www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

Archives of Acoustics Online First September 26, 2024 doi: 10.24425/aoa.2024.148811

Research Paper

Combined Evaluation of Room Acoustic Descriptors in Different Structural Configurations via ODEON Simulations and Artificial Neural Networks

Eriberto Oliveira DO NASCIMENTO*, Paulo Henrique Trombetta ZANNIN

Laboratory of Environmental and Industrial Acoustics and Acoustic Comfort
Federal University of Paraná – UFPR
Curitiba, PR, Brazil

*Corresponding Author e-mail: eriberto@ufpr.br

(*received August 26, 2023; accepted June 18, 2024; published online September 26, 2024*)

This study evaluated the combined sensitivity analysis of several room acoustic descriptors: reverberation time (T30), center time (Ts), early decay time (EDT), definition (D50), clarity (C50), useful-to-detrimental sound ratio (U50), and speech transmission index (STI); and also it assessed how these descriptors responded jointly to different acoustic-structural factors. The first-order factors were background noise (A), acoustic ceiling tile sound absorption coefficient (B), confinement (C), and occupancy (D), along with their interaction effects. A novel method is proposed for this joint evaluation of sensitivity factors. This method involves in situ measurements and an unreplicated 2^4 factorial design, which has been validated by ODEON software. The significance of input factors is determined using artificial neural networks (ANN) and the modified profile method (MPM), validated by multiple linear regression (MLR). Three significant correlation groups are identified at $p < 0.05$: group 1 (EDT, T30, Ts), group 2 (C50, D50), and group 3 (U50, STI). The ceiling material sound absorption (B) is found to affect reverberation (groups 1 and 2), while background noise (A) impacts STI and U50. A weak correlation is found between D50 and STI. These results are confirmed by the MLR and MPM methods.

Keywords: speech transmission index; reverberation time; artificial neural networks; room acoustics; ODEON simulation.

1. Introduction

Classical approaches to acoustics are aimed at characterizing optimal classroom design for speech intelligibility based on reverberation time (RT) (Lochner, Burger, 1960). In this context, other descriptors have been developed to consider exogenous effects such as background noise (BGN) to quantify the acoustic quality of rooms (Kang *et al.*, 2023). As a result, the speech transmission index (STI) was the leading objective descriptor designed to measure speech intelligibility (SI), and one of the driving forces for the development of STI was the recognition of the influence of BGN on intelligibility (Houtgast *et al.*, 1980). In this work, SI was assessed using the STI descriptor.

Many other descriptors have been proposed to quantify acoustic fields in classrooms; most of them are based on sound propagation and its energy ratios by applying the diffuse field theory. These descriptors are: definition (D50), clarity (C50), center time (Ts), early decay time (EDT), and useful-to-detrimental sound ratio (U50). Additional descriptors, such as noise criteria (NC) curves and signal-to-noise ratio (SNR), are correlated with strictly subjective aspects of speech and sound perception (Bradley, 2011).

Many studies have sought to quantify the possible correlations between these parameters and the STI. For example, such studies included analyses of the relationship between STI and SNR (Bradley *et al.*, 1999), STI and RT (Mikulski, Radosz, 2011; Rennies *et al.*, 2014), STI and energy descriptors (EDT, Ts, C50, D50) (Sato *et al.*, 2006; Ansay, Zannin, 2016), and STI and U50 (Bradley *et al.*, 2003; Sato *et al.*, 2012; 2016; Choi, 2017a; 2017b).

However, most of these studies have focused on quantifying the correlations between individual con-

struction design factors and STI, often examining only one aspect at a time. Furthermore, these studies have generally found that STI levels were low, often failing to meet the minimal threshold values of acoustic standards. Several studies have linked these low STI values to high BGN, which may be due to factors such as low sound absorption coefficients of ceiling panels, inadequate acoustic insulation, poor placement of sound-diffusion panels, or unfavorable building façade conditions (Visentin *et al.*, 2018; Bistafa, Bradley, 2001; Sala, Rantala, 2016; Secchi *et al.*, 2017).

Therefore, the combined evaluation of acoustic descriptors is essential to determine how they are correlated and how room design and occupancy affect them. This paper seeks to contribute to the body of knowledge on this subject by quantitatively determining how the STI, T30 (reverberation time), EDT, C50, D50, Ts, and U50 respond jointly to the following factors: (A) background noise, (B) ceiling sound absorption; (C) confinement (if the room's windows and doors are open or closed), and (D) occupancy (if the room is or is not occupied by people, i.e., an audience), and by proposing a new method for this evaluation. An extensive comparison of results reported in the literature was also carried out.

Artificial intelligence in the form of artificial neural networks (ANNs) was used to ascertain the sensitivity of construction and occupancy conditions on acoustic parameters. This identification should lead to numerous benefits, contributing to improve the development of better classroom acoustic designs.

2. Materials and methods

2.1. In situ measurements

The speech transmission index measurements were taken in five classrooms by the indirect STI method, following the International Electrotechnical Commission 60268-16 standard (IEC, 2011) procedures. The BGN was measured for five minutes based on the equivalent continuous sound level (L_{eq}) in empty rooms with closed doors and windows to reduce external noise interference, using the B&K 2260 sound level meter. The NC value was then calculated as specified by American National Standard (ANSI, 2008).

The STI was measured using the following instruments: DIRAC room acoustics software (B&K type 7841), version 5.0, installed on a Sony VAIO notebook, Audio Interface ZE-0948 data acquisition board; Behringer FBQ800 equalizer; Lab. Gruppen LAB 300 amplifier; B&K 4227 mouth simulator, and a B&K 2260 real-time sound analyzer as the receiver microphone. According to IEC (2011) standard Sec. 7.2, the signal spectrum at the output of the B&K 4227 mouth simulator was equalized using the maximum length sequence (MLS) signal with a Pink + Blue filter. The equalizer gains were then adjusted to the reference operational speech level of 60 dB, with a tolerable error of ±1 dB, in the octave bands of 125 Hz and 8 kHz.

The descriptors T30, EDT, C50, D50, and Ts, were measured according to International Organization for Standardization [ISO] (2008), using the impulse response method. The e-sweep signal generated by DIRAC 5.0 software was used as an excitation signal. A B&K 4296 dodecahedral sound source was placed at a height of 1.5 m from the floor and more than 1.2 m from the walls. Five measurements were recorded with the receiver (B&K 2260 sound analyzer) in different positions in each classroom.

2.2. Classroom simulations

In this paper, five university classrooms were studied in the Federal University of Paraná, located in southern Brazil. Table 1 describes the maximum dimensions of the classrooms and their volumes. All the measured data, ODEON virtual models, sound absorbing materials, and classroom photos are available in the dataset reference provided by Do Nascimento and Zannin (2023). ODEON software is widely used worldwide for predicting room acoustics parameters, for consulting, and for academic research. ODEON uses the image source method combined with a modified ray tracing algorithm for acoustic simulations. In this study, ODEON was used to simulate STI, T30, EDT, C50, D50, Ts, and U50 values based on various structural acoustic factors, such as A, B, C, and D.

Table 1. Dimensions of the measured classrooms.

Room	Width [m]	Length [m]	Height [m]	Volume [m^3]
1	12.40	13.40	2.75	358.11
2	10.00	8.70	2.95	236.06
3	11.14	7.57	3.91	355.16
4	11.50	5.53	4.20	289.27
5	13.99	13.99	5.84	1200.77

Classroom modeling and validation was performed using ODEON v. 11 software (Rindel, 2012), which specifies that acoustic descriptors must be simulated in accordance with the guidelines (ISO, 2008; IEC, 2011) for in situ measurements. ISO (2008) establishes that T30, EDT, C50, D50, and Ts must be measured using an omnidirectional sound power level; hence, the Omni.SO-8 sound source was selected from the ODEON library. STI and U50 were simulated using the BB93_NORMAL_NATURAL.SO-8 sound source equalized at 60 dB. ODEON allowed for the weighting of male and female speech spectra using the virtual sound source, resulting in a total of three STI measures: STI_{male}, STI_{female}, and the non-gender filtered STI. A 20 cm × 20 cm calculation meshgrid was placed at a height of 1.20 m from the floor, i.e., the average height of the ears of listeners sitting in typi-

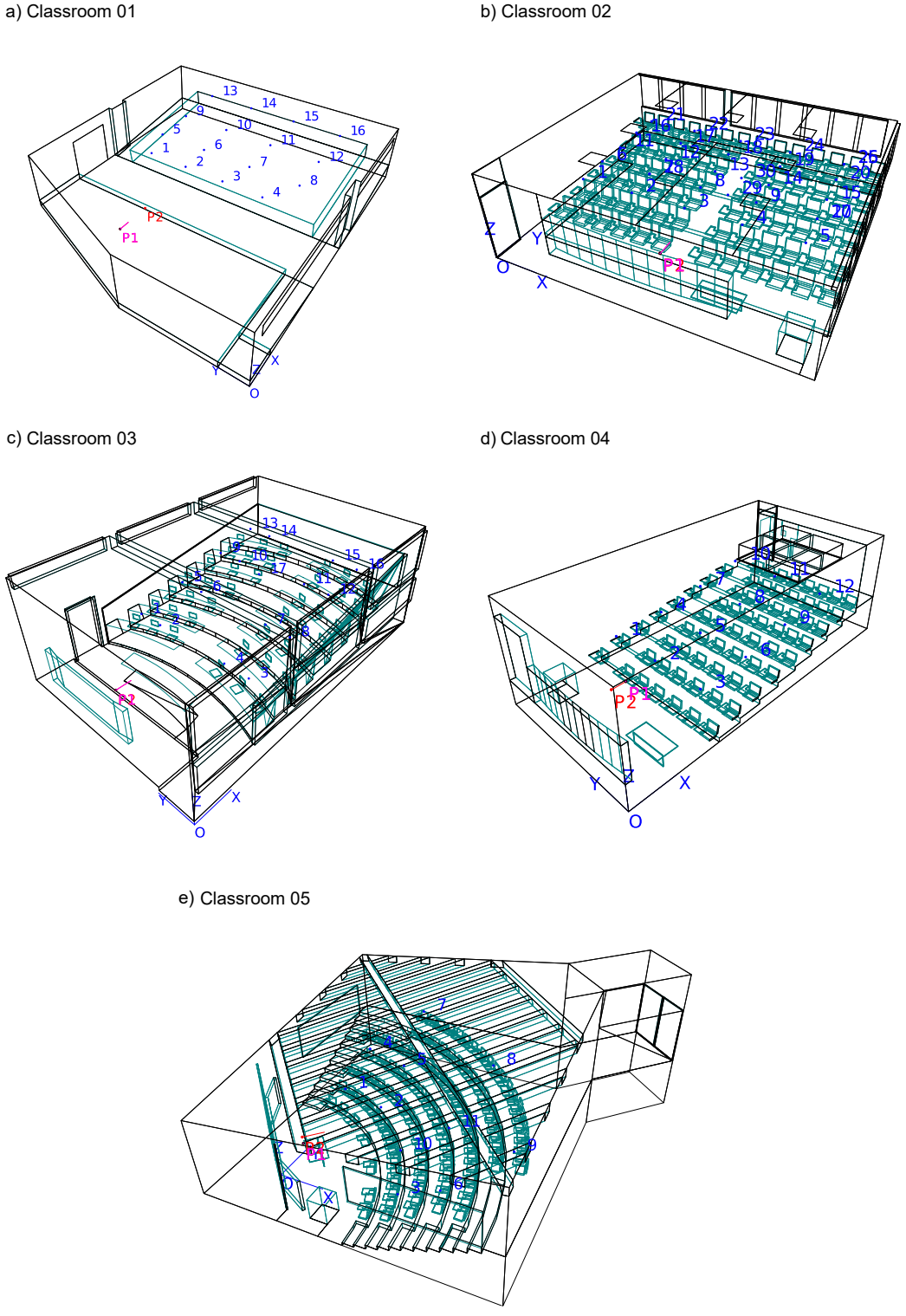

Fig. 1. Classrooms modeled using ODEON v. 11. The blue dots indicate the receiver positions, while the sound source is situated in the typical teacher position at the front of the class.

cal chairs. Figure 1 depicts the virtual models of the evaluated classrooms.

The receiver's placement in the virtual classrooms was chosen based on two criteria, to ensure full spatial coverage to accurately capture the room impulse response (RIR) and to mimic the listener's perspective. The simulations were validated globally for each room using the root mean squared percentage error (RMSPE) – Eq. (1) – calculated between the measured and simulated values for descriptors T30 and STI in

their respective measurement positions in each classroom:

$$\text{RMSPE} = 100\%\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\widehat{y}_i - y_i}{y_i}\right)^2}, \tag{1}$$

where n is the number of evaluated points, and $\widehat{y}_i$ and y_i are the simulated (S) and measured (M) values, respectively, the mean percentage error (MPE) is $(100\%/n)\sum_{i=1}^{n}(\widehat{y}_i - y_i)/y_i$. Table 8 lists these errors for each classroom and for the T30 and STI descriptors.

To assess the goodness-of-fit of the validation in terms of simulated and measured values, the RMSPE was used as the deviation metric, because it is directly comparable to the just noticeable difference (JND) values, which were 5 % and 3 % for T30 and STI, respectively. Studies by Bradley (2011) and Christensen *et al.* (2014) reported JND values more than five-fold higher than those of the current study. Therefore, a 10 % threshold for RMSPE was established as the decision criterion for validation, corresponding to 3 JND for STI and 2 JND for T30.

Each classroom was then calibrated by interactively fine-tuning the sound absorption coefficients of the various surface areas until the 10 % RMSPE threshold was reached. Lastly, the calibration curve for the acoustic model was built by line fitting of the measured versus simulated data, and the Pearson correlation coefficient was calculated. After experimentally validating a virtual room design created with ODEON, new conditions have been developed through the design of experiments (DOE), and a joint analysis was made to determine how the room acoustic parameters varied in these new conditions.

2.3. Creation of training dataset

Montgomery (2012) described the DOE as a method to determine how modifying certain controllable factors (inputs) affects the system's response variables (outputs). The DOE expresses the relationships between the controllable factors (x) and responses (y), through a standard multilinear model, whose interactions are given in Eq. (2):

$$y = \beta_0 + \sum_{j=1}^{k}\beta_j x_j + \sum\sum_{i<j}\beta_{ij}x_i x_j + \epsilon, \tag{2}$$

where β_0 is the model's linear intercept coefficient, β_j is the regressor coefficient of the main factors x_j, β_{ij} represents the coefficients of the regressors for the effects of the combinatorial interaction between the main factors $x_i x_j$, and k is the number of factors. A 2^k DOE was applied in this work, with $k = 4$ controllable factors and two levels, ranging from minimum (−1) to maximum (+1). The 2^k factorial design with a range of −1 to +1 was chosen for its efficiency in screening factors and its representation of the standardized effect, which simplifies analysis and interpretation. Table 2 describes the factors and their levels.

Table 2. DOE levels and factors.

Level	Controllable factors – natural scale			
	Noise criteria	Ceiling absorption	Confinement	Occupancy
Lowest value	15	0.10	Open	Present
Highest value	40	0.90	Closed	Absent
	Controllable factors – encoded scale			
	x_1	x_2	x_3	x_4
	A	B	C	D
Lowest value	−1	−1	−1	−1
Highest value	+1	+1	+1	+1

Within the context of the DOE (see Table 2), the transition of variables from −1 to +1 in the encoded or natural space has significant implications for multiple linear regression (MLR) models. A positive shift from −1 to +1 signifies an ascent towards the higher end of the experimental range, indicating a positive influence on the evaluated response variable. Conversely, a negative shift denotes a descent towards the lower range of the response, suggesting a negative impact.

The simulations with ODEON software were used to emulate classrooms and their respective acoustic parameters as a function of four controllable factors (see Table 2). The combinatorial experimental conditions were based on the factors called experimental runs, which correspond to the design matrix lines x_{ij} listed in Table 3.

The effect of room occupancy, i.e., audience, was estimated by assigning sound absorption coefficients to the equivalent absorption surface area of the audience in ODEON classroom model, based on the literature (see Table 4).

Specifically, for a full 2^4 factorial design, the equivalent linear model with 3rd and 4th-order interactions is shown in Eq. (3):

$$\begin{aligned} y_{k,ij} = {} & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 \\ & + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 \\ & + \beta_{34} x_3 x_4 + \beta_{123} x_1 x_2 x_3 + \beta_{124} x_1 x_2 x_4 \\ & + \beta_{134} x_1 x_3 x_4 + \beta_{234} x_1 x_2 x_3 + \beta_{1234} x_1 x_2 x_3 x_4, \end{aligned} \tag{3}$$

here, the room acoustics parameters are frequency-dependent and are written in octave bands to express the response of a given parameter. Therefore, a nested output ($y_{k,ij}$) was used in each classroom. Let k (= T30, EDT, C50, D50, Ts, STI) indicate the room acoustic parameters, i (= 1 to 16) represent the experimental runs, and j (= 63 Hz, 125 Hz, 250 Hz, 500 Hz,

Table 3. Design matrix used in simulations with ODEON software.

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
Run	x_1	x_2	x_3	x_4	x_1x_2	x_1x_3	x_1x_4	x_2x_3	x_2x_4	x_3x_4	$x_1x_2x_3$	$x_1x_2x_4$	$x_1x_3x_4$	$x_1x_2x_3$	$x_1x_2x_3x_4$
1	–1	–1	–1	–1	1	1	1	1	1	1	–1	–1	–1	–1	1
2	1	–1	–1	–1	–1	–1	–1	1	1	1	1	1	1	–1	–1
3	–1	1	–1	–1	–1	1	1	–1	–1	1	1	1	–1	1	–1
4	1	1	–1	–1	1	–1	–1	–1	–1	1	–1	–1	1	1	1
5	–1	–1	1	–1	1	–1	1	–1	1	–1	1	–1	1	1	–1
6	1	–1	1	–1	–1	1	–1	–1	1	–1	–1	1	–1	1	1
7	–1	1	1	–1	–1	–1	1	1	–1	–1	–1	1	1	–1	1
8	1	1	1	–1	1	1	–1	1	–1	–1	1	–1	–1	–1	–1
9	–1	–1	–1	1	1	1	–1	1	–1	–1	–1	1	1	1	–1
10	1	–1	–1	1	–1	–1	1	1	–1	–1	1	–1	–1	1	1
11	–1	1	–1	1	–1	1	–1	–1	1	–1	1	–1	1	–1	1
12	1	1	–1	1	1	–1	1	–1	1	–1	–1	1	–1	–1	–1
13	–1	–1	1	1	1	–1	–1	–1	–1	1	1	1	–1	–1	1
14	1	–1	1	1	–1	1	1	–1	–1	1	–1	–1	1	–1	–1
15	–1	1	1	1	–1	–1	–1	1	1	1	–1	–1	–1	1	–1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 4. Sound absorption coefficient of materials used to emulate room occupancy.

Room	Occupancy – audience area	ODEON number	Reference	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
1, 2, 3, 5	Audience on heavily upholstered seats	11007	Beranek, Hidaka (1998)	0.72	0.80	0.86	0.89	0.90	0.90	0.90
4	Audience on wooden chairs, 2/m^2	11004	Meyer *et al.* (1964)	0.24	0.40	0.78	0.98	0.96	0.87	0.87

1 kHz, 2 kHz, 4 kHz, and 8 kHz) indicate the octave bands.

Consequently, to obtain the $y_{i,kj}$ output used to train the ANNs in each classroom, 51 responses were generated in each run, 48 corresponding to the descriptors $T30_{63\,Hz–8\,kHz}$, $EDT_{63\,Hz–8\,kHz}$, $C50_{63\,Hz–8\,kHz}$, $D50_{63\,Hz–8\,kHz}$, and $Ts_{63\,Hz–8\,kHz}$ in the octave bands at 63 Hz, 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz, while the other three responses were STI_{male}, STI_{female}, and the STI in nominal values.

2.4. Artificial neural networks

The fundamental objective of ANNs is to produce generalizations through a training process (Russell, Norvig, 1996). Generalization refers to estimates for patterns not used in training. Training occurs when the weights of the connections of an ANN's neurons are adjusted to minimize a loss function. Hence, an ANN is a set of nested functions, with regard to, $Y_n = W_nX_{n-1}$, that have undergone nonlinear pointwise transformations, resulting in a generalized model, $X_n = \Phi(Y_n)$, where W_n is the weight matrix of layer n, and X_{n-1} represents the input patterns for a previous layer $n-1$. The generalized multilayer model is expanded, as shown in Eq. (4):

$$Y_n = \Phi_n\Big[W_n\Phi_{n-1}\big[W_{n-1}\Phi_{n-2}\left[\cdots\Phi_1[W_1X]\right]\big]\Big]. \tag{4}$$

The nonlinear transformations performed by Φ are called activation functions. As proposed by LeCun *et al.* (1998), the weight update rule depends on a loss function that, for a pattern (p) compares the ground truth value (D^p) with the neural networks' output model, given $M(X_n,W)^P$, using the residual error metric, as shown in Eq. (5):

$$E^P = \frac{1}{2}\Big(D^p - M(X_n,W)^P\Big)^2, \tag{5}$$

where E^P is the residual sum of squares about the $M(X_n,W)^P$, which is the ANN output value estimated as a function of the weight matrices (W), and the input patterns (X_n). Accordingly, neural network training can be interpreted as an optimization paradigm, in which the weights are adjusted as the loss function is minimized. Notably, in regression problems, a widely used loss function is the mean square error (MSE) shown in Eq. (6):

$$\mathrm{MSE}_{\mathrm{train}} = \frac{1}{P}\sum_{p=1}^{P} E^P, \tag{6}$$

where P stands for the total number of samples. Moreover, the weight update rule can be applied using various approaches, such as the gradient-based learning paradigm, e.g., the stochastic gradient descent, given in Eq. (7):

$$W(t) = W(t-1) - \eta\frac{\partial E}{\partial W}. \tag{7}$$

In this work, the error backpropagation learning algorithm was employed to update the weight matrices. Equations (8)–(10) portray the generic approach to backpropagation learning, which is interpreted as the automatic differentiation in the reverse mode (Baydin *et al.*, 2017):

$$\frac{\partial E^P}{\partial Y_n} = \Phi'(Y_n)\frac{\partial E^P}{\partial X_n}, \tag{8}$$

$$\frac{\partial E^P}{\partial W_n} = X_{n-1}\frac{\partial E^P}{\partial Y_n}, \tag{9}$$

$$\frac{\partial E^P}{\partial X_{n-1}} = W_n^T\frac{\partial E^P}{\partial Y_n}. \tag{10}$$

In this work, ANNs were applied as universal functional approximators to map the relationship between inputs and outputs described in Subsec. 2.2.

2.4.1. Input and target pre-processing

The ANN model's input-output training pairs were based on the DOE method described in Table 3, which served not only to create an optimal experimental design but also as an augmentation method. According to the DOE runs, augmentation refers to the virtual response values for the acoustic parameters obtained through ODEON software.

Firstly, an ANN was trained individually for each acoustic parameter. Although the same input matrix was used in all the ANN models, only the output variable varied according to the evaluated acoustic parameters. As a result, a single training sample consisted of a 15-input feature vector and a scalar output. The input feature vector was derived from the DOE model inputs, A, B, C, D, AB, AC, AD, BC, BD, CD, ABC, ABD, ACD, BCD, and ABCD, and the scalar target corresponded to the reduced single band acoustic parameter. The $Y_{16,8}$ octave band output matrix was dimensionally reduced to the $Y_{16,1}$ octave single band vector encoded in the first principal subspace, using principal component analysis – PCA (Jolliffe, Cadima, 2016). Thus, the entire training set for one classroom was composed of an $X_{16,15}$ input matrix and a $Y_{16,1}$ output vector.

The PCA worked by deflating redundant information through multicollinearity reduction, using eigenvector information as a threshold measure. For example, for an $X_{m\times n}$ dataset, where m is the number of observations of n variables, the PCA diagonalizes X through its variance-covariance matrix, given as $S = (1/N-1)\,XX^{\mathrm{T}}$, where N is the number of observations, $n = 1, 2, ..., N$. The eigendecomposition of S results in $Su_k = \lambda_k u_k$, thus determining the eigenvectors u_k and their eigenvalues λ_k. The sorted eigenvalue λ_k informs the cumulative explained variance in the original X dataset (Jolliffe, 2011).

Secondly, Russell and Norvig (1996) recommend applying the min-max scaling range −1 to 1, in the training dataset. The input data set was then assigned as the design matrix, which was already in a scaled form; hence, only the target variables underwent scaling through Eq. (11):

$$y_{ij} = a + \frac{(d_{ij} - \min(d_{ij}))(b-a)}{\max(d_{ij}) - \min(d_{ij})}, \tag{11}$$

where y_{ij} is the j scaled target feature, i.e., j (= T30, EDT, C50, D50, Ts, STI), $i = 1, 2, ..., 16$ is the i-th sample of the j feature, simulated d_{ij} stands for the non-scaled feature, and a and b correspond to −1 and +1, which represent the transformed range.

The hyperbolic tangent activation function, which has the same image range as min-max scaling, was employed here. Figure 2 shows the assignment of a training sample, with the input, $x_{1,15}$ (vector), and output

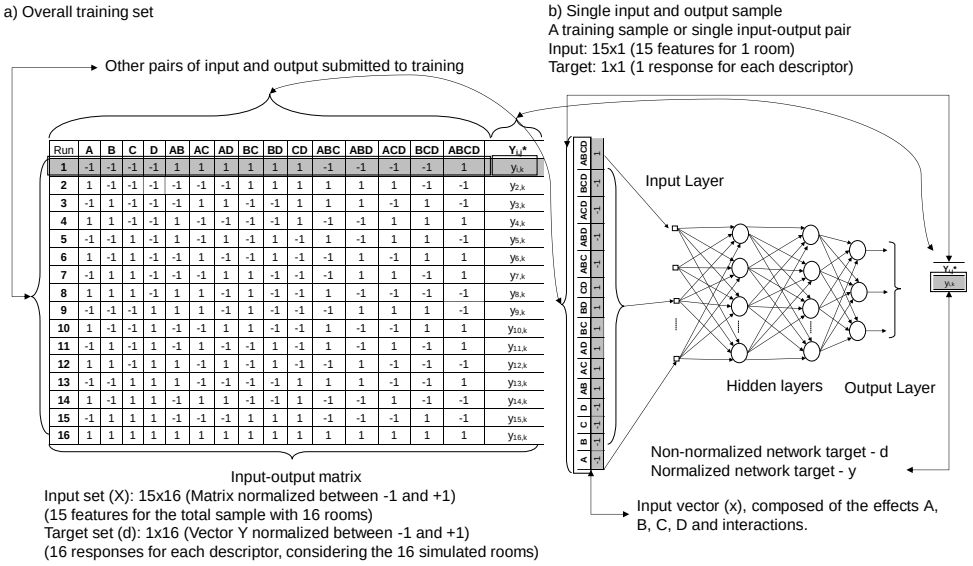

Fig. 2. Attribution of the input/output relationship in training the network.

$y_{1,1}$ (scalar) in the ANN and its relationship with the DOE for a classroom.

The algorithm described in Fig. 2 was used for each of the five classrooms where measurements were taken, and 16 virtual rooms were generated corresponding to the curves (runs). Thus, a total of 80 virtual classrooms were created using ODEON software. As mentioned earlier, the $X_{16,15}$ matrix was then adopted as the ANN input, while the target for subsequent ANN training was the $Y_{PCA_{16.6}}$ vector. Lastly, the ANN was trained with an input composed of 15 features, which were the main factors A, B, C, D and their respective 2nd order (AB, AC, AD, BC, CD), 3rd order (ABC, ABD, ACD, BCD) and 4th order (ABCD) interactions, and only one scalar as output.

Regarding the training dataset, a correlation matrix was evaluated in two ways: (a) by intra-correlations between the responses for the same descriptor, in the octave bands, and (b) the intercorrelations between the different descriptors T30, Ts, EDT, D50, U50, and STI in their octave bands.

2.4.2. Artificial neural network topological structures

The multilayer perceptron (MLP) architecture was utilized to train the ANN. The architecture of a network determines how artificial neurons are arranged and how signals flow between them (LeCun *et al.*, 2015). The term network topology indicates the number of neurons contained in these hidden layers. The ANN model was employed as a regression problem for multidimensional function approximation, leveraging the universal approximation theorem. This approach enabled the sensitivity analysis of the effects of controllable factors on the ANN models, as described in Subsec. 2.5. Table 5 outlines the design of the ANN applied in this study.

ANNs are subject to overfitting, which occurs when they approximate the training data too closely, causing them to lose their ability to generalize. Therefore, early stopping and holdout were employed to reduce the possibility of overfitting, according to the heuristics proposed by Piotrowski and Napiorkowski (2013).

Having trained the ANN, the network inference phase began, after the synaptic weights had already been adjusted. Thus, the inference of a network consisted of applying its input values (X) and calculating its estimated output (y). The quality of ANN training was evaluated by comparing the target (d) and the estimated neural network (y), using Pearson's R^2 correlation coefficient and the MSE as performance metrics. Lastly, an ANN was trained for each of the outputs $T30_{PCA}$, EDT_{PCA}, $C50_{PCA}$, $D50_{PCA}$, Ts_{PCA}, and STI_{PCA}.

The optimized response from the 50 independent training sessions of the neural network, called the average equivalent network, was calculated for each topology. All the computational implementations were developed in MATLAB®.

2.5. Artificial neural networks input variable sensitivity

In this work, the modified profile method (MPM) (Gevrey *et al.*, 2003; Do Nascimento, Oliveira, 2016) was applied to determine the relative significance of the input variables. The MPM calculates the significance of the ANN input variables, using the angular coefficient of the linear regression of the profile curve. Hence, the significance rating for the 15-input vector, A, B, C, D, AB, AC, AD, BC, BD, CD, ABC, ABD, ACD, BCD, and ABCD was calculated as the average of the effects in the five evaluated classrooms. Based on the non-replicated DOE, the significance of the MPM effects was estimated by transforming the angular coefficients into the z-score.

Multiple linear regression (MLR) was used to benchmark the MPM results. MLR was applied directly into the design matrix to compare the sensitivity of each input variable in the output. The sensitivity was taken as the regressor values shown in Eq. (2). The regressors were estimated using the least squares method, which yields the following equation, $\widehat{\beta} = \left(X^{\mathrm{T}}X\right)^{-1} y$, where $\mathbf{X}$ is the design matrix shown in Table 3, X^{T} is the transpose design matrix, $(\cdot)^{-1}$ is the inverse matrix operator, y is the output variable under analysis,

Table 5. Configurations of the neural network design.

Architecture	MLP
Number of inputs/number of outputs	15/1
Topology 1	MLP 15-5-5-1
Topology 2	MLP 15-10-10-1
Topology 3	MLP 15-15-15-1
Topology 4	MLP 15-20-20-1
Topology 5	MLP 15-30-30-1
Topology 6	MLP 15-35-35-1
Training algorithm	Error backpropagation, optimized by Levenberg–Marquardt
Activation functions in hidden layers	Hyperbolic tangent function
Activation function in the output layer	Linear function

y (= T30, EDT, C50, D50, Ts, STI), one output at a time. The estimated model is $\widehat{y} = X\,\widehat{\beta}$, and the residual values are $e = y - \widehat{y}$. The quality of the regression was evaluated using the sum of residual squares, given as $SS_r = y^{\mathrm{T}}y - \widehat{\beta}^{\mathrm{T}}X^{\mathrm{T}}y$ and through Pearson's squared linear correlation coefficient (R^2), usually known as the determination coefficient.

3. Results and discussion

Figure 3 shows the values of the measured descriptors. As can be seen, the hypothesis proposed by Montgomery (2012), who recommended that the DOE should be carried out in the most differentiated conditions possible, was met since the descriptors showed high statistical variability.

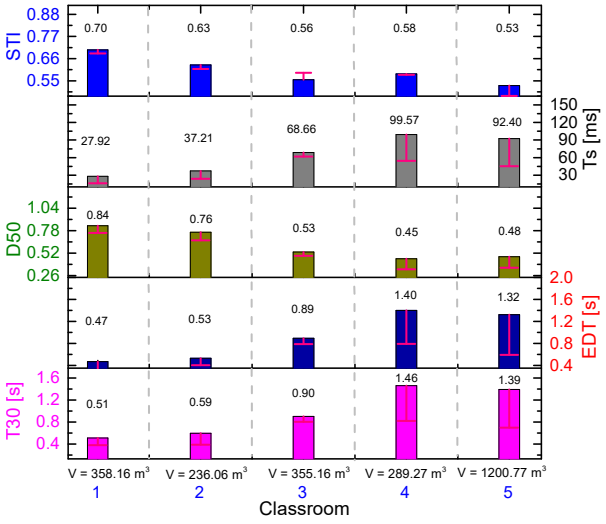

Fig. 3. Mean STI, T30, EDT, D50, and Ts values measured in the classrooms.

Table 6 lists the measured background noise levels and the respective NC of the classrooms. The NC was used as an input for ODEON software, as indicated in Fig. 2.

Table 7 lists the average measured values of SNR in the octave bands. The STI values for each classroom were calculated according to the IEC (2011) standard Annex M: adjustments to measured STI and STIPA results for simulation of occupancy noise and different speech levels. The BGN level is listed in Table 6, and the input sound source signal level was equalized at L_{eq} 60 dB, as described in Sec. 2. The corresponding mean STI values measured in classrooms 1 up to 5 were 0.70 ± 0.06, 0.63 ± 0.02, 0.56 ± 0.02, 0.58 ± 0.02, and 0.53 ± 0.03, respectively.

According to Subsec. 2.2, the classroom simulations were validated by comparing the most significant percent error values through the calibration curve and the R^2 value, as shown in Fig. 4.

Table 8 describes the generalization of the procedure highlighted in Fig. 4 to calculate validation errors pertaining to the other evaluated classrooms.

Using the validated virtual classroom models, new simulations were generated corresponding to the conditions illustrated in Fig. 2, which comprised the 15 combinations between factors A, B, C, and D. Figure 5 illustrates the distribution of the responses of descriptors T30, EDT, Ts, D50, U50, and STI as a function of each DOE run, using the results of the classroom two as an example.

More than 95 % of the variance was explained solely by the first principal component through the PCA method, as was also the case in the other classrooms. All the multi-band information contained in Fig. 5 was therefore condensed in Fig. 6.

Figure 7 shows the correlation matrix for the descriptors EDT, T30, Ts, D50, C50, U50, and STI obtained after applying the PCA.

As can be seen, Fig. 7 and Table 9 indicate the correlations and significance between the groups. Thus,

Table 6. Background noise level [dB] measured in the classrooms and noise criteria (NC) rating.

Room	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	NC
1	34.5	30.0	23.2	21.4	21.1	19.6	17.0	23
2	48.3	38.6	32.2	32.0	29.9	23.0	19.5	35
3	48.0	41.0	40.2	38.9	30.0	25.3	24.9	30
4	48.9	42.6	34.6	28.9	25.1	21.9	21.4	32
5	50.4	44.9	36.6	31.9	29.3	22.8	18.0	34

Table 7. Mean measured SNR values.

Room	SNR, mean ± standard deviation						
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
1	18.61 ± 4.62	25.61 ± 1.75	22.39 ± 0.68	19.22 ± 3.44	11.83 ± 2.43	8.11 ± 2.95	-3.61 ± 2.66
2	21.28 ± 2.67	20.48 ± 0.92	19.48 ± 0.59	17.44 ± 1.23	10.96 ± 1.59	6.04 ± 1.51	-9.32 ± 2.21
3	15.94 ± 3.73	20.19 ± 2.64	19.13 ± 3.01	18.31 ± 3.34	13.63 ± 3.03	10.13 ± 2.50	-1.44 ± 1.93
4	8.00 ± 3.28	16.17 ± 3.69	20.67 ± 3.26	23.17 ± 2.66	21.33 ± 2.77	20.83 ± 1.64	15.92 ± 1.62
5	14.09 ± 2.43	22.55 ± 3.08	21.91 ± 2.30	19.36 ± 1.80	11.09 ± 2.02	9.27 ± 1.85	-2.82 ± 1.83

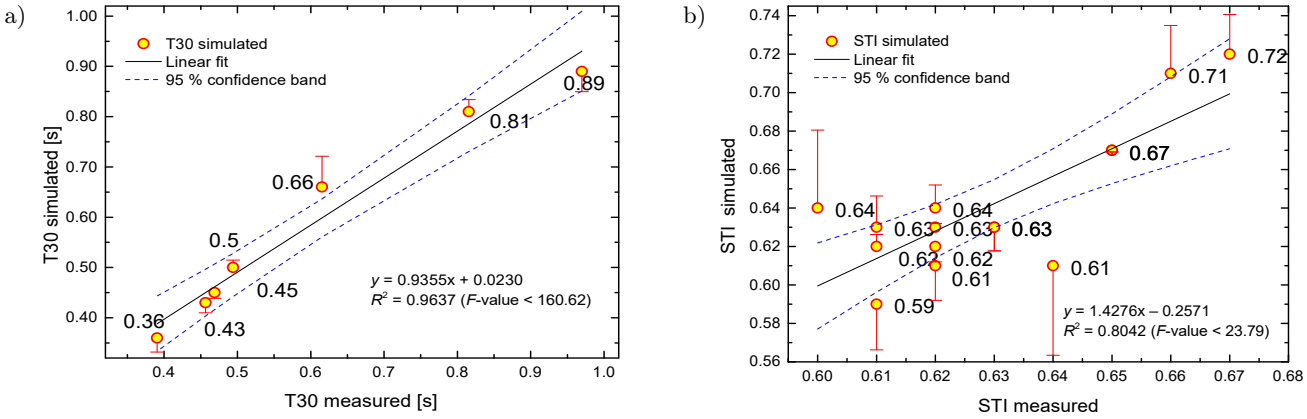

Fig. 4. Calibration curve of: a) T30; b) STI in classroom 2. The error bar is the residual value, i.e., the difference between the measured value and the corresponding value estimated by the linear fit.

Table 8. Forecast accuracy validation metrics for ODEON acoustic model via MPE, RMSPE and correlation.

Room	Estimation type	STI		T30 [s]								
		Mean	RMSPE, r^2	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	Mean	RMSPE, r^2
1	Measured	0.70	7.14 %, 0.47	0.78	0.64	0.45	0.37	0.44	0.46	0.46	0.51	6.03 %; 0.96
	Simulated	0.72		0.75	0.62	0.42	0.40	0.43	0.47	0.41	0.50	
	MPE	1.33 %	–	4.03 %	2.67 %	7.18 %	7.02 %	3.04 %	1.73 %	10.87 %	–	–
2	Measured	0.63	3.99 %, 0.80	0.97	0.82	0.62	0.49	0.47	0.46	0.39	0.60	5.77 %; 0.96
	Simulated	0.64		0.89	0.81	0.66	0.50	0.45	0.43	0.36	0.59	
	MPE	2.00 %	–	8.22 %	0.67 %	7.27 %	1.21 %	4.00 %	5.81 %	7.81 %	–	–
3	Measured	0.56	5.03 %, 0.57	0.71	0.83	0.92	1.02	1.02	0.92	0.81	0.89	6.35 %; 0.91
	Simulated	0.54		0.77	0.81	0.89	0.95	0.98	0.84	0.74	0.85	
	MPE	3.48 %	–	8.33 %	2.90 %	3.13 %	6.61 %	3.52 %	8.48 %	8.14 %	–	–
4	Measured	0.58	3.88 %; 0.32	2.62	2.08	1.58	1.08	1.07	1.06	0.86	1.29	4.91 %; 0.99
	Simulated	0.59		2.68	1.99	1.60	1.01	0.99	1.03	0.80	1.24	
	MPE	0.43 %	–	2.43 %	4.42 %	1.54 %	6.12 %	7.34 %	2.63 %	6.52 %	–	–
5	Measured	0.53	6.63 %; 0.77	2.60	2.11	1.31	1.01	1.01	0.95	0.77	1.39	4.19 %; 0.99
	Simulated	0.55		2.54	2.10	1.22	0.98	0.98	0.88	0.78	1.35	
	MPE	5.01 %	–	2.13 %	0.50 %	6.71 %	3.09 %	2.91 %	7.29 %	1.38 %	–	–

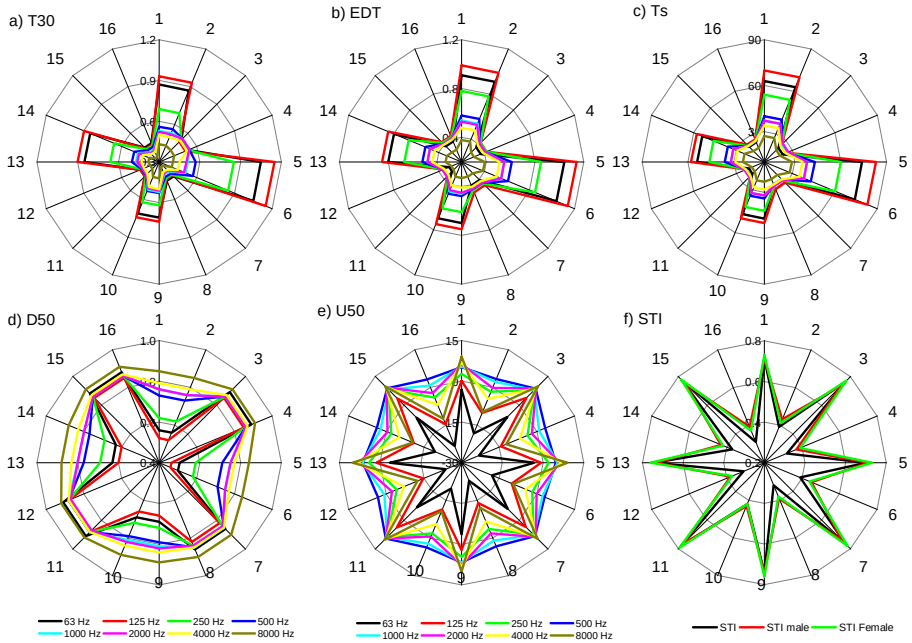

Fig. 5. Simulation of descriptors for each run for classroom 2.

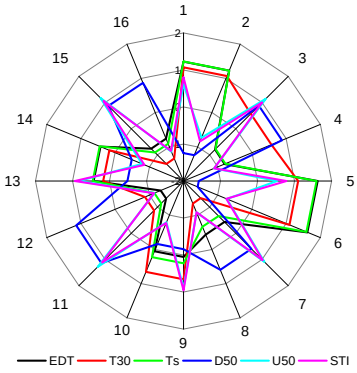

Fig. 6. Dimensional reduction of the acoustic descriptors.

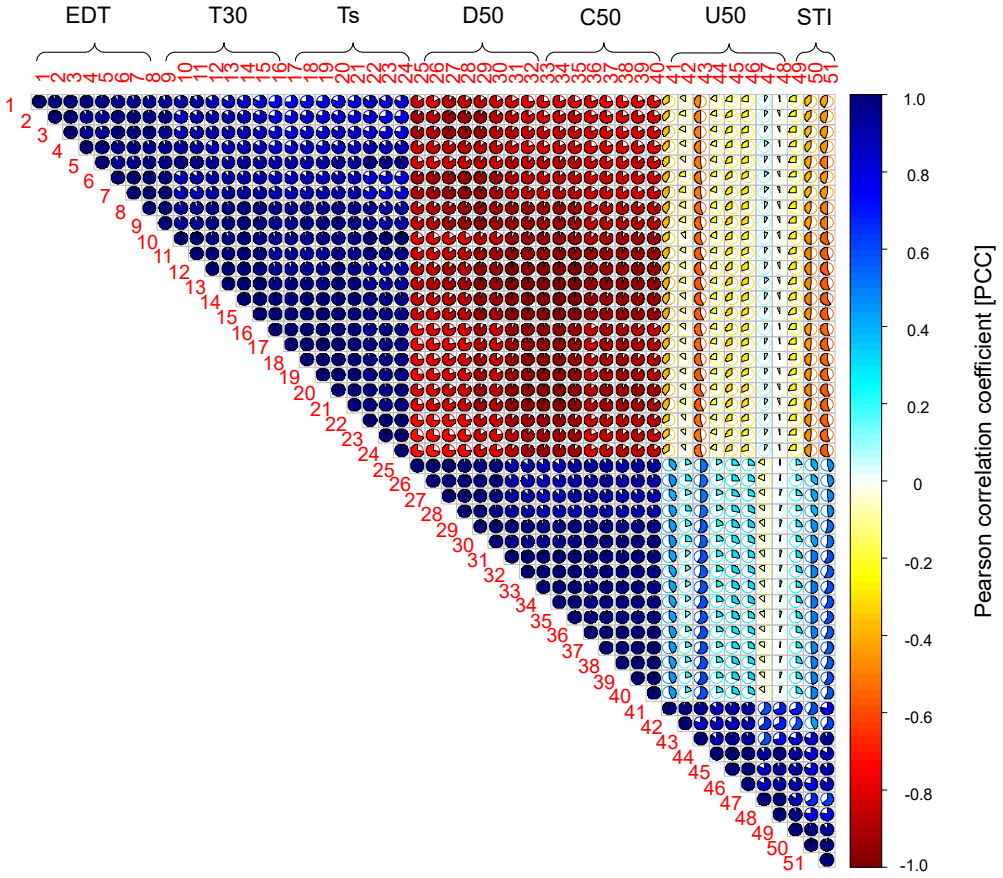

Fig. 7. Correlation matrix plot between all the acoustic descriptors.

Table 9. Correlations between the evaluated descriptors and their significance.

	EDT [s]	T30 [s]	Ts [ms]	D50	C50 [dB]	U50 [dB]	STI
EDT [s]	1						
T30 [s]	0.947*	1					
Ts [ms]	0.996*	0.948*	1				
D50	−0.963*	−0.902*	−0.965*	1			
C50 [dB]	−0.914*	−0.863*	−0.912*	0.977*	1		
U50 [dB]	−0.273*	−0.254*	−0.266*	0.267*	0.257*	1	
STI	−0.281*	−0.252*	−0.278*	0.269*	0.256*	0.975*	1

Note: a 2-tailed test of significance is used; *correlation is significant at the 0.05 level.

upon analyzing the correlation matrix, three groups were observed, identified by the dark blue samples, i.e., the 1st (EDT, T30, and Ts), the 2nd (C50 and D50), and the 3rd (U50 and STI).

In the benchmarking literature, Tang (2008) made a joint evaluation of the speech-related acoustic descriptors using regression models. The author reported the following correlations:

STI (D50) r^2 = 0.893 (+); STI (C80) r^2 = 0.916 (+);

STI (Ts) r^2 = 0.907 (−); STI (RT) r^2 = 0.903 (−).

For comparison, the following correlations were extracted from Table 9:

STI (D50) r^2 = 0.269 (+); STI (C50) r^2 = 0.256 (+);

STI (Ts) r^2 = 0.278 (−); STI (RT) r^2 = 0.252 (−).

Thus, it can be inferred that the effects followed the same vein.

Additionally, a literature review by Minelli *et al.* (2022) found that the acoustic descriptors are significantly affected by the occupation condition of built environments (factor D), i.e., unoccupied or occupied. This finding is in line with ours, which indicated that the entire acoustic response was altered in various ways by a combined change in construction factors (A, B, C, and D).

Likewise, Croce *et al.* (2023) and Choi (2020) independently established relationships between STI, U50, and RT and assessed the influence of construction factors. Choi (2020) determined that, quantitatively, C50(125–4 kHz) is more strongly correlated with U50(125–4 kHz) r^2 = 0.824 than with STI (0.627). Hence, the results of linear regression indicated that C50, specifically, plays a more significant role in increasing r^2 values of both U50 and STI, outweighing the impact of the background noise component (factor A). This conclusion can be seen in the last two rows of Table 10, where the encoded effects on U50 and STI were the same (−)A, (+)B, (−)AB, and (+)D, and as proxy, the correlation STI(U50) was 0.975 (+), while the other correlations were lower: C50(U50) 0.257 (+), C50(STI) 0.269 (+).

The r^2 between U50(125–4 kHz) and C50(125–4 kHz) is 0.824, while that between STI and C50(125–4 kHz) is 0.627. This indicates that the correlation between STI and C50 is not much stronger, which resembles the result described in Table 9, i.e., C50(STI) 0.269 (+). In an extended discussion about the role of early reflections on C50 and their implications for speech intelligibility, Prodi *et al.* (2022) reported the same findings. This was highlighted in the current work by the MPM, which separated C50 and D50 from the STI and U50 group.

Graphically, to assess the significance rating, Fig. 8 illustrates the effects of A, B, C, and D, and their respective interactions in the classrooms on the evaluated descriptors, using the MPM.

In Fig. 8, note the spatial uniformity of the significance rating, which is why the discussion was gener-

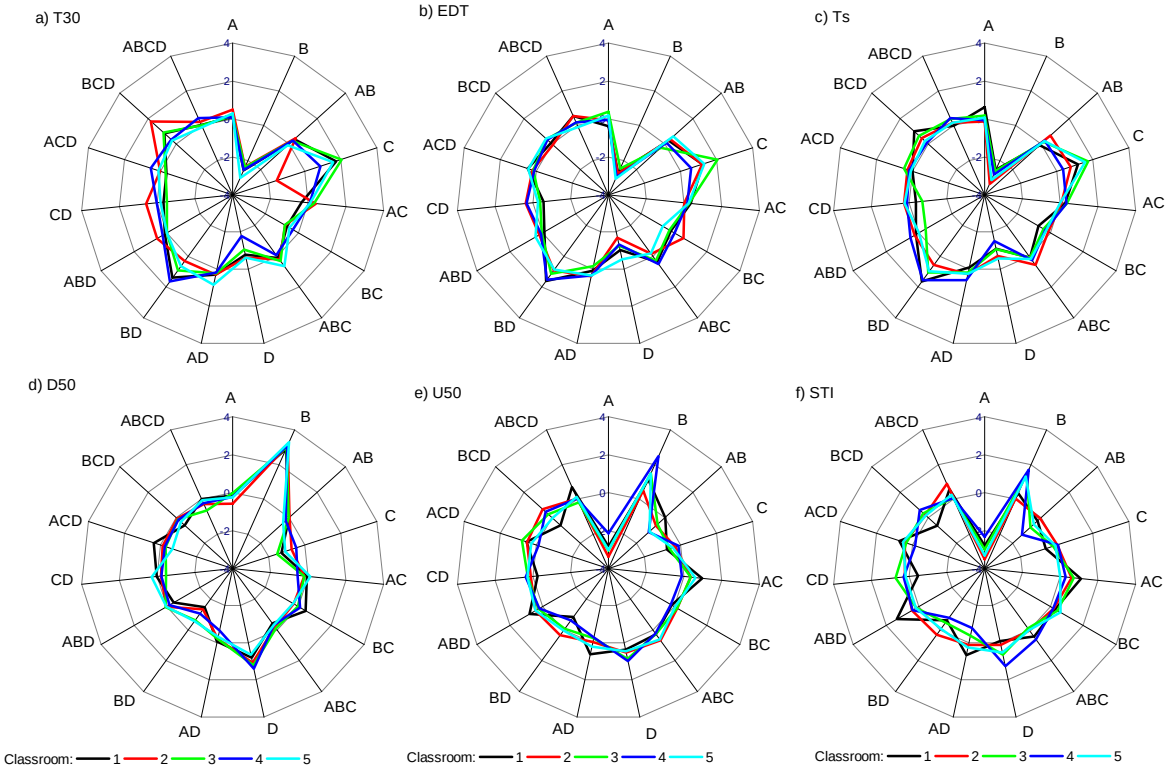

Fig. 8. Comparison of the significance rating in the classrooms.

Table 10. Comparison of the absolute effects on the descriptors: MPM versus MLR.

Group	Parameter	Rating and effects				
			1	2	3	4
1	EDT [s]	Coded effect	(−)B	(+)C	(+)BD	(−)D
		MPM magnitude	2.902	1.242	1.212	1.202
		MLR magnitude	2.963	1.348	1.133	1.222
	T30 [s]	Coded effect	(−)B	(+)BD	(−)D	(+)C
		MPM magnitude	2.728	1.210	1.076	0.834
		MLR magnitude	2.858	1.492	1.175	1.155
	Ts [ms]	Coded effect	(−)B	(+)BD	(−)D	(+)C
		MPM magnitude	2.944	1.236	1.146	1.134
		MLR magnitude	2.968	1.365	1.221	1.084
2	D50	Coded effect	(+)B	(−)C	(−)BD	(+)D
		MPM magnitude	3.096	1.096	1.074	1.044
		MLR magnitude	3.168	1.076	1.048	1.046
3	U50 [dB]	Coded effect	(−)A	(+)B	(−)AB	(+)D
		MPM magnitude	2.906	1.288	0.754	0.704
		MLR magnitude	3.347	0.910	0.598	0.582
	STI	Coded effect	(−)A	(+)B	(+)D	(−)AB
		MPM magnitude	3.072	0.996	0.632	0.582
		MLR magnitude	3.340	0.904	0.631	0.565

alized. Moreover, it was found that the groups in the effect analysis were the same as those in the correlation analysis. Table 10 quantifies the significance ratings of the impact shown in Fig. 8. The four main effects in each group are highlighted with a (+) sign, indicating an increase in the descriptor, while the (−) sign indicates a decrease in the descriptor as the effect varies from −1 to +1.

As can be seen in Table 10, the acoustic parameters were given the same rating by both MPM and MLR, matching the construction design factors. Likewise, it was found that the rating was identical for each acoustic parameter, confirming that the MPM showed excellent convergence when benchmarked with the MLR.

In the first group (EDT, T30, and Ts), the most important factors were: (−)B, (+)BD, (−)D, and (+)C. The increase in absorption (B) from −1 to +1, implied a decrease in EDT, T30, and Ts. The same finding was reported by Beranek (2006).

The second group (C50, D50) responded, in terms of magnitude, as (+)B, (−)C, (−)BD, and (+)D. The increase in absorption (B), from −1 to +1 implied an increase in D50. Similarly, a positive correlation was found between D50 and both U50 and STI, and these results are promising, since D50 is strongly associated with STI (Bradley *et al.*, 2003; Croce *et al.*, 2023; Choi, 2020). It should be noted that the higher the D50, the greater the intelligibility, according to the relationship described by Bradley *et al.* (2003).

In the third group (U50 and STI), the first effect with the highest impact on intelligibility was background noise (A). In contrast, the second most significant effect was sound absorption (B). As a result, the increase in background noise from −1 to +1 caused a decrease in STI, while sound absorption alone did not exert a significant effect. Thus, the importance ratio of effect A on B was approximately three-fold. The factor A explained most of the variation in STI. A similar finding has been reported in several studies (Bistafa, Bradley, 2000; Rennies *et al.*, 2014; Sato *et al.*, 2012; Leccese *et al.*, 2018).

It is essential to note that while this work provided a comprehensive understanding of the relationships between various acoustic descriptors and building design, there is room for improvement to enhance the consistency and reliability of results. Firstly, the study focused on a specific set of factors (A, B, C, and D) and their influence on acoustic parameters, potentially overlooking other variables that could contribute to the acoustic environment. Moreover, some limitations of this work must be recognized, including the potential for unaccounted variables or interactions that could influence the acoustic descriptors, as well as reliance on specific models and methods (MPM and MLR) that may not detect all the nuances of the acoustic environment. To improve the reliability of the results, further studies involving larger sample sizes or different types of classrooms would be advisable to validate the findings. Furthermore, incorporating more advanced modeling techniques or considering additional variables, such as room geometry or occupants' positions, could provide a more comprehensive picture of the relationships between acoustic descriptors and building design factors. Lastly, the study's reliance on correlational analyses means that other unmeasured variables could potentially lead to confusion about the observed relationships, underscoring the need for additional research to explore these relationships.

4. Conclusions

In conclusion, this work involved an in-depth comparison of acoustic parameters using both MPM and MLR methods, revealing a high degree of convergence between the two methods. This analysis, guided and benchmarked by a review of the literature, provides a comprehensive understanding of the complex relationships between various acoustic descriptors and building design.

The correlation analyses of U50, STI, and C50, as discussed in the context of (Choi, 2020) and validated in the current study, elucidate the intricate interaction between these predictors of speech intelligibility. The observed predominance of C50 in influencing both U50 and STI values, as described in Table 10, supports Choi's (2020) quantitative findings, and underscores the pivotal role of room acoustic parameters in enhancing speech intelligibility.

As for the significance of controllable factors, it was concluded that the sound absorption, the factor (B), interferes more strongly in reverberation-related descriptors (groups 1 and 2), while the background noise, the factor (A), strongly interferes with the STI. The factor (C) was the one that interfered the least with the acoustic descriptors in general. Moreover, it was shown that D50 responds to the acoustic conditions of the classroom in the same way as EDT, Ts, and T30 do. The only difference is that RT decreases and D50 increases in response to an increased classroom sound absorption. Surprisingly, the combined 3rd and 4th order interactions had negligible effects in the classrooms of this study.

Acknowledgments

This study was partly financed by the Brazilian research funding agency CAPES (Federal Agency for the Support and Improvement of Higher Education) under the Finance Code 001.

References

1. American National Standard (2008), *Criteria for evaluating room noise*, Standard ANSI S12.2:2008.
2. Ansay S., Zannin P.H.T. (2016), Using the parameters of definition, D50, and reverberation time, RT, to investigate the acoustic quality of classrooms, *Canadian Acoustics/Acoustique Canadienne*, **44**(4): 6–11.
3. Baydin A.G., Pearlmutter B.A., Radul A.A., Siskind J.M. (2017), Automatic differentiation in machine learning: A survey, *Journal of Machine Learning Research*, **18**(1): 5595–5637.
4. Beranek L.L. (2006), Analysis of Sabine and Eyring equations and their application to concert hall audience and chair absorption, *The Journal of the Acoustical Society of America*, **120**(3): 1399–1410, doi: 10.1121/1.2221392.
5. Beranek L.L., Takayuki H. (1998), Sound absorption in concert halls by seats, occupied and unoccupied, and by the hall's interior surfaces, *The Journal of the Acoustical Society of America*, **104**(6): 3169–3177, doi: 10.1121/1.423957.
6. Bistafa S.R., Bradley J.S. (2000), Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics, *The Journal of the Acoustical Society of America*, **107**(2): 861–875, doi: 10.1121/1.428268.
7. Bistafa S.R., Bradley J.S. (2001), Predicting speech metrics in a simulated classroom with varied sound absorption, *The Journal of the Acoustical Society of America*, **109**(4): 1474–1482, doi: 10.1121/1.1354199.
8. Bradley J.S. (2011), Review of objective room acoustics measures and future needs, *Applied Acoustics*, **72**(10): 713–720, doi: 10.1016/j.apacoust.2011.04.004.
9. Bradley J.S., Reich R.D., Norcross S.G. (1999), On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility, *The Journal of the Acoustical Society of America*, **106**(4): 1820–1828, doi: 10.1121/1.427932.
10. Bradley J.S., Sato H., Picard M. (2003), On the importance of early reflections for speech in rooms, *The Journal of the Acoustical Society of America*, **113**(6): 3233–3244, doi: 10.1121/1.1570439.
11. Choi Y.-J. (2017a), Comparison of two types of combined measures, STI and U50, for predicting speech intelligibility in classrooms, *Archives of Acoustics*, **42**(3): 527–532, doi: 10.1515/aoa-2017-0056.
12. Choi Y.-J. (2017b), Predicting classroom acoustical parameters for occupied conditions from unoccupied data, *Applied Acoustics*, **127**: 89–94, doi: 10.1016/j.apacoust.2017.05.036.
13. Choi Y.-J. (2020), Evaluation of acoustical conditions for speech communication in active university classrooms, *Applied Acoustics*, **159**: 107089, doi: 10.1016/j.apacoust.2019.107089.
14. Christensen C.L., Koutsouris G., Rindel J.H. (2014), Estimating absorption of materials to match room model against existing room using a genetic algorithm, [in:] *Forum Acusticum 2014*, pp. 7–12, doi: 10.13140/2.1.1588.8647.
15. Croce P., Leccese F., Salvadori G., Berardi U. (2023), Proposal of a simplified tool for early acoustics design stage of classrooms in compliance with speech intelligibility thresholds, *Energies*, **16**(2): 813, doi: 10.3390/en16020813.
16. do Nascimento E.O., de Oliveira L.N. (2016), Sensitivity analysis of cutting force on milling process using factorial experimental planning and artificial neural networks, [in:] *IEEE Latin America Transactions*, **14**(12): 4811–4820, doi: 10.1109/TLA.2016.7817015.
17. do Nascimento E.O., Zannin P. (2023), Deep learning applied to speech transmission index prediction: Simulations and measurements, *Harvard Dataverse*, doi: 10.7910/DVN/RZRUTT.
18. Gevrey M., Dimopoulos I., Lek S. (2003), Review and comparison of methods to study the contribution

of variables in artificial neural network models, *Ecological Modelling*, **160**(3): 249–264, doi: 10.1016/S0304-3800(02)00257-0.

19. Houtgast T., Steeneken H.J.M., Plomp R. (1980), Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function. I. General Room Acoustics, *Acta Acustica united with Acustica*, **46**(1): 60–72.
20. International Electrotechnical Commission (2011), *Sound system equipment. Part 16: Objective rating of speech intelligibility by speech transmission index* (IEC Standard No. 60268-16).
21. International Organization for Standardization (2008), *Acoustics – Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms* (ISO Standard No. 3382-2), https://www.iso.org/standard/36201.html.
22. Jolliffe I. (2011), Principal component analysis, [in:] *International Encyclopedia of Statistical Science*, Lovric M. [Ed.], Springer, Berlin, Heidelberg, doi: 10.1007/978-3-642-04898-2_455.
23. Jolliffe I.T., Cadima J. (2016), Principal component analysis: A review and recent developments, *Philosophical Transactions of the Royal Society A*, **374**(2065): 20150202, doi: 10.1098/rsta.2015.0202.
24. Kang S., Mak C.M., Ou D., Zhou X. (2023), Effects of speech intelligibility and reverberation time on the serial recall task in Chinese open-plan offices: A laboratory study, *Applied Acoustics*, **208**: 109378, doi: 10.1016/j.apacoust.2023.109378.
25. Leccese F., Rocca M., Salvadori G. (2018), Fast estimation of Speech Transmission Index using the Reverberation Time: Comparison between predictive equations for educational rooms of different sizes, *Applied Acoustics*, **140**: 143–149, doi: 10.1016/j.apacoust.2018.05.019.
26. LeCun Y, Bengio Y. Hinton G. (2015), Deep learning, *Nature*, **521**: 436–444, doi: 10.1038/nature14539.
27. LeCun Y., Bottou L., Bengio Y., Haffner P. (1998), Gradient-based learning applied to document recognition, [in:] *Proceedings of IEEE*, **86**(11): 2278–2324, doi: 10.1109/5.726791.
28. Lochner J., Burger J. (1960), Optimum reverberation time for speech rooms based on hearing characteristics, *Acta Acoustica*, **10**: 394–399.
29. Meyer E., Kunstmann D., Kuttruff H. (1964), On the measurements of sound absorption of audiences [in German: Über einige Messungen zur Schallabsorption von Publikum], *Acta Acoustica*, **14**(2): 119–124.
30. Mikulski W., Radosz J. (2011), Acoustics of classrooms in primary schools – Results of the reverberation time and the speech transmission index assessments in selected buildings, *Archives of Acoustics*, **36**(4): 777–793, doi: 10.2478/V10168-011-0052-6.
31. Minelli G., Puglisi G.E., Astolfi A. (2022), Acoustic parameters for learning in classroom: A review, *Building and Environment*, **208**: 108582, doi: 10.1016/j.buildenv.2021.108582.
32. Montgomery D.C. (2012), *Design and Analysis of Experiments*, 8th ed., John Wiley & Sons, Inc.
33. Piotrowski A.P., Napiorkowski J.J. (2013), A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling, *Journal of Hydrology*, **476**: 97–111, doi: 10.1016/j.jhydrol.2012.10.019.
34. Prodi N., Pellegatti M., Visentin C. (2022), Effects of type of early reflection, clarity of speech, reverberation and diffuse noise on the spatial perception of a speech source and its intelligibility, *The Journal of the Acoustical Society of America*, **151**: 3522–3534, doi: 10.1121/10.0011403.
35. Rennies J., Schepker H., Holube I., Kollmeir B. (2014), Listening effort and speech intelligibility in listening situations affected by noise and reverberation, *The Journal of the Acoustical Society of America*, **136**: 2642–53, doi: 10.1121/1.4897398.
36. Rindel J.H. (2012), ODEON APPLICATION NOTE – ISO 3382-3 Open plan offices. ODEON – Room Acoustic Software.
37. Russell S., Norvig P. (1996), Artificial intelligence – A modern approach, [in:] *The Knowledge Engineering Review*, doi: 10.1017/S0269888900007724.
38. Sala E., Rantala L. (2016), Acoustics and activity noise in school classrooms in Finland, *Applied Acoustics*, **114**: 252–259, doi: 10.1016/j.apacoust.2016.08.009.
39. Sato H., Morimoto M., Miyagawa Y., Suzuki Y. (2016), Relationship between speech intelligibility and objective measures in sound fields with a discrete long-path echo, *The Journal of the Acoustical Society of America*, **140**(4): 3193, doi: 10.1121/1.4970046.
40. Sato H., Morimoto M., Wada M. (2012), Relationship between listening difficulty rating and objective measures in reverberant and noisy sound fields for young adults and elderly persons, *The Journal of the Acoustical Society of America*, **131**: 4596–4605, doi: 10.1121/1.4714790.
41. Sato H., Nishikawa Y., Sato H., Morimoto M. (2006), The relation between speech transmission index, clarity, and reverberation time and listening difficulty in the impulse response database of AIJ, *The Journal of the Acoustical Society of America*, **120**(5): 3321, doi: 10.1121/1.4781197.
42. Secchi S. *et al.* (2017), Effect of outdoor noise and façade sound insulation on indoor acoustic environment of Italian schools, *Applied Acoustics*, **126**: 120–130, doi: 10.1016/j.apacoust.2017.05.023.
43. Tang S.K. (2008), Speech-related acoustic parameters in classrooms and their relationships, *Applied Acoustics*, **69**(12): 1318–1331, doi: 10.1016/j.apacoust.2007.08.008.
44. Visentin C., Prodi N., Cappelletti F., Torresin S., Gasparella A. (2018), Using listening effort assessment in the acoustical design of rooms for speech, *Building and Environment*, **136**: 38–53, doi: 10.1016/j.buildenv.2018.03.020.