

PAPER • OPEN ACCESS

# Objective-Subjective Sound Quality Correlation Performance Comparison of Genetic Algorithm Based Regression Models and Neural Network Based Approach

To cite this article: Reza Jamali *et al* 2021 *J. Phys.: Conf. Ser.* **2041** 012015

View the [article online](#) for updates and enhancements.

## You may also like

- [Sound quality prediction of vehicle interior noise and mathematical modeling using a back propagation neural network \(BPNN\) based on particle swarm optimization \(PSO\)](#)  
Enlai Zhang, Liang Hou, Chao Shen et al.
- [Impact of biased scores on ranking in bipartite competition networks and inference of modular structure via generalized modularity](#)  
Gyuhyeon Jeon and Juyong Park
- [Einstein's Jury: The Race to Test Relativity](#)  
Jürgen Ehlers



The Electrochemical Society  
Advancing solid state & electrochemical science & technology

242nd ECS Meeting

Oct 9 – 13, 2022 • Atlanta, GA, US

Abstract submission deadline: **April 8, 2022**

Connect. Engage. Champion. Empower. Accelerate.

**MOVE SCIENCE FORWARD**



Submit your abstract



# Objective-Subjective Sound Quality Correlation Performance Comparison of Genetic Algorithm Based Regression Models and Neural Network Based Approach

Reza Jamali<sup>1,\*</sup>, Gianmarco Battista<sup>1</sup>, Milena Martarelli<sup>1</sup>, Paolo Chiariotti<sup>2</sup>,  
Deepti Shriram Kunte<sup>3</sup>, Claudio Colangeli<sup>3</sup>, Paolo Castellini<sup>1</sup>

<sup>1</sup>Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche,  
Via Brecce Bianche, 60100 Ancona, Italy

<sup>2</sup>Department of Mechanical Engineering, Politecnico di Milano, 20156 Milan, Italy

<sup>3</sup>Siemens Industry Software NV, Interleuvenlaan 68, 3001 Leuven, Belgium

E-mail: \*r.jamali@staff.univpm.it

**Abstract.** Every product is growingly being evaluated in terms of acoustic characteristics. The most accurate way to rate sound quality is by performing jury tests; however, jury tests require a lot of time and human resources. To overcome this problem, jury tests results can be correlated to objective sound quality metrics owing to the fact that objective metrics could be easily obtained from sound data. In this study, advanced techniques for feature identification are explored to correlate objective metrics to subjective perception retrieved from jury tests. The data set refers to the interior noise of a regional propeller aircraft. Artificial Neural Network and two regression models (i.e. linear and quadratic regression models) have been chosen to predict subjective metrics according to the objective data. To obtain the optimized model parameters for the regression models, a Genetic Algorithm has been used as optimization strategy. In each modelling, 85 percent of sound sample data have been utilized to perform the model and remaining 15 percent have been reserved for testing the models. The results showed that the Artificial Neural Network can provide better prediction.

## 1. Introduction

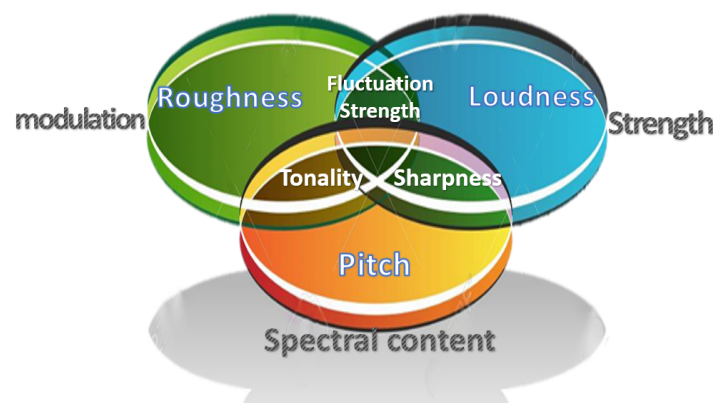
Acoustic pleasantness is becoming a crucial part of the product design process. Today, customers do not look only into product functionality specifications, but more often demand high-quality sound in addition to other aspects. Products' noise ultimately becomes a major decision maker and contributes to brand perceptions in many more products. This turns sound into an essential engineering design parameter [1].

Traditional processing and objective sound quality metrics offer a wide and detailed perspective of judging the sound of a product. However, the objective metrics don't consider how customers appreciate the product's sound in the overall product experience. Jury tests aim at establishing a clear relation between sound and other perceived characteristics, like robustness, luxury, etc. Jury tests consist of performing subjective survey during which sound data are supplied to jurors who are asked to answer to specific questions related to the sounds they are listening to. The information provided by jurors are gathered and analyzed to perform for instance pleasantness or annoyance score and will be used to rate products in terms of sound quality performance. Jury tests are time-consuming and require significant efforts, in terms of costs and human resources. Contrarily, objective sound quality metrics are directly calculated from the recorded noise and don't involve people as jurors. A schematic of sound quality metrics is shown in Figure 1 [2, 3].



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Correlating objective sound quality metrics and subjective rating (jury test results) therefore represent an important task in products development. This justify the high interest this topic has received in the scientific acoustics community. Dedene, et al. utilized multiple regression methods to assess a parametric model that is able to predict the subjective sound quality metrics performed by jurors by means of objective metrics [4]. Wang and Subic studied side mirror vehicle sound quality and correlate the subjective rating with objective metrics using two different mathematical models. Sound pressure level, roughness, and tonality were the main sound quality metrics that have been used in their correlation algorithms [5]. Li et al. developed an Artificial Neural Network (ANN) to estimate subjective sound quality metrics using objective metrics in hub permanent-magnet synchronous motors. They achieved average error rate of 3.97% with weight analysis [6].



**Figure 1.** Objective sound quality metrics

Sound quality investigations are a fundamental part also of the development process in aircraft design. If on the one hand it is important to lower the exterior noise emission component in aircrafts, on the other hand it is a key topic to improve the sound quality characteristics of aircraft's cabins to improve the passengers comfort. To evaluate noise emission in real-time for aerospace engineering applications, Vecchio et al. employed both experimentally measured sound data and a virtual model that predicts noise production for newer components [7]. Janssens et al. calculated sound quality metrics by means of synthesizing sound data simulated for aircraft interior and exterior noises, and validated the simulation results by jury tests' data [8]. Gurun and Sheth conducted semantic differential listening tests to evaluate the perception of cabin noises and correlate the subjective listening test results with objective metrics by using principal components regression [9]. Then, Brindisi and Concilio brought ANN into action to predict human judgment about aircraft sound quality from environmental parameters [10]. Lopes et al. conducted a jury test to evaluate jurors' perception of aircraft interior noise and deployed an ANN to model the subjective assessment of in-cabin recorded noise for propeller aircrafts according to the objective sound quality metrics. Finally, they acquired data from an aircraft interior cabin and conduct a jury test to evaluate jurors' perception of aircraft's interior noise. Then, he performed a correlations between objective and subjective results by ANN, developed a virtual passenger model that is able to directly predict human perception from the sound samples [11, 12]. Then, Deepti performed a model for prediction Annoyance according to sound quality metrics for washing machine and aircraft cabin noises using regression models and ANN [13].

The main goal of this paper is to investigate the effectiveness of exploiting Genetic Algorithms (GAs) and ANN for correlating subjective sound quality rating and objective sound quality metrics. Linear and quadratic regression models and ANNs have been adopted to perform the prediction, and their performances were compared and evaluated statistically to find the best performance.

## 2. Case study

The aircraft cabin noise data that have been used in this study belong to a measurement campaign performed by Siemens Industry Software NV on a regional propeller aircraft in the frame of the Brite/Euram project entitled "ASANCA" (figure 2). Among 81 seats that this aircraft has, 30 microphones measured data from 30 different seats while the aircraft was flying in a cruise condition. Jury tests were then performed to assess the human perception of the aircraft's interior noise (pleasantness in contrast to annoyance). The resulting data set available consists of 30 time histories, out of which sound quality metrics were calculated, and subjective test results [7].



**Figure 2.** The studied regional propeller aircraft [14]

## 3. Methodology

To investigate possible correlation approaches between subjective sound quality ratings and objective sound quality metrics, noise data were analysed using linear and quadratic regression models and Artificial Neural Network. Data analysed were split in input and output sets, with the purpose of predicting the output parameters from the input ones. For the case study three objective metrics, namely Articulation Index, Loudness, and Modulation Depth, were considered as input, while Annoyance was considered as an output parameter [13].

### 3.1. Linear regression model

An essential and popular statistical modelling tool that can find data relations is regression. Linear regression uses a linear model approach to show the connection between a response variable and explanatory variables. In this case, it is assumed that one or more explanatory variables that are relatively independent can predict response variable in terms of linear relations but can not predict non-linearities. As it can be seen in Eq. 1, the general equation for the linear regression model is first-order polynomial. Eq. 2 expresses the number of coefficient factors a linear regression model has with  $i$  input. [15, 16].

$$Z = k_0 + \sum_{j=1}^i k_j X_j + error \quad (1)$$

$$CF = i + 1 \quad (2)$$

Where,  $X$ ,  $Z$ ,  $k_j$  and  $CF$  are respectively input parameters, output parameter, coefficient factors and number of coefficient factors.

### 3.2. Quadratic regression model

Non-linear regressions are able to find out eventual non-linear connections between variables. A quadratic regression model is a non-linear regression that shows the response variable as a function of explanatory variables using a second-order polynomial. Eq. 3 demonstrate the general equation for quadratic regression model [17].

$$Z = k_{0,0} + \sum_{j=1}^i K_{j,0}X_j + \sum_{j=1}^i \sum_{k=1}^i K_{j,k}X_jX_k + error \quad (3)$$

Where,  $X$ ,  $Z$ , and  $K_{i,j}$  respectively are input parameter, output parameter, and coefficient factors. A quadratic regression model, therefore, has  $i + a + 1$  coefficient factors for  $i$  input and one output. where:

$$a = \frac{(i+1)!}{2!(i-1)!} \quad (4)$$

### 3.3. Genetic Algorithms

Genetic algorithms (GAs) [18] belong to a family of computational models and inspired by the evolution theory of living creatures. These kinds of algorithms are able to preserve some potential solutions or candidate solutions or even possible hypotheses in a Chromosome-like structure. Usually, GAs are used as optimization machines in which the cost function is optimized in a process similar to evolution in biology science. In other words, GA evolves to the optimized parameter without assessing all the possible solutions (possible coefficients). Given the capability of GAs to deal with non-linear problems, this strategy is exploited to optimize the fit of the linear and quadratic regression models.

$$R^2 = 1 - \sum_{j=1}^p (z_j - d_j)^2 \quad (5)$$

$$z = f(x, k_0, k_1, \dots, k_i) \quad (6)$$

Eq. 5 shows the equation for R-squared value, where  $p$  is the number of input-output data sets,  $z_j$  is the predicted value for the  $j$ th output, and  $d_j$  is the  $j$ th output value. Eq. 6 indicates the general equation for output of linear and quadratic regression models ( $z$ ), which means  $z$  depends on Coefficient factors ( $k_0, k_1, \dots, k_i$ ) and the input values ( $x$ ). Afterwards, by substitution of input-output values, and  $z$ , the R-squared would just depend on the coefficient factors, as it is shown in Eq. 7:

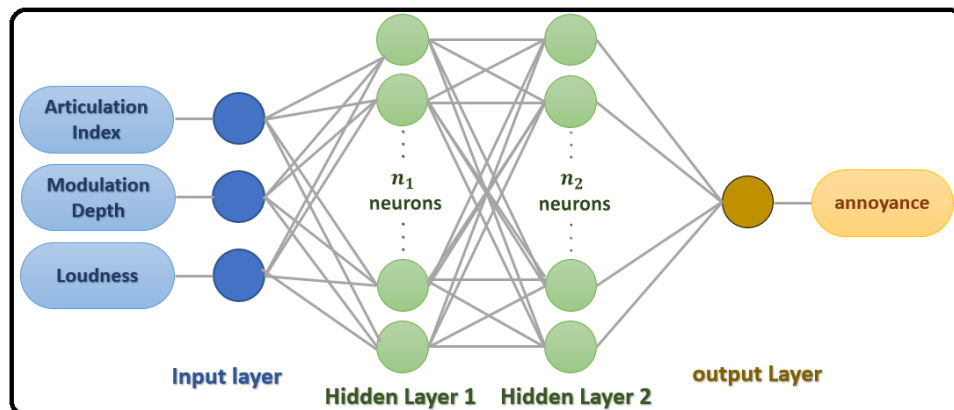
$$R^2 = 1 - \sum_{j=1}^p (f(x_j, k_0, k_1, \dots, k_i) - d_j)^2 = g(k_0, k_1, \dots, k_i) \quad (7)$$

The GA was then given the task of minimizing  $[-g(k_0, k_1, \dots, k_i)]$  as a cost function. Therefore, the cost function is a function of coefficient factors, in which GA will determine the optimized coefficient factors. It is also worth noting that the coefficient factors' values were limited to an interval as small as possible to expedite the optimization process, and the initial size of the population was set at 3000 for the linear model and 2800 for the quadratic regression model [19].

### 3.4. Artificial Neural Network

Artificial Neural Networks have been widely used in several fields to identify both linear and non-linear relations between different variables. However, their applications to acoustic problems has been gained attention mainly in this last decade [20].

A fully connected Multilayer Perceptron (MLP) model [21, 22] receiving Articulation Index, Modulation Depth and Loudness as input, and providing Annoyance as output was exploited. The model was trained using feed forward backpropagation. The model has two hidden layers, and the networks



**Figure 3.** The neural network arrangement and design

**Table 1.** ANN specifications

Parameters	values
Network type	Multilayer Perceptron
Activation Function	Tangent Sigmoid
Number of hidden layer	2
Neurons of first hidden layer	17 neurons
Neurons of second hidden layer	7 neurons
Number of Coefficient Factors	157

were trained by Feed Forward Backpropagation algorithm. A schematic of the network is shown in figure 3. Given this model structure, the number of coefficient factors can be calculated according to Eq. 8:

$$CF = i * m + m * n + n * j \quad (8)$$

Where  $i$  and  $j$  are the number of inputs and outputs of the network and  $m$  and  $n$  respectively are the number of neurons in the first and second hidden layer. Table 1 represent specifications of the ANN utilized in this paper.

#### 4. Result and discussions

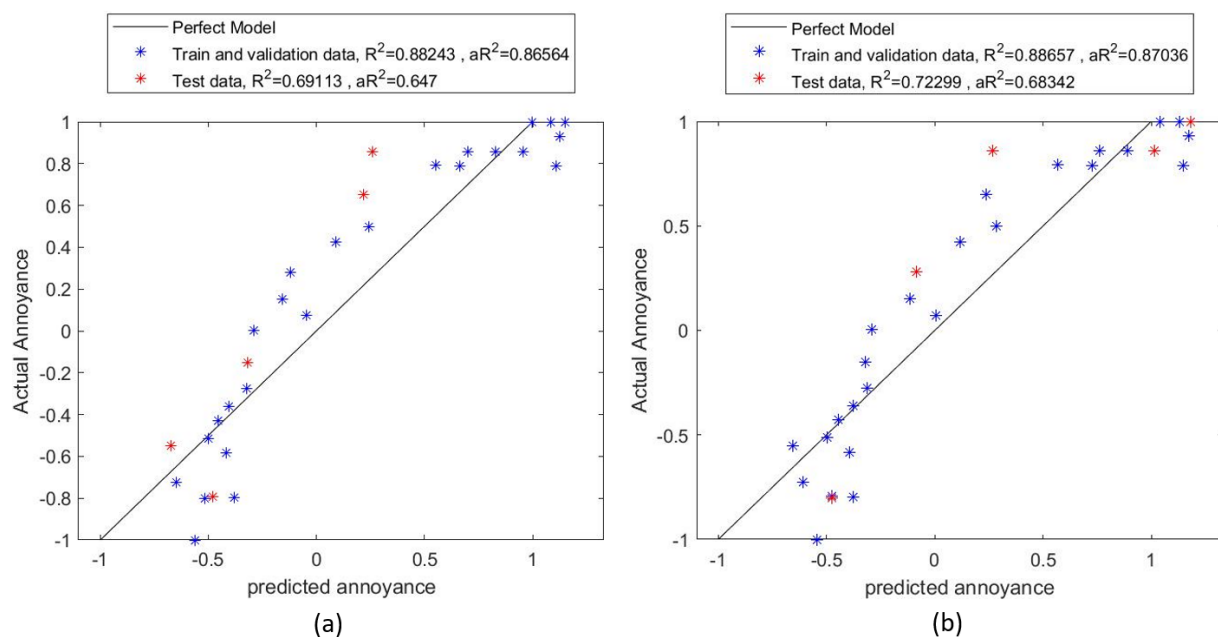
Table 2 briefly reports the parameters used to correlate the objective metrics selected and the Annoyance level obtained from the subjective evaluations of the jury test. The whole dataset available consists in 30 different data. Among this dataset, 25 data were used as training data and the remaining 5 data were used as test data. The three aforementioned models, namely linear, quadratic and ANN-based regression models, were then tested on these data and results compared to identify the strategy that suits best for the correlation purpose. The R-squared and R-squared values were adopted as judging values for identifying the best regression model.

The Annoyance prediction results over the actual Annoyance that are shown in figure 4, figure 5, and figure 6 are respectively performed by linear, quadratic regression and ANN. R-squared and adjusted R-squared are displayed in each figure for train and test data sets. The output of the training data is represented with blue stars, while the test output data are shown with red stars. The line  $y = x$

represents the ideal correlation so that as long as the data points are closer to this line, it can be inferred that the model predicts better. Therefore, as perceived by figure 4, figure 5, and figure 6 the Neural network has the best prediction performance among the two regression models in terms of R-squared and adjusted R-squared. Likewise, the quadratic regression model has better prediction performance than linear regression.

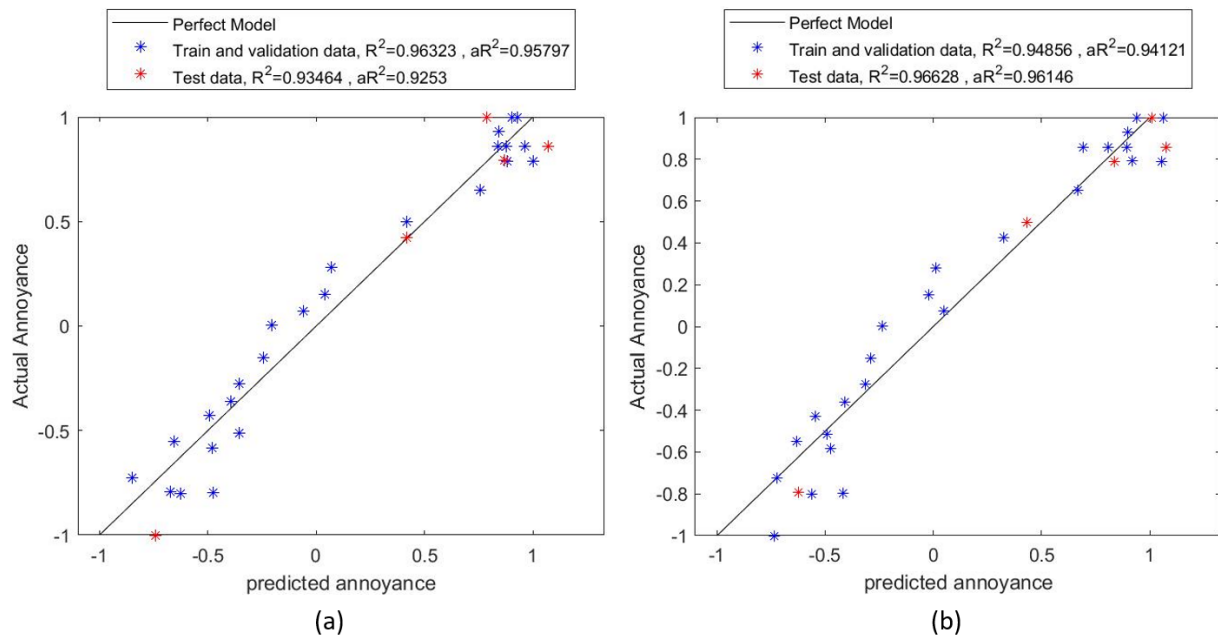
**Table 2.** Effective parameters of the modellings

Parameters	Values	Description
Input Parameters	Objective sound quality metrics	Articulation Index Modulation Depth Loudness
Output parameter	Subjective sound quality metrics	Jury tests result (Annoyance)
Gross data	30 sets	
Number of data for train and validation	25 sets	
Number of data for test the models	5 sets	

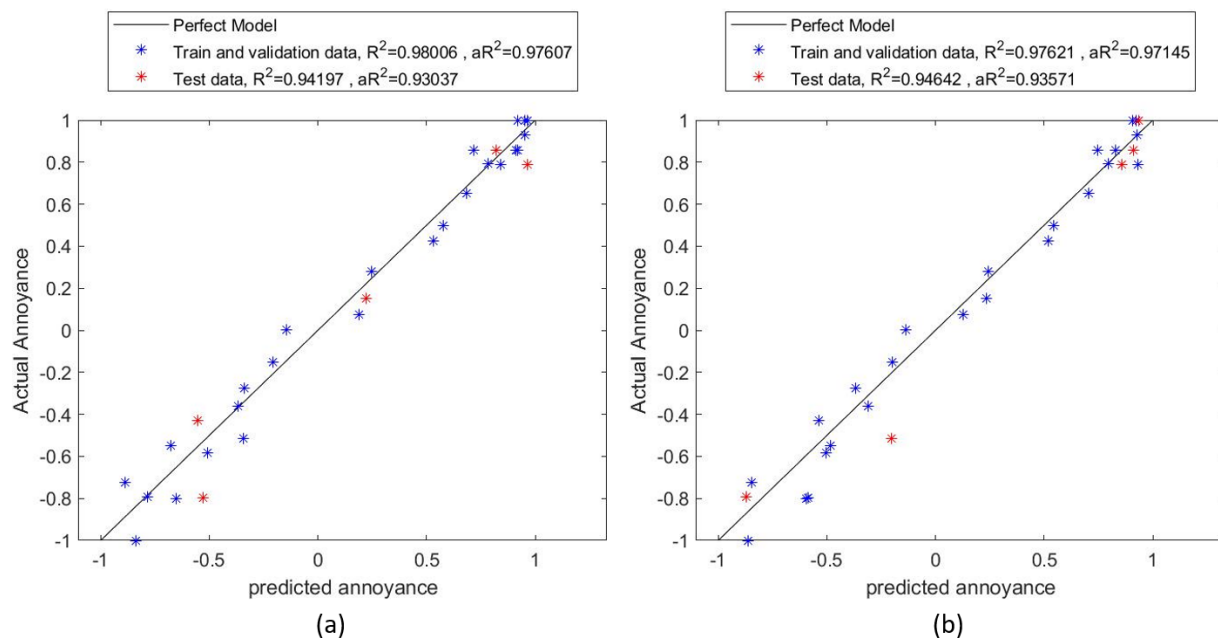


**Figure 4.** Prediction results calculated by GA optimized linear regression model; (a) and (b) are performed using different permutations of test and train data sets.





**Figure 5.** Prediction results calculated by GA optimized quadratic regression model; (a) and (b) are performed using different permutations of test and train data sets.



**Figure 6.** Prediction results calculated by ANN; (a) and (b) are performed using different permutations of test and train data sets.

The only difference between the calculation leads to the results represented in figure 4a and figure 4b is selecting the permutation for test data set among total data , and likewise the differences between

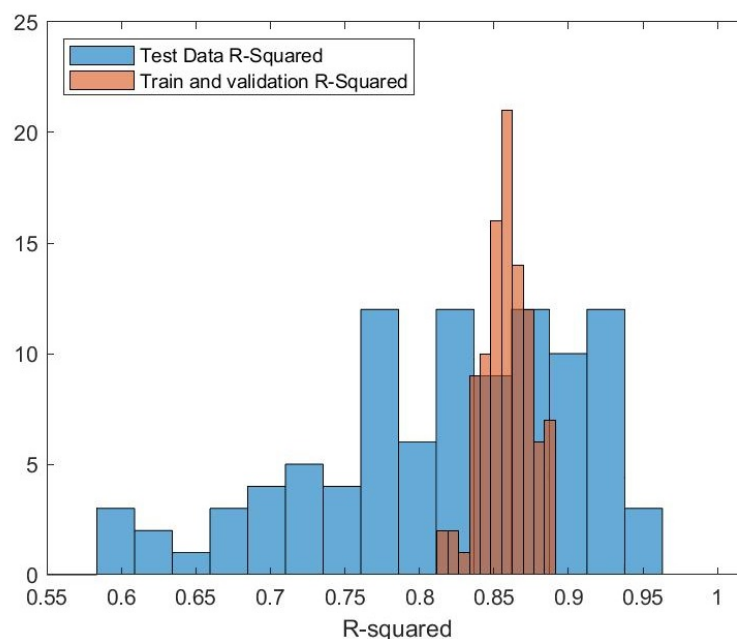


figure 5a and 5b, or figure 6a and 6b . As discussed before, a permutation of 5 data among total 30 data sets are picked and reserved for test the models, and the rest 25 data are used for train and validation of the models. According to figure 4, figure 5, and 6 the result highly depends on choosing test and training data set in which different permutations give different results. Thus, it can be deduced that the size of the total data set is relatively small and limited (30 measured points), and the models adjust overly to the training data set in which the models are not able to recognize the overall relationship between the input and output.

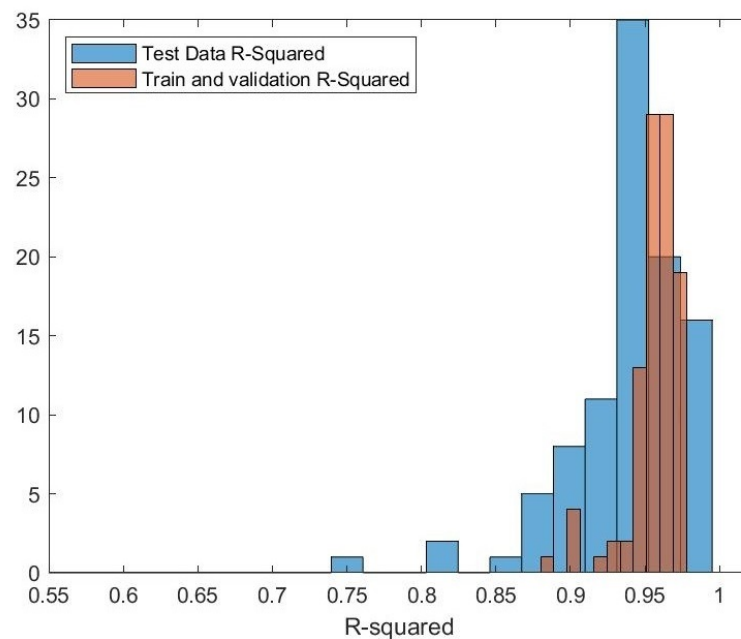
According to the fact that the results are dependent on selecting the permutation of the dataset for training and testing the models, a multi-run approach was followed in which the calculations were repeated for different random permutations of test and training data sets. A total of 100 different permutations of the data constituting the training and test datasets were performed. The three models were then applied on each different permutation and results obtained from the multi-run approach were statistically analyzed.

Figure 7, figure 8, and figure 9 represent the histograms respectively for linear and quadratic regression models and ANN. As it can be seen in figure 7 that shows the histogram for the linear model's outputs, the R-squared values for test data is spread between 0.55 to 0.95, and for train and validation data is spread between 0.8 to 0.9, which means the result is changing so much according to change in permuting the data sets. According to figure 8 the R-squared values for testing the quadratic regression model are between the interval of [0.75 1] and for the train of quadratic model are inside an interval of [0.88 0.98]. Compared to the linear model, the quadratic model has a higher and narrower distribution of R-squared, that means the quadratic model is respectively more efficient and reliable than the linear model in this concept. figure 9 shows the histogram of the multiple outputs belong to ANN. Accordingly, R-squared value for train and validation is bounded into an interval of [0.96 0.98], and for test data set is bounded to [0.87 1]. ANN has the higher R-squared values among the three methodologies that means it can predict the output better.

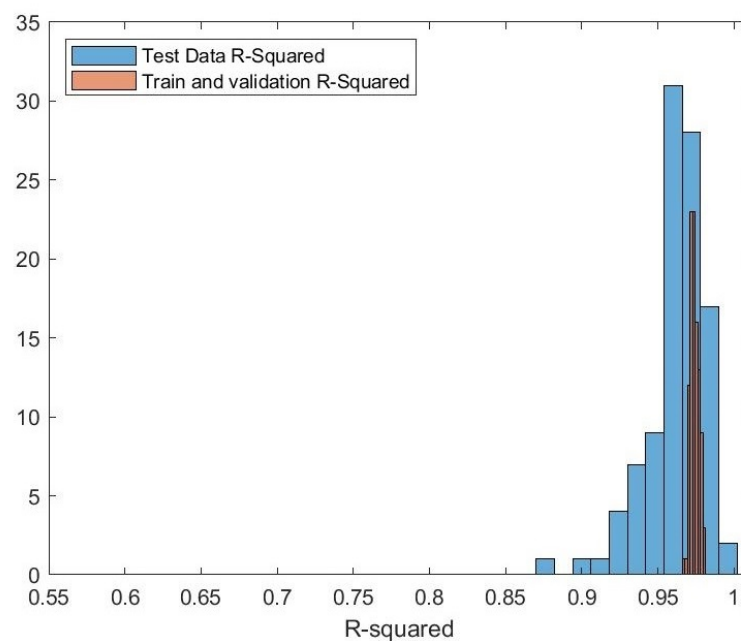
Table 3 illustrates the statistical assessment of the multi-run approach results. Mean value and standard deviation were calculated and represented for train and test results of the models. The mean



**Figure 7.** Histogram chart representing multi-run approach results for the linear regression model



**Figure 8.** Histogram chart representing multi-run approach results for the quadratic regression model



**Figure 9.** Histogram chart representing multi-run approach results for ANN

R-squared values belong to the test result of linear and quadratic regression models and ANN are respectively 0.8, 0.93, 0.96 that denote that ANN is the best model to reduce the prediction error and

**Table 3.** Statistical assessment of the results for multi-run approach; a higher mean value of R-squared implies better prediction Performance, whereas a smaller Standard Deviation indicates more prediction reliability.

	Train and validation R-squared		Test data R-squared	
	Mean Values	Standard Deviation	Mean Values	Standard Deviation
Linear Model	0.85	0.02	0.8	0.11
Quadratic Model	0.95	0.02	0.93	0.06
Artificial Neural Network	0.97	0.003	0.96	0.02

have a better prediction. Moreover, the quadratic regression model has higher R-squared values that signify a better prediction performance than the linear model because the quadratic model is able to tackle non-linearities. The relation between subjective and objective sound quality metrics is highly non-linear and depend on multiple parameters. Accordingly, the subjective rating can not be represented simply by sound quality metrics. ANN got the best result because ANN is the best model for tackling complex non-linear behaviours.

An essential parameter that plays an important role in the performance of the models is the number of coefficient factors. According to Eq. 2, Eq. 4, and Eq. 8 the linear and quadratic regression models and ANN respectively have 4, 10, and 157 coefficient factors. Increasing the number of coefficients enlarges the complexity of the models and their capacity to cope with complexity of the phenomenon. As the relation between subjective and objective sound quality metrics is complex and non-linear, the larger number of coefficient factors has enhanced the prediction performances of the model.

## 5. Conclusion

Correlation between subjective sound quality evaluation and objective sound quality metrics has been performed using advanced fitting methods, i.e. linear and quadratic regression models and ANN for a regional propeller aircraft cabin noises. In addition, the coefficient factors were optimized by the GA. The following conclusions can be therefore drafted.

- The small size of the dataset available make results highly dependent on the selection of the data adopted for the training and testing phases. This suggests that the whole regression problem should be faced adopting a statistical approach performing several permutations of the data involved (multi-run approach).
- Comparison between prediction performances of linear and quadratic regression models shows higher mean R-squared values and lower standard deviation for the quadratic model than linear model. It can be inferred that the quadratic regression model predicts better and more reliable because the quadratic model can deal better with non-linearities.
- ANN possessed the highest mean R-squared value and least standard deviation among the models. It means Neural Network works better than regression models in terms of prediction performances and is more reliable. Therefore, ANN was able to get along non-linear and complex concepts one step ahead.

## References

- [1] Colangeli C, Chiariotti P, Battista G, Castellini P, Janssens K. Clustering inverse beamforming for interior sound source localization: application to a car cabin mock-up. In 6th Berlin Beamforming Conference 2016 (No. 2016-D9, p. 17).
- [2] Claudi L, Arnesano M, Chiariotti P, Battista G, Revel GM. A soft-sensing approach for the evaluation of the acoustic comfort due to building envelope protection against external noise. *Measurement*. 2019 Nov 1;146:675-88.
- [3] Battista G, Chiariotti P, Martarelli M, Castellini P, Colangeli C, Janssens K. 3D Acoustic Mapping in Automotive Wind Tunnel: Algorithm and Problem Analysis on Simulated Data. *Applied Sciences*. 2021 Jan;11(7):3241.
- [4] Dedene L, Valgaeren R, Van Overmeire M, Guillaume P. Correlation of subjective sound perception of exhaust systems with sound quality metrics. In Society for Experimental Mechanics, Inc, 16th International Modal Analysis Conference. 1998 Feb 2 (Vol. 2, pp. 1007-1013).
- [5] Wang X, Subic A. Psychoacoustic modelling of vehicle side mirror power-fold actuator noise characteristics. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*. 2011 Jun;225(6):1419-29.
- [6] Ma C, Li Q, Liu Q, Wang D, Gao J, Tang H, Sun Y. Sound quality evaluation of noise of hub permanent-magnet synchronous motors for electric vehicles. *IEEE Transactions on Industrial Electronics*. 2016 May 13;63(9):5663-73.
- [7] Vecchio A, Polito T, Janssens K, Van Der Auweraer H. Real-time sound quality evaluation of aircraft interior noise. *Acta Acustica*. 2003;89:S53.
- [8] Janssens K, Vecchio A, Van der Auweraer H. Synthesis and sound quality evaluation of exterior and interior aircraft noise. *Aerospace Science and Technology*. 2008 Jan 1;12(1):114-24.
- [9] Gurun NT, Sheth HN. Sound Quality of Aircraft Cabin for VIP and Business Jets. In Noise Control and Acoustics Division Conference 2018 Aug 26 (Vol. 51425, p. V001T01A007). American Society of Mechanical Engineers.
- [10] Brindisi A, Concilio A. Passengers' Comfort Modeling Inside Aircraft. *Journal of aircraft*. 2008 Nov;45(6):2001-8.
- [11] Lopes B, Colangeli C, Janssens K, Mroz A, Van der Auweraer H. Neural network models for the subjective and objective assessment of a propeller aircraft interior sound quality. In INTER-NOISE and NOISE-CON congress and conference proceedings 2019 Sep 30 (Vol. 259, No. 5, pp. 4124-4135). Institute of Noise Control Engineering.
- [12] de Sá Lopes BO. Sound Quality Predictive Models: Annoyance in the Interior of a Propeller Aircraft.
- [13] KUNTE DS. Sound Quality Prediction Using Neural Networks, KTH Royal Institute of Technology, 2020.
- [14] <https://www.flickr.com/photos/24874528@N04/18649688613>
- [15] Boyko AA, Kukartsev VV, Tynchenko VS, Korpacheva LN, Dzhioeva NN, Rozhkova AV, Aponasenko SV. Using linear regression with the least squares method to determine the parameters of the Solow model. In *Journal of Physics: Conference Series* 2020 Jul 1 (Vol. 1582, No. 1, p. 012016). IOP Publishing.
- [16] Speckmayer P, Höcker A, Stelzer J, Voss H. The toolkit for multivariate data analysis, TMVA 4. In *Journal of Physics: Conference Series* 2010 Apr 1 (Vol. 219, No. 3, p. 032057). IOP Publishing.
- [17] Fajri M, Astuti S. Comparison of Kernel regression model with a polynomial regression model on financial data. In *Journal of Physics: Conference Series* 2021 (Vol. 1763, No. 1, p. 012017). IOP Publishing.
- [18] Stepanov LV, Koltsov AS, Parinov AV, Dubrovin AS. Mathematical modeling method based on genetic algorithm and its applications. In *Journal of Physics: Conference Series* 2019 Apr 1 (Vol. 1203, No. 1, p. 012082). IOP Publishing.
- [19] Vahidi-Moghaddam A, Mazouchi M, Modares H. Memory-augmented system identification with finite-time convergence. *IEEE Control Systems Letters*. 2020 Jun 23;5(2):571-6.
- [20] Wanto A, Zarlis M, Hartama D. Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves in the Predicting Process. In *Journal of Physics: Conference Series* 2017 Dec 1 (Vol. 930, No. 1, p. ).
- [21] Yegnanarayana B. Artificial neural networks. PHI Learning Pvt. Ltd.; 2009 Jan 14. 012018). IOP Publishing.
- [22] Vahidi-Moghaddam A, Mazouchi M, Modares H. Learning Dynamics System Models with Prescribed-Performance Guarantees using Experience-Replay. In 2021 American Control Conference (ACC) 2021 May 25 (pp. 1941-1946). IEEE.