{\text{400 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{125 Hz, 5 min}}$
$L_{\text{250 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{8 kHz, 5 min}}$	$L_{\text{4 kHz, 5 min}}$	$L_{\text{10 kHz, 5 min}}$	-	$L_{\text{eq, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{eq, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{31.5 Hz, 5 min}}$	$L_{\text{500 Hz, 5 min}}$	$L_{\text{250 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$
$L_{\text{315 Hz, 5 min}}$	$L_{\text{315 Hz, 5 min}}$	$L_{\text{315 Hz, 5 min}}$	$L_{\text{315 Hz, 5 min}}$	$L_{\text{10 kHz, 5 min}}$	$L_{\text{10 kHz, 5 min}}$	-	-	$L_{\text{315 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	$L_{\text{200 Hz, 5 min}}$	$L_{\text{63 Hz, 5 min}}$	DP	$L_{\text{63 Hz, 5 min}}$	$L_{\text{630 Hz, 5 min}}$	$L_{\text{400 Hz, 5 min}}$	$L_{\text{400 Hz, 5 min}}$

Table 5
Classification performance (*F*-measure) of each machine-learning algorithm implemented for categories 1, 2, 3 and 4.

	Category 1				Category 2				Category 3				Category 4			
	MLP	SMO_PN	SMO_RBF	SMO_PuK	MLP	SMO_PN	SMO_RBF	SMO_PuK	MLP	SMO_PN	SMO_RBF	SMO_PuK	MLP	SMO_PN	SMO_RBF	SMO_PuK
CFS_Bw	0.89	0.90	0.89	0.90	0.76	0.78	0.78	0.80	0.81	0.86	0.94	0.92	0.85	0.87	0.87	0.87
CFS_Fw	0.89	0.90	0.89	0.90	0.76	0.78	0.78	0.80	0.81	0.86	0.94	0.92	0.85	0.87	0.87	0.87
CFS_Bi	0.89	0.90	0.89	0.90	0.76	0.78	0.78	0.80	0.81	0.86	0.94	0.92	0.85	0.87	0.87	0.87
CFS_LFS	0.89	0.90	0.89	0.90	0.76	0.78	0.78	0.80	0.81	0.86	0.94	0.92	0.85	0.87	0.87	0.87
CFE_Bw	0.86	0.87	0.88	0.85	0.66	0.68	0.76	0.73	0.73	0.74	0.89	0.89	0.80	0.75	0.85	0.86
CFE_Fw	0.89	0.88	0.87	0.89	0.75	0.76	0.76	0.79	0.88	0.92	0.92	0.94	0.84	0.89	0.89	0.89
CFE_Bi	0.87	0.88	0.89	0.89	0.73	0.76	0.77	0.79	0.88	0.90	0.92	0.94	0.84	0.89	0.89	0.89
CFE_LFS	0.88	0.90	0.89	0.90	0.76	0.77	0.78	0.79	0.90	0.89	0.91	0.92	0.89	0.89	0.87	0.89
ReliefF_k10	0.90	0.91	0.90	0.90	0.79	0.83	0.81	0.82	0.91	0.93	0.94	0.94	0.88	0.92	0.90	0.91
ReliefF_k10w	0.88	0.87	0.88	0.87	0.76	0.78	0.79	0.77	0.89	0.93	0.95	0.95	0.85	0.87	0.90	0.89
ReliefF_k15	0.89	0.92	0.89	0.87	0.79	0.83	0.80	0.77	0.88	0.92	0.95	0.95	0.86	0.90	0.89	0.89
ReliefF_k15w	0.88	0.87	0.88	0.89	0.76	0.78	0.79	0.80	0.89	0.93	0.95	0.94	0.85	0.87	0.90	0.88
ReliefF_k20	0.90	0.92	0.89	0.87	0.78	0.83	0.80	0.77	0.90	0.92	0.95	0.95	0.86	0.90	0.89	0.89
ReliefF_k20w	0.88	0.87	0.88	0.89	0.76	0.78	0.79	0.80	0.89	0.93	0.95	0.94	0.85	0.87	0.90	0.88
mRMR_MaxRel	0.86	0.87	0.87	0.90	0.75	0.74	0.76	0.79	0.83	0.87	0.90	0.89	0.82	0.88	0.89	0.86
mRMR_MID	0.89	0.89	0.90	0.89	0.76	0.78	0.79	0.78	0.88	0.93	0.93	0.93	0.84	0.90	0.88	0.88
mRMR_MIQ	0.90	0.89	0.89	0.89	0.80	0.79	0.80	0.81	0.91	0.91	0.93	0.92	0.90	0.92	0.94	0.95

inherent in ANN algorithms, and thus, achieves better classification performance (Hur & Lim, 2005).

Regarding the feature selection techniques considered in this work, under similar conditions in reducing the model complexity (11–13 input variables selected), it could be stated that attribute evaluation algorithms (ReliefF and mRMR) achieve better classification results than subset evaluation algorithms (CFS and CFE). The merit function of ReliefF and mRMR ensures a better search across the whole search space, selecting the most influential attributes in discriminating among categories and reducing redundancy. Thus, ReliefF_k10 and mRMR_MIQ allow the best classification results, on the basis of the two statistical indicators used (*F*-measure and Kappa statistics). The subsets of input variables selected by these two techniques are different. ReliefF_k10 selected all the environment variables included in Table 1 (with the sole exception of TD), and $L_{eq,\ 5\ min}$ and $L_{315\ Hz,\ 5\ min}$. As for mRMR_MIQ, besides selecting 6 temporal-, road-, and geometrical-related variables, $L_{eq,\ 5\ min}$ and other 6 low-frequency-bands sound level descriptors are chosen. However, in both cases, environmental variables have the highest values in the merit function. These results, along with the outcomes shown in Fig. 2, point out the set of environment variables considered in this work as highly influential on classifying urban locations as to HV and MM traffic content. In any case, the use of a combination of environmental variables (Table 1) along with $L_{eq,\ 5\ min}$ and low-frequency-bands sound level descriptors as input variables in SMO-based models achieves high performance in classifying urban locations according to the percentage of HV and MM in circulation.

Taking in mind the above considerations, we can suggest a whole procedure for aiding the process decision-making for environmental noise impact assessment:

(i) Use of a sensor platform for continuous noise monitoring. Information on the acoustical descriptors identified in this paper should be gathered for each urban location. An analysis of the urban agglomeration should be performed in order to assess the necessity of extending the noise monitoring system to new urban locations.
(ii) Characterization of the urban location using the set of environment variables selected in this paper (environmental variables (Table 1) along with L_{eq} and low-frequency-bands sound level descriptors)
(iii) Using information from (i) and (ii), the developed model classifies the urban location in one of the four categories found in this paper. We suggest applying SMO-based models described in this paper to obtain an accurate classification.
(iv) Detection of road-traffic-noise related problems. First, the dominant land use in the considered urban location should be identified, i.e. residential, health and education, commercial/leisure or industrial. In Table 6 it is showed the decision matrix on the basis of the dominant land use and the urban location category.
(v) Corrective measures. Table 7 shows a set of proposed corrective measures (Torija et al., 2012) for each urban location category. Although some corrective measures are suggested, a thorough analysis would be required in order to establish the most appropriate corrective measure for each problem detected.

Finally, we should take into account two practical considerations regarding this research and the applicability of the obtained results:

(a) For addressing the classification of urban locations based on road-traffic content, the model developed in this paper uses acoustical descriptors as input data. For this reason, the suggested classifier requires the absent of loud noise sources other than road-traffic-noise. The presence of loud non road-traffic-noise sources could lead to misleading results in the classification process. However, this seems not to be an issue in urban agglomerations due to the significant dominance of road-traffic-noise in urban sound environment. On the other hand, the classifier presented in this paper has been trained and tested for a typical medium-sized Mediterranean city. The application of this classification model to other type of urban agglomeration would require the update (re-train) of the algorithm in order to make it able to learn new urban configurations and/or road-traffic conditions.
(b) The model presented in this paper classifies with high accuracy urban locations based on the road-traffic content. In the relevant literature, there can be found several environmental noise models (Genaro et al., 2010; Givargis & Karimi, 2010 Cammarata et al., 1995; Kumar, Nigam & Kumar, 2014; Nedic et al., 2014; Torija & Ruiz, 2015; Torija et al., 2012) which are aimed at estimating the sound-pressure-level from a number of input variables (mostly related to road-traffic). In this work, the approach is different, and from a set of environment variables (for the characterization of the urban location) and a number of acoustical descriptors (Leq, 5 min and low-frequency-bands sound level) a given urban location is classified in a category based on the road-traffic content. This approach allows the user to gather information about the dominant noise source, but also about the number of loud events (HV and/or MM) which has been pointed out as a key factor in explaining road-traffic-noise annoyance (Bartels et al., 2015; Guski, 1999).

5. Conclusions

This paper examines the use of machine learning methods to induce knowledge in expert noise monitoring systems to obtain a reliable classification of urban areas on the basis of their traffic content. In this context, it proposes several machine learning algorithms and features selection methods adapted to this problem to test their behavior and so suggesting the best alternatives to use it. We have shown the viability of this concept since the application of this classifier can offer valuable information to establish measures against road-traffic-noise.

In environmental applications, it is of great interest to design an expert system aimed to help urban planners to classify urban locations based on their traffic composition and consequently controlling noise pollution. The circulation of heavy vehicles and motorbikes/moped causes an important negative impact on the surrounding environment and on the exposed population (Table 3). In light of the results obtained in this research, the application of machine-learning algorithms achieves high performance in the classification of urban locations into the 4 identified categories on the basis of their content in heavy vehicles and motorbikes/moped. In reference to the best classification algorithms for this problem, although MLP-based models provide good classification results, they were significantly outperformed by the SMO-based classification models ($p \leq 0.05$). Moreover, with the same number of input variables selected, attribute evaluation algorithms obtained better classification performances than subset evaluation algorithms. Thus, the subsets of input variables selected by two ReliefF and mRMR feature selection algorithms (ReliefF_k10 and mRMR_MIQ) reach the highest classification performances (weighted average *F*-measure around 0.88–0.89, and Kappa statistics around 0.82–0.83).

In addition, the set of environment variables considered in this work has been identified as a key factor in the classification of urban location according to traffic content. Along with these environment variables, the low-frequency sound levels and the L_{eq} descriptor are found as influential variables to be considered in this

Please cite this article as: 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 (2016), http://dx.doi.org/10.1016/j.eswa.2016.01.011

Table 6
Decision matrix for action against road-traffic-noise.

Land use	Category 1	Category 2	Category 3	Category 4
Residential	Corrective measures only with complaints from the population	Corrective measures only with complaints from the population	Corrective measures	Corrective measures
Health and Education	Corrective measures only with complaints from the public	Corrective measures	Corrective measures	Corrective measures
Commercial/Leisure	No action	Corrective measures only with complaints from the public	Corrective measures only with complaints from the public	Corrective measures
Industrial	No action	No action	Corrective measures only if legal standards exceeded	Corrective measures only if legal standards exceeded

Table 7
Set of corrective measures suggested for each urban location category based on road-traffic content.

Urban location category	Corrective measures
1	- Development of urban mobility plans. - Pedestrianization of urban locations. - Promoting non-motorized mobility.
2	- Development of urban mobility plans. - Pedestrianization of urban locations. - Promoting non-motorized mobility. - Encourage public transport use.
3	- Fostering the replacement of light vehicles by hybrid/electric cars. - Promoting non-motorized mobility. - Development of urban mobility plans. - Encourage public transport use. - Setting more restrictive speed limits.
4	- Development of urban mobility plans. - Encourage public transport use. - Promoting the replacement of urban buses fleet by hybrid/electric vehicles. - Prohibition or restriction of traffic of heavy vehicles. - Design and planning of new routes for heavy vehicles. - Inclusion of a thorough inspection of acoustic emission within the regular technical inspections programs for motorised vehicles. - Identification (and banning if considered) of motorized vehicles which exceed established acoustic emission limits. - Minimization of slopes in urban roads.

classification problem. Therefore, in terms of the applicability of the presented classification models, there is a need, not only for a description of the sound environment, but also an appropriate characterization of the environment (temporal period, road conditions, speed, and geometry of the locations). With the use of such input variables higher performance in the classification based on traffic content is achieved.

Since the obtained classification results are promising, this work suggests a whole procedure in the discussion section to help urban planners to face this problem. Based on the content in heavy vehicles and motorbikes and the other environmental input variables, the implementation of the model developed in this research allows an accurate automatic classification of the urban locations. In a second stage, the information provided by the implementation of the developed classifier can be used to establish actions to address road-traffic-noise-related problems in urban environments, and thus, reduce both the exposure sound levels and the reported people annoyance.

Acknowledgments

This work is funded by the University of Malaga and the European Commission under the Agreement Grant no. 246550 of the seventh Framework Programme for R & D of the EU, granted within the People Programme, «Co-funding of Regional, National and International Programmes» (COFUND). Moreover, this work is partially supported by the "Campus de Excelencia Internacional BIOTIC Granada" (CIE BioTic) of Spain under project P-CP-27 and by the Ministerio de Economía y Competitividad of Spain under project TEC2012-38883-C02-02.

References

Anifowose, F., Labadin, J., & Abdulraheem, A. (2015). Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines. *Applied Soft Computing, 26*, 483–496.
Babisch, W., Pershagen, G., Selander, H., Gouthuijs, D., Breugelmans, O., Cadum, E., et al. (2013). Noise annoyance – A modifier of the association between noise level and cardiovascular health? *Science of the Total Environment, 452-453*, 50–57.
Babisch, W. (2006). Transportation noise and cardiovascular risk: Updated review and synthesis of epidemiological studies indicate that the evidence has increased. *Noise & Health, 8*, 1–29.
Barkana, B. D., & Uzkent, B. (2011). Environmental noise classifier using a new set of feature parameters based on pitch range. *Applied Acoustics, 72*, 841–848.
Bartels, S., Márki, F., & Müller, U. (2015). The influence of acoustical and non-acoustical factors on short-term annoyance due to aircraft noise in the field – The COSMA study. *Science of the Total Environment, 538*, 834–843.
Blanco, J. C., & Flindell, I. (2011). Property prices in urban areas affected by road traffic noise. *Applied Acoustics, 72*, 133–141.
Braun, M. E., Walsh, S. J., Homer, J. L., & Chuter, R. (2013). Noise source characteristics in the ISO 362 vehicle pass-by noise test: literature review. *Applied Acoustics, 74*, 1241–1265.
Brink, M. (2011). Parameters of well-being and subjective health and their relationship with residential traffic noise exposure — a representative evaluation in Switzerland. *Environment International, 37*, 723–733.
Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining & Knowledge Discovery, 2*, 121–167.
Caciari, T., Rosati, M. V., Casale, T., Loreti, B., Sancini, A., Riservato, R., et al. (2013). Noise-induced hearing loss in workers exposed to urban stressors. *Science of the Total Environment, 463–464*, 302–308.
Calixto, A., Diniz, F. B., & Zannin, P. H. (2003). The statistical modeling of road traffic noise in an urban setting. *Cities, 20*, 23–29.
Cammarata, G., Cavalieri, S., & Fichera, A. (1995). A neural network architecture for noise prediction. *Neural Networks, 8*, 963–973.
Can, A., Dekoninck, L., Rademaker, M., Van Renterghem, T., De Baets, B., & Botteldooren, D. (2011a). Noise measurements as proxies for traffic parameters in monitoring networks. *Science of the Total Environment, 410-411*, 198–204.
Can, A., Rademaker, M., Van Renterghem, T., Mishra, V., Van Poppel, M., Touhafi, A., et al. (2011b). Correlation analysis of noise and ultrafine particle counts in a street canyon. *Science of the Total Environment, 409*, 564–572.
Chinchor, N. (1992). MUC-4 Evaluation metrics. In *Proceedings of the fourth message understanding conference (MUC-4)*, VA, USA (pp. 22–29).
Cho, S., Lim, B., Jung, J., Kim, S., Chae, H., Park, J., et al. (2014). Factors affecting algal blooms in a man-made lake and prediction using an artificial neural network. *Measurement, 53*, 224–233.
Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational & Psychological Measurement, 20*, 37–46.
De Coensel, B., Botteldooren, D., De Muer, T., Berglund, B., Nilsson, M. E., & Lercher, P. (2009). A model for the perception of environmental sound based on notice-events. *Journal of the Acoustical Society of America, 126*, 656–665.
Dechter, R., & Pearl, J. (1985). Generalized best-first search strategies and the optimality of A*. *Journal of the Association for Computing Machinery, 32*, 505–536.
Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics & Computational Biology, 3*, 185–205.
Ekici, B. B. (2014). A least squares support vector machine model for prediction of the next day solar insolation for effective use of PV systems. *Measurement, 50*, 255–262.
Feng, Q., Zhang, J., Zhang, X., & Wen, S. (2015). Proximate analysis based prediction of gross calorific value of coals: a comparison of support vector machine, alternating conditional expectation and artificial neural network. *Fuel Processing Technology, 129*, 120–129.

Foraster, M., Deltell, A., Basagaña, X., Medina-Ramón, M., Aguilera, I., Bouso, L., et al. (2011). Local determinants of road traffic noise levels versus determinants of air pollution levels in a Mediterranean city. *Environmental Research, 111*, 177–183.

Fyhri, A., & Klboe, R. (2009). Road traffic noise, sensitivity, annoyance and self-reported health — a structural equation model exercise. *Environment International, 35*, 91–97.

Genaro, N., Torija, A., Ramos-Ridao, A., Requena, I., Ruiz, D. P., & Zamorano, M. (2010). A neural network based model for urban noise prediction. *Journal of the Acoustical Society of America, 128*, 1738–1746.

Gnana Sheela, K., & Deepa, S. N. (2013). Neural network based hybrid computing model for wind speed prediction. *Neurocomputing, 122*, 425–429.

Givargis, Sh., & Karimi, H. (2010). A basic neural traffic noise prediction model for Tehran's roads. *Journal of Environmental Management, 91*, 2529–2534.

Guetlein, M., Frank, E., Hall, M., & Karwath, A. (2009). Large scale attribute selection using wrappers. In *Proceedings of the ieee symposium on computational intelligence and data mining,* Nashville, USA (pp. 332–339).

Guski, R. (1999). Personal and social variables as co-determinants of noise annoyance. *Noise & Health, 3*, 45–56.

Hájek, P., & Olej, V. (2012). Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. *Ecological Informatics, 12*, 31–42.

Hall, M. A., & Smith, I. A. (1997). Feature subset selection: a correlation based filter approach. In *Proceedings of the international conference on neural information processing and intelligent, information systems* (pp. 855–858). Singapore: Springer.

Haykin, S. (1999). *Neural networks. A comprehensive foundation* (2nd ed.). New York: Prentice Hall.

Hur, Y., & Lim, S. (2005). Customer churning prediction using support vector machines. In Jun Wang, Xiaofeng Liao, & Zhang Yi (Eds.), *Advances in neural networks – ISNN 2005 (Part II)*.

Ising, H., & Kruppa, B. (2004). Health effects caused by noise: evidence in the literature from the past 25 years. *Noise & Health, 6*, 5–13.

Jiang, X., Zhang, L., & Chen, X. (2014). Short-term forecasting of high-speed rail demand: a hybrid approach combining ensemble empirical mode decomposition and gray support vector machine with real-world applications in China. *Transportation Research Part C: Emerging Technologies, 44*, 110–127.

Kang, S., & Cho, S. (2014). Approximating support vector machine with artificial neural network for fast prediction. *Expert Systems with Applications, 41*, 4989–4995.

Kassomenos, P., Vogiatzis, K., & Bento Coelho, J. L. (2014). Critical issues on environmental noise: editorial. *Science of the Total Environment, 482-483*, 399.

Kononenko, I. (1994). Estimating attributes: analysis and extensions of Relief. In *Proceedings of the european conference of machine learning,* Catania, Italy (pp. 171–182).

Kumar, P., Nigam, S. P., & Kumar, N. (2014). Vehicular traffic noise modeling using artificial neural network approach. *Transportation Research Part C: Emerging Technologies, 40*, 111–122.

Lafdani, E. K., Nia, A. M., & Ahmadi, A. (2013). Daily suspended sediment load prediction using artificial neural networks and support vector machines. *Journal of Hydrology, 478*, 50–62.

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., & Fang, D. (2014). Improving rail network velocity: a machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies, 45*, 17–26.

Liu, H., & Setiono, R. (1996). A probabilistic approach to feature selection – A filter solution. In *Proceedings of the 13th international conference on machine learning,* Bari, Italy (pp. 319–327).

Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies, 54*, 187–197.

Márquez-Molina, M., Sánchez-Fernández, L. P., Suárez-Guerra, S., & Sánchez-Pérez, L. A. (2014). Aircraft take-off noises classification based on human auditory's matched features extraction. *Applied Acoustics, 84*, 83–90.

Mena, R., Rodríguez, F., Castilla, M., & Arahal, M. R. (2014). A prediction model based on neural networks for the energy consumption of a bioclimatic building. *Energy & Buildings, 82*, 142–155.

Mercer, A., Dyer, J., & Zhang, S. (2013). Warm-season thermodynamically-driven rainfall prediction with support vector machines. *Procedia Computer Science, 20*, 128–133.

Muzet, A. (2007). Environmental noise, sleep and health. *Sleep Medicine Reviews, 11*, 135–142.

Nedic, V., Despotovic, D., Cvetanovic, S., Despotovic, M., & Babic, S. (2014). Comparison of classical statistical methods and artificial neural network in traffic noise prediction. *Environmental Impact Assessment Review, 49*, 24–30.

Paviotti, M., & Vogiatzis, K. (2012). On the outdoor annoyance from scooter and motorbike noise in the urban environment. *Science of the Total Environment, 430*, 223–230.

Pirrera, S., De Valck, E., & Cluydts, R. (2010). Nocturnal road traffic noise: a review on its assessment and consequences on sleep and health. *Environment International, 36*, 492–498.

Platt, J.C. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines. Technical Report MSM-RT-98-14 Microsoft Research.

Salvador, S., & Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *Proceedings of the 16th IEEE international conference on tools with artificial intelligent (ICTAI04),* Florida, USA (pp. 576–584).

Sánchez-Pérez, L. A., Sánchez-Fernández, L. P., Suárez-Guerra, S., & Carbajal-Hernández, J. J. (2013). Aircraft class identification based on take-off noise signal segmentation in time. *Expert Systems with Applications, 40*, 5148–5159.

Tan, G., Yan, J., Gao, C., & Yang, S. (2012). Prediction of water quality time series data based on least squares support vector machine. *Procedia Engineering, 31*, 1194–1199.

Torija, A. J., & Ruiz, D. P. (2015). A general procedure to generate models for urban environmental-noise pollution using feature selection and machine learning methods. *Science of the Total Environment, 505*, 680–693.

Torija, A. J., Ruiz, D. P., & Ramos-Ridao, A. (2014). A tool for urban soundscape evaluation applying support vector machines for developing a soundscape classification model. *Science of the Total Environment, 482-483*, 440–451.

Torija, A. J., Ruiz, D. P., & Ramos-Ridao, A. F. (2013). Application of a methodology for categorizing and differentiating urban soundscapes using acoustical descriptors and semantic-differential attributes. *Journal of the Acoustical Society of America, 134*, 791–802.

Torija, A. J., & Ruiz, D. P. (2012). Using recorded sound spectra profile as input data for real-time short-term urban road-traffic-flow estimation. *Science of the Total Environment, 435-436*, 270–279.

Torija, A. J., Ruiz, D. P., Alba-Fernandez, V., & Ramos-Ridao, A. (2012). Noticed sound events management as a tool for inclusion in the action plans against noise in medium-sized cities. *Landscape & Urban Planning, 104*, 148–156.

Torija, A. J., Ruiz, D. P., & Ramos-Ridao, A. (2012). Use of back-propagation neural networks to predict both level and temporal-spectral composition of sound pressure in urban sound environments. *Building & Environment, 52*, 45–56.

Torija, A. J., Genaro, N., Ruiz, D. P., Ramos-Ridao, A., Zamorano, M., & Requena, I. (2010). Priorization of acoustic variables: environmental decision support for the physical characterization of urban sound environments. *Building & Environment, 45*, 1477–1489.

Uzkent, B., Barkana, B. D., & Yang, J. (2011). Automatic environmental noise source classification model using fuzzy logic. *Expert Systems with Applications, 38*, 8751–8755.

Vapnik, V. N. (1998). *Statistical learning theory*. New York: Wiley.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer.

Xu, L., & Liu, S. (2013). Study of short-term water quality prediction model based on wavelet neural network. *Mathematical & Computer Modelling, 58*, 807–813.

Wu, J., Long, J., & Liu, M. (2015). Evolving RBF neural networks for rainfall prediction using hybrid particle swarm optimization and genetic algorithm. *Neurocomputing, 148*, 136–142.

Yadav, A. K., & Chandel, S. S. (2014). Solar radiation prediction using artificial neural network techniques: a review. *Renewable & Sustainable Energy Reviews, 33*, 772–781.

Yadav, A. K., Malik, H., & Chandel, S. S. (2014). Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. *Renewable & Sustainable Energy Reviews, 31*, 509–519.

Yaïci, W., & Entchev, E. (2014). Performance prediction of a solar thermal energy system using artificial neural networks. *Applied Thermal Engineering, 73*, 1346–1357.

Zaheeruddin, V. K. J. (2008). An expert system for predicting the effects of speech interference due to noise pollution on humans using fuzzy approach. *Expert Systems with Applications, 35*, 1978–1988.

Zaheeruddin, V. K. J. (2006). A fuzzy expert system for noise-induced sleep disturbance. *Expert Systems with Applications, 30*, 761–771.

Zeng, J., & Qiao, W. (2013). Short-term solar power prediction using a support vector machine. *Renewable Energy, 52*, 118–127.

Zhu, J. Z., Cao, J. X., & Zhu, Y. (2014). Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections. *Transportation Research Part C: Emerging Technologies, 47*, 139–154.