Applied Acoustics 138 (2018) 147–155

Contents lists available at ScienceDirect

Applied Acoustics

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

Traffic noise and pavement distresses: Modelling and assessment of input parameters influence through data mining techniques

Elisabete F. Freitas[a,*], Francisco F. Martins[b], Ana Oliveira[a], Iran Rocha Segundo[a], Hélder Torres[a]

[a] CTAC, School of Engineering, Department of Civil Engineering, University of Minho, Portugal
[b] ISISE, School of Engineering, Department of Civil Engineering, University of Minho, Portugal

ARTICLE INFO

Keywords:
Tyre-pavement noise
Acoustic and psychoacoustic indicators
Pavement distresses
Data mining
Support vector machines
Artificial neural networks

ABSTRACT

Traffic noise affects greatly health and well-being of people, consequently the knowledge and control of the factors affecting it is very important. In this study models to predict tyre-pavement noise acoustic and psychoacoustic indicators based on type of pavement, texture, pavement distresses and speed were developed and used to assess the importance of each factor. By applying data mining techniques, in particular artificial neural networks and support vector machines, models with good predictive capacity of both acoustic and psychoacoustic noise indicators were obtained, constituting a precious tool to reduce the tyre-pavement noise. Moreover, the proposed models allowed for the assessment of the influence of the input parameters controlling noise such as: type of pavement, texture, speed and pavement distresses for the first time. It was found that pavement distresses and, as expected, speed influence strongly tyre-pavement noise. In this way it is clearly shown that preventive maintenance of road pavements by authorities, which eliminates distresses, can have an important effect on tyre-road noise, promoting the well-being of the populations.

1. Introduction

The high population growth rate in urban and metropolitan areas has led to an exponential increase in car traffic in these areas leading to a substantial increase in noise. This noise is a major concern for populations given its negative impact on their health. Traffic noise may affect the mental health and sleep quality [1–3]. In addition, it is a risk factor for hearing, cardiovascular diseases and diabetes [4–7]. Taking into account these concerns of the populations the car manufacturers have been reducing the noise of the motors to very low levels, becoming more significant the noise caused by the tyre. Therefore, it is very important to study the tyre-pavement noise.

The tyre-pavement noise is influenced by a number of factors, namely driver behaviour, tyre characteristics, pavement surface characteristics and climate [8]. The speed of the vehicle has a strong influence on annoyance regardless the type of pavement, as well as the traffic composition, where higher densities correspond to higher annoyance rates [9]. However, with successive vehicle pass-bys and climatic variations, after a certain period of time, road pavements start to develop different types of distresses or pathologies, such as cracks and alligator cracking, rutting, potholes, ravelling, among others [10].

In the period of use up to the development of the first distresses, the tyre/road noise increases with different rates depending on the pavement type. Some studies report very important increases of noise levels in the first years of use [11,12] caused by the wearing of the tyres that removes the asphalt film from the aggregates, changes in texture, clogging and stiffness.

Although the surface characteristics are considered one of the influent factors in tyre-pavement noise, there are no studies relating the existing pathologies on the surface with tyre-pavement noise.

The existing pathological conditions on the surface of the pavement, besides causing discomfort to the drivers and increasing accident risk, appear to influence road traffic noise due to the perceptible intensification of tyre vibrations, which is expected to increase the auditive discomfort of road users. In order to demonstrate what is currently perceived by road users, a detailed acoustic study of distressed pavements is essential, therefore psychoacoustic indicators should be considered. There are three key factors associated with this type of studies: type of pavement, traffic speed and level of pavement distress. In this context, the aim of this study was to develop a tyre-pavement noise prediction model, with the traffic speed, the type of pavement and the existing pathological conditions on the pavement surface as inputs and as outputs the equivalent sound pressure level in decibels (dB) (L_{eq}), A-weighted equivalent mean sound pressure level (L_{Aeq}) and the A-weighted maximum sound pressure level (L_{Amax}) and also the psychoacoustic indicators such as Loudness, Roughness and Sharpness. In

* Corresponding author.
E-mail address: efreitas@civil.uminho.pt (E.F. Freitas).

https://doi.org/10.1016/j.apacoust.2018.03.019
Received 12 August 2017; Received in revised form 3 December 2017; Accepted 19 March 2018

addition, it was intended to evaluate the influence of pavement distresses in the traffic noise.

For this purpose, data mining (DM) techniques were used. DM is a step of Knowledge Discovery in Databases (KDD) aiming at the extraction of useful patterns from observed data. The KDD process can be resumed in five main steps: data selection, pre-processing, transformation, DM and interpretation [13]. There are several DM algorithms. Artificial Neural Network (ANN), Support Vector Machines (SVM) and Multiple Regressions (MR) were used in this study.

2. Literature review

Early models of traffic noise prediction appeared in the 1950s [14]. Since then many models have been developed. Also, in the last few years there has been interest in developing predictive tyre-road contact noise models. As this article is focused on the application of data mining techniques to assess the influence of distresses on tyre-road noise, only the models related with these techniques will be treated here. Readers are encouraged to consult Refs. [14–17] that constitute a relevant contribution to the literature review. Most of the models presented in the literature are based on linear relationships between the considered parameters. Therefore, these models fail to capture the complex relationships between the involved parameters. Hence, the development of models based on data mining techniques.

Cammarata et al. [18] proposed an instrument for modelling and filtering urban noise by using neural networks. They used a learning vector quantization (LVQ) network as a filter of wrong measurements before the use of a backpropagation network (BPN) for the prediction of the sound pressure level, L_{eq}. They adopted as input parameters the number of cars, the number of motorcycles, the number of trucks, the average height of the buildings facing the road, and the width of the road. The results were compared with classical solutions to noise prediction and validated using data belonging to several medium and small-size towns. They point out the versatility of the BPN and its good capacity to predict the sound pressure level.

Nedic et al. [19] applied artificial neural networks for the prediction of traffic noise descriptor, L_{eq}. The number of light motor vehicles, the number of medium trucks, the number of heavy trucks, the number of buses and the average traffic flow speed were the input variables. It was shown that the artificial neural networks can be a useful tool for the prediction of noise with sufficient accuracy and that ANN model has much better capabilities to predict traffic noise level than any other statistical methods.

Kumar et al. [20] developed an ANN model to predict the highway noise descriptors, 10 Percentile exceeded sound level (L_{10}) and equivalent sound level (L_{eq}). The model input parameters are total vehicle volume/hour, percentage of heavy vehicles and average vehicle speed. Results obtained with ANN approach were compared with regression analysis and with the field measurement. They concluded that ANN approach is better than regression analysis and constitutes a powerful technique for traffic noise modelling.

Garg et al. [21] developed two ANN models to predict equivalent continuous sound level (L_{Aeq}) and 10 Percentile exceeded sound level (L_{10}) generated due to traffic noise. Eight input parameters, denominated number of two-wheelers, three-wheelers, cars, medium commercial vehicles, buses, trucks, average speed of heavy vehicles, and average speed of light vehicles were considered. The conclusions indicated that proposed models are able to produce accurate predictions of hourly sound levels in the urban environment.

Singh et al. [22] developed models to predict the equivalent sound level, L_{eq}, based on soft computing methods, namely, Generalized Linear Model, Decision Trees, Random Forests and Neural Networks. The input variables include the traffic volume per hour, percentage of heavy vehicles and average speed of vehicles. The Random Forest model gave the best results and the potential of using this method for traffic noise with accuracy and stability is highlighted by the authors.

Hamad et al. [23] employed artificial neural network technique to model L_{eq} in a city with known hot climate, namely Sharjah City in United Arab Emirates. They used the following inputs: Distance from the edge of the road in meters, hourly light-vehicle volume, hourly heavy-vehicle volume, average speed in km/h and roadway temperature. They ran several ANN models and compared the best-performing ANN models with two conventional models. In general, results showed that ANN models outperformed the conventional models.

Bravo et al. [24] presented a methodology, which allows to train an ANN model properly, in order to predict the willingness to pay (WPT) range to reduce road traffic noise annoyance within a given population. They performed a socio acoustic survey that collects the WTP of the respondents and adopted as input variables characteristics such as environmental noise perception of the respondents, modelled day-night noise exposure level (*LDN*) at the facade of their dwellings, and the respondents' demographic and socioeconomic status. The developed model predicts, with precision and accuracy, a range for willingness to pay from subjective assessments of noise, a modelled noise exposure level, and both demographic and socio-economic conditions.

From real-time acoustic analysis of tyre/road noise, Alonso et al. [25] proposed an asphalt status (dry/wet) classification system using support vector machines. They reported very high success rates. Freitas et al. [17] modelled tyre-road noise measured by the close proximity method (CPX) with Speed, Temperature, Aggregate size, Mean Profile Depth and Damping at 800 Hz and 2000 Hz as inputs. The support vector machine and artificial neural network algorithms showed high quality predictive performances.

Masino et al. [26] applied support vector machines to predict different types of road surfaces from tyre noise. Their approach took multiple features based on the power spectral density (PSD) of time series data of the tyre cavity sound under vehicle operation as input. Tests of the classifier revealed an accuracy above 90%.

3. Materials and methods

3.1. Road stretches

A total of 21 road stretches with different distresses, inserted in 6 national roads, were selected: 6 stretches in thin Gap Graded Asphalt (GGA), 8 stretches in Asphalt Concrete (AC) and 7 stretches in Gap Graded Asphalt Rubber (CGAR). The distresses chosen were the alligator cracking (high severity), cracking (medium severity) and ravelling, shown in Figs. 1–3, which are typical of urban areas. Also, a stretch of each pavement without distresses was considered for reference. In Table 1 are presented all the combinations used. At least two pass-bys were done in each stretch over each distress. The sounds registered by both microphones were included in the analysis.

Fig. 1. Example of pavement with alligator cracking.

Fig. 2. Example of pavement with cracking.

Fig. 3. Example of pavement with ravelling.

Table 1
Selected pavements and respective distresses.

Pavement type	Road	Distress			
		Alligator cracking	Cracking	Ravelling	Without distress
AC	EN207-4	x			x
	EN105	x		x	x
	EN206 (1)		x	x	
	EN206 (2)	x			
	EN14 (1)		x		x
GGAR	EN14 (2)			x	
	EN103 (1)		x		x
	EN103 (2)	x		x	
	EN310 (1)	x	x		
GGA	EN310 (2)	x			
	EN310 (3)		x	x	x

3.2. Data acquisition

The method used to acquire tyre-pavement noise was the Close Proximity Method (CPX) described in ISO 11819-2: 2000: “Acoustics – Measurement of the influence of road surfaces on traffic noise – Part 2: Close Proximity Method”. For the acquisition of the noise generated by the tyre-pavement interaction, two Free-field ½ Type 4190 microphones were connected to the Pulse Type 3560-C portable platform using cables AO-0419, all from Brüel & Kjær. The Pulse platform was powered by a portable battery, and connected through a network cable to a laptop computer. The tests were done with a Continental

Fig. 4. Microphones mounted in the test tyre.

ContiEcoContact3 195/65-R15 tyre applied to a light vehicle to represent normal road traffic noise. The arrangement of the microphones is shown in Fig. 4. As only two microphones were used, they were placed on the side of the tyre in a 45° orientation, 0.20 m away from the tyre and its centre, and 0.10 m from the pavement surface. The microphones were connected to the Pulse platform, which in turn was connected to the portable computer with which sound acquisition was controlled through Brüel & Kjær's Labshop 14.1.1 software.

The macrotexture is defined by the deviation of a pavement surface from a true planar surface with the characteristic dimensions along the surface of 0.5 mm to 50 mm. The macrotexture was also measured on the wheel path in each stretch with a High Speed Profilometer.

3.3. Sound files manipulation

The measured noise by the CPX method was recorded and partially processed in the Labshop software, where the sound pressure level in dB (A) and dB was obtained in addition to waveform sound files. These files were listened to and analysed. The extraction of the 5-s excerpts was followed through the Audacity program. With sound files of equal size, it was possible to extract the acoustic indicators, such as equivalent and maximum sound pressure levels, and psychoacoustic indicators such as Loudness, Sharpness and Roughness, thus allowing a more detailed analysis of the sound. The indicators were extracted using the Psysound3 application executed in the Matlab software 7.13.0.564.

3.4. Data base

For the application of the DM techniques a database with 204 records was built, in which the variables were the type of pavement (AC, GGA and GGRA), existing pathologies (without distress, cracking, alligator cracking and ravelling), the texture, speed and acoustic and psychoacoustic indicators, taken from the PsySound3 application.

The input parameters were type of pavement, pathologies, texture and speed. The type of pavement and pathologies were labelled as categorical variables. The first with 3 categories, each one designating the type of pavement, - AC, GGA and GGRA -; and the second with 4

categories - without distress, cracking, alligator cracking and ravelling. The sound generated by the contact of the tyre with the distressed surface was measured over the distress and the duration of the measurement was fixed. In this way each degradation was isolated and the time of contact with road users was exactly the same, which is essential for the consistency of the dependent variables. Therefore, it was not possible to treat pathologies as numerical variable. Other approaches could be done, for example the length or area of each distress could be considered. In such a case, long segments of the road which would include distressed and not distressed parts should be used. In these conditions, the inference of the effect of each distress on noise would be much more complex, perhaps impossible, even with a big data base.

The texture of the surface of the pavement can be described by several texture descriptors. The Mean Profile Depth (MPD), described in EN ISO 13473-1, used to describe macrotexture, recognized as an important factor on noise generation and, therefore, included in several tyre-road noise models was selected to characterize the texture. This variable as well as speed are numerical variables. Since these variables have close orders of magnitude they were not normalized.

The output parameters were the acoustic indicators equivalent sound pressure level in decibels (dB) (L_{eq}), A-weighted equivalent mean sound pressure level (L_{Aeq}) and the A-weighted maximum sound pressure level (L_{Amax}). Psychoacoustic indicators such as Loudness, Roughness and Sharpness were also used. These indicators were all extracted from the CPX sound files that were acquired close to the car wheel. It must be emphasised that only one output parameter per analysis was used. Since psychoacoustic indicators are not generally used in predicting models, a brief definition of these parameters is given below.

Loudness is the attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from soft to loud. The 'loudness level' of a sound is defined as the sound pressure level of a 1 kHz tone in a plane wave and frontal incident that is as loud as the sound; its unit is "phon". The sone scale is based on the observation that a 10 phon increase in a sound level is most often perceived as a doubling of loudness. Sharpness is a measure of the high frequency content of a sound (over 1100 Hz), the greater the proportion of high frequencies the 'sharper' the sound [27]. A sound of sharpness 1 acum is defined as a narrow band noise one critical band wide at a centre frequency of 1 kHz having a level of 60 dB [27]. Roughness is a complex effect which quantifies the subjective perception of rapid fluctuations (15–300 Hz) in the sound received by auditory filters [27]. The unit of measure is the asper. One asper is defined as the roughness produced by a 1000 Hz tone of 60 dB which is 100% amplitude modulated at 70 Hz [27].

Table 2 presents some statistical data of the numerical parameters used in the analyses. The coefficient of variation of texture (MPD), speed and Loudness are quite similar which means that they have analogous variability. *Roughness* presents the highest variability and L_{eq} the lowest. L_{eq}, L_{Aeq} and L_{Amax} have similar variability.

Table 2
Basic descriptive statistics of the parameters used in database.

Parameters	Min	Mean	Max	Standard Deviation	Coefficient of variation (%)
MPD (mm)	0.9	1.41	2.3	0.339	23.993
Speed (km/h)	30	48.30	66	12.787	26.472
Leq (dB)	89.14	100.01	108.29	4.500	4.499
LAeq (dB(A))	79.31	91.55	99.52	4.725	5.161
LAmax (dB(A))	79.7	92.30	100.76	4.785	5.184
Loudness (sone)	35.69	70.92	104.03	16.563	23.356
Roughness (asper)	0.09	0.227	0.56	0.087	38.469
Sharpness (acum)	3.1	4.494	5.96	0.601	13.377

3.5. Data mining

In this study the data mining process was applied to predict tyre-pavement noise descriptors. Artificial neural networks, support vector machines and multiple regressions were applied. All the experiments were conducted using the R statistical environment and the RMiner library [28] which presents a set of functions that facilitates the use of DM algorithms both in classification and regression tasks. This study is related with regression tasks.

The ANN is a computational technique based on the architecture of the human brain. The ANN architecture is defined according the way the artificial neurons are linked to each other. The neurons communicate with each other via signals sent through the links. Each liaison has an associated weight, $w_{i,j}$ (i and j are neurons or nodes), and each neuron has an activation function that introduces a non-linear component. This study used a logistic activation function f given by $1/(1+e^{-x})$ and the following general equation:

$$\hat{y} = w_{o,0} + \sum_{j=I+1}^{o-1} f\left(\sum_{i=1}^{I} x_i w_{j,i} + w_{j,0}\right) w_{o,i} \quad (1)$$

where x_i are the input parameters or nodes, i is the number of input parameters and o is the output parameter. In this study it was adopted the most common ANN architecture, the multilayer perceptron (feed forward network) [29] with one hidden layer of HN hidden nodes. The grid search of the number of hidden nodes HN was {0,2,4,6,8,10,12,14,16,18,20}.

The SVM technique was initially developed to classification tasks by Cortes and Vapnik [30]. The application of SVM to regression problems was possible after the introduction of the ε-insensitive loss function [31]. This method uses a nonlinear mapping to transform the input data into a multidimensional feature space by using a nonlinear mapping ϕ, which is normally unknown. SVM presents theoretical advantages over ANN, such as the absence of local minima in the learning phase which means that the model always converges to the optimal solutions [32]. After this transformation the SVM finds the best hyperplane of linear separation inside the feature space. The nonlinear mapping depends on a kernel function *k(x,x')*. This work uses the following kernel function:

$$k(x,x') = e^{(-\gamma \cdot \|x-x'\|^2)}, \ \gamma>0 \quad (2)$$

The performance of the regression is affected by the kernel parameter, γ, a penalty parameter, C, and the width of the ε-insensitive zone. To limit the searching space, C was considered equal 3 and it was used a heuristic for ε [33]: $\varepsilon = \hat{\sigma}/\sqrt{N}$, where, $\hat{\sigma} = 1.5 \times \sum_{i=1}^{N} (y_i - \hat{y}_i)$, $\hat{y}_i$ is the value predicted by a 3-nearest neighbour algorithm and N the number of examples. Therefore, the search space was limited to the input values of γ which in this study were $\{2^{-15},2^{-13},2^{-11},2^{-9},2^{-7},2^{-6},2^{-5},2^{-4},2^{-3},2^{-2},2^{-1},2^{0},2^{1},2^{2},2^{3}\}$.

To test the predictive capacity of the data mining techniques the dataset was divided in two subsets. One, composed by 80% registers, was used to train the model, and other, composed by the remaining 20% registers, to test the model. With the training set a 10-fold cross validation was used where the data was divided in ten subsets of equal size. Nine subsets were used to adjust the model whereas the remaining subset was used to test the model. This process was repeated until all the subsets had been tested. Ten runs of this process were carried out on this study. After fitting the model with the training dataset its future performance was assessed with the unseen dataset (*i.e.* 20% register).

There are several metrics to assess the performance of the regression models. This study uses the mean absolute deviation (*MAD*), the root mean squared error (*RMSE*) and the coefficient of determination (R^2). These metrics are defined as follow:

$$MAD = \frac{1}{N} \times \sum_{i=1}^{N} |y_i - \hat{y}_i| \quad (3)$$

Table 3
Mean values of the metrics obtained in the cross validation process.

	L_{eq}			L_{Aeq}			L_{Amax}		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.558	0.655	0.728	0.694	0.679	0.805	0.687	0.684	0.812
MAD	2.233	1.943	1.748	1.827	1.922	1.537	1.918	1.918	1.609
RMSE	2.971	2.754	2.331	2.570	2.816	2.053	2.672	2.858	2.070

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \quad (4)$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}(y_i - \overline{y}) \times (\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i - \overline{y})^2} \times \sqrt{\sum_{i=1}^{N}(\hat{y}_i - \overline{\hat{y}})^2}} \right)^2 \quad (5)$$

where N denotes the number of examples, y_i the real value, $\hat{y}_i$ the value estimated by the model, $\overline{y}$ the mean of the real values and $\overline{\hat{y}}$ the mean of the estimated values.

The lower the *MAD* and *RMSE* values the better the predictive capacity of the model. On the contrary, a higher value of R^2 corresponds to its better performance.

To obtain the relative importance of the input parameters in the model it is necessary to perform a sensitivity analysis where each input parameter is ranged from its lowest value to its highest value while the remaining input parameters keep their mean values [34]. The importance of an input parameter is proportional to the variance induced by it in the model output. Therefore, higher variance corresponds to higher importance.

4. Results and discussion

4.1. Acoustic indicators

After training the three DM techniques for acoustic indicators, the errors and the coefficient of determination were obtained in the cross validation process (Table 3). The best results were obtained with the SVM model. MR models to predict L_{Aeq} and L_{Amax} have a better performance than ANN models. However, as it can be seen in Figs. 5–7, the fitting of the models with the complete training dataset increased

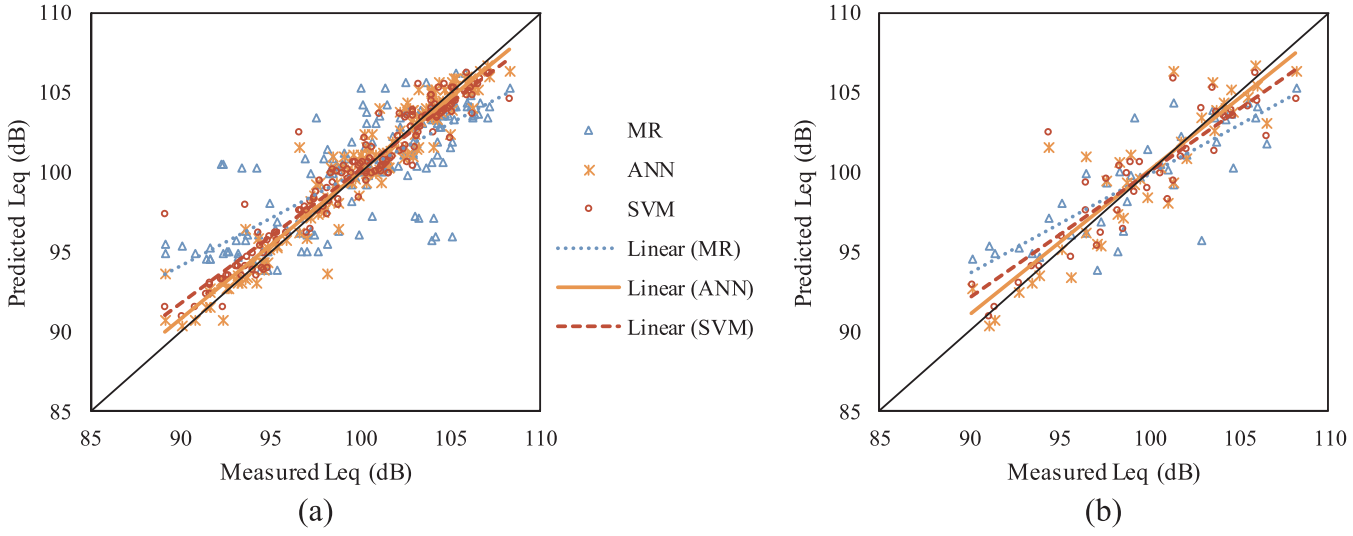

Fig. 5. Performance of the MR, ANN and SVM models in the L_{eq} prediction using: (a) training dataset; (b) testing dataset.

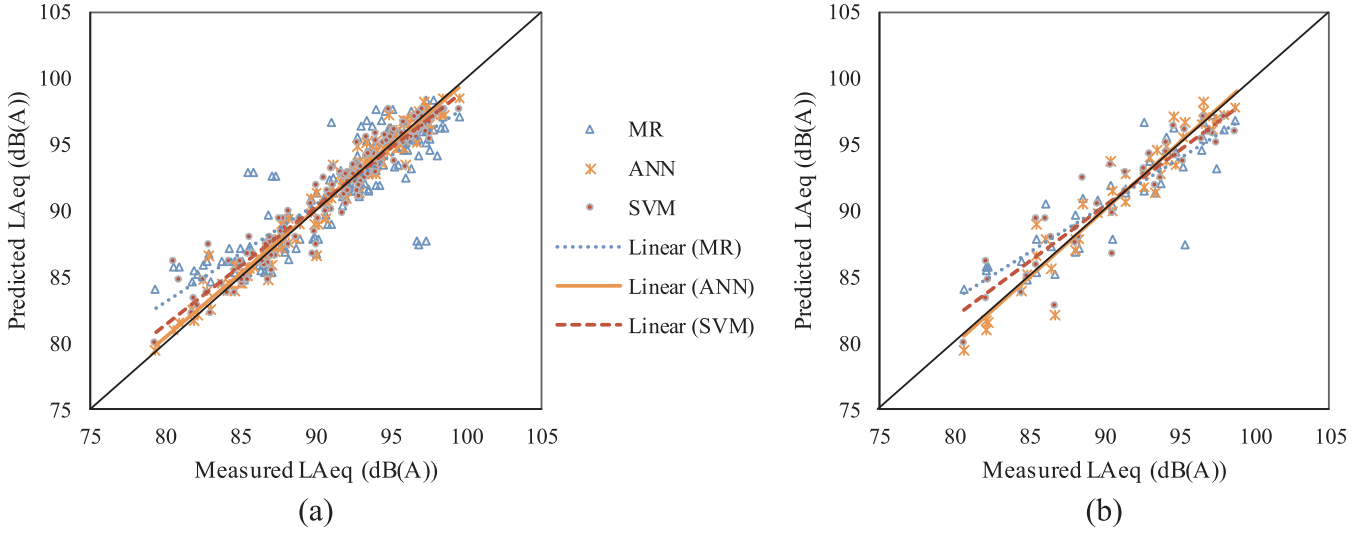

Fig. 6. Performance of the MR, ANN and SVM models in the L_{Aeq} prediction using: (a) training dataset; (b) testing dataset.

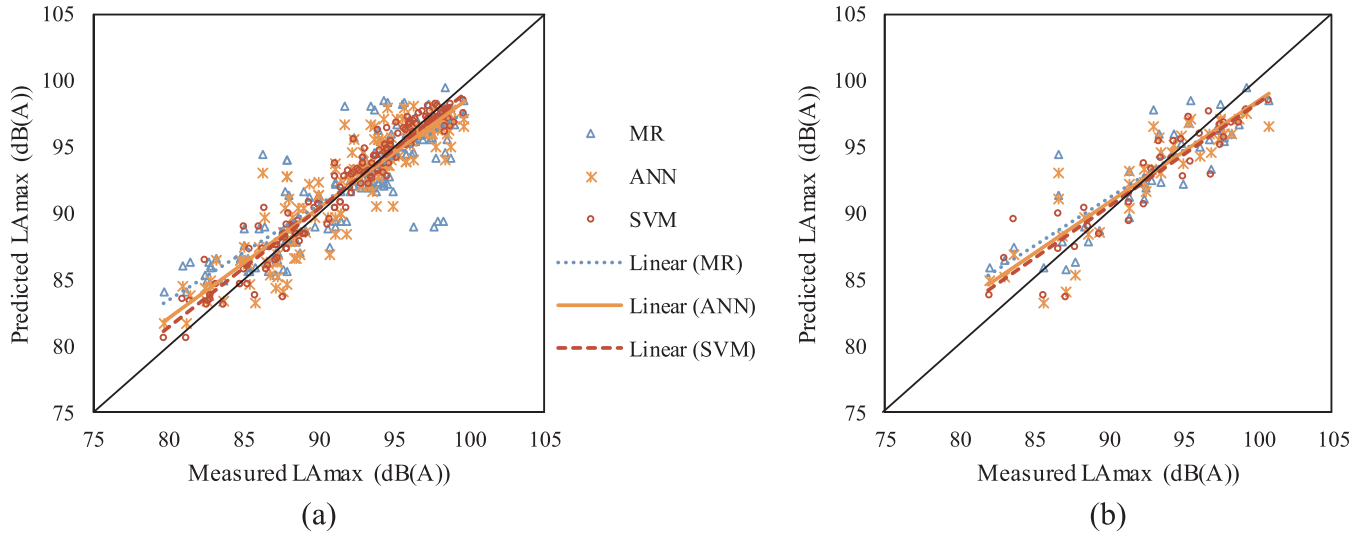

Fig. 7. Performance of the MR, ANN and SVM models in the L_{Amax} prediction using: (a) training dataset; (b) testing dataset.

considerably the performance of all models, mainly the ANN models. The metric values and slope and y intercept of the trend lines extracted from Figs. 5–7, presented in Tables 4 and 5, confirm these performance improvements. In fact, ANN models have the best performance in predicting L_{Aeq} and L_{eq}, and SVM model remains the best to predict L_{Amax}.

The importances of the input parameters attributed by the different models are presented in Table 6. The high importance assigned to speed is according to the expected. An exception to this trend is the ANN model that assigns an importance less than 20% to speed. However, the SVM predictive model of L_{eq}, with lower performance, attributes the greatest importance to the distresses. For the L_{Aeq} and L_{Amax} indicators, in opposition to the ANN models, the SVM models assign a consistent and great importance to distresses, nonetheless consider speed more important. These models assign little relevance to the type of pavement and texture. In summary, using the acoustic indicators the parameters that most influence the tyre-pavement noise are speed and distresses, the latter reaching about 20% for L_{Aeq} and L_{Amax} indicators and 37% for L_{eq}.

Table 4
Metrics and slope and y intercept obtained from Figs. 5–7 using training dataset.

	L_{eq}			L_{Aeq}			L_{Amax}		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.602	0.928	0.916	0.728	0.964	0.929	0.723	0.837	0.940
MAD	2.114	0.863	0.960	1.717	0.631	0.970	1.799	1.533	0.920
RMSE	2.814	1.196	1.338	2.419	0.878	1.250	2.513	1.930	1.189
Slope	0.602	0.930	0.841	0.728	0.965	0.893	0.723	0.837	0.904
y Intercept	39.80	7.062	15.95	24.90	3.263	9.850	25.55	15.06	8.921

Table 5
Metrics and slope and y intercept obtained from Figs. 5–7 using testing dataset.

	L_{eq}			L_{Aeq}			L_{Amax}		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.691	0.809	0.780	0.776	0.924	0.860	0.750	0.793	0.837
MAD	2.027	1.437	1.554	1.778	1.118	1.419	1.838	1.756	1.514
RMSE	2.570	2.058	2.147	2.372	1.455	1.859	2.404	2.170	1.945
Slope	0.626	0.901	0.791	0.715	1.022	0.845	0.722	0.763	0.779
y Intercept	37.20	9.887	20.83	25.98	−2.03	14.25	26.02	22.03	20.35

Table 6
Importance of the input variables in the evaluation of the acoustic indicators (%).

	L_{eq}			L_{Aeq}			L_{Amax}		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
Type of pavement	0.81	12.47	17.19	1.59	15.11	0.02	4.02	2.11	4.25
Distresses	4.17	27.72	37.56	0.35	38.90	20.18	0.78	2.49	19.81
Texture	9.20	8.33	18.33	7.71	28.56	1.79	11.68	3.68	5.04
Speed	85.82	51.48	26.92	90.35	17.43	78.01	83.52	91.72	70.90

Table 7
Mean values of the metrics obtained in the cross validation process for the psychoacoustic indicators Loudness, Roughness and Sharpness.

	Loudness			Roughness			Sharpness		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.669	0.677	0.814	0.089	0.051	0.093	0.655	0.588	0.770
MAD	6.652	7.732	5.690	0.062	0.065	0.060	0.254	0.280	0.216
RMSE	9.465	9.867	7.098	0.085	0.095	0.085	0.353	0.404	0.288

4.2. Psychoacoustic indicators

Likewise, after training the three DM techniques for the psychoacoustic indicators, the errors and the coefficient of determination were obtained in the cross validation process (Table 7). The best results for all psychoacoustic indicators were obtained with SVM models. However, the poor capacity of all models to predict Roughness and the fair capacity of ANN models to predict the psychoacoustic indicators is denoted by their low coefficients of determination and higher *MAD* and *RMSE*.

As with acoustic indicators, the performance of the adjusted models with all training records has improved considerably. As it can be seen in Figs. 8–10 and confirmed in Tables 8 and 9, SVM models continue to have de best performance to predict the psychoacoustic indicators. The ANN model performs similarly to the SVM model in predicting Loudness but denotes difficulty in capturing the non-linear relationships between the variables to predict roughness and sharpness. This is evidenced by the fact that they provide the same results as linear MR models.

The importances of the input parameters attributed by the different models are presented in Table 10. As the models cannot predict Roughness accurately this psychoacoustic indicator was disregarded. Previous works based on common prevision techniques also failed to model roughness. The independent variables selected might not explain this sound measure and other variables should be used in future works. Accordingly, only the importance given by SVM model to the input parameters to predict Loudness and Sharpness was considered in the analysis. In this framework, speed is the most important parameter followed by distresses. These parameters summed up approximately 98% of importance. The importance of distresses reached 36% for Loudness. For sharpness, the results were very close to the ones obtained for the acoustic indicators with 23% of importance for de distresses. In this context, Loudness is the most sensitive indicator to the

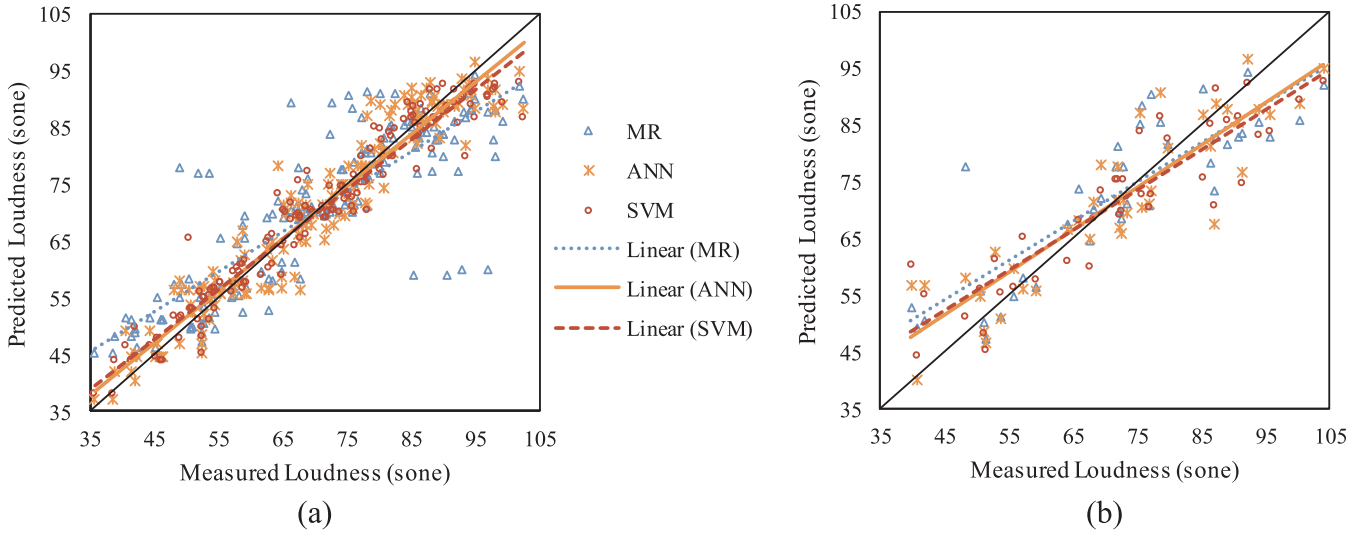

Fig. 8. Performance of the MR, ANN and SVM models in the Loudness prediction using: (a) training dataset; (b) testing dataset.

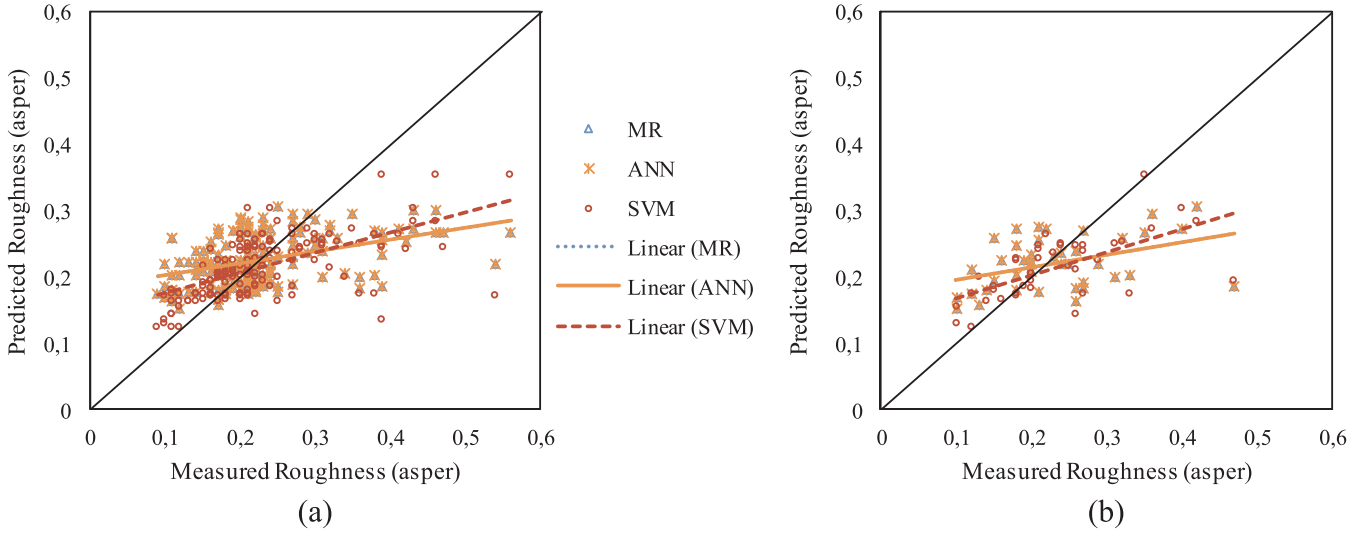

Fig. 9. Performance of the MR, ANN and SVM models in the Roughness prediction using: (a) training dataset; (b) testing dataset.

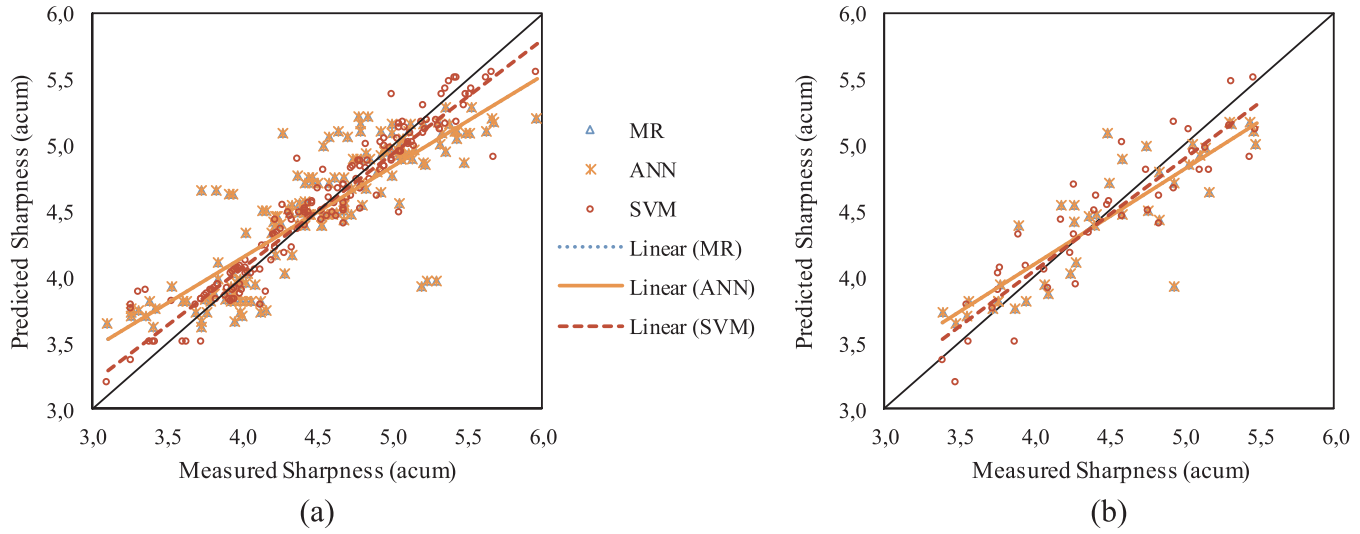

Fig. 10. Performance of the MR, ANN and SVM models in the Sharpness prediction using: (a) training dataset; (b) testing dataset.

Table 8
Metrics and slope and y intercept obtained from Figs. 8–10 using training dataset.

	Loudness			Roughness			Sharpness		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.704	0.919	0.945	0.179	0.179	0.365	0.693	0.693	0.921
MAD	6.268	3.642	3.113	0.058	0.058	0.049	0.239	0.239	0.125
RMSE	8.938	4.691	4.022	0.079	0.079	0.071	0.333	0.333	0.172
Slope	0.704	0.917	0.880	0.179	0.179	0.304	0.693	0.693	0.867
y Intercept	20.91	5.842	8.180	0.186	0.186	0.146	1.383	1.383	0.605

Table 9
Metrics and slope and y intercept obtained from Figs. 8–10 using testing dataset.

	Loudness			Roughness			Sharpness		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
R^2	0.716	0.801	0.811	0.171	0.171	0.389	0.758	0.758	0.833
MAD	7.233	5.954	5.933	0.061	0.061	0.047	0.234	0.234	0.204
RMSE	9.084	7.571	7.607	0.079	0.079	0.070	0.297	0.297	0.246
Slope	0.677	0.751	0.708	0.190	0.190	0.343	0.724	0.724	0.852
y Intercept	24.32	17.48	20.20	0.177	0.177	0.134	1.194	1.194	0.641

Table 10
Importance of the input variables in the evaluation of the psychoacoustic indicators (%).

	Loudness			Roughness			Sharpness		
	MR	ANN	SVM	MR	ANN	SVM	MR	ANN	SVM
Type of pavement	3.23	2.80	0.37	1.16	1.16	22.57	3.80	3.80	0.05
Distresses	1.24	15.78	36.76	26.29	26.29	19.50	1.44	1.44	23.13
Texture	6.67	11.81	1.88	68.94	68.94	33.80	0.00	0.00	2.15
Speed	88.86	69.61	60.99	3.61	3.61	24.13	94.76	94.76	74.67

presence of distresses on the surface of pavements among the reliable SVM acoustic and psychoacoustic models (L_{Aeq}, L_{Amax}, Loudness and Sharpness indicators).

5. Conclusions

The exponential rise of the number of vehicles on roads have been increasing the traffic noise levels. Exposure to traffic noise is a risk factor for the health of people and has a strong impact on life quality. For these reasons it is necessary to develop models to predict accurately traffic noise and to extract information about the relative influence of factors affecting its main component, tyre-road noise. These models are fundamental to support for the implementation of noise abatement measures.

Previous works address mainly environmental noise and the variables used to predict this type of noise such as traffic. In this work the

noise generated by the contact between one reference tyre and the pavement surface was modelled. In this case the factors used to explain the noise are not the same. They respect to the road surface characteristics and condition and to vehicle operation (speed). Therefore, this approach is new. The models are used to predict tyre-road noise acoustic and psychoacoustic indicators, based on pavement texture, vehicles speed, and pavement distresses (no distress, cracking, alligator cracking and revelling). In the case of distresses, the objective was to confirm its influence in the tyre-road noise and this was achieved.

Despite the general good results, the applicability of the developed models is constrained by the fact of being used only 3 types of pavements. The pavements selected are widely used in Europe, nevertheless they do not cover the existing variety of pavement surfaces. Another limitation arises from the classification of pavement distresses that might be different from country to country. A general model could be developed by using data from different countries covering most of the pavement surfaces used in a certain region, for example in Europe, and a harmonized distress classification method. The real impact of distressed roads on environmental noise can be easily assessed by introducing the estimated tyre/road noise values on environmental noise models, such as CNOSSOS, which is complementary to the road ageing effect already generally known.

The main results that can be drawn from this study are described in the following items:

- ANN models gave slightly better results than SVM models to predict *Leq* and *LAeq* whereas SVM model was better in predicting *LAmax*.
- ANN and SVM models have a similar behaviour to predict loudness but SVM is quite better to predict sharpness.
- None of the developed models were able to predict the roughness indicator properly, other factors need to be investigated;
- Speed and distresses are the most influential factors in tyre-road noise. The high importance of the speed was already expected. However, in relation to distresses for the first time their relative importance to noise could be measured.
- Loudness was the most sensitive indicator to the presence of distresses on pavement surface.

This work succeeded in assessing the important contribution of 3 very common types of roads degradation on tyre-road noise. This should be enough to compel road maintenance managers acting earlier. In this way they can reduce noise, and contribute to the population wellbeing, and at the same time it is possible to reduce costs by practicing a preventive maintenance policy.

Acknowledgements

This work was partly financed by FEDER funds through the Competitivity Factors Operational Programme - COMPETE and by national funds through FCT – Foundation for Science and Technology within the scope of the projects POCI-01-0145-FEDER-007633 and PEst-OE/ECI/UI4047/2014.

References

[1] Sygna K, Aasvang GM, Aamodt G, Oftedal B, Krog NH. Road traffic noise, sleep and mental health. Environmental Res 2014;2010(131):17–24.
[2] Frei P, Mohler E, Röösli M. Effect of nocturnal road traffic noise exposure and annoyance on objective and subjective sleep quality. Int J Hyg Environ Health 2014;217(2–3):188–95.
[3] Ohrström E. Sleep disturbances caused by road traffic noise – studies in laboratory and field. Noise Health 2000;2(8):71–8.
[4] Sørensen M, Andersen ZJ, Nordsborg RB, Becker T, Tjønneland A, Overvad K. Long-term exposure to road traffic noise and incidente diabetes: a cohort study. Environ Health Perspect 2013;121:217–22.
[5] Yongbing S, Martin WH. Noise induced hearing loss in China: a potentially costly public health issue. J Otol 2013;8(1):51–6.
[6] Babisch W. Updated exposure-response relationship between road traffic noise and coronary heart diseases: a meta-analysis. Noise Health 2014;16(68):1–9.
[7] Recio A, Linares C, Banegas JR, Díaz J. Road traffic noise effects on cardiovascular, respiratory, and metabolic health: an integrative model of biological mechanisms. Environment Res 2016;146:359–70.
[8] Sandberg U, Ejsmont J. Tyre/Road noise reference Book, Informex SE – 59040. Sweden: Kisa; 2002.
[9] Freitas E., Mendonça C., Santos JA., Murteira C., Ferreira JP., Traffic noise abatement: how different pavements, vehicle speeds and traffic densities affect annoyance levels, Transportation research (Part D): Transport and environment, 2012; 17:321–326.
[10] Miller JS, and Bellinger WY. Distress identification manual for the long-term pavement performance program (No. FHWA-RD-03-031) 2003.
[11] Licitra G, Cerchiai M, Teti L, Ascari E, Fredianelli L. Durability and variability of the acoustical performance of rubberized road surfaces. Appl Acoust 2015;94:20–8.
[12] Gardziejczyk W. The effect of time on acoustic durability of low noise pavements–The case studies in Poland. Transport Res Part D: Transport Environment 2016;44:93–104.
[13] Fayyad U, Piattetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data. Commun ACM 1996;39(11):27–34.
[14] Steele C. A critical review of some traffic noise prediction models. Appl Acoust 2001;62:271–87.
[15] Quartieri J, Mastorakis N, Iannone G, Guarnaccia C, D'Ambrosio S, Troisi A, et al. A review of traffic noise predictive models. The 5th WSEAS International Conference on Applied and Theoretical Mechanics. Puerto De La Cruz, Canary Islands; 2009.
[16] Garg N, Maji S. A critical review of principal traffic noise models: strategies and implications. Environ Impact Assess Rev 2014;46:68–81.
[17] Freitas E, Tinoco J, Soares F, Costa J, Cortez P, Pereira P. Modelling tyre-road noise with data mining techniques. Archives Acoust 2015;40(4):547–60.
[18] Cammarata G, Cavalieri S, Fichera A. A neural network architecture for noise prediction. Neural Net 1995;8(6):963–73.
[19] Nedic V, Despotovic D, Cvetanovic S, Despotovic M, Babic S. Comparison of classical statistical methods and artificial neural network in traffic noise prediction. Environmental Impact Assess Rev 2014;49:24–30.
[20] Kumar P, Nigam SP, Kumar N. Vehicular traffic noise modeling using artificial neural network approach. Transportation Res Part C 2014;40:111–22.
[21] Garg N, Mangal SK, Saini PK, Dhiman P, Maji S. Comparison of ANN and analytical models in traffic noise modeling and predictions. Acoust Aust 2015;43:179–89.
[22] Singh D, Nigam SP, Agrawal VP, Kumar M. Vehicular traffic noise prediction using soft computing approach. J Environmental Manag 2016;183:59–66.
[23] Hamad K, Khalil MA, Shanableh A. Modeling roadway traffic noise in a hot climate using artificial neural networks. Transportation Res Part D 2017;2017(53):161–77.
[24] Bravo-Moncayo L, Naranjo JL, García IP, Mosquera R. Neural based contingent valuation of road traffic noise. Transportation Res Part D 2017;50:26–39.
[25] Alonso J, Lopez JM, Pavón I, Recuero M, Asensio C, Arcas G, et al. On-board wet road surface identification using tyre/road noise and Support Vector Machines. Appl Acoust 2014;76:407–15.
[26] Masino J, Pinay J, Reischl M, Gauterin F. Road surface prediction from acoustical measurements in the tire cavity using support vector machine. Appl Acoust 2017;125:41–8.
[27] Zwicker E, Fastl H. Psychoacoustics: Facts and Models. New York: Springer; 1999.
[28] Cortez P. Data mining with neural networks and support vector machines using the R/rminer Tool. In: Perner P, editor. Advances in Data Mining. Proceedings of 10th Industrial Conference on Data Mining, lecture notes in artificial intelligence 6171. Berlin: Springer; 2010, p. 572–83.
[29] Haykin S. Neural networks - a compreensive foundation. 2nd ed. New Jersey: Prentice-Hall; 1999.
[30] Cortes C, Vapnik V. Support vector networks. Mach learning 1995;20(3):273–97.
[31] Smola A, Scholkopf B. A tutorial on support vector regression. Stat Comput 2004;14(3):199–222.
[32] Tinoco J, Correia AG, Cortez P. Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time. Construction Buil Mater 2011;25:1257–62.
[33] Cherkassy V, Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Net 2004;17(1):113–26.
[34] Kewley R, Embrechts M, Breneman C. Data strip mining for the virtual design of pharmaceuticals with neural networks. IEEE Trans Neural Net 2000;11(3):668–79.