inter.noise
HAMBURG 2016

Estimation of Electric Shaver Sound Quality using Artificial Neural Networks

Serkan Atamer[1]; M. Ercan Altinsoy[2]

Institute of Acoustics and Speech Communication, Chair of Acoustics and Haptics, TU Dresden

ABSTRACT

For competitive markets, sound quality is an important attribute of a particular product that effects end-user preferences which discriminates rival products. Traditional sound quality assessments include subjective jury testing which is both expensive and time consuming. However, contemporary studies in that field mostly focus on developing new methodologies that can replace subjective jury testing, mainly artificial neural networks (ANNs). Main reason underlying this search is to compensate the shortcomings of jury testing.
In this study, artificial neural networks are used to predict annoyance estimations of electric shaver/trimmer sounds. Psychoacoustic parameters obtained from different shaver/trimmer recordings by using different psychoacoustical models and subjective annoyance estimations gathered from jury testing are used as inputs and outputs of a neural network, respectively. With the correlations between input and output values obtained from jury testing, neural networks are trained and correlation rates are evaluated. Some of the sound samples are used for training the neural network while the rest of the data are used to verify the accuracy of the artificial neural network. Conclusive remarks are included at the end of the study related to idea of replacing the jury testing with neural networks with possible strengths and shortcomings.

Keywords: Psychoacoustics, sound quality, annoyance, electric shaver sound, neural networks

1. INTRODUCTION

Soundscape of our living environments are mostly disturbed or altered by the noise emitting from household devices. During the operation, feedbacks from the equipment – not only auditory but sometimes visual and tactile – provide information to the end user about the operational condition of this equipment in use. Hence, besides being annoying, sound can also be an information carrier in most of the cases. Sound of an equipment gives the end user a perception of quality, durability and liability of that particular product. For that reason, sound quality studies of household devices become more important as they help producers to gain advantage in market.

Electric shaving and trimming devices are being used nearly every day by the most of the society. Different brands with different designs in market are competing with each other to take the advantage. As the sensitivity of noise arises each year, sound emitting from electric shavers are becoming an important aspect in market behavior.

So far, sound quality estimations and annoyance and pleasantness indexing are performed by subjective jury evaluations representing the consumer profile and linear regression models between psychoacoustical parameters and jury evaluations (1–3). Even though that method gave reliable results so far, has a disadvantages related to practical limitations on the number of sounds that can be evaluated in each group and linearity of the dataset (4). Moreover, those studies assume that there is a relation between some set of parameters and annoyance estimations, which can easily be representable mathematically, mostly in the form of a linear regression. However, decision making might be a phenomena which is difficult to properly formulate the model representing relation or set of rules. To overcome those shortcomings, alternate methods are being developed every day. Recently, artificial neural network models (ANN) are gaining popularity in mostly automotive engineering field and fair developments are obtained in different sound quality studies so far. ANN is used in the work of Lee et

[1] serkan.atamer@tu-dresden.de
[2] ercan.altinsoy@tu-dresden.de

al. (5) for objective evaluation of the rumbling sound within passenger cars and to obtain a rumbling index. Pellegrini and Baars also researched the reliability of using ANNs to represent jury response (6). Lee and his colleagues also used artificial neural networks for obtaining booming index in cars (7). Wang et al. also described wavelet pre-processing neural network model (WT-NN) for sound quality predictions of nonstationary vehicle interior noise (8). Yildirim and Eski also used hybrid neural networks for evaluating sound quality of cars during road test (9). Lee et al. (10) used also ANNs to obtain a metric for quantification of axle gear whining sounds in SUVs. Wang et al. described good correlation between calculated values of psychoacoustic indices and evaluated annoyance values for vehicle noise in their work (11). Same procedure is also being followed in the work of Duan et al. (12). Tendency in the literature shows a great potential of using intelligent systems in sound quality estimations. Main aim of this study is to investigate the opportunities that ANNs provide in sound quality estimations and implement that for the sound quality estimations of electric shaver sounds.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are developed by imitating the biological structure and operating condition of a biological neuron. There is a vast literature in terms of artificial neural networks. They are mostly used in different studies to obtain complex curve fitting, data clustering, speech recognition, image recognition purposes etc. Within this study, artificial neural networks are used to mimic human decision making process in terms of annoyance estimations. Main assumption behind this study is that there is a – either linear or nonlinear – pattern behind the human annoyance estimation process, as long as the context and the borders of the estimation set is kept well defined. For that reason, the process is going to be nothing but a curve fitting approach. ANN tailored in this study is mostly efficient for data fitting procedure, since the main assumption of this work is that there is a understandable and definable mapping between some set of parameters and the annoyance estimations. As common sense implies and to keep the descriptive set as relative and consistent to each other as possible, psychoacoustical metrics are used as inputs for the neural network. In each trial, there is only one output of the neural network: annoyance estimations.

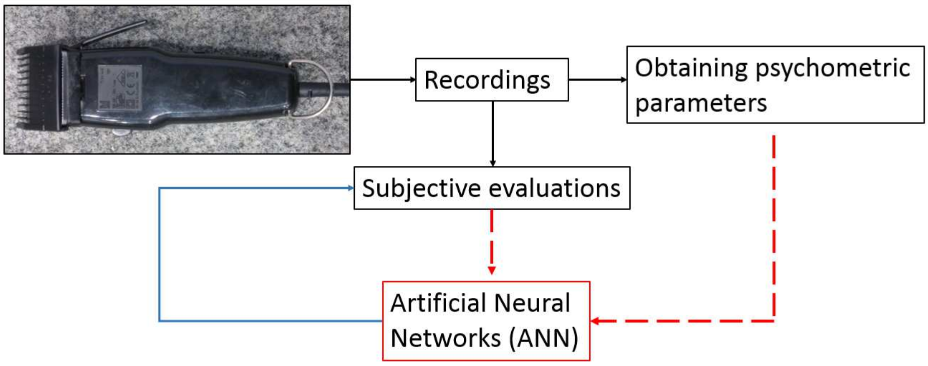

Figure 1 – Schematic of the workflow

Within this study, a two-layer feed-forward neural network is used. Sigmoid functions are used as activation functions in hidden neurons. Number of hidden neurons are taken as 10. Bayesian Regularization method is used as training algorithm. Different training states are used to find the optimum performance of the neural network. Psychoacoustical metrics are used for inputs of the ANN, however number and type of the input parameters are also evaluated for performance.

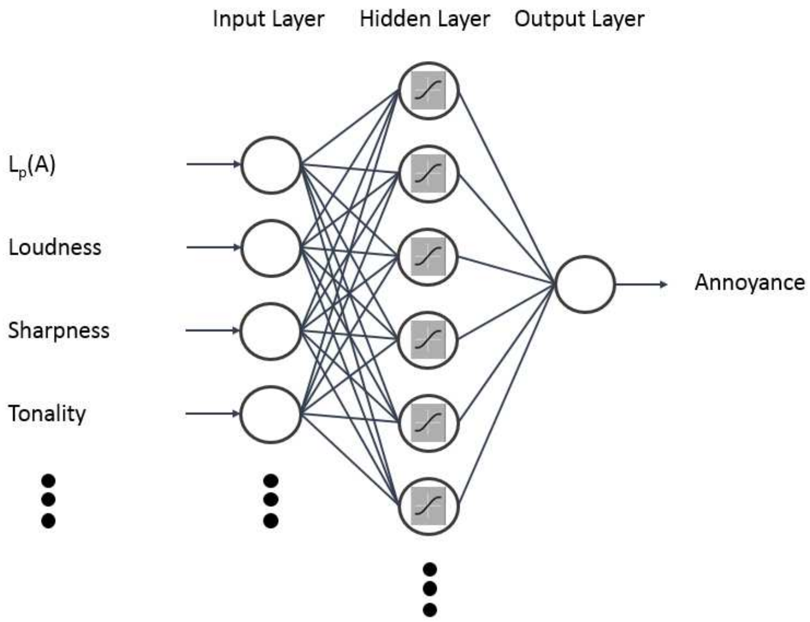

Figure 2 – Structure of the ANN used for annoyance estimations

3. STIMULI

Within this study, 27 different stimuli are used. All of them are real recordings and no synthesized stimuli is used. Recordings are performed binaurally in an anechoic environment. Equipment sounds are recorded when the equipment are in their idle running, no contact mode. Contact with skin surface changes the sound emission characteristics of the equipment. However, basic consumer selection process is selected as a case study, in which the unit is not in contact with skin and equipment is nearly 15–20 cm far away from the listener's ear, directly in front of the listener's head. Equipment are connected to the power supply during recording, since different charging conditions might change the rotating speed of the equipment resulting in different sound emissions. Figure 2 shows 4 example spectrograms used in this study. During recordings, SQuadriga frontend of Head Acoustics is used with the Head Recorder software. Analyses are performed in ArtemiS Software to calculate psychoacoustical metrics. Recording takes 10 seconds and reliable 5 seconds of the stimuli is used for evaluation purposes. Spectrograms show that shaver sounds are quite stationary, contains high tonal components, some of them positioned so close to each other especially high frequency range, changing the perception as a broadband noise. Also most of the stimuli have a fundamental tonal component around 100 Hz.

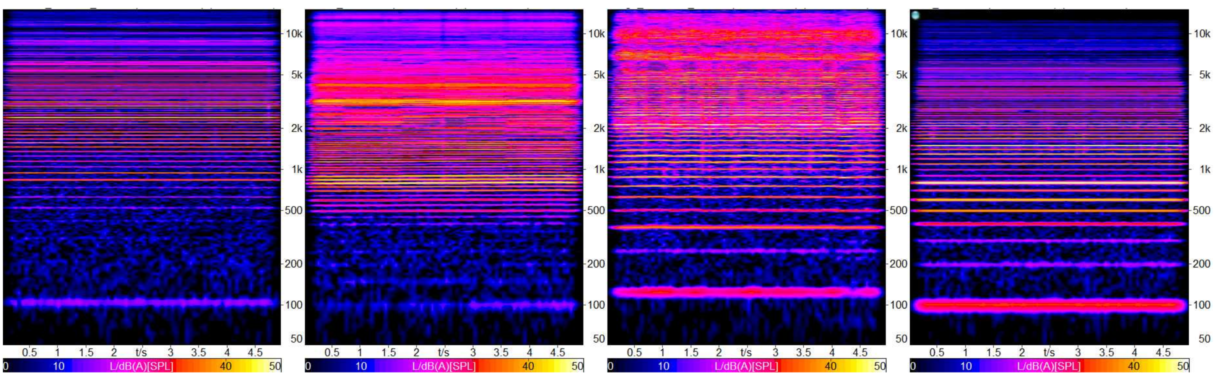

Figure 3 – Examples' spectrograms of 4 different stimuli (Spectrum size 4096, A-weighted, Hanning Window)

4. SUBJECTIVE EVALUATIONS

4.1 Subjects, Experimental Procedure and Setup

Binaurally recorded sound samples are presented to the 12 participants, 2 women and 10 men aged between 22 and 53, through Sennheiser HD600 headphones. Experiments are conducted in a sound attenuating room. Stimuli were presented in random order and 5 random stimuli were presented before the test as sample stimuli. Every stimuli is presented twice, to check inter-individual validity. The subjects then asked to evaluate the annoyance of the sounds on a quasi-continuous scale (from 0 to 100) with equidistance neighboring categories (not at all, slightly, moderately, very, extremely). A graphical user interface in Matlab, is shown in Figure 4, was used for evaluation of the experiments.

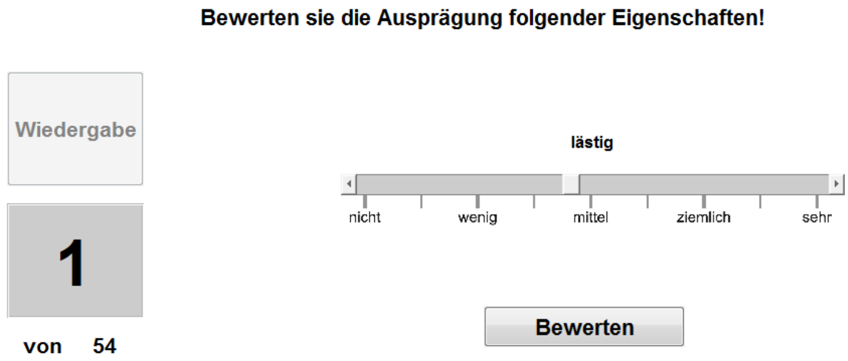

Figure 4 – Matlab GUI used for subjective evaluations (annoyance)

4.2 Results

Annoyance estimations of the 27 stimuli is presented in a box plot in Figure 5. Results are averaged for each test subject and for each repetition of the stimuli. Median values are shown in red and mean values are the mid points of the represented boxes. Edge of the boxes are the 25th and 75th percentiles of the population, and extreme data points (outliers) are plotted individually as red plus signs.

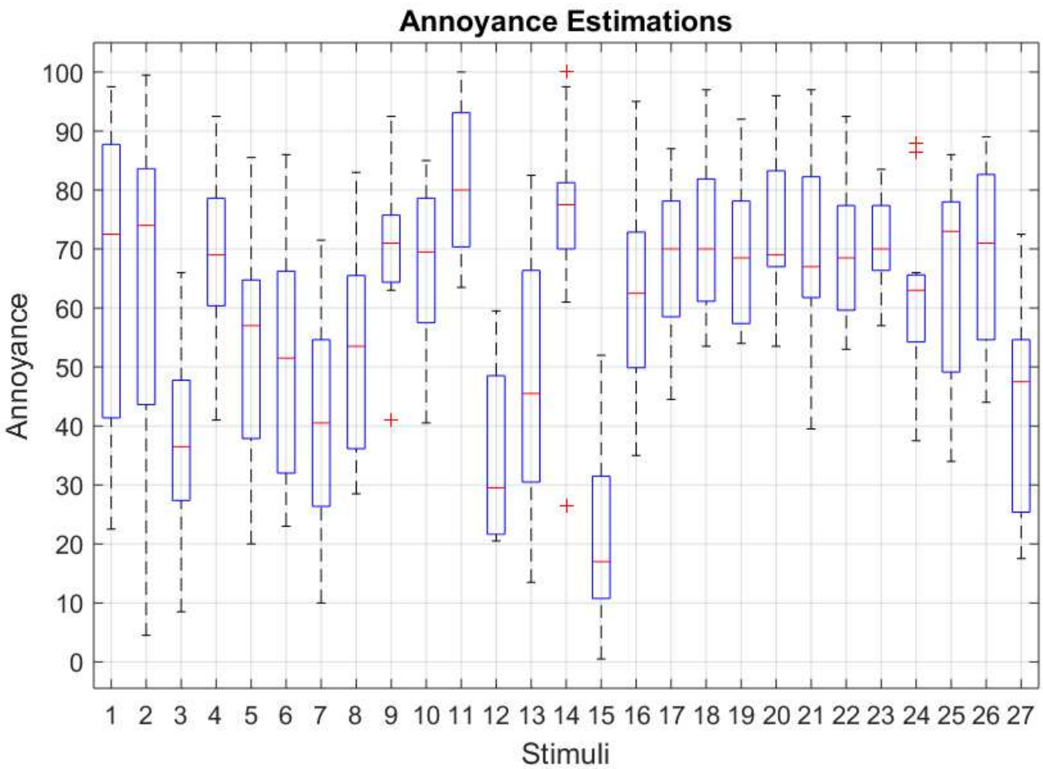

Figure 5 – Annoyance ratings of the shaver/trimmer sounds

5. PSYCHOACOUSTICAL SOUND METRICS

Psychoacoustical metrics are calculated in Head ArtemiS software. For the left and right ear recordings parameters are calculated and mean values are obtained for single value estimations. Considered acoustical variables are A-weighted sound pressure level, loudness, sharpness, tonality and roughness. Loudness calculations are based on ISO 532B standard including FFT; sharpness calculations are based on Aures model, tonality calculations are including 50% overlapping. For roughness calculations, in order to obtain reasonable results, first 0.5 seconds are not considered.

Table 1 – Calculated psychoacoustical parameters and annoyance estimations

Stimuli	Level (A) [dB(A)]	Loudness (FFT/ISO 532B) [Sone]	Sharpness (Aures) [Acum]	Tonality [Tu]	Roughness [Asper]	Annoyance (Evaluation)
1	62.5	12.45	3.59	0.207	1.335	65.7
2	63.2	12.95	3.71	0.235	1.445	63.4
3	50.0	5.52	2.73	0.934	0.942	37.9
4	59.3	11.35	3.43	0.880	1.965	68.6
5	55.1	8.53	2.96	1.035	3.065	53.0
6	54.8	8.47	3.04	1.023	3.170	49.8
7	54.7	8.43	3.02	1.020	2.785	42.8
8	54.2	8.41	3.08	1.050	2.850	53.1
9	55.7	9.77	3.57	0.938	0.509	69.7
10	57.2	10.85	3.29	0.828	1.215	67.2
11	64.0	15.05	3.41	0.455	1.935	81.6
12	56.5	8.67	1.88	1.050	0.983	33.6
13	55.2	8.58	2.89	1.065	2.730	47.8
14	58.8	12.70	3.20	1.140	1.110	74.6
15	54.4	7.24	3.38	1.176	1.205	21.6
16	59.1	11.30	3.70	1.075	1.585	62.6
17	59.4	11.50	3.76	1.085	1.715	68.5
18	59.4	11.65	3.76	1.090	1.775	71.1
19	60.7	11.60	4.40	1.014	2.265	69.0
20	60.9	11.75	4.46	1.006	2.115	73.9
21	61.6	12.60	4.27	0.980	2.420	70.3
22	61.8	13.00	4.25	0.986	2.345	69.5
23	61.6	12.90	4.16	0.968	2.220	71.0
24	57.0	9.38	3.50	0.891	1.290	63.1
25	57.4	9.61	3.66	0.797	1.325	64.5
26	58.2	10.20	3.76	0.759	1.565	69.8
27	54.1	7.55	2.37	0.943	0.501	42.0

6. ANNOYANCE ESTIMATIONS BASED ON ANN

For tailoring a neural network, there are different parameters to consider, those are can be described as:

1. Function type within the cells
2. Training methods that neural network use
3. Size of the network
4. Input parameters
5. Number of training data

Within this study, function type within cells are taken as sigmoid functions, training method is selected as Bayesian Regularization and number of hidden neurons are selected as 10. Besides those three parameters, effect of the other two is analyzed. One of the other aspect that needs to be avoided in neural networks is overlearning or overfitting. Overfitting would cause neural network to have poor performance and give biased results for different input sets. For that particular reason input parameters need to be selected carefully so that no overlearning or *mislearning* occurs.

6.1 Correlation Analyses

Before selecting the input parameters for ANN, correlation between psychoacoustical metrics and annoyance estimations need to be obtained. In order to understand that effect, a correlation study is performed between potential input parameters and annoyance estimations obtained from subjective evaluations. Results shown in Table 2 suggest that there is a strong correlation between A-weighted sound pressure level and annoyance as well as loudness and annoyance. Also correlation between sharpness and annoyance is also quite high. On the other hand, it is obvious that tonality and roughness (calculated) are not correlated with the annoyance estimations obtained from subjects. It should be noted that, due to the definitions of those two parameter, – A-weighted level and loudness – correlations might overlap each other and necessary cautions need to be taken if both of them are going to be selected as input parameters. Also the correlation coefficient between sharpness and annoyance is in the middle range so that its effect is also considerable.

Table 2 – Correlation between annoyance estimations and calculated acoustic/psychoacoustic parameters

R^2	Level (A)	Loudness (FFT/ISO 532B)	Sharpness (Aures)	Tonality	Roughness
Annoyance	0.5667	0.7053	0.477	0.0981	0.0046

6.2 Selecting the input data

For selecting the input data to train and simulate ANN, 20 stimuli is selected to train the network and the remaining 7 is used for estimating the performance of the ANN. Number of neurons in hidden layer is selected as 10. Three cases are considered:

1. A-weighted levels, loudness and sharpness are taken as input
2. Loudness and sharpness are taken as input
3. A-weighted levels and sharpness are taken as input

Figure 6 shows the results of those three sets. Lines show the results obtained from ANN for three different case definitions described above and the boxes represent the annoyance estimations obtained from listening tests with standard deviation and median values. It is understood that, using only A-weighted sound pressure levels and sharpness levels cannot mimic the annoyance estimations especially for the stimuli number 3. The other cases shows reasonable agreement between results. For that reason, input parameters are taken as A-weighted sound pressure levels, loudness and sharpness.

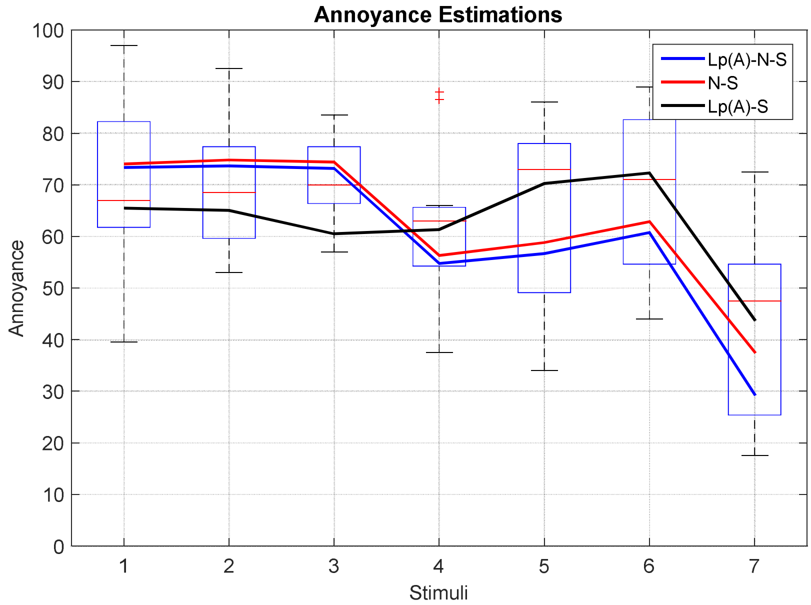

Figure 6 – Results obtained from ANN, compared to subjective evaluations, for last 7 stimuli

6.3 Results

Final neural network architecture is described as 3-input-1-output system, A-weighted sound pressure level, loudness and sharpness as inputs and annoyance as output. Moreover, one of the most important points is dividing the data set into training and validation segments. For understanding the whole system behavior, data division and training stimuli selection process is performed in a random manner; that means, in each training process, 13 data set is chosen randomly (nearly 50% of the data) to train the neural network. Training is repeated until the best regression values are obtained hence minimizing the error between target values and estimated values from neural network. With that randomized procedure, biased errors which might occur due to the division of the data set is prevented. For the final neural network architecture, all of the input sets are given as inputs to the network and results are shown in Figure 7. It can be shown that, even for the stimuli which is showing quite a different trend than the remaining data set, ANN is capable of mimic the annoyance estimation process.

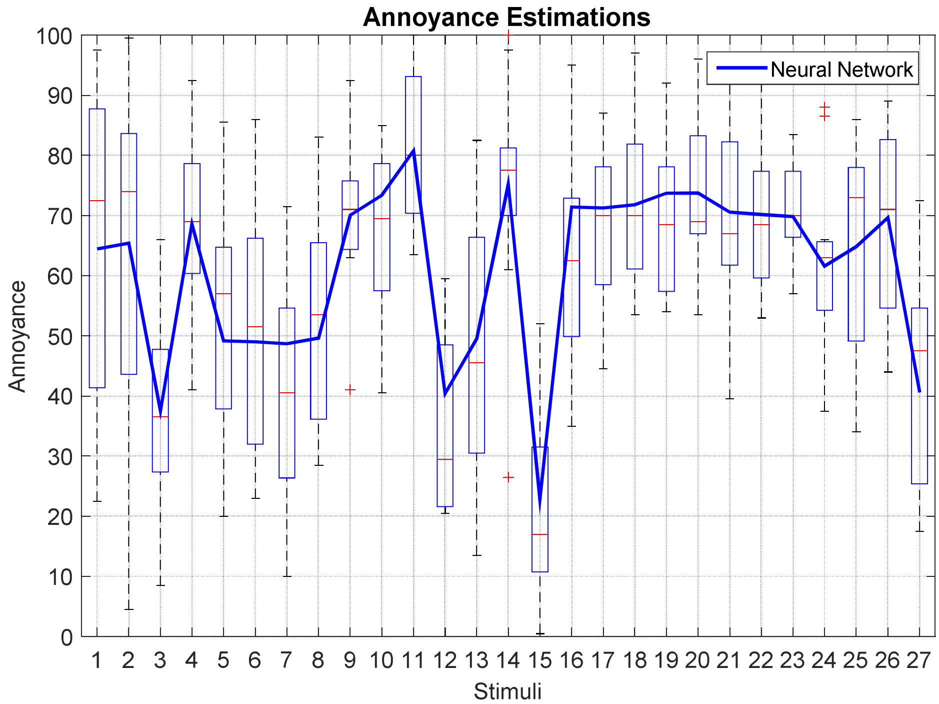

Figure 7 – Results obtained from ANN, compared to subjective evaluations, (inputs: L_p (A), N and S)

7. CONCLUSION

In this study, annoyance estimations of electric shaving/trimming devices are obtained by subjective jury testing and results are tried to obtain also with a neural network design. 27 different stimuli obtained from different shavers are used during study. Stimuli are quite stationary, having a high tonal density in high frequency range and have a fundamental tonal component around 100 Hz nearly for all examples.

Binaural recordings then presented to the 12 subjects to understand their annoyance estimations. All stimuli are presented twice and inter-individuality is controlled for each subject. Difference between two evaluations for a particular equipment and for a subject was not more than 30 points in annoyance scale for all equipment. Results, then, averaged for each repetition and for each subjects to obtain annoyance estimation of shaving devices.

Calculation of acoustic parameters are performed using Head ArtemiS software, and for the binaural recordings, parameters are calculated for both ears. For simplification, right and left ear components are averaged to obtain a single value estimation for each equipment, since these single values are used as inputs for ANN. For the future studies, especially orientation of sound source in space is important, left and right ear calculations might be considered separate as inputs for neural network estimations.

Lastly, a neural network is tailored to mimic annoyance estimations of the subjects, and pre-studies suggested that A-weighted sound pressure levels, loudness and sharpness values are selected as inputs for neural network system. By having a randomized division of data into training and validation segments, optimum neural network is obtained giving best regression values. All data sets then are given as inputs for the finalized neural network to estimate the performance of the network. Results show quite strong correlations between annoyance estimations obtained from test subjects and neural network. Neural networks are showing high potential in sound quality studies and further examples should be studied to understand different aspects and challenges.

REFERENCES

1. Altinsoy, M. Ercan, Gül, Metin and Kuyumcuoglu, Aleks. "Sound Quality of Household Appliances for Life Quality–An Investigation on Tumble Dryer Sound Quality."
2. Altinsoy, M. Ercan., G. Kanca, and H. T. Belek. "A Comparative on the Study Sound Quality of Wet-And-Dry Type Vacuum Cleaners." (1999).
3. Altinsoy, M. Ercan. "Knocking sound as quality sign for household appliances and the evaluation of the audio-haptic interaction." Haptic and Audio Interaction Design. Springer Berlin Heidelberg, 2012. 121-130.
4. Pietila, Glenn, and Teik C. Lim. "Intelligent systems approaches to product sound quality evaluations–A review." Applied Acoustics 73.10 (2012): 987-1002.
5. Lee, Sang-Kwon, Byung-Soo Kim, and Dong-Chul Park. "Objective evaluation of the rumbling sound in passenger cars based on an artificial neural network". Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 219.4 (2005): 457-469.
6. Pellegrini, Claudio de, and Edmar Baars. "Sound Quality Evaluation of Hermetic Compressors Using Artificial Neural Networks." (2006).
7. Lee, Sang-Kwon, Tae-Gue Kim, and Usik Lee. "Sound quality evaluation based on artificial neural network." Advances in Natural Computation. Springer Berlin Heidelberg, 2006. 545-554.
8. Wang, Y. S., et al. "Sound-quality prediction for nonstationary vehicle interior noise based on wavelet pre-processing neural network model." Journal of Sound and Vibration 299.4 (2007): 933-947.
9. Yıldırım, Şahin, and Eski, Ikbal. "Sound quality analysis of cars using hybrid neural networks." Simulation Modelling Practice and Theory 16.4 (2008): 410-418.
10. Lee, Hyun-Hoo, Sung-Jong Kim, and Sang-Kwon Lee. "Design of new sound metric and its application for quantification of an axle gear whine sound by utilizing artificial neural network." Journal of mechanical science and technology 23.4 (2009): 1182-1193.
11. Wang, Yansong, Gongqi Shen, and Hui Guo. "Modelling for sound annoyance evaluation of vehicle noise based on neural network." Przeglad Elektrotechniczny (Electrical Review), vol (2012): 303-306.
12. Duan, Zutong, Yansong Wang, and Yanfeng Xing. "Sound Quality Prediction of Vehicle Interior Noise under Multiple Working Conditions Using Back-Propagation Neural Network Model." Journal of Transportation Technologies5.02 (2015): 134.