Investigation the Capability of Neural Network in Predicting Reverberation Time on Classroom

Musli Nizam Yahya*1, Toru Otsuru2, Reiji Tomiku2, Takeshi Okozono1

1Graduate School of Engineering, Oita University, 700 Dannoharu, Oita, 870-1192, Japan.
2Department of Architecture and Mechatronics, Architecture Course, Faculty of Engineering, Oita University, 700 Dannoharu, Oita, 870-1192, Japan.

*Corresponding mail: musli_nizam2000@yahoo.com

Abstract

The purpose of this paper is to investigate the capability of neural network in predicting a classroom’s reverberation time. A classroom in Oita University was chosen as a sample to obtain the virtual data (reverberation time) based on 20 types of sound absorptions coefficients using Finite Element Method (FEM) and Sabine equation. The capability of FEM has shown that it is able to simulate virtual data in each location of a classroom. To develop a neural network model, virtual data (721 data) was taken from FEM for the learning process. The assessment was made by using testing subset (20% from 721 data) to verify the performance. The testing’s means square error (MSE) was $3.7751 \times 10^{-4}$ and correlation coefficient ($R^2$) was 0.992 approximately to 1. The optimum network used was 4 hidden nodes. Extended assessment was made using the unseen data (35 data) and it showed that neural network prediction was approximately close to the actual data with MSE is $4.154 \times 10^{-4}$. Basically, the capability of reverberation time prediction using neural network is shown in this paper.

Keywords: Classroom, Reverberation time, Neural network, Finite element methods (FEM).
1.0 INTRODUCTION

Classrooms are places where teaching and learning processes happen. Thus, to obtain better speech intelligibility it is recommended that reverberation time for the classrooms should be below 1s [1]. Three factors bring effect to reverberation time: the volume of the room (size and ceiling height), proportions (shape), and material used on the walls, floor, and ceiling absorption sound energy [2][3]. Hodgson [4] studied intensively on surface of classroom mentioned that there were eight clusters in the classroom which are hard surface, paneled surface, glued-on acoustic tiles, suspended acoustical ceiling, carpeted surfaces, upholstered seats, porous absorbers and Helmholtz-resonator. Based on those factors, there were several theories created to predict reverberation time. The classical theory that was proposed by Sabine and Eyring [5] for diffuse sound level failed to give accurate prediction when the sound absorption coefficient was non-homogenous and different cubic shape [6]. Extended studies created many kinds of calculation methods because of the weaknesses in Sabine and Eyring equation. Therefore, the simulation and numerical method have been chosen to recover it. Nannariello and Fricke [7] came out with a research using neural network to predict reverberation time, \((RT_{125-500} \text{ and } RT_{500-1000})\) in large room. They wanted to prove the capability of neural network as a method that can be used for early stage in building designing. Assessment was made to compare Sabine and Eyring equation, ray tracing and neural network. It seems that neural network was able to be used in predicting reverberation time and gave better result than existing methods. In the first part of their second paper, Nannariello and Fricke [8] focused on predicting strength factors (dB), \(G\) value \((G_{\text{low}(125-250), G_{\text{mid}(500-1000)}, G_{\text{avg}(125,250,500 \text{ and } 1000)}\) and \(G_{p(125,250,500 \text{ and } 1000)})\). They investigated the sound distribution in enclosure whether the energy transmitted was deficient at some frequency. Enclosures used were in different sizes (rectangular, fan and shoebox), type position of sound source and receivers considered. The analysis indicated that, neural network can be trained and tested using numerical data. Subsequently, at the second part [9] of the paper, neural network was successfully used to predict acoustical attributes of concert hall (strength factor (dB), \(G\) value), \(C_{80}\) (clarity factor), and \(LF\) (lateral energy fraction) using complex data (noisy and poorly measured data). Both papers analyzed the ability of neural network by evaluating conventional method and computer technique (ray-tracing-ODEON). In predicted speech level in British Columbia University classrooms, Hodgson [10] also used neural network model, however, the Hopkins-Stryker equation and Barron’s revised model is applied for assessment. They investigated the capability of prediction of sound propagation at various positions in empty classrooms based on the room volume, room area and distance and location of sound source to listener.

Besides neural network, ray-tracing is another alternative tool to predict acoustic performance. Nannariello et al. [8][9][10] used ray tracing method to generate simulation data for neural network database and as an assessment in comparing it with neural network. Billon et al. [6] predicted the reverberation time in high absorbent room and one of the models used was ray tracing model. Arianna Astolfi et al. [11] used Sabine and Eyring equation and Hodgson equation to compare with ray tracing in order to predict the acoustical characteristics in small classroom. Bistafa and Bradley [1] analyzed analytical methods (7 formulas; Sabine, Eyring, Millington, Cremer’s absorption, Kuttruff’s absorption exponent, Fizroy’s absorption exponent, and Arau-Puchade) and ray-based computer programs (ODEON and Raynoise) on classroom. Overall, most of the papers found out that ray tracing is able to be used as a prediction tool.
Previous literatures have proposed several methods in obtaining reverberation time in the enclosure. Most of the papers compared a variety of prediction methods. In this case, this study uses FEM analysis because many researchers [12][13][14] verified that FEM is able to be applied in acoustic. The benefits of using FEM are: to provide a simulation form, have the ability to analyze non-homogeneous material and can preserve a very dynamic, linear or nonlinear data [15]. The important thing is FEM is reliable in low frequency (125 kHz to 1 kHz) rather than ray tracing (2 kHz to 16 kHz) but it will be confirmed in the next stage.

The objective of this study is to investigate the capability in predicting reverberation time in the classroom using neural network. In creating neural network model, it needs a set of data to be learned. To develop a set of data (reverberation time), FEM is used because of the reliability to simulate reverberation time in 500 Hz. At this stage FEM analysis will be compared with Sabine equation (Eq. (1)) to identify the advantage of FEM. Then, the simulation data will be fed into neural network for a learning process to recognize the data characteristics for prediction

\[ RT_{\text{sabine}} = \frac{0.161V}{S \bar{\alpha}} \]  

(1)

where \( V \) is room volume, in m\(^3\), \( \bar{\alpha} \) is average absorption co-efficient, and \( S \) is the total room surface area, in m\(^2\).

2.0 SIMULATION DATA (VIRTUAL DATA)

FEM is used to create the simulation data using FORTRAN [16] and is run using super computer linked from Kyushu University. Only one classroom is selected from Oita University where the dimension of the classroom is 7.08 x 6.09 x 3.02 m illustrated in Fig. 1. About six surfaces of sound absorption coefficients are considered; table (seat), floor, ceiling, doors, windows and wall. Each of them will be combined in different sound absorption coefficient in order to investigate the reverberation time distribution of each location. The sound absorption coefficients are taken from a literature reference [17] and Table 1 shows the sound absorption coefficient at 500 Hz [18]. FEM arguments can be expressed based on the combination of sound absorption coefficients as follows;

\[ RT (1 \rightarrow 7)_i = f_{\text{FEM}} (\alpha f_n, \alpha c_n, \alpha w_n, \alpha wd_n, \alpha s_n, \alpha d_n) \]  

(2)

where \( RT \) is reverberation time for location 1 to 8. \( i \) is the number of \( RT \) (\( i = 1,2,3,\ldots, 721 \)). \( \alpha f \) is floor sound absorption coefficient, \( \alpha c \) is ceiling sound absorption coefficient, \( \alpha wd \) is window sound absorption co-efficient, \( \alpha s \) is seat sound absorption coefficient and \( \alpha d \) is door sound absorption co-efficient. A \( n \) is type of sound absorption coefficient (\( n = 1,2,3 \)). Around 721 arguments are created to be used in neural network as input data. This will be discussed in the next sub topic.

<table>
<thead>
<tr>
<th>surface</th>
<th>code</th>
<th>Name of material</th>
<th>500 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>floors</td>
<td>fl</td>
<td>wood floor (parquet or flooring on stud)</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 1. List of Sound Absorption coefficient
<table>
<thead>
<tr>
<th>Material</th>
<th>Description</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>f2</td>
<td>pile carpet 10 mm thick</td>
<td>0.2</td>
</tr>
<tr>
<td>f3</td>
<td>cork floor tile (3-4 inch thick)-glue down</td>
<td>0.08</td>
</tr>
<tr>
<td>c1</td>
<td>1/4 inch fiberglass ceiling tile</td>
<td>0.88</td>
</tr>
<tr>
<td>c2</td>
<td>linear wood ceiling</td>
<td>0.7</td>
</tr>
<tr>
<td>c3</td>
<td>acoustic ceiling tile panel 1&quot;</td>
<td>0.42</td>
</tr>
<tr>
<td>w1</td>
<td>brick, bare concrete surface</td>
<td>0.02</td>
</tr>
<tr>
<td>w2</td>
<td>mortar smooth finishing, plaster, marble</td>
<td>0.02</td>
</tr>
<tr>
<td>w3</td>
<td>cloth finishing on concrete floor</td>
<td>0.03</td>
</tr>
<tr>
<td>wd1</td>
<td>glass 24oz operable window, closed</td>
<td>0.4</td>
</tr>
<tr>
<td>wd2</td>
<td>window glass (in wooden frame)</td>
<td>0.18</td>
</tr>
<tr>
<td>wd3</td>
<td>Plexiglas for illuminating</td>
<td>0.2</td>
</tr>
<tr>
<td>cr1</td>
<td>velvet curtain (with no drape)</td>
<td>0.13</td>
</tr>
<tr>
<td>cr2</td>
<td>velvet curtain draped to half (100mm air space)</td>
<td>0.55</td>
</tr>
<tr>
<td>cr3</td>
<td>500mm air space</td>
<td>0.5</td>
</tr>
<tr>
<td>s1</td>
<td>plywood chair</td>
<td>0.02</td>
</tr>
<tr>
<td>s2</td>
<td>theatre chair upholstered with moquette</td>
<td>0.3</td>
</tr>
<tr>
<td>s3</td>
<td>person sitting on upholstered chair</td>
<td>0.4</td>
</tr>
<tr>
<td>d1</td>
<td>solid timber door</td>
<td>0.06</td>
</tr>
<tr>
<td>d2</td>
<td>glass 10mm thickness</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Figure 1. Classroom dimension.
2.1. Setting of FEM

The number of elements and nodes of all the surfaces in the classroom (volume = 130.2 m³) including 16 seats/tables should be recognized before using FEM. Here, GiD9 software (1) is used to generate the mesh of surfaces using the rule: the value of \( \lambda/d > 4.8 \) (\( \lambda \) is acoustic wavelength, \( d \) is nodal distance) to get the elements’ and nodes’ numbers (2). This is shown in Fig. 2 and Table 2. The information on number of elements and nodes are important because these numbers will be applied in FEM calculation.

<table>
<thead>
<tr>
<th>Overall</th>
<th>Ceiling</th>
<th>Floor</th>
<th>Window</th>
<th>Door</th>
<th>Table</th>
<th>Wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of elements</td>
<td>60499</td>
<td>3078</td>
<td>2986</td>
<td>459</td>
<td>504</td>
<td>5752</td>
</tr>
<tr>
<td>No. of Nodes</td>
<td>515717</td>
<td>3190</td>
<td>3228</td>
<td>520</td>
<td>570</td>
<td>5944</td>
</tr>
</tbody>
</table>

Figure 2. Classroom mesh

2.2. Implementation of FEM (Time Domain Analysis)

The standard of finite element procedure is based on the principle of minimum of total potential energy applied to the three dimension sound field. The discretized matrix equation for the sound field in the frequency domain can be expressed as follows [20];

\[
\left(K + i\omega C - \omega^2 M\right)p = \ell \omega^2 u_n W
\]  

(3)

where \( M, C \) and \( K \) are acoustic mass, dissipation and stiffness matrices, respectively. Besides that, \( i, p, \rho, \omega, u_n \) and \( W \) are respectively imaginary unit (\( i^2 = -1 \)), sound pressure vector, the air density angular frequency, the particle velocity and distribution vector. By assuming \( \dot{} \) and \( \ddot{} \) are first order and second order derivatives in time, the semi discrete equation in time domain can be evaluated using equation (3) as shown below. To calculate the equation (3) the Newmark \( \beta \) [21] scheme is used to solve step by step as shown in equation (4) and (5).

\[
M \ddot{p} + C \dot{p} + Kp = \ell u_n W
\]  

(4)
\[ P_{n+1}^t = P_n^t + \Delta t \left( 1 - 2 \beta_H P_n^t + 2 \beta_H P_{n+1}^t \right) \]  
\[ \dot{P}_{n+1} = \dot{P}_n + \Delta t \left( 1 - 2 \gamma_H \dot{P}_n^t + \gamma_H \ddot{P}_{n+1}^t \right) \]  

where \( n \) and \( \Delta t \) is respectively denotes the time step counter, the time step value. \( \beta_H \) and \( \gamma_H \) are integration parameters. While \( P' \) is the vector of unknowns at the finite element node, \( \dot{P} \) and \( \ddot{P} \) are first and second derivative with respect to time. From the analyses, the simulation data will be fed into a neural network for learning process.

### 3.0 NEURAL NETWORK ANALYSIS

Neural network is an artificial intelligence which is inspired from the human brain. It can learn, generalize and organize data and store the processes similar with a brain. It can also recall all the storing knowledge similar with input data and recognize the pattern even though the pattern has never been presented before [22].

In this paper we use Multilayer Perceptron (MLP) network and train using back-propagation algorithm with Leverberg-Marquardt training (LMT) method. This training method is chosen because it is faster than other methods even though needs more memory [23] [24]. The MLP architecture has three layers of network, which are input layer, hidden layer and output layer as shown in Fig. 3. Each layer includes number of nodes to create network connections. These connections can be calculated using equations as follows;

\[ Y_{net} = \sum_{i=1}^{N} P_i W_i + W_o \]  
\[ Y_i = \frac{1}{1 + e^{-Y_{net}}} \]  
\[ E = \frac{1}{2} \sum_{i=1}^{N} (T_i - Y_i)^2 \]

where \( E \) is an error, \( T_i \) is an output desired and \( Y_i \) is the prediction, \( N \) is number of output, \( Y_{net} \) is the summation of weight inputs, \( P_i \) is the input node, \( W_i \) is weight coefficient of each input nodes, \( W_0 \) is bias and \( n \) is number of input nodes [25]. The learning process is started from the Eq. (7), where \( W_i \) and \( W_o \) will update continuously using MLT. Then sigmoid function will modify the \( Y_{net} \) using Eq. (8). From the Eq. (9), the learning will stop when the minimum error is achieved. Thus, this process flow is called back propagation.
3.1 Neural Network Development

To design the network, this paper uses eight parameters taken from the FEM simulation data, which are $\alpha_{\text{ceiling}}$, $\alpha_{\text{door}}$, $\alpha_{\text{window}}$, $\alpha_{\text{wall}}$, $\alpha_{\text{seat/table}}$, $x$ axis for receiver point and $y$ axis for receiver point. Each parameter has its own range and neural network will learn it based on that range as in Table 3. From the number of parameters, the network argument can be expressed as follows;

$$\text{Predicted RT}_s = f_{\text{NN}}[\alpha_{\text{floor}}, \alpha_{\text{ceiling}}, \alpha_{\text{walls}}, \alpha_{\text{windows}}, \alpha_{\text{seat}}, \alpha_{\text{door}}, x\text{ axis}, y\text{ axis}]$$

(10)

Table 3. Range of data

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_{\text{floor}}$</th>
<th>$\alpha_{\text{ceiling}}$</th>
<th>$\alpha_{\text{Walls}}$</th>
<th>$\alpha_{\text{windows}}$</th>
<th>$\alpha_{\text{seat}}$</th>
<th>$\alpha_{\text{door}}$</th>
<th>Receiving Point (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>0.08</td>
<td>0.42</td>
<td>0.02</td>
<td>0.13</td>
<td>0.02</td>
<td>0</td>
<td>x-axis 0.74 y-axis 0.79</td>
</tr>
<tr>
<td>max</td>
<td>0.2</td>
<td>0.88</td>
<td>0.03</td>
<td>0.55</td>
<td>0.4</td>
<td>0.18</td>
<td>x-axis 6.08 y-axis 5.29</td>
</tr>
</tbody>
</table>

Before learning process started, a set of data should be normalized between 0.1 and 0.9 in order to simplify the data. The normalization equation is shown as follows (3);

$$x_{\text{new}} = 0.1 + \left[ 0.8 \times \frac{x_{\text{old}} - x_{\min}}{x_{\max} - x_{\min}} \right]$$

(11)

where $x_{\text{old}}$ is the old value, $x_{\text{new}}$ is the new value, $x_{\min}$ and $x_{\max}$ are the minimum and maximum of the original data range. Since the minimum-maximum normalization is a linear transformation, it can preserve all relationships of the data values exactly.
To help in designing network’s architecture, this study uses MATLAB [27] which provides the neural network toolbox. The architecture of the models consists of three types of nodes; which are input, hidden and output node shown in Fig. 4. The number of input and output nodes is 8 and 1, which represents the input parameter (8 parameters) and output parameter (reverberation time) respectively. While, for the hidden nodes, there are about 2 to 20 nodes suggested. Thus, to get the optimum network (architecture), it can be obtained using trial and error method by connecting all those nodes (etc: [i, h, o] for input, hidden and output nodes; [8, 6, 1], [8, 10, 1], [8, 9, 1] or [8, ..., 1]) but only one of combination may offer a good result.

The optimum network can be obtained when learning process is done. Nevertheless during the process, ‘over training’ may occur and this will affect the results. To avoid this from happening, early stopping (4) method is required by separating the data into three subsets; train subset (60% of 721 raw data), validate subset (20% of data) and test subset (20% of data). Train subset is usually used for learning processes, validate subset is validated the process and the test subset is recognized the prediction performance. The optimum model selected by comparing mean square errors (MSE) and correlation coefficients (R^2) of each argument for performance assessments. The optimum network will provide number of hidden nodes, and it combines with the input and output nodes to become a set of neural network architecture or prediction model.

Figure 4. Neural network architecture

4.0 RESULTS AND DISCUSSION

As mentioned earlier, neural network will be learning when it receives a set of data. In this study we used Sabine equation and FEM to create a set of data (simulation data) to be applied in neural network learning. The comparison between them will be made to define which simulation data gives the better results. The advantages of these methods can be found by that comparison.
Fig. 5 shows the example of comparison between Sabine equation and FEM based on the data used in Table 4. Generally, most of the results indicate Sabine equation only simulates one reverberation time for each argument. In contrast, with a same argument, FEM can present simulations data based on the receiver’s point (RP 1, RP 2, RP 3, RP 4, RP 5, RP 6 and RP 7). This is because FEM calculates each of the classroom surface area in detail depending on the number of elements, nodes and sound absorptions coefficient, while Sabine equation only accounts on the average of surfaces’ sound absorption coefficient. The advantages of using FEM are; it can be simulated reverberation time and able to visualize the sound propagation, thus we can look at the sound wave moving in a classroom based on time. Unfortunately, the disadvantage of using FEM is it needs several hours until a day to calculate the FEM arguments. The sound propagations visualization illustrated in Fig. 6a, 6b and 6c to show the transformation of sound propagation in time (second) is $0.12 \times 10^{-1}$ s, $0.19 \times 10^{-1}$ s and $0.22 \times 10^{-1}$ s, respectively.

Figure 5. Example comparison between FEM and Sabine equation.

Figure 6a

Figure 6b

Figure 6c

Figure 6. The transformation of sound wave propagation based on time using FEM.
Table 4. Example data for FEM and Sabine equation

<table>
<thead>
<tr>
<th>No.</th>
<th>α floor</th>
<th>α Ceiling</th>
<th>α walls</th>
<th>α windows</th>
<th>α seat</th>
<th>α door</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.11</td>
<td>0.88</td>
<td>0.02</td>
<td>0.4</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>0.11</td>
<td>0.42</td>
<td>0.03</td>
<td>0.2</td>
<td>0.4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>0.88</td>
<td>0.02</td>
<td>0.4</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
<td>0.88</td>
<td>0.02</td>
<td>0.4</td>
<td>0.02</td>
<td>0.06</td>
</tr>
</tbody>
</table>

4.1 Neural Network using FEM’s simulation data

Neural network is learned using data [60% train data, 20% validate data, 20% test data] in order to implement the early stopping. To learn this data, the learning process will run repeatedly to verify the characteristics of data and at the same time the number of hidden node changes automatically. This process will stop until a network meets the minimum error (approximately desired data) and the number of hidden nodes appears. From that process, there are 4 hidden nodes with testing MSE 3.7751x10^-4 is indicated, and it means that optimum network is [8, 4, 1]. To confirm the optimum network is able to predict accurately, the assessment is made using R². In fig. 7 shows the percentage of accuracy of testing is 99.4%. Other than that, the training data and validation R² also gives good result (99.5% training and 99.5% validation). Overall, from the analysis, it could be confirmed that optimum network has a potential to predict reverberation time because the value of error is small and the R² is close to 1.

The capability of the neural network model is done using the unseen data to investigate the credibility of prediction. This phase is the additional assessment after using the test data. Unseen data is a set of data, which is not used for a learning process and neural network does not recognize the characteristics. About 35 data are used and by using the optimum network (network architecture: [8, 4, 1]) the prediction will be produced. The prediction from neural network than will be compared with the desired value (unseen data reverberation time). Referring to Fig. 8, the results verify that the neural network predicted value is approximately close to the desired value. To prove that, the MSE between these two values is 4.154×10^-4. Therefore, it could be concluded that neural network model is able to give a good agreement in predicting classroom’s reverberation time.
There are two problems that can be resolved; (1). In this case the classical method simulation is not really detailed. Here, FEM is the simulation model that simulates the reverberation time at non-diffuse classroom in every location of the classroom (2) FEM requires more time to simulate and analyze a classroom with variety of sound absorption coefficient combination. Neural network can be suggested as an alternative model that could bring results approximately to FEM in a few minutes.

![Comparison desired value with neural network prediction data.](image)

**5.0 CONCLUSION**

This paper presents the advantages of using FEM compared to classical method. FEM is capable to predict the reverberation time in every space of the classroom while the classical method only provides a single value. The analysis also shows that FEM is capable to visualize the sound propagation. Using a set of data from FEM, the neural network develops a prediction model. The credibility of neural network prediction has been confirmed using unseen data. It indicates that neural network could give a good agreement and is suitable to predict the reverberation time. This paper is still a preliminary work and needs further improvement. The next study may focus on;

- Adding the number of classroom with difference size,
- Implement the variety of sound absorption coefficient,
- Increasing the number of receiver position where the location is selected randomly,
- Using ray tracing to compare with FEM, neural network and measurement to investigate prediction credibility of those models.
6. REFERENCES


