ORIGINAL ARTICLE

A Comparative Study on Interior Acoustic Comfort Level of Compact Cars using Data Mining Approach

K. A. Azuddin[1], A. K. Junoh[1*], Z. Mohamed[2] and M. T. Abdul Rahman[3]

[1]Institute of Engineering Mathematics, Universiti Malaysia Perlis, 02600, Kampus Pauh Putra, Arau, Perlis.
[2]Faculty of Mechanical Engineering, Universiti Malaysia Pahang, 26600 Pekan Pahang.
[3]School of Mechatronics Engineering, Universiti Malaysia Perlis, 02600, Kampus Pauh Putra, Arau, Perlis.

ABSTRACT – Vehicle acoustic comfort is one of the ergonomic measurement criteria that are essential for car occupants. Furthermore, interior cabin noise of a car may affect the driver's concentration when driving. This study is to investigate the noise comfort level of car interior on several compact cars. The objective is to measure interior cabin noise for all three cars and then to compare their acoustic comfort level using subfield data mining approach. A deduction was made to rate the best car among the three in term of acoustic comfort. The interior cabin noise was obtained for the cases where engine speed is varied while the cars are in stationary and moving condition. The noise was assessed according to pre-determined subjective and objective criteria. The sound quality parameters was assessed by regression analysis. In subjective assessment, the recorded noise was evaluated based on jury assessment. Then, the data mining approach was implemented to illustrate the noise level. The collected noise data were divided into five clusters through hierarchical clustering method. To assess the accuracy of noise data clusters, the method of k-nearest neighbours was performed and the results show a high accuracy rate (> 95%). Finally, the interior noise of the three cars was compared by using the analysis of variation. The vehicle acoustic comfort index was produced for the three cars tested in this study. In addition, the acoustic quality among the three cars is presented using anova. Annoyance index of the three cars was generated using data mining method. From the results, Axia car model has the best acoustic comfort compared to the other two cars by objective evaluation. By subjective evaluation, Axia car model recorded the lowest level of annoyance.

INTRODUCTION

Driving comfort is an essential criterion for customers when buying cars [1]. As such, car interior acoustic quality is one aspect that is usually considered for classifying driving comfort among academic and industry researchers. This leads to extensive effort done by car manufacturers to improve the car's interior noise comfort. Likewise, a significant number of studies and experiments have been done on car interior noise [2]. Driving comfort can be subjective due to varied perception among individuals who tested the car, even in the controlled environment. Paulraj et al. [3] stated that driving comfort can be categorised into several aspects: vibration, acoustic, and motion occur inside a vehicle. The driving/ride comfort can be divided into four domains. They are interior noise, steering wheel vibration, seat vibration and common control in the motion of the vehicle. The transmission of car vibration to occupant comes from the seat, steering wheel and controls.

There are two major sources of interior noise, the structural noise which comes from the vibration of engine, gearbox and suspension and the airborne noise as a result of car body aerodynamic noise as well as engine combustion noise. Other noise sources come from the air inlet and the cooling fan, while the powertrain and the exhaust system can be regarded as a combination of structural and airborne noise [4]. Meanwhile, Krajnak and Grega have classified the numerous noise sources of the automobile into exhaust noise, induction noise, other car accessories that can generate noise, and noise that originated from the engine [5].

Power transmission from various origins such as the engine and wheels also contributed to the noise and vibration in cars. These multiple causes are the gear transmission, tyre-road interaction, drive and propeller shaft, and body aerodynamics. In transmission gear, inappropriate bending dynamics of both torsional and gear tooth and winding features of shafts result to the generation of noise as well [6]. Besides, the rising or falling of the engine speed is a major provider to the generated noise inside of the vehicle cabin. Certain engine speeds would produce harmonics that are more audible than other speeds.

In addition, tire-road contact can be the cause of interior noise due to the vibration generated. A moving or vibrating object generates noise, where the vibration transferred to the occupant's ear from the original location it is produced [7]. The tyre-road noise is affected by road roughness, road discontinuities, bumps, and surface irregularities, [8]. Type of

---

road is an important aspect where it influences the contribution of the generated noise and vibration in the interior car cabin (Figure 1). The production of noise and vibration are higher when driving on a rougher road surface.

According to Genuit [9], engineers in the automotive industry have been working continuously to design cars that can provide the best quality in term of cabin interior noise. One of the challenges is, that the interior acoustic comfort assessment for a vehicle cannot be done only by considering the external airborne noise but also the noise generated from the inside of the car. The relation between the sound quality and vibration caused by the engine, transmission and tire-road contact have been analysed by using trend's significance levels [10]. Throughout the research, they realised that the generation of noise and vibration are related and influenced by the acoustic comfort intensity in the car cabin. The authors have proposed an assessment method model on the acquired sound and vibration data by implementing the clustering and classification method.

Recently, most of the efforts carried out in introduced different simulation and investigation methods based on data acquisition and analysis. These researches aimed to identify the sources that mainly influence the generation of noise and thereby developing the related algorithm and methods for evaluating the optimal design of the interior cabin. Typically, various factors trigger the noise and vibration inside the vehicle's cabin. However, the contribution level of each factor varies. Among these factors are the engine transmission, the tire-road interaction, types of road, structure-borne vibration, and carpet types. Besides, vibrations in an engine are the result of the responding mechanism used for changing the energy from linear into rotary motion [8]. The forces generating the engine vibration are interchanging, rotational forces and combustion [11], which generate the structure-borne noise. In other words, the noise is generated by vibration at the mounting point. The noise by the structure bone path is the bonding between the structural resonance and cabin acoustic modes and hence is typically modal dominating at low frequencies below 500Hz [12]. Such as the rings used to hold the exhaust pipe can cause structural vibration in the exhaust systems. The engine transmission causes these vibrations, and it travels to the car body frame through drives shafts. Zhang et al. [13] investigated the noise generated from structure-borne of vehicles using the hybrid approach of boundary element technique and finite element technique. From the result, the vibration of the girder-plate bridge was the cause of the noise. The research also shows that the speed of a moving vehicle affected the sound pressure components with frequencies lower than 10 Hz.

Various methods have been implemented to study the noise and vibration classification in a car. In one study, a data mining approach has been proposed to classify different types of carpets based on the materials and location placed in the car [14]. The data mining technique is used to obtain the connections or uniform patterns among groups in large relational databases [15, 16]. In another study, regression analysis with multivariate linear was employed to describe the connection between the variation of surface features and the noise level changes and to display the spectral bands of sound coefficients and the texture effect where partial least square (PLS) was used [17]. Monte Carlo simulation is another method that is very useful for analysing indeterminate circumstances and providing probabilistic analysis of diverse states [18]. For assessing vehicle interior acoustic comfort, the Monte Carlo simulation is employed by randomly producing a large number of samples by observing the statistical trends that take place in the data set. The multivariate normal (MVN) distribution was considered as the best technique to be applied in the Monte Carlo simulation since it allows univariate normal distribution that has the probability density function to be generalised [19].

(a) pavement road   (b) highway road

Figure 1. Types of tested road surfaces.

Data mining is an interdisciplinary subfield of computer science [20]. Sequential data are composed of a set of sequences, where each sequence is an ordered list of discrete events or elements [21]. Data mining techniques are described with a special emphasis on classification techniques as one important supervised learning technique. Bioinformatics tools in the field for medical applications, especially in medical microbiology, were discussed. The authors present WEKA software as a tool of choice to perform classification analysis for different kinds of available data. Uses of WEKA data mining tools for biological applications such as genomic analysis and medical applications such as diabetes are discussed. Data mining offers novel tools for medical applications for infectious diseases; it can help in identifying the pathogen and analysing the drug resistance pattern. For non-communicable diseases such as diabetes, it provides excellent data analysis options for analysing large volumes of data from many clinical studies [22]. In this study, the aims are to produce an acoustical comfort index for each three models of compact cars, measure the level of noise in the cabin by employing the data mining approach, and compare the interior noise between the three models of compact cars in objective evaluation and subjective evaluation.

METHODOLOGY

Experimental Setup

The experiment started with the development of a measurement system using LabView software. Data acquisition took place after the coding was developed. The device used to record interior noise is NI cDAQ-9174 and NI-9234 by National Instruments while the sensor is GRAS array microphone, as shown in Figure 2(a) and 2(b). The sensor has a wide useful frequency range reaching up to 20 kHz, a large dynamic range topping at around 135 dB and sensitivity were set at 50 mV/Pa. The apparatus is depicted in Figure 3. The system was developed to record the car interior noise for 10 seconds and then to store it in the computer disk. The recorded noise was saved in wav format. The system operational flow is illustrated in Figure 4 and Figure 5.

The data acquisition process began by acquiring the data in a stationary state. Then, the process continued for the non-stationary state. Acquiring data in the non-stationary state was divided into three conditions (pavement road, urban road and highway road). The collected data then were assessed by juries in subjective evaluation. Data analyses were divided into the data mining stage and data comparison stage. The interior noise measurements were conducted on three local compact cars model, namely Myvi, Viva, and Axia. They are all equipped with automatic transmission. A few criteria must be met during the car selection where the selected cars must never be caught into an accident, and were never been into a major modification that changes the car's default specifications such as equipped with a turbocharger, loud exhaust, body frame and others. The research flow is summarised in Figure 6.

(a) NI cDAQ-9174 and NI-9234   (b) Noise sensors (microphones)

Figure 2. Data acquisition system.
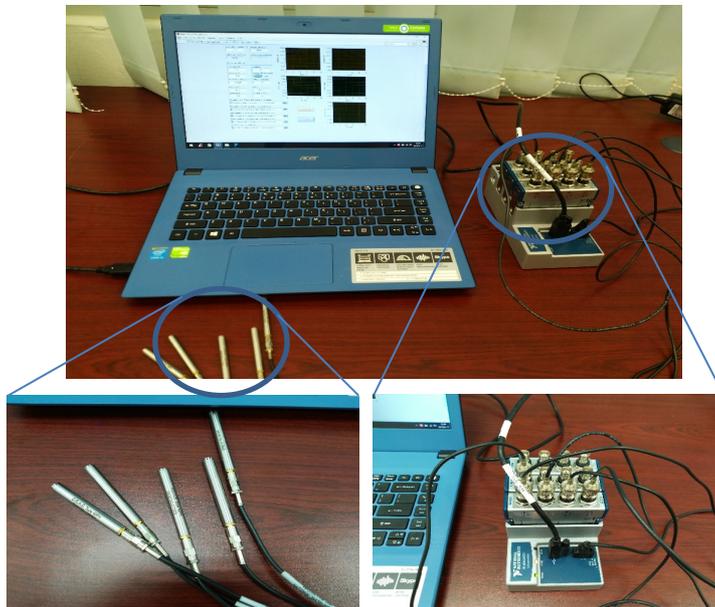
Figure 3. The noise recording system setup.

Figure 4. The developed LabVIEW system process flow.
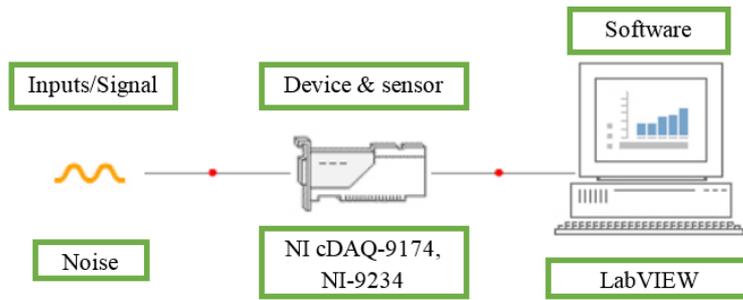
Figure 5. System flow illustration.
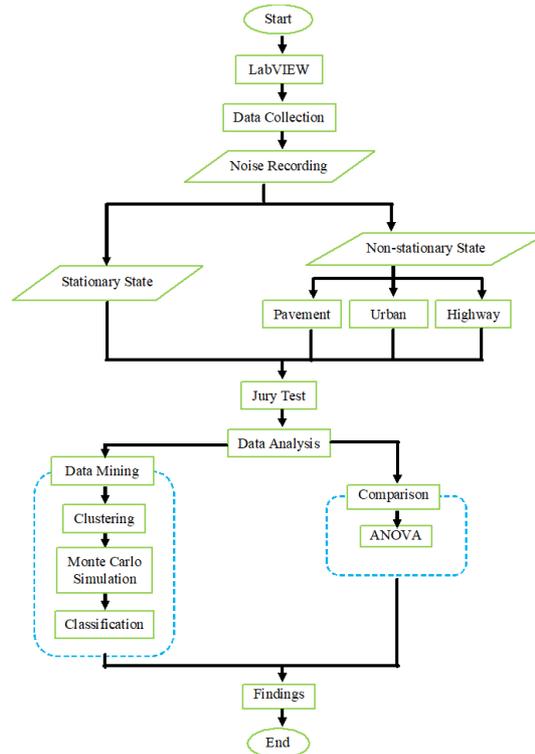
Figure 6. Methodology flowchart.

Data Collection

Data collection was conducted in two situations, the stationary state and non-stationary state. The noise for each car was recorded one at a time for all three car models. The data acquisition process began by assembling the measurement device and sensors (microphones) in the car. Five measurement microphones were attached to the car. The microphones were attached to the cars as follows: one on the dashboard centre, two on the driver window (both sides), two on the front passenger window (both sides). The locations of the sensors are illustrated as in Figure 7.

The interior car noise measurements for the stationary state of all the three cars were conducted at the same place or environment. The place to carry out the measurement tests needs to be somewhere quiet to avoid unnecessary background noise. The noise was recorded for 10 seconds at various engine speeds. During the 10 seconds recording period, the engine speed was carefully maintained at a constant speed. The interior noise was recorded starting at the engine speed of 1000 rpm. The process was repeated with the increment 250 rpm until 3000 rpm. For non-stationary state, the interior noise measurements were done using the same procedures as the stationary state, and only they were conducted on three different road types (pavement road, highway road, and urban road) as in Figure 8(a) to 8(c). Another difference was that the measurements on the paved road and urban road were done up to the engine speed of 2250 rpm while on the highway road were done up to 3000 rpm.
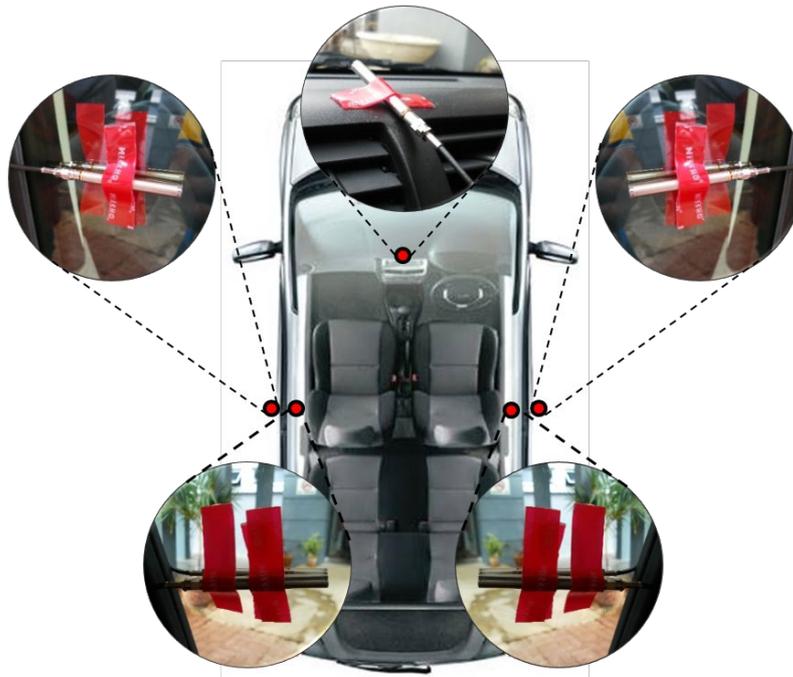
Figure 7. The location of the measurement microphones.
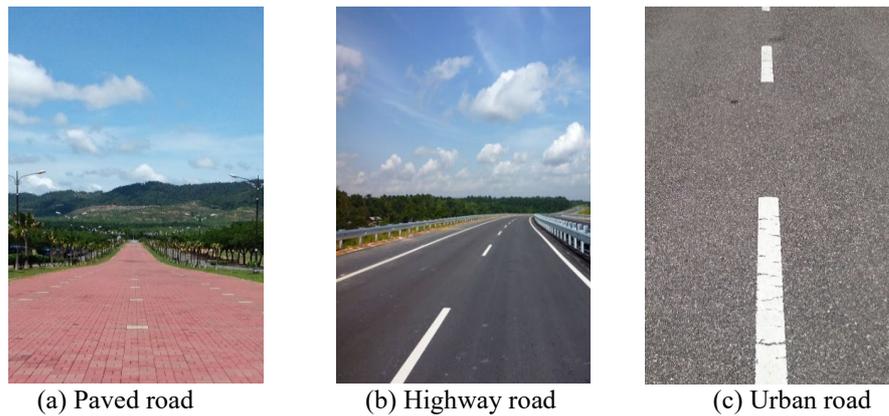
| (a) Paved road | (b) Highway road | (c) Urban road |

Figure 8. Different road types.

Jury Test

After conducting the noise recording for all three cars, the noise recordings were used for subjective assessment. In this subjective assessment, five individuals were selected as juries. The juries must be someone who does not have hearing problems. In the evaluation process, these evaluators will need to listen to the recorded noise sound for each car using a headphone. The listening process needs to be conducted in a quiet close room to avoid background noise. Before the juries began evaluating the noise, they were first exposed to two extreme noises which were the lowest and the loudest noise. This step was done to let the juries aware of the reference noise level. The juries then evaluated the noise based on the annoyance level, as shown in Table 1.

Table 1. Level of noise annoyance

Annoyance Level	Annoyance circumstances
1	Very pleasant
2	Pleasant
3	Marginal
4	Annoying
5	Very annoying

After the tests have been done for all five juries, the average of the annoyance level for every engine speed was obtained, and these average values are known as Vehicle Acoustical Comforts Factor (VACF). Then a multiple regression

analysis was performed against the obtained VACF with the sound quality parameter at each engine speed. The regression analysis was done for all cars and conditions. Vehicle Acoustical Comforts Index can be expressed as:

$$\alpha = \partial_\gamma * \gamma + \partial_\lambda * \lambda + \partial_\rho * \rho + \varepsilon \tag{1}$$

Where, ∂_γ is the coefficient for loudness sound quality metrics, ∂_λ is the coefficient for sharpness sound quality metrics, ∂_ρ is the coefficient for sound pressure sound quality metrics, ρ is the sound pressure in a noise recording (Pa), γ is the sound quality Loudness metrics in a noise recording (sone), λ is the sound quality Sharpness metrics in a noise recording (acum) and ε is constant.

In the regression analysis, the p-value is the indicator to show the level of significance among the involved factors in the equation. In this case, any factor which has the p-values that are equal and less than 0.05 were assumed as statistically significant. If there is any factor that is not significant in Eq. (1), it was then excluded from the equation. After eliminating the insignificance sound parameter factors, a new level of annoyance was obtained by using Eq. (1). To increase the efficiency of the equation, it is necessary to repeat the regression analysis by excluding the non-significance sound quality factors.

Data Mining

Clustering

The data mining is divided into three major parts; the first part is the clustering process where the clustering tools that were used in this study are the hierarchical clustering. Before the clustering process was done, the sound quality parameters were extracted from each recorded noise using the B&K Sound Quality software. Here, the sound quality parameters extracted from the noise are the loudness, sharpness, and the sound pressure.

After the parameters of sound quality from the noise were extracted, the recorded noise was clustered into five categories. The first class is labelled as the most pleasant, followed by lighter annoying classes, while the fifth class is the most annoying noise (Table 1). After the noise had been clustered, the noise dose in each cluster class was presented. The method that was used in hierarchical clustering was the average method. The details of the process are shown as follows:

i.   begin by creating or pilling all the components data in one single group of data and assign each component data into its own cluster.
ii.   the closest pair of cluster is found and it is merged into a single cluster.
iii.   the distance from each cluster to a new cluster is computed.
iv.   distance between all pairs of component data in different cluster were measured and averages all the distances.
v.   the smallest average distance between one component data in a cluster to another component data in the different cluster was used to combine into one cluster.
vi.   repeat the process until all component are considered.

Monte Carlo Simulation

The next part of data mining is the Monte Carlo simulation. Generally, Monte Carlo simulation can be described as a process to generate random data in a huge amount. After obtaining the mean and standard deviation for each sound quality parameter in each cluster class, the values were used to generate substantial random data. The new random data were produced in Microsoft Excel, and these random data were generated with respect to each cluster class, the mean and standard deviation of each sound quality. The random data were generated by using the normal distribution approach. The number of random data that was generated is 1000 for every sound quality parameter and for every cluster class. Altogether, 5000 random data were generated from five clusters of class. The definition of generating random data can be explained as follows:

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \in \mathbb{R}^N$. The expected of x, $E(x) = \begin{bmatrix} E(x_1) \\ E(x_2) \\ \vdots \\ E(x_N) \end{bmatrix} \in \mathbb{R}^N$, and the $\sum(x)$ is the covariance, where

$\sum = E\left((x - \widehat{\mu})(x - \widehat{\mu})^T\right) \in \mathbb{R}^{N \times N}$ and $\widehat{\mu}$ is the average variable, where $\widehat{\mu} = E(x) \in \mathbb{R}^N$. The generation of the random data from x until N can be expressed as Eq. (2).

$$f\left(x \mid \widehat{\mu}, \sum(x)\right) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} . e^{\left\{ -\frac{(x-\widehat{\mu})^T \Sigma^{-1}(x-\widehat{\mu})}{2} \right\}} \tag{2}$$

Classification

The classification conducted after random data were generated by Monte Carlo simulation. Here, the classification tool used was K-Nearest Neighbour (KNN). KNN is a process that uses supervised learning to classify new data according to the majority of the K-nearest neighbour in a specific category. KNN classification was used to obtain the classification accuracy of the data. The process of the classification is explained as follow:

i.   Determine the set of data that needed to classify and make it as an input. As in this study, it is the new generated random data from Monte Carlo Simulation. Denote the input as $D = \{(x_1, c_1), ..., (x_N, c_N)\}$ and the new component to classify as $x = (x_1, ..., x_n)$.

ii. Calculate the amount of data for each label class. For our case, it is 1000 for each sound quality parameter and each cluster class.

iii. Arrange the input data accordingly from small to large.

iv. Choose the nearest component to the K for $x : D_x^K$

v.   Allocate the most x class within $x : D_x^K$ into a group

Comparison

Three result groups were referred for comparing the interior noise between these three cars. For the objective evaluation, sound pressure level generated inside the car cabin was compared for all cases. The sound pressure values were obtained through equation received by conducting the regression analysis. The second result is the comparison through subjective evaluation where the level of annoyance from the interior noise is considered as the factor to compare the cars. The level of annoyance is obtained from the VACI value that was generated in the jury test. The comparison for the first and second result groups is performed by implementing the ANOVA test. Lastly, the last result is based on the amount of noise sorted in each cluster class for each car.

The first part of the result is to produce an acoustical comfort index for each car. The second part is the result is the measurement of noise in the cabin through the data mining approach by clustering and classification method. Finally, the last segment is the comparison of the interior noise between the three cars by referring to the result of ANOVA.

RESULT AND DISCUSSION

The results were acquired after extracting the sound quality from the recorded noise by using the B&K sound Quality software. The sound quality from the recorded noise are sound pressure, loudness, and sharpness. The sound quality data were arranged according to the type of cars and the state of the cars. First, the data for all three cars in a stationary state as shown in Table 2. From the table, it shows that the noises of the tested cars increase as the engine speed increase. Notice that at some point, the value of the sound metric is sometimes higher or lower at certain engine speed. This could possibly be due to the undesired inference from the echoed noise of cars itself with the surrounding environment or it could also due to the poor performance of the cars itself where it sometimes generated unusual vibration which led to higher noise. Another possible factor was due to the computer used during the experiment, where the computer starts to heat up and to cause the cooling fan to operate louder than usual.

Simple regression analysis was conducted separately to test the significance of the sound pressure RMS with loudness and between sound pressures RMS with sharpness. The result of the significance is shown as p-value in Table 3. After it is confirmed that the predictor variables are significance with the dependent variables, a multiple regression analysis are applied to generate an equation, where the sound pressure RMS is the dependent variable while loudness and sharpness are the predictors variable. Through the regression analysis, acoustical comfort index is obtained as Eq. (3). The constants and the weighted values for coefficients for each car at each condition are presented in Table 3.

$$\rho = \partial_\gamma * \gamma + \partial_\lambda * \lambda + \varepsilon \tag{3}$$

Where, ρ is the sound pressure rms in a noise recording (Pa), ∂_γ is the coefficient for loudness sound quality metrics, ∂_λ is the coefficient for sharpness sound quality metrics, γ is the sound quality loudness metrics in a noise recording (sone), λ is the sound quality sharpness metrics in a noise recording (acum) and ε is constant.

The jury assessment results were obtained for getting the values of Vehicle Acoustical Comfort Factor (VACF) (in Table 4 and Table 6). In the tables, as the engine speed increase, the jury assessment also increase, at most cases, but there are also some case where a subject responded with a high level of annoyance when exposed to slower engine speed, here it can be said that subjective assessment is affected by the jury feeling or maybe the jury was confused when responding. The VACF is the average values of the evaluation towards the exposed recorded noise. This VACF was used in regression analysis to obtain Vehicle Acoustical Comfort Index (VACI). Here the assessment of the juries is made based on the level of annoyance related to the noise that was exposed to the juries. The level of annoyance can be explained as in Table 5. After obtaining the VACF value, a multiple regression analysis is performed to produce an equation. The

VACF was the dependent variable while the loudness, sharpness and sound pressure RMS sound quality were the predictors variable. Then, a vehicle acoustical comfort index (VACI) will be produced based on Eq. (3).

Table 2. Sound quality value for all driver interior sensors in a stationary state.

		Axia driver interior sensor			Myvi driver interior sensor			Viva driver interior sensor	
Speed	Sound Pressure RMS (Pa)	Loudness (sone)	Sharpness (acum,Z)	Sound Pressure RMS (Pa)	Loudness (sone)	Sharpness (acum,Z)	Sound Pressure RMS (Pa)	Loudness (sone)	Sharpness (acum,Z)
1000	422	757	0.737	523	1450	0.936	705	1150	0.801
1000	425	748	0.719	453	1300	0.911	809	1170	0.806
1000	421	778	0.715	865	1250	0.858	890	1040	0.777
1250	330	933	0.726	851	2820	1.37	1550	2450	1.27
1250	330	906	0.713	860	2910	1.39	1510	2400	1.26
1250	297	898	0.722	833	2970	1.36	1460	2150	1.18
1500	282	1100	0.77	585	2950	1.4	1930	3190	1.42
1500	401	1240	0.83	844	2890	1.38	1770	3070	1.42
1500	385	1180	0.812	571	2890	1.38	1760	3170	1.43
1750	301	1080	0.77	729	1230	0.824	2570	5040	1.62
1750	216	1100	0.773	239	1120	0.782	2520	5030	1.62
1750	228	1120	0.783	221	1110	0.776	2620	5120	1.62
2000	294	1360	0.877	519	2910	1.39	582	2340	1.25
2000	258	1350	0.874	714	3120	1.43	617	2320	1.25
2000	254	1360	0.878	715	3000	1.40	655	2380	1.27
2250	307	1330	0.866	512	2820	1.37	873	2870	1.37
2250	282	1340	0.87	870	2950	1.39	762	2700	1.33
2250	326	1360	0.88	875	2950	1.39	835	2750	1.35
2500	544	1810	1.07	427	3110	1.42	453	1530	0.952
2500	553	1730	1.03	887	3090	1.42	525	1570	0.966
2500	559	1810	1.07	862	2990	1.40	540	1500	0.956
2750	425	2030	1.15	1460	5680	1.66	680	2710	1.34
2750	429	2070	1.16	1350	5590	1.65	721	2730	1.35
2750	422	2040	1.15	1360	5650	1.66	673	2750	1.35
3000	499	2240	1.22	1870	7210	1.78	541	1960	1.14
3000	459	2270	1.23	1870	7230	1.77	547	1930	1.12
3000	369	2200	1.21	1840	7330	1.78	551	1970	1.14

Table 3. Constant and coefficient values of the obtained index for all cars at all state.

	Axia			Myvi			Viva		
Condition	∂_γ	∂_λ	ε	∂_γ	∂_λ	ε	∂_γ	∂_λ	ε
Stationary	- 0.842	2576	- 788	0.306	- 535	569	0.913	- 1923	1142
Highway	1.17	- 2034	766	2.27	- 18963	21938	1.59	- 6638	6469
Urban	1.26	- 2533	1179	2.09	- 10156	8767	1.77	- 6438	5412
Pavement	1.12	- 2200	1163	1.17	- 5034	4757	1.65	- 5766	5151

In the multiple regression analysis, the P-value for each car at each condition is referred. Any sound quality factor which is not significant was excluded from the equation. Here, the results for any significance factor for all cars in all states are presented in Table 7. Here, at a certain state, some sound metric is excluded from the equation since in the regression analysis, those parameters are insignificance and if included in the equation it is unreliable. From the equations that were obtained for each car and in each state, the value for VACI can now be obtained by substituting the sound quality metrics value into the related equation. Here, lower VACI value means that the car has better acoustical comforts while higher VACI value is the opposite of it. This depends on the assessment that was performed by the juries toward the recorded noise with respect to the level of annoyance (Table 5). In other words, 'low level of annoyance' means that the evaluated noise annoyed the jury less than the 'higher level of annoyance'.

The proposed formulas of Vehicle Acoustical Index for each car at different condition and roads are given in Table 3. Meanwhile, the values of VACI for each car in each state are presented in Table 11, and the table shows that the values of the obtained VACI increase as the engine speed of the tested cars increase and rougher road surface also contributed in producing higher VACI values. Notice the data in pavement and urban road for engine speed 2500 rpm and above were

left blanked due to that at that speed, it is due to the safety issue; therefore the urban and paved road were tested until an engine speed of 2250 rpm only.

Table 4. Sample of results of the jury assessment for all cars in stationary state.

Car	Speed	Number of Subjects					VACF
		1	2	3	4	5	
Axia	1000	1	1	3	1	1	1.4
	1250	1	1	1	1	2	1.2
	1500	2	1	2	1	2	1.6
	1750	1	2	2	1	2	1.6
	2000	1	1	2	2	1	1.4
	2250	2	3	3	1	1	2.0
	2500	1	3	2	1	3	2.0
	2750	3	3	3	1	3	2.6
	3000	3	3	3	2	3	2.8
Myvi	1000	1	2	2	1	1	1.4
	1250	1	2	4	1	2	2
	1500	2	2	2	1	2	1.8
	1750	2	2	3	1	2	2.0
	2000	2	2	3	1	2	2.0
	2250	1	2	4	1	2	2.0
	2500	3	5	5	2	4	3.8
	2750	5	5	5	3	4	4.4
	3000	5	4	5	3	4	4.2
Viva	1000	1	3	3	1	2	2.0
	1250	3	3	3	2	2	2.6
	1500	3	4	3	2	2	2.8
	1750	3	3	4	2	2	2.8
	2000	3	4	4	2	2	3.0
	2250	4	4	4	3	3	3.6
	2500	4	4	5	3	3	3.8
	2750	5	4	5	4	4	4.4
	3000	4	5	5	3	4	4.2

Table 5. Level of annoyance used by the juries to evaluate the noise.

Most Pleasant	Pleasant	Marginal	Annoying	Most Annoying
1	2	3	4	5

Table 6. Sample of results of the jury assessment for all cars in the non-stationary state (highway).

Car	Speed	Number of Subjects					VACF
		1	2	3	4	5	
Axia	1000	1	1	3	1	2	1.6
	1250	1	1	2	1	3	1.6
	1500	1	1	3	1	4	2.0
	1750	2	2	4	2	4	2.8
	2000	3	2	3	2	4	2.8
	2250	3	3	4	4	5	3.8
	2500	4	3	5	4	4	4.0
	2750	4	4	4	3	4	3.8
	3000	4	3	3	3	4	3.4
Myvi	1000	2	1	4	2	3	2.4
	1250	3	2	5	2	4	3.2
	1500	3	2	5	2	4	3.2
	1750	4	3	5	3	5	4.0
	2000	4	4	4	4	5	4.2
	2250	4	4	5	5	5	4.6
	2500	5	4	5	5	5	4.8
	2750	5	5	5	5	5	5.0
	3000	5	5	5	5	5	5.0
Viva	1000	2	1	3	2	3	2.2
	1250	2	1	3	1	3	2.0
	1500	3	2	4	3	4	3.2
	1750	3	3	4	2	5	3.4
	2000	3	3	3	3	4	3.2
	2250	4	3	4	4	4	3.8
	2500	4	4	3	4	4	3.8
	2750	4	3	3	3	4	3.4
	3000	5	3	3	4	4	3.8

Meanwhile, for the results of clustering approach, the means value for each sound quality metrics in each cluster class are presented in Table 9, while the standard deviations for each cluster class for each sound quality metrics are presented in Table 10. From the result of clustering, the total amount of noise is sorted in each cluster class. The most pleasant class of noise is those with the least annoying or the lowest loudness. From the clustering result, the amount of noise clustered in each class for each car is obtained and presented in Table 12.

Table 7. Significance factor in multiple regressions for all cars in each state.

Cars	Stationary	Highway	Urban	Pavement
Axia	Sharpness	Sound pressure RMS	Loudness	Loudness, sharpness, sound pressure RMS
Myvi	Loudness	Loudness, sharpness, sound pressure RMS	Sound pressure RMS	Sharpness
Viva	Loudness, sharpness	Sharpness	Sharpness	Loudness, sharpness

Table 8. Generated vaci (α) equation for all cars in each state.

	Stationary
Axia	$\alpha = -0.674 + 2.76\,\lambda$
Myvi	$\alpha = 0.974 + 0.000492\,\gamma$
Viva	$\alpha = 17.7 + 0.00259\gamma - 16.7\lambda$
	Highway
Axia	$\alpha = 0.419 + 0.00169\,\rho$
Myvi	$\alpha = -38.6 + 0.00124\rho - 0.00266\gamma + 32.8\,\lambda$
Viva	$\alpha = -5.27 + 5.46\,\lambda$
	Urban
Axia	$\alpha = 0.749 + 0.000733\gamma$
Myvi	$\alpha = 1.1 + 0.00104\,\rho$
Viva	$\alpha = -3.14 + 4.43\lambda$
	Pavement
Axia	$\alpha = -7.17 + 0.00174\rho - 0.00187\,\gamma + 9.38\lambda$
Myvi	$\alpha = -7.43 + 0.000121\,\gamma + 6.73\,\lambda$
Viva	$\alpha = -7.33 + 6.70\,\lambda$

Table 9. Amount of noise and means for each sound quality metrics in each class

Cluster Class	Count	Sound Pressure RMS (Pa)	Loudness (sone)	Sharpness (acum, Z)
Most Pleasant	38	468.03	1220	0.84
Pleasant	102	1126.58	2712.5	1.32
Marginal	112	2924.09	4679.02	1.59
Annoying	6	1625	6448.33	1.72
Most Annoying	12	6326.67	7075.83	1.71

Table 10. Amount of noise and standard deviation for each sound quality metrics

Cluster Class	Count	Sound Pressure RMS (Pa)	Loudness (sone)	Sharpness (acum, Z)
Most pleasant	38	201.018	230.38	0.081
Pleasant	102	487.594	458.479	0.112
Marginal	112	814.316	674.845	0.052
Annoying	6	260.519	886.892	0.066
Most annoying	12	1034.297	470.849	0.027

From Table 12, it can be determined that Axia is the car that produces the lowest noise inside the car cabin among the three cars that were tested in all conditions. It is shown that Axia has the highest total noise counted in class 1 (most pleasant noise). The results obtained also show that in all the tested experiments, Axia is the only model that produces noise up to class 3 (marginal noise level), compared to Myvi and Viva model. Meanwhile the car that produces the highest noise in the car cabin during all the tested experiment is Viva. This can be seen in the Table 12 where in cluster class 5 (most annoying class), Viva has the highest count of '9' while Myvi only have '3'. In this experiment, Viva is the oldest version follow up by Myvi and Axia is the newest version developed by the manufacturer, therefore, it is logical to assumed that due to the degradation of performance over time, it could be the reason why Viva produced highest noise

among the other cars, while Axia produced lowest noise could probably because the manufacturer have done some noise optimisation to improved their product other than because the car is newer than the other two cars.

Table 11. VACI value for each car in every state.

Speed	Stationary			Highway			Urban			Pavement		
	Axia	Myvi	Viva	Axia	Myvi	Viva	Axia	Myvi	Viva	Axia	Myvi	Viva
1000	1.36	1.69	2.25	1.35	2.49	2.21	1.61	2.04	2.22	1.38	2.28	1.85
	1.31	1.61	2.27	1.28	2.32	2.16	1.60	1.97	1.95	1.58	2.14	1.85
	1.30	1.59	2.25	1.68	2.53	2.10	1.59	1.88	2.22	1.95	2.21	1.92
1250	1.33	2.36	2.50	1.73	3.41	2.10	1.95	2.05	2.57	2.05	2.80	3.26
	1.29	2.41	2.36	1.64	3.87	2.10	2.00	2.06	2.57	1.86	2.73	3.19
	1.32	2.44	2.46	1.59	3.44	1.99	1.90	2.23	2.75	2.06	2.81	3.12
1500	1.45	2.43	2.97	2.79	3.50	3.25	2.97	2.18	3.64	3.44	3.04	3.32
	1.62	2.40	2.84	2.28	3.39	3.25	2.89	2.26	3.59	3.55	2.88	3.39
	1.57	2.40	2.76	2.65	3.19	3.36	2.97	2.33	3.68	3.82	3.18	3.32
1750	1.45	1.58	3.18	2.79	4.14	3.47	3.34	3.09	3.86	3.74	3.34	3.99
	1.46	1.53	3.23	2.70	3.87	3.47	3.40	2.94	3.95	3.96	3.27	4.06
	1.49	1.52	3.23	2.62	3.73	3.47	3.48	3.07	3.99	4.05	3.35	4.06
2000	1.75	2.41	3.35	3.12	4.54	3.52	3.77	4.23	4.04	4.54	4.63	3.99
	1.74	2.51	3.16	2.72	3.99	3.41	3.86	4.12	4.04	4.23	4.71	3.93
	1.75	2.45	3.56	3.19	4.21	3.52	3.81	4.24	4.04	4.40	4.64	3.93
2250	1.72	2.36	3.16	3.26	4.56	3.41	3.73	4.52	4.13	4.42	5.03	4.40
	1.73	2.43	3.36	3.78	4.27	3.47	3.93	4.15	3.99	4.15	5.05	4.40
	1.75	2.43	3.30	3.19	4.81	3.52	4.01	4.62	4.04	4.49	5.02	4.40
2500	2.28	2.50	3.54	4.07	4.55	3.47						
	2.17	2.49	3.35	3.95	4.29	3.52						
	2.28	2.45	3.53	3.75	4.64	3.52						
2750	2.50	3.77	4.35	3.36	4.82	3.68						
	2.53	3.72	4.35	3.60	4.73	3.68						
	2.50	3.75	4.12	3.48	5.19	3.68						
3000	2.69	4.52	4.19	3.63	5.04	3.68						
	2.72	4.53	4.47	3.43	5.62	3.74						
	4.61	7.15	3.46	3.88	4.99	5.01						

Table 12. Number of noise sorted in each cluster class for each car.

Cluster Class	Axia	Myvi	Viva
Most Pleasant	26	6	6
Pleasant	34	35	33
Marginal	30	40	42
Annoying	0	6	0
Most Annoying	0	3	9

Later, the KNN classification approach is performed to test the accuracy of the generated cluster classes in the previous stage. Ten trials were done wherein each trial, the KNN calculated the accuracy in percentage for every five cluster classes and repeated for ten iterations. The average accuracy for each trial and iteration are presented in Table 15. It can be said that the cluster class are highly accurate, considering the accuracies are all above 95%.

The comparison between the three models of cars which are Axia, Myvi, and Viva are shown accordingly by the state/condition of the cars starting with the stationary state, highway, urban and paved road. Firstly, the objective evaluation is compared based on the predicted sound pressure level by Eq. (3) for each car in each state. Secondly, the subjective evaluation is compared in the aspect of noise annoyance. In this case, VACI values are referred for each car model in each state. The comparison is made with respect to the amount of data in each cluster class.

Figure 9 shows a comparison of the three cars in stationary and driven condition on highway road. Figure 9(b) shows that Axia produced the lowest noise compared to other models during highway test. In overall, Viva model produced the highest noise in all states during the recording. In Figure 9(a) and 9(b) as well, looking at Axia's data the noise while driving on pavement road, urban road and highway road increase consecutively. This means that the rougher the road surface, the higher the noise that generated inside the car cabin. Here the roughest road surface is the paved road. Hence, it can be said that Axia has the best interior acoustical comforts among these three cars where it also produces the lowest sound pressure level. This is followed by MyVi and Viva.

Table 13. Accuracy test with KNN classification approach.

Iteration	Average accuracy for each trial (%)									
	1	2	3	4	5	6	7	8	9	10
1	98.06	97.89	98.2	98.18	98.2	98.16	98.16	98.18	98.32	98.15
2	98.2	98.17	98.1	98.25	98.16	98.43	98.08	98.11	98.26	98.41
3	98.57	98.42	98.43	98.46	98.56	98.61	98.55	98.57	98.71	98.48
4	98.47	98.48	98.46	98.52	98.68	98.58	98.63	98.44	98.63	98.75
5	98.56	98.46	98.58	98.67	98.74	98.67	98.58	98.57	98.41	98.54
6	98.76	98.57	98.64	98.66	98.63	98.55	98.65	98.53	98.65	98.74
7	98.6	98.64	98.67	98.76	98.51	98.61	98.56	98.71	98.57	98.62
8	98.64	98.75	98.55	98.62	98.75	98.74	98.82	98.72	98.61	98.68
9	98.7	98.64	98.61	98.6	98.6	98.77	98.73	98.61	98.6	98.63
10	98.63	98.58	98.73	98.61	98.63	98.67	98.78	98.8	98.61	98.81

(a)   (b)
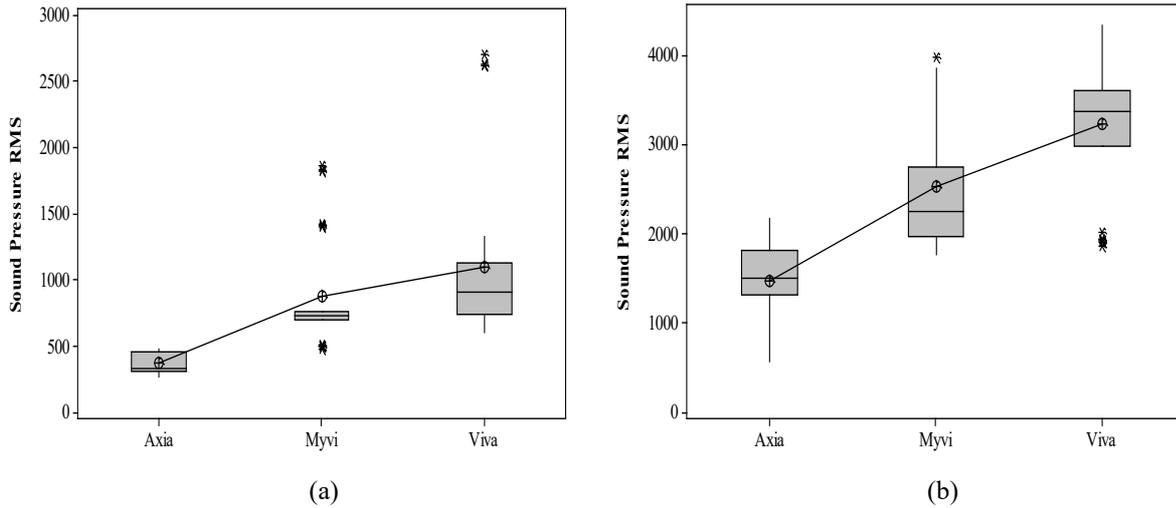
Figure 9. Comparison of sound pressure at (a) stationary state and (b) highway road.

Figure 10(a) and 10(b) graph compare for subjective evaluation. Here, the comparison is made based on the level of annoyance that was obtained using the generated VACI equation for each car in each state. Based on the figures, it shows that Axia is the car that has the best acoustical comfort. From the objective evaluation, Axia generated the lowest sound pressure level in the interior cabin. The subjective evaluation indicates that Axia has the lowest level of annoyance in all conditions tested in this study.

(a)   (b)
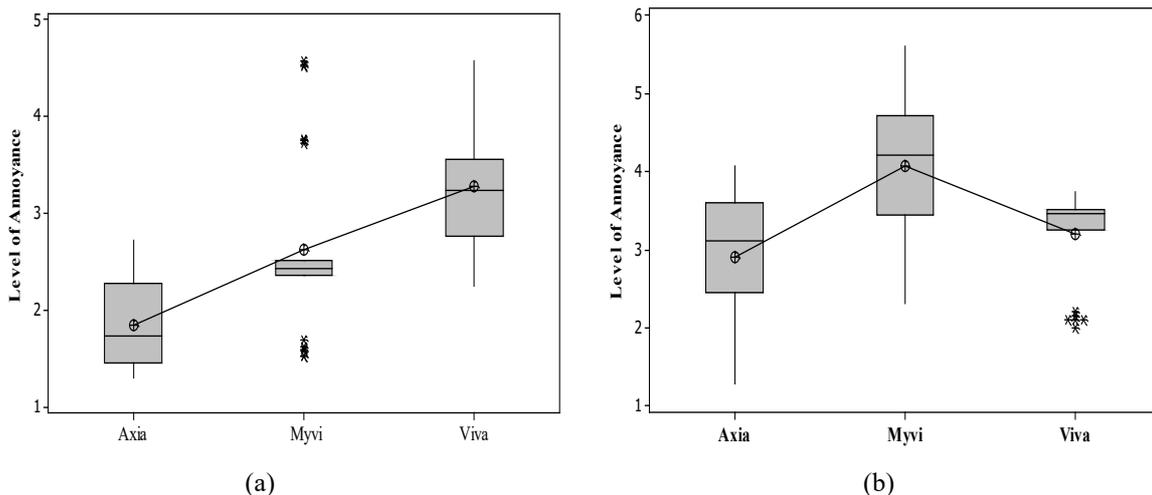
Figure 10. Comparison of the level of annoyance at (a) stationary and (b) non-stationary state (highway).

Based on clustering results, cluster class 1 is considered 'most pleasant noise', class 2 is 'pleasant noise', class 3 is 'marginal noise', class 4 is 'annoying noise', and class 5 is 'annoying noise'. In Figure 11, it can be seen that Axia has the highest amount of noise counted for cluster class 1 where the rest lies in cluster class 2 and 3. This means that in term of the quieter cabin, Axia is better than Myvi and Viva. It can be said also that there are not many differences between Myvi and Viva in term of the amount of noise categorised in each cluster class. However looking at the amount of noise in class 5 (the most annoying level), the amount of noise counted for Myvi is only '3' while for Viva it is '9'. Therefore, Viva is seen as having the lowest interior acoustic comfort among the three cars (Figure 11).
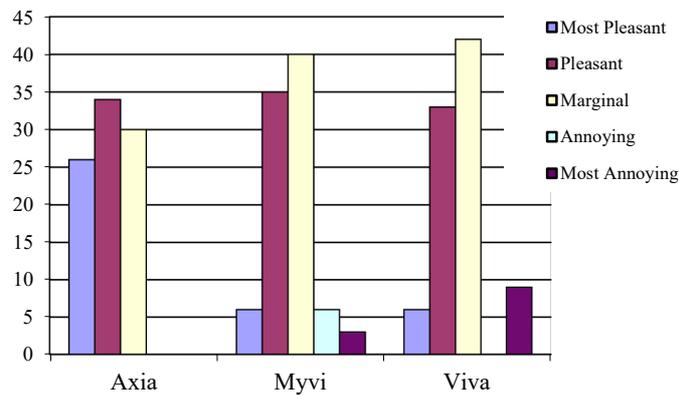
Figure 11. The amount of data in each cluster class for each model of cars.

Based on the results, there are some situations where the noise generated unaccustomed as it supposed to respond with the engine speed due to external disturbance like the sound of the surrounding vehicles. When driven on the highway road, and uneven surface, echo noise reflected by the surrounding building surface. The car itself created unusual vibration or from the computer's cooling fan. However, data mining approach can still be considered as one of the methods to study vehicle interior noise.

CONCLUSION

From the experiment and data analysis, vehicle acoustic comfort index was produced for the three cars used in this study. This index may be used by engineers or researchers as a guide to improving vehicle cabin noise. In addition, the acoustic quality among the three cars is presented using ANOVA. Annoyance index of the three cars is generated using the data mining method. From the results, Axia car model has the best acoustic comfort by objective evaluation. By subjective evaluation, it recorded as the lowest level of annoyance and produced the 'low noise level' rating among the three. Consistently ranked second to Axia in terms of the noise comfort criteria is the car from Myvi brand. However, Myvi ranked last in the subjective evaluation, possibly because of aerodynamic noise as this car run faster than Viva at similar engine speed. Since the test was done by using engine speed as one of the parameters, the difference in gear ratio leads to the different vehicle speed. Another reason for Myvi to perform poorly in the annoyance index is due to other perceived noise during the driving test. This other noise was suspected to originate from the suspension. This bias factor could be eliminated if one to use newer cars in the test. In the same way, as the sample of cars was limited the said sound pressure level and annoyance index may differ if one to use a larger sample. However, the technique used to measure and evaluate vehicle interior noise in this study is found to be feasible.

REFERENCES

[1]   Da Silva MG. Measurements of comfort in vehicles. Measurement Science and Technology. 2002; 13(6):R41.

[2]   David SK, Saeb AT, Rafiullah M, Rubeaan K. Classification Techniques and Data Mining Tools Used in Medical Bioinformatics. In: Big Data Governance and Perspectives in Knowledge Management 2019; pp. 105-126. IGI Global.

[3]   Deulgaonkar VR, Kallurkar SP, Mattani AG. Review and diagnostics of noise and vibrations in automobiles. International Journal of Modern Engineering Research 2011; 1(2):242-6.

[4]   Dilla EM, Wang X, Subic A. In-situ noise level mapping for vehicle floor carpet. International Journal of Vehicle Noise and Vibration 2009; 5(3):193-204.

[5]   Genuit K. Vehicle interior noise—combination of sound, vibration and interactivity. Sound and Vibration. 2009; 43(12):8.

[6]   Jacob SG, Ramani RG. Evolving efficient classification rules from cardiotocography data through data mining methods and techniques. European Journal of Scientific Research 2012; 78(3):468-80.

[7]   Junoh AK, Mohd Nopiah Z, Muhamad W, Azman WZ, Nor MJ. Multi objective optimization of noise and vibration in passenger car cabin by using goal programming approach. Advanced Materials Research 2012; 383: 976-983.

[8]   Krajniak J, Grega R. Application of apparatus–noiseness analyser on laboratory conditions. Zeszyty Naukowe. Transport/Politechnika Śląska 2014.

[9]   Lee JW, Giraud-Carrier C. Automatic selection of classification learning algorithms for data mining practitioners. Intelligent Data Analysis 2013; 17(4):665-78.

[10]   Li M, Van Keulen W, Ceylan H, van de Ven M, Molenaar A. Influence of changes in surface layer properties on tire/pavement noise. Noise Control Engineering Journal 2013; 61(4):417-24.

[11]   Liptai P, Moravec M, Lukacova K. Qualitative evaluation and identification of defects in the automotive industry using visualization techniques. Center for Quality 2014.

[12]   Nopiah ZM, Junoh AK, Ariffin AK. Noise annoyance fuzzy index in passenger car cabin. International Journal of Vehicle Noise and Vibration 2013; 9(3-4):216-33.

[13]   Nopiah ZM, Junoh AK, Ariffin AK. Vehicle interior noise and vibration level assessment through the data clustering and hybrid classification model. Applied Acoustics 2015; 87:9-22.

[14]   Nopiah ZM, Junoh AK. A review on data mining approaches for clustering, classifying and optimizing the level of sound and vibration in vehicle cabin. International Information Institute (Tokyo). Information. 2015 Aug 1; 18(8):3693.

[15]   Paulraj MP, Andrew AM, Yaacob S. Classification of interior noise comfort level of proton model cars using artificial neural network. In: International Conference on Advanced Computing and Communication Systems 2013; pp. 1-7.

[16]   Petrova E, Pauwels P, Svidt K, Jensen RL. In search of sustainable design patterns: Combining data mining and semantic data modelling on disparate building data. In: Advances in Informatics and Computing in Civil and Construction Engineering 2019; pp. 19-26.

[17]   Putra A, Munir FA, Juis CD. On a simple technique to measure the airborne noise in a car interior using substitution source. International Journal of Vehicle Noise and Vibration 2012; 8(3):275-87.

[18]   Rahman MM, Ahmed CF, Leung CK. Mining weighted frequent sequences in uncertain databases. Information Sciences. 2019; 479:76-100.

[19]   Raychaudhuri S. Introduction to Monte Carlo simulation. In: 2008 Winter Simulation Conference 2008; pp. 91-100.

[20]   Yongman L, Tusheng L. Engine noise of identification using an improved blind sources separation algorithm. In: Third International Conference on Natural Computation 2007; pp. 759-763.

[21]   Zhang H, Xie X, Jiang J, Yamashita M. Assessment on transient sound radiation of a vibrating steel bridge due to traffic loading. Journal of Sound and Vibration 2015; 336:132-49.

[22]   Zuo S, Yan J. Experimental analysis for the interior noise characteristics of the fuel cell car. In: 2006 IEEE International Conference on Vehicular Electronics and Safety 2006; pp. 241-245.