UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND THE ENVIRONMENT

Institute of Sound and Vibration Research

Assessment of a hybrid numerical approach to estimate sound wave propagation in an enclosure and application of auralizations to evaluate acoustical conditions of a classroom to establish the impact of acoustic variables on cognitive processes

by

Luis Alberto Tafur Jiménez

Thesis for the degree of Doctor of Philosophy

July 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT

Institute of Sound and Vibration Research

Thesis for the degree of Doctor of Philosophy

ASSESSMENT OF A HYBRID NUMERICAL APPROACH TO ESTIMATE SOUND WAVE PROPAGATION IN AN ENCLOSURE AND APPLICATION OF AURALIZATIONS TO EVALUATE ACOUSTICAL CONDITIONS OF A CLASSROOM TO ESTABLISH THE IMPACT OF ACOUSTIC VARIABLES ON COGNITIVE PROCESSES

Luis Alberto Tafur Jiménez

In this research, the concept of auralization is explored taking into account a hybrid numerical approach to establish good options for calculating sound wave propagation and the application of virtual sound environments to evaluate acoustical conditions of a classroom, in order to determine the impact of acoustic variables on cognitive processes. The hybrid approach considers the combination of well-established Geometrical Acoustic (GA) techniques and the Finite Element Method (FEM), contemplating for the latter the definition of a real valued impedance boundary condition related to absorption coefficients available in GA databases. The realised virtual sound environments are verified against real environment measurements by means of objective and subjective methods. The former is based on acoustic measurements according to international standards, in order to evaluate the numerical approaches used with established acoustic indicators to assess sound propagation in rooms. The latter comprises a subjective test comparing the virtual auralizations to the reference ones, which are obtained by means of binaural impulse response measurements. The first application of the auralizations contemplates an intelligibility and listening difficulty subjective test, considering different acoustic conditions of reverberation time and background noise levels. The second application studies the impact of acoustic variables on the cognitive processes of attention, memory and executive function, by means of psychological tests.

Contents

List of tables

List of figures

DECLARATION OF AUTHORSHIP

I, Luis Alberto Tafur Jiménez

declare that the thesis entitled

Assessment of a hybrid numerical approach to estimate sound wave propagation in an enclosure and application of auralizations to evaluate acoustical conditions of a classroom to establish the impact of acoustic variables on cognitive processes

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Some results previous to the ones presented in this thesis were presented in the conference Reproduced Sound 2012.

Signed:

Date: February 2016

Acknowledgements

I would like to thank Dr. Keith Holland and Dr. Takashi Takeuchi, whose professional expertise and supervising have been immeasurably valuable. Special thanks to Takashi for letting take with me the OPSODIS Sound bar when I had to go back to my home country.

I am also grateful to my colleagues and students in the San Buenaventura University, without your help the experimental work would not have been possible. Many thanks to Luz Magnolia, Daniel (rest in peace), Nataly, Deimer, Esteban, Jonathan y Santiago.

Many thanks to my classmates in the ISVR, Greg, Paul, Roland, Dani, Hessam and Sandra.

A huge thank you to my family, Cami and Paola for their support despite all the time I have dedicated to this thesis.

Abbreviations

BEM	Boundary Element Method
BFR	Binaural Frequency Responses
BIR	Binaural Impulse Response
CTC	Cross Talk Cancellation
CUDA	Compute Unified Device Architecture
DOFs	Degrees of Freedoms
EDC	Energy Decay Curve
EDT	Early Decay Time
FBEM	Fast Multiple Boundary Element Method
FDTD	Finite Difference Time Domain
FMM	Fast Multipole Method
FE	Finite Elements
FEM	Finite Elements Method
FFT	Discrete Fourier Transform
GA	Geometrical Acoustics
HRIR	Head Related Room Impulse Responses
HRTF	Human Related Transfer Function
ISM	Image Source Model
OPSODIS	Optimal Source Distribution
PDEs	Partial Differential Equations
RFR	Room Frequency Responses
RIR	Room Impulse Response
RT	Ray Tracing
RTC	Randomized Tail-corrected Cone-tracing
SNR	Signal-to-Noise Ratio
VR	Virtual Reality
WDF	Wave Digital Filter
WFS	Wave Field Synthesis

1. Introduction

1.1 Auralization system definition and hybrid models

A virtual sound environment or Auralization is the process of audibly rendering the sound field created by a source in a space, in order to reproduce the binaural hearing experience at a given position. Binaural hearing refers to the human listening capability based on the use of both ears in order to get a perception of the direction from which a sound comes. According to Fastl & Zwicker (2007), binaural hearing is related to the human ability to process and correlate the sounds coming to each ear, which generates psychoacoustical effects given by the differences of both signals in terms of time and level. The concept of auralization was transformed with the fast development of computers in the 1990s, when the idea of using a small work-station to model the sound field in a room with the purpose of auralizing anechoic material was introduced by Lehnert & Blauert (1992) and Kleiner et al (1993). The first considered auralization as a second stage in a binaural room-simulation, related only with the reproduction of the anechoic signals. In contrast, Kleiner defined auralization as a process considering room simulation and aural event generation, where the sound field prediction could be created by means of computer modelling or by using acoustic scale models. Eventually, Kleiner introduced the concept of the auralization system, in which three steps are required in order to create a virtual sound environment: the first one is the calculation of the room impulse responses (RIR), the second is the digital signal processing to convolve the anechoic material with the RIR, and the last is the presentation, where the signals are reproduced for the listener via headphones or by loudspeakers. More recently and following the same line as Kleiner, Vorländer (2008) defined auralization as a technique consisting of three stages: sound generation, sound transmission and sound reproduction. The first stage describes the procedures to characterize the sources to be auralized. Sound transmission involves the methods used to estimate the RIR and the last phase includes the reproduction system and the signal processing required to convolve the output of the first two stages.

As mentioned by Vorländer (2008), each of these stages requires a relevant knowledge in different disciplines and hence, each one has a challenge for the

researchers in order to find the best procedure that contributes to create a more accurate virtual sound environment. For instance, sound generation involves an understanding of electroacoustic and anechoic recording techniques to characterise a specific source. For sound transmission it is important to know the different possibilities offered by the numerical methods in time and frequency domains in order to simulate the sound propagation in a room; and for the sound reproduction, a previous background in signal processing and psychoacoustics is required to recognize the implications of listening to the auralizations with a specific sound reproduction technology. Besides the inherent background attached to each stage, it is important to bear in mind that these phases cannot be observed independently. In fact, the three of them are interacting with each other and the use of a particular methodology in one of the stages will determine which approaches should be used in the other two. In other words, the way these three variables converge will dictate how convincing the whole experience is. This research is focused on sound transmission and sound generation stages, taking into consideration the numerical methods to estimate a sound propagation and the 3D sound reproduction techniques exposed in the next paragraphs.

In the sound transmission stage, the numerical methods used to model a sound field in an environment can be classified in two main groups. The first group consists of the techniques based on Geometrical Acoustics (GA) theory. The methods that can be found in this group are the Ray Tracing (RT) (Krokstad, et al., 1968), the Image Source Model (ISM) (Allen & Berkley, 1979) and the Hybrids. The most commonly used and accepted methods in architectural acoustic design coming from this group are the Hybrids, which combine RT and ISM in a stand-alone algorithm. The second group is made up of the methods based on wave equation theory, which might be divided in two subgroups. In the first subgroup are the techniques established in the frequency domain, where the most widely known are the Finite Elements Method (FEM) (Heckl, 1992), the Boundary Elements Method (BEM) (Terai & Kawai, 1991) and the recent Fast Multiple Boundary Element Method (FBEM) (Wu, et al., 2012). In the second subgroup are the techniques based in the time domain, where the most recognized is the Finite Difference in Time Domain (FDTD) (Yee, 1966). Each group of techniques presents advantages and drawbacks that are mainly related to their principle and the size of the model. In GA, there is a simple principle stating that a sound wave

can be replaced by a sound ray. This idea eases the analysis of reflections and the manipulation of the source directivity. However, at the same time this assumption ignores the diffraction phenomena and the phase relations. For this reason, the limitation in the low and mid frequency range is given by the size of the environment to be modelled, where only accurate results can be achieved when the wavelength is small compared to the dimensions of the space. On the other hand, the wave-based methods do not have drawbacks related to the physics, as these are methods created to solve the wave equation; hence, physical characteristics of the sound wave propagation such as diffraction, interference, and diffusion can be modelled. The limitation of wave based methods is given by the cost in terms of computational effort. In the frequency domain techniques, the number of degrees of freedom to solve is related to the size of the model and the highest frequency of interest, restricting the use of these methods to the low frequency range in most of the cases. Although in the time domain one simulation is enough to obtain broadband results, there are other disadvantages such as the grid stability, the geometrical dispersion, and the difficulty to model complex geometries, which limits the use of this technique to the low and mid frequency range.

In the last stage of creating an auralization, the 3D sound reproduction technologies can also be classified into two main systems: headphones and loudspeakers. The headphone systems are well developed and they have been the most practical solution to reproduce binaural auralizations so far. However, as indicated by Vorländer (2008), these systems have some disadvantages that reduce significantly the realism of the auralizations. First, the hearing sensation is affected by the unnatural obstruction of the ear and the wearing discomfort of the transducers. In addition, there is a problem called “lateralization” or “in head localization”, where the sources are located incorrectly and unconsciously inside the head. Finally, although a proper equalization might solve partially the problem of lateralization, each headphone system requires a different one, which makes the implementation of these systems more complex. The other alternative is the use of loudspeaker arrangements, what according to Vorländer should be divided in two subgroups: the first subgroup is based on binaural technology and the second on sound field technology. The idea behind the binaural technology is to use the same head related signals utilized for headphone reproduction, and reproduce them through specific loudspeaker

arrangements providing the appropriate audio signal processing applied, in order to get the binaural sound field in the listener position. Some examples of binaural loudspeaker systems are the Stereo Dipole, the Four-speaker system (Krebber, et al., 2000), and the more recent Optimal Source Distribution (OPSODIS) (Takeuchi & Nelson, 2002). In the second subgroup of loudspeaker arrangements, the principle is to reproduce the incident sound field around a listening point or "sweet-spot" (Vorländer, 2008). The two main points of this technology are the Wave Field Synthesis (WFS) (Berkhout, 1988) and the Ambisonics technique (Gerzon, 1976). In terms of sound reproduction, the sound field approach could represent an advantage if more than one listener is required at the same time. However, the accuracy of the 3D sound reproduction depends on the number of speakers used, what complicates the set up and the signal processing required in this last stage.

As mentioned in the last two paragraphs, the main difficulties to create convincing auralizations were described in the stages of sound transmission and sound reproduction. In the sound transmission phase, it was observed how a pragmatic approach of replacing a sound wave by a sound ray has made the GA the most widely used and convenient techniques to model the sound wave propagation for mid and high frequencies in a room. Similarly, the numerical methods based on the wave equation seem to be the best solution to model the low frequency sound field in an enclosure, also having the potential to estimate all the audible frequency range in the near future, when the computational effort is not a limitation anymore. For the reproduction of an auralization, loudspeaker systems give the impression to offer a more convincing experience for the listener since unnatural obstruction of the ear canal or "in head localization effect", is not present as in headphone systems. In this case, sound field and binaural technologies are able to provide a realistic 3D sound reproduction.

The importance of an auralization lies in the possibility of recreating the acoustics of a non-existing environment. This concept generates a number of potential applications for the auralizations in many fields, as was stated by Kleiner (1993) over 20 years ago. Nevertheless, the main application of the auralizations has always been in the architectural acoustics field, where the assessment of the acoustic performance of a space a priori offers a powerful tool to designers and engineers. Therefore, the application of numerical methods to estimate the sound propagation in a room takes an important role,

given that it provides the possibility of estimating acoustic indicators to assess a non-built enclosure. Taking into consideration the numerical methods explained previously, a hybrid model combining GA and a wave equation based method seems to be the best option to estimate a sound wave propagation considering all the audibly frequency range.

Hybrid models to estimate a sound wave propagation in an enclosure combining GA and a wave equation based method have been developed by different authors (Southern, et al., 2013) (Aretz, et al., 2009) (Mahesh, et al., 2005) (Granier, et al., 1995). A common practice in these exercises is given by the use of filters and a crossover frequency in order to combine the results obtained by each numerical approach. A typical difficulty in the implementation of a numerical method relates to the definition of the acoustic boundary conditions. In this sense, Aretz (2009) indicates that to obtain good agreement between measured and simulated sound fields, a FE model requires a realistic representation of the room geometry, the source and the boundaries.

Regarding the latter, considering that the required input data for GA and FEM numerical simulations differ significantly, Aretz & Vorländer (2010) consider that material parameter databases of the latter can not be extended to FE models. Aretz (2009) proposed a way to specify the impedance in FE simulations for both, extended and local reaction materials. The former understood as the materials in which waves can travel freely along a surface, such as sheets of glass, metal or plywood, and the latter as the absorbent materials in which wave motion parallel to the surface within the medium is strongly attenuated by viscous dissipation. For extended reaction materials, Aretz (2009) defined a real impedance valued, corresponding to the average absorption coefficient obtained from reverberation time measurements. For local reaction materials such as porous absorbers materials, the author used a complex impedance data derived from theoretically based models. Considering Aretz (2009) approach regarding the definition of impedance boundary conditions in FE simulations and bearing in mind the main application of an auralization consisting of recreating the acoustics of a non-existing environment, in this research the hypothesis considers that the acoustic impedance in FE models is to be defined as a real valued according to GA material parameter databases. This hypothesis is researched taking into account that the rooms to be investigated have surfaces mostly with extended reaction characteristics.

Taking the last into account, the novel contribution of this thesis to the field is based on the sound wave propagation study within two rooms having different conditions in terms of size, shape and purpose. In both rooms, numerical simulations are carried out by means of GA and the FEM, the first to estimate the sound propagation in all the frequency range and the latter to estimate the low frequency sound field. Hence, the objective evaluation of the simulated sound fields is focused on the low frequency range, assessing the accuracy of determining the acoustic boundary conditions in FEM simulations according to GA material parameter databases. This evaluation considers the consequences in the sound field estimates by comparing with acoustic measurements of the rooms the results in terms of time domain room transfer functions, frequency domain room transfer functions and room acoustic parameters. The analysis of the measured and simulated sound fields indicates that Aretz (2009) approach gives successful results in small rooms, though, the estimates in the larger room point out that research to characterise the boundary conditions in the FEM is still needed.

The numerical outcomes of the smaller room were taken into consideration to create auralizations by means of GA and by means of a hybrid approach combining FEM and GA. A contribution to the field is given by the creation of the auralizations applying binaural loudspeaker systems in the sound reproduction stage. The auralizations created by the application of the hybrid approach were compared to the ones obtained by GA methods, having as a reference the auralizations created by means of binaural impulse response measurements. A novel design of a subjective test to evaluate auralizations allowed to corroborate the accuracy of Aretz (2009) approach for the smaller room, which had been evidenced in the objective evaluation of the sound propagation estimates. The subjective test considered different acoustic sources, sound messages and source-receiver combinations, with the purpose of evaluating variables such as localization of the source, sense of space and loudness at low and high frequencies.

1.2 Auralizations applied to assess acoustical conditions of classrooms

There is evidence that auralizations have been used to assess real and virtual rooms by means of intelligibility tests since (1981) when Kleiner analysed the Gothenburg Town Theatre. Later, Jorgensen et al (1991) developed an audible simulation system to judge the speech intelligibility of large rooms. A validation on the use of auralizations to assess speech intelligibility subjectively was confirmed by (Peng) in 2005, whose predicted Binuaral Room Impulse Responses's (BIR) results presented a reasonable correlation in comparison with measured BIR results. According to Peng´s work, the use of auralizations to subjectively assess the intelligibility of a space is a methodology offering controlled conditions in order to repeat the experiment as many times as required. Moreover, the auralizations might be applied to evaluate the acoustic performance of a room after the implementation of a hypothetical acoustic design. An aspect that has not been studied in detail consists of the application of auralizations in intelligibility tests taking into account background noise, although there are examples of the use of auralizations allowing the influence of external sound transmitted through the interior of the room. In this sense, Yang & Hodgson (2006) applied auralizations to assess by means of subjective tests several acoustic conditions given by changes in voice levels, absorption surfaces and background noise levels. It is important to take into consideration that the reproduction systems used in the experiments mentioned included headphones or stereo loudspeaker arrays.

The application of auralizations to assess acoustical conditions of classrooms have focused on considering as independent variables the reverberation time and the background noise levels, in order to evaluate intelligibility and listening difficulty (Yang & Hodgson, 2006) (Ljung & Kjellberg, 2009) (Ljung, et al., 2009). There is no evidence in the literature of the application of auralizations to assess these variables in Spanish language. Therefore, this project aims to assess the current acoustic performance of a classroom in terms of intelligibility and listening difficulty, to propose a hypothetical acoustic design in order to control the reverberant noise and applying the subjective intelligibility tests for both conditions with the purpose of measuring the impact of the acoustic measures. In this case, a novel contribution to the field is given by the application of

auralizations to assess intelligibility and listening difficulty in Spanish language applying binaural loudspeaker systems in the sound reproduction stage. Moreover, a novel approach considering binaural sound recordings is implemented in order to include background noise in the auralizations.

There is evidence of the impact of noise and reverberation on cognitive tasks (Dockrell & Shield, 2006) (Ljung, et al., 2009) (Ali, 2013), which corroborates the importance of these acoustical variables in a classroom. The possibility of having the acoustical variables of reverberation time and background noise as independent variables in an experiment provides controlled conditions to assess the impact of these on teaching and learning processes. In this sense, another application to be explored in this research involves the evaluation of the acoustic variables' impact on cognitive processes such as attention, memory and executive function, by the application of psychological tests. The importance of such experiments is given by the application of auralizations to control the independent variables of background noise and reverberation time, taking into account that the use of this technology to do these subjective tests in Spanish language has not been reported in the literature yet. The analysis of the psychological test results illustrate a novel contribution in the field regarding the underestimated impact of reverberation on cognitive processes of attention, memory and executive function.

1.3 Outline of the present thesis

This thesis presents an auralization system with three main inputs: creation, evaluation and application. The creation considers in the transmission stage the use of a hybrid model, which contemplates the combination of GA and FE models. In the reproduction stage, a 3D binaural loudspeaker technology is used to reproduce the virtual sound environments. To evaluate objectively and subjectively the accuracy of these approaches, auralizations were created for a Meeting Room and a Classroom to study the sound wave propagation within them. For the subjective evaluation, BIR measurements are used to create the reference auralizations for the assessment of the virtual sound environments, analysing aspects as source localization, reverberance, warmth and brightness. The objective evaluation consists of the comparison of the acoustical parameters obtained from the measured and simulated RIR, according to ISO standard 3382 procedures (2009). The application approach considers the assessment of the

acoustic conditions of a classroom and the evaluation of the acoustic variables' impact on cognitive processes. The first case consists in determining the current acoustical conditions of a classroom by means of RIR and BIR measurements, with the purpose of designing an acoustic treatment to meet the acoustic standards according to local regulation. Afterwards, an intelligibility subjective test is implemented in order to compare the current conditions against the situation considering the acoustic design, all by means of auralizations in order to determine the impact of the acoustic treatment. The second application considers the creation of auralizations with specific acoustic conditions of reverberation time and background noise levels, in order to assess the influence of these variables on the cognitive processes of attention, memory and executive function. The conceptual approach of the present thesis is synthesized in Figure 1.1, where the main aspects of the auralization system are exposed.

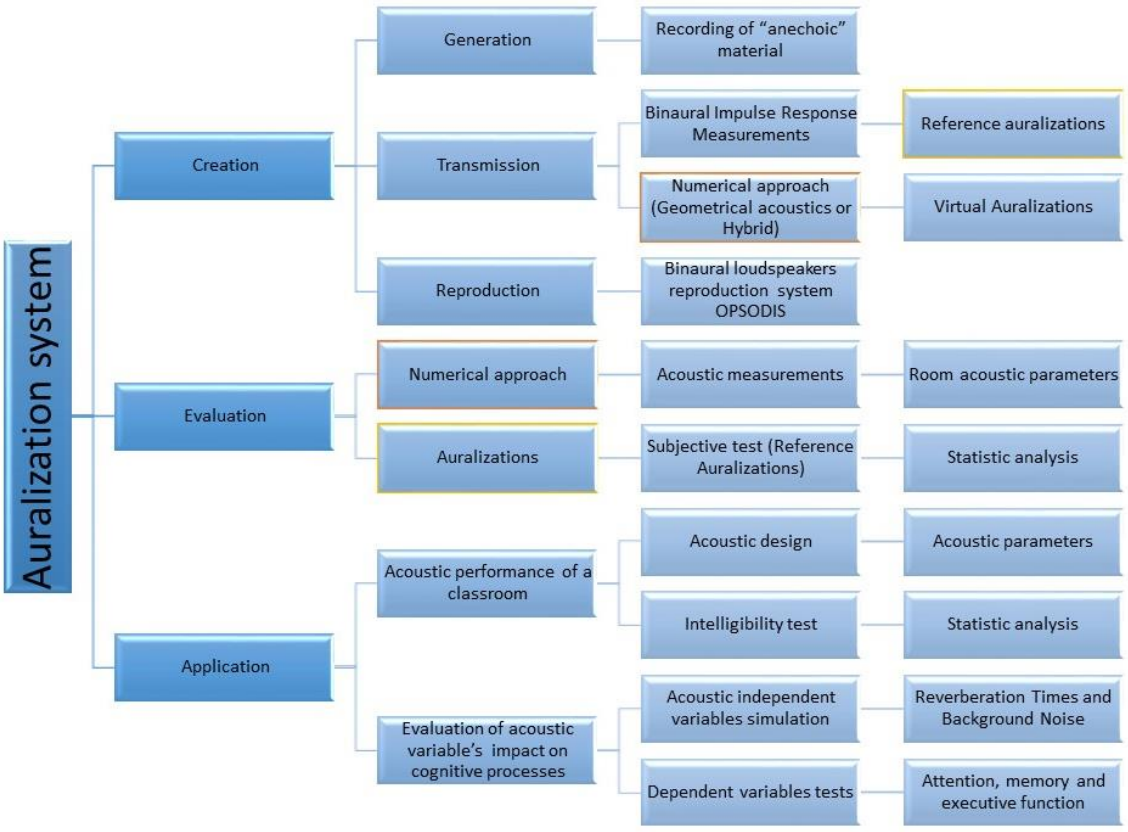


Figure 1.1: Overview of the conceptual approach used in the thesis.

The organization of the thesis is given by the literature review in chapter two, which provides the background of the auralizations systems, a review of hybrid models, the effects of noise on cognitive processes and the applications of virtual sound environments to assess acoustical conditions in classrooms.

Chapter three considers the theoretical foundations of the methods used to create the auralizations. Chapter four describes the acoustic measurement procedures and results obtained for the meeting room and classroom. Next chapter details the numerical modelling applied to obtain the RIR and BIR for both rooms, including an objective assessment of the results and a subjective evaluation of the meeting room auralizations. Chapter six explores the application of the auralization system to evaluate the acoustical conditions of a classroom and the cognitive processes such as attention, memory and executive function. For the first part, the methods to create auralizations with specific acoustical conditions, the analysis of implementing an acoustic design to improve the classroom interior acoustic performance and the results of the subjective assessment of intelligibility and listening difficulty are presented. The second part includes an examination of the acoustic performance standards considered for a learning space, a description of the psychological tests to assess cognitive processes and the results and discussion of the subjective tests implementation. Finally, chapter seven details the main conclusions of this research and outlook on possible future work.

2. Auralization system review

This chapter presents a literature review describing the background of the auralizations systems, the stages to create an auralization, the applications of virtual sound environments to assess acoustical conditions in classrooms and the effects of noise on cognitive processes.

2.1 Auralization systems background

The history of auralizations systems using computer simulations began in the 1990s when Lehnert & Blauert (1992) and Kleiner et al (1993) created virtual sound environments estimating the RIR by means of GA modelling. The only difference between the methods used by the authors was in the concept. For Lehnert, a binaural room-simulation system was what Kleiner considered as an auralization system. In fact, Lehnert described auralization just as a stage in a binaural room-simulation made using GA and reproduced via headphones. On the other hand, Kleiner understood auralization as a process that required the completion of three stages in which different approaches could be used. The first stage, not only mentioned the scale models as an alternative to calculate the RIR, but also described a number of numerical methods to predict the sound field in a room. In this sense, the wave-related limitations of the ISM and RT techniques were clear for Kleiner, where the diffraction and scattering phenomena are ignored. At that time, he suggested frequency domain wave based methods such as FEM and BEM to model the sound propagation in small rooms at low frequencies, taking into account their constraint in terms of computational effort. In the digital signal processing stage, the convolution with anechoic material was implemented according to the reproduction system to be used; when a binaural reproduction was required via headphones or stereo loudspeakers array, RIR were approximated in a post-processing process to BIR in order to have the source and receiver responses for both ears. When a multiple-loudspeaker array was used, a multi-channel convolution was applied, where the number of channels was dependent of the number of speakers. In the last stage, there were three scenarios considering the audio system and the space to be used to reproduce the auralizations. To reproduce binaural signals, headphone systems and stereo loudspeaker arrays with crosstalk cancellation in an anechoic chamber were considered. For the headphone systems, the

problems of in-head localization and back–front confusion were identified. The last scenario considered a large anechoic room and full range loudspeakers as many signals to be convolved.

During the last decades, the fast development of *Virtual Reality* systems has created the need for more realistic auralizations, changing the perception of how a virtual sound environment should be generated. This idea and the advanced techniques developed for each of the processes involved in the creation of an auralization were Vorländer´s motivation (2008) to redefine the stages creating a virtual sound environment in: sound generation, sound transmission and sound reproduction. The main difference with Kleiner et al (1993) approach is the inclusion of the sound generation stage, where there are two main factors for the recording technique in order to obtain the dry audio signal to be auralized, the analogue to digital conversion process and the acoustics of the room. Another difference is the merge of the digital signal processing process and the presentation of convolved material in the sound reproduction stage, which is given by the fast development of processor's speed, memory space, convolution devices and 3D audio systems. As it was mentioned before, all of the stages to create an auralization present significant development covering a wide number of disciplines, hence the most relevant advances of each phase will be reviewed in the next section taking into account the dependent relationship between them and the complexity of implementation on each one.

2.2 Stages to create an auralization

In order to create a more realistic auralization, it is important to understand the techniques used in each stage and the implications of implementing a specific method in the other two phases. This section offers a brief overview of the most widely used methods applied to create virtual sound environments.

2.2.1 Sound Generation

There are two main recording techniques for auralization purposes, single and multichannel. The first one is the most recognized and commonly used method for single musical instruments, human voice or loudspeakers; consisting of a recording made at a specific point taking into account the direction of the main radiation (Vorländer, 2008). In fact, there are commercial room acoustics codes

such as CATT-Acoustics (CATT, 2007) and ODEON (Vigeant, et al., 2011) providing libraries of anechoic recordings obtained by this approach, with several instruments and female and male speech. The other alternative is the use of a multichannel recording technique (Otondo & Rindel, 2004), which uses a set of microphones in both horizontal and vertical planes, in order to get directivity patterns for both planes (see Figure 2.1).

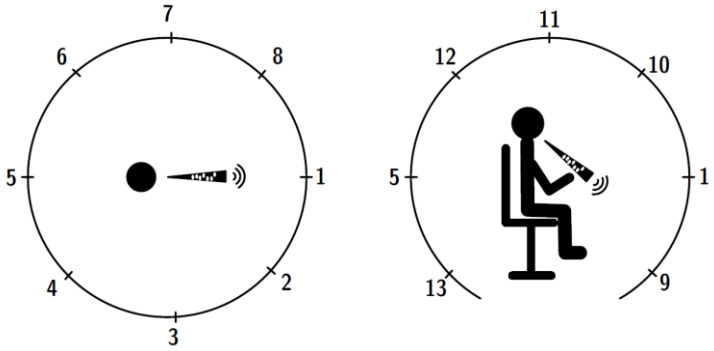


Figure 2.1. Multichannel source recording set up (Otondo & Rindel, 2004).

The application of this method suggests that different approaches must be used in the sound transmission and sound reproduction stages. The first case means that independent room simulations must be run for each angular segment of the source and the results superposed at the receiver position. This procedure is not difficult to implement in a GA simulation, however, in a wave based numerical method it not only involves a more complex process to characterize a sound source, it also means that at least thirteen simulations must be driven for each source-receiver combination in order to get a RIR. In the sound reproduction stage, the use of this technique indicates that a convolution process must be applied for each RIR-anechoic recording combination. Moreover, an additional audio signal procedure, dependent of the 3D sound reproduction system selected, must be applied to obtain the final audio signals to be auralized.

The advantages of the multichannel recording technique are given by an improvement of the quality in the reproduced sound (Otondo & Rindel, 2004). According to the authors of this method, this approach provides a more realistic representation of the source in an auralization, especially if tonal directional instruments are simulated. Similarly, Vigeant et al, (2011) in their subjective study evaluating this technique, confirmed that the realism of an auralization increases in comparison with the single microphone set up. Another aspect to take into account is that in both studies, a pairwise comparison between

techniques was used in the subjective tests, having no real reference to judge realism or timbre.

Although an improvement in the definition of a source in an auralization is perceived when the multichannel recording technique is used, there are some relevant aspects before thinking of implementing this technique in an auralization system. The method is still unsuitable for wideband instruments with transients and more omnidirectional directivities (Otondo & Rindel, 2004). Nevertheless, the multichannel recording technique seems like the most appropriate option when a group of instruments is to be simulated, in order to create an auralization with solo instruments, the single channel recording technique seems to be the most suitable option.

2.2.1.1 Other methods to capture/synthesize instrumental/anechoic material with directionality

According to Pollow et al (2009), in order to enhance the authenticity of auralizations, the impulse response of the enclosure should be obtained with a source that radiates with the same directivity pattern as the original sound source. Since the frequency content of each source may vary (for example, musical instruments), it is required to perform simultaneous measurements with microphone arrangements and a post-processing of the captured signals. The authors developed a spherical arrangement (approximate diameter 4.2 m) with 32 microphones distributed evenly surrounding the source (instrumentalist), using the Ambisonics recording technique (based on an arrangement of microphones placed as close as possible to each other). The measurement was initially performed in a hemi-anechoic chamber, but given the influence on the measurements of the floor reflections, the measurements took place again in an anechoic chamber.

There are several possibilities for the processing of data. One of them is to calculate the spectrum of each capture for each microphone arrangement. The combined spectral-spatial information can be interpolated in a continuous sphere using the decomposition of spherical harmonics. A major limitation is when the sound radiated is not emitted from the centre of the arrangement, since the different arrival times of the microphones produce phase shift. This phenomenon is known as spatial aliasing, and it is caused by a poor spatial resolution of the distribution of microphones (Pollow, et al., 2009). One way to

minimize the problem is to set aside the phase information and use only the magnitudes, dealing only with the spherical wave propagation. Nachbar et al (2010) used spherical arrays by means of 64 microphones for measurements. These arrangements capture the direct sound of an instrument in 64 discrete directions of radiation, thus represent an approximately complete acoustic imaging of the instrument. This implementation uses the hyper interpolation method, which provides an accurate representation of the signals recorded on the microphone positions performing interpolation between positions.

Implementing a multiband approach (multi-pass), the directivity of a frequency dependent source can be examined according to Sheaffer et al (2011). Using first-order differential sources of low complexity, radiation patterns were generated for various frequency band simulations, which approximately correspond to the directivity characteristics of a speaker. In the study, the source is located in the centre of the domain and 72 receivers were located at a radial distance around the source. Six simulations were performed with corresponding data to the octave bands of 125-4000 Hz. The impulse responses obtained were filtered using a Butterworth filter of order 6, and the data were combined to form a set of frequency dependent room impulse responses, each corresponding to a different radial position around the source. This set of room impulse responses were further filtered in third-octave band frequencies in order to examine how the intermediate frequencies were interpolated in the filtering process (Sheaffer, et al., 2011). Directivity patterns were automatically interpolated based on their close values. According to the authors, the method cannot be used in real time applications, since the processing cannot be performed simultaneously. Furthermore, to avoid spatial aliasing, the functions of the source are limited by bands resulting in excitation signals of different shapes which cannot be conventionally added (Sheaffer, et al., 2011).

2.2.2 Sound transmission

This section offers a brief overview of the most widely used numerical methods applied to create virtual sound environments. The numerical techniques to be described are classified in two main groups: the methods based on Geometrical Acoustics (GA) principles and the ones based on wave equation theory. Special attention is focused on the advantages and disadvantages of the techniques that

have been used to simulate a sound field in an enclosed space and the ones that have potential to be part of an auralization system.

2.2.2.1 GA numerical methods

This group includes the techniques based in the GA theory, where a sound wave is replaced by the concept of a sound ray. The methods to be reviewed are the Ray Tracing (RT), the Image Source Model (ISM) and the hybrid methods.

2.2.2.1.1 Ray Tracing (RT)

The RT has been the most accepted and developed numerical method in architectural acoustics since (1968), when Krokstad et al calculated for the first time an acoustical room response by means of this technique. In this study they were able to estimate the reverberation time, understood as time needed for the sound to disappear after the sound source has been turned off, to analyse the detailed early reflection history and to include directional information. In (1977), Wayman & Vanyo predicted the reverberation decay of a classroom and compared the numerical results with the analytical and measured decays, showing a reasonable similitude. At the end of this decade, Krokstad & Strøm (1979) modelled a more complex room called the "Hjertnes multi-purpose hall". In this work, they were able to carry out a very detailed analysis of reflections, to assign different absorption factors to each room surface in order to create echograms in several receiver positions and to investigate the sound energy distribution in the room. An echogram is defined as the time history representation of direct sound and reflections for a given source, receiver position and frequency indicating sound levels and saving information regarding the angle of arrival for each discrete signal. The authors also identified some limitations inherent to the technique; they identified the limited frequency range in GA where the results are valid, as they stated: "Diffraction effects at low frequencies and diffusion effects at high frequencies cannot, therefore, be investigated" (Krokstad & Strøm, 1979).

Over the next decades, the development of the RT technique continued with its main authors using it, not only to estimate acoustic parameters of built rooms, but also to implement this numerical method as an acoustic design tool for non-constructed spaces such as concert halls (Strøm, et al., 1985), (Strøm, et al., 1986). On the other hand, more limitations of the technique were enunciated

during this decade. According to Kulowski (1982), the sound intensity calculation was not accurate due to the stochastic principle of RT. In addition, the interference effect was lost due to the lack of phase information and hence standing waves were not able to be calculated (Krokstad, et al., 1983). In the 1990s, Hodgson (1991) introduced the scattering reflection coefficient in a RT algorithm in order to simulate the diffuse reflection effect and to improve the geometric sound field estimation. Following that, the errors obtained in the estimation of some acoustic parameters were related to the number of rays hitting the receptor, which has a direct correlation to the volume size of the receiver (Giner, et al., 1999). This idea was confirmed by the same authors in 2001 (Giner, et al.), when a relationship between the number of rays and the estimation of the principal acoustic parameters was identified. Finally, in (2003) a new method to improve the sound ray detection was proposed by Xiangyang et al. In this work, the authors concluded that to obtain more accurate results in a RT algorithm, the receiver size should be defined considering the dimensions of the room, the initial number of rays and the distance from source to receiver.

Savioja & Svensson (2015) state that in the ray tracing methods the fundamental principle is to emit rays from a sound source, reflect and record valid paths. The rays can be emitted from the source according to a predefined distribution or in random directions through Monte Carlo sampling. However, the random directions with very few rays introduce fluctuation in the results and therefore, the use of an appropriate distribution should be preferred. If a source has a directivity pattern, the distribution of rays can be weighted according to it; it means that for an omnidirectional pattern, the directions of the lines are evenly distributed around the source point. There are two ways to obtain uniformly distributed directions of the vector around the source (Kulowski, 1985). The first called deterministic uses the algebraic formula given by the regular network of points on a sphere surrounding the source. These points are the vertices of the vectors. The second method is called statistical, consisting of vertices randomly distributed on the surface of the sphere. In order to calculate the point coordinates, two values are chosen randomly along with the elevation angles and azimuth. If the number distribution is uniform, it is assumed that the distribution of the direction of the vectors is also uniform. In the deterministic method the distribution of points on the sphere is not uniform until all the points of the regular network have been considered. Therefore, if the number of rays

is too small to obtain reliable results, a new network with a larger number of points must be created and calculations must be repeated from the beginning.

According to Savioja & Svensson the next step is to trace each ray propagation so that each time a ray hits a wall, a reflection occurs and the ray continues in a new direction. Simultaneously, intersections of rays and detectors are recorded in order to trace the paths of reflections arriving at the receiver (Savioja & Svensson, 2015). The reflection algorithm consists of three parts: the determination of the reflective surface, calculating the direction of the reflected ray and the energy of the reflected ray (Kulowski, 1985). To determine the reflective plane when the shape of the enclosure is convex (no acoustic shadows inside are formed), the planes which are not pierced by the ray are removed first. When the angle between the normal vector to a plane directed into the enclosure and the ray vector is obtuse, the distance between the sound source and the pierce point is calculated. The shortest distance is found by checking all the planes that approximate to the shape of the enclosure. This distance corresponds to the plane that reflects the ray. When the enclosure has a concave shape the intersections are given by the number of planes facing inwards the enclosure. It is important to remove the intersections inside to determine the shortest distance and find the reflecting plane. With the minimum distance found, the reflection point coordinates are calculated as well as the coordinates of the source image with the projection distance from the source to the reflecting plane. The direction of the reflected ray is obtained by substituting the coordinates of the ray direction vector and the source position, with the coordinates of the vector given by the reflected ray and the reflection point respectively. At each reflection the angle of incidence of the ray is known, allowing the decrease of the energy of the reflected ray to be considered proportional to the physical absorption coefficient. If there are numerous incident rays and their directions are uniformly distributed in the hemisphere in front of each limiting plane of the enclosure, a sound field can be considered diffuse. This allows the use of diffuse field absorption coefficients. In addition, the sound energy attenuated by air absorption is estimated taking into consideration the frequency and the distance travelled by the ray.

According to Vorländer, some systematic uncertainties that cause deviations in the simulations are caused by deficiencies in their algorithms and the modelling approach (Vorländer, 2013). A suitable polygon for the representation of the

enclosure is essential for room acoustic simulations using geometrical acoustics. The surface elements, usually polygons, should be large compared with the wavelength in order to cover the entire audible frequency range. This is not possible in practice. The results may be incorrect due to a very low level of detail; however, if the detail level is very high, it will result in unnecessarily long calculation time.

The stability of a solution by ray tracing, according to Savioja & Svensson, increases as the number of rays increases so that the results of the method converge in a scenario in which the specular reflections are considered (Savioja & Svensson, 2015). The number of rays should be discussed separately for stochastic ray tracing and for deterministic (Vorländer, 2013). To find image sources using ray tracing, an enough number of rays must be used so that no reflection is discarded. This means that the path leads to an approximation of the exact result, where some reflection paths can be ignored because a finite number of rays must be used (Savioja & Svensson, 2015). In practice it is difficult to know the number of rays needed to obtain reliable results. These will depend on the geometry and materials of the enclosure, so it is not possible to know the amount of required rays before performing a simulation. Vorländer (2013) showed that the number of required rays is proportional to the volume and area of the absorption space. In any case, to vary the number of rays provides a good basis for assessing the validity of an implementation (Savioja & Svensson, 2015). If done correctly the results should converge to a single solution when the number of rays is increased. Similarly the results should not be affected if in the enclosure geometry a surface is divided into convex polygons such as triangles.

2.2.2.1.2 Image Source Model (ISM)

There is evidence of ISM implementation to estimate RIRs of rectangular rooms back in (1979), when Allen & Berkley developed their own code to apply the method in different applications. The authors applied the RIRs of several source-receiver combinations to simulate reverberation on speech, to estimate the reverberation time of the rooms and to calculate the critical distances. The last defined as the distance from the source at which the sound level given by the direct field and the reverberant sound field are similar. In the first application, a number of sentences recorded in anechoic conditions were convolved with the RIRs in order to evaluate the reverberation effects in what, it might be considered

was the first auralization created by using ISM. According to the researchers, a close alignment with theoretical results was found in the reverberation time and critical distance calculations since the ISM method provides a good understanding of how the direct sound and the early reflections, reaches a specific receiver in time. Likewise, the drawbacks of the technique were identified by the authors, who assumed angle independent impedance and frequency independent absorption coefficients in the model due to the difficulty of implementation and the additional computation effort.

One of the main limitations of this technique is the difficulty of modelling image sources in irregular geometries, where there are invisible places in the room for the image sources. This situation is caused by the extension of the wall and the limited solid angle, which can be seen in Figure 2.2, where the receiver R2 is not visible for the source S1 in a second order reflection. This means that for a particular receiver position, the number of image sources that must be validated grows exponentially according to the order of the reflection. In the 1980s, the method was improved in order to model more complex geometries as arbitrary polyhedral rooms (Borish, 1984) (Lee & Lee, 1988). For Borish (1984), this improvement meant the possibility of applying the technique not only to estimate fundamental acoustical parameters, but also to evaluate the shapes in concert hall designs and to create audible simulations. However, in both cases the computation time continued to be an issue. For this reason, Lee (1988) developed a faster algorithm improving the efficiency of the ISM. Nevertheless, the difficulties of including the directional characteristics of sources, sound scattering and diffraction effects of reflectors remained.

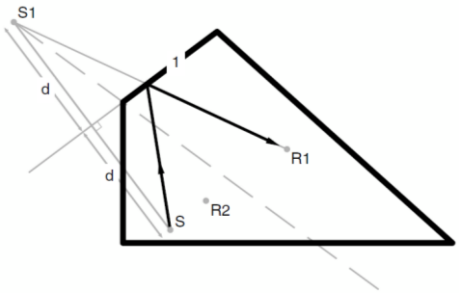


Figure 2.2: ISM construction showing a real source S, a virtual source S1 and receivers R1 and R2 (D'Antonio, 2001).

In an effort to overcome one of the disadvantages of the ISM, Mechel (2000) proposed to model the convex corners presented in a room as monopole sources. As is stated by Mechel (2002), the corner source model is required to

simulate the diffraction phenomena, when the room has important convex corners. However, the author recognised that its analysis and implementation are more complicated than the normal ISM. In the following years, no significant developments were seen in this method and its use was limited to estimate RIRs and Energy Decay Curves (EDC) (Lehmann, et al., 2007) (Lehmann & Johansson, 2008), having the same drawbacks mentioned above. In another attempt to improve the computational process of the ISM, Hachabiboglu (2008) designed an algorithm to reduce the number of early reflections to be used in binaural room auralizations. Although the subjective tests showed that this reduction did not affect the spatial qualities of the auralizations, it is important to bear in mind that simple geometries and fixed frequency-independent absorption coefficients were used in this study.

The basis of the image source method is to reflect a sound source against all surfaces in a model, thus resulting in a set of source images, which are reflected back against all surfaces (Savioja & Svensson, 2015). This process is repeated until a termination condition is satisfied such as a length of response or order of reflection. The resulting image sources can be considered as secondary sources and each represents a reflection such that the distance from the image source to the receiver corresponds to the actual length of the path of reflection within the enclosure.

The number of image sources to a determined order of reflection depends on the number of surfaces. In the case of a rectangular enclosure, some image sources are degenerated, hence, the actual number is smaller. The first order image sources are generated from the reflections on the walls of the enclosure. The second order image sources are generated from successive reflections of the corresponding combinations of reflecting walls (Lee & Lee, 1988). Higher-order image sources can be generated in the same way. The result of a calculation of image source can be seen as a hierarchical tree of image sources including the sound source as the root, and each branch represents an image source (Savioja & Svensson, 2015). In the traditional image source method, the image source tree is built by iterating over all surfaces for a given order of reflection before entering the next higher order reflection. It can also be done by building each branch of the tree to the higher order reflection before proceeding to the next image source. The impulse response can be constructed as the sum of the contributions from all image sources in the tree. Directivities

of the sound sources can be considered reflecting directivity patterns of the sources in addition to the location of the sources.

A technique for calculating the image sources is through the coordinate transformation, using a transformation matrix of the image coordinates to the originals (Lee & Lee, 1988). The method uses the concept of homogenous coordinates, through which a point in the three-dimensional space is represented by (x, y, z, 1) and the transformation of a matrix of 4×4. When an image source is established, the corresponding ray path can be computed by linking the image source to the receiver and determining the intersection points of the connecting line with the corresponding plane mirror. To prevent invalid paths in the model, it is needed to verify the validity of the calculated ray path. There are two cases where invalid paths can be presented. The first occurs when the reflection points are located in the outer region of the reflecting wall. For this, it is needed to test if an intersection point is within the boundary of the reflective wall. In order to carry out this, vectors are formed from the intersection point to each of the vertices on the boundary of the plane mirror. The cross product of successive pairs of these vectors is always the orthogonal vectors to the reflective wall. If each of these normal vectors point in the same direction, then the intersection point is inside the boundary of the reflective wall. Otherwise it is on the outside, and such a ray path is invalid. The second case occurs when there is an obstructed path of rays, that is, the enclosure has obstructive walls. For this, the obstruction verification is performed for each segment of the ray path if an intersection occurs within the limits of an obstructive plane and also within the limits of ray segment, the latter is judged to be obstructed and therefore, the complete ray path is invalid.

To include arbitrary polygonal structures, it is required to perform several additional checks (Savioja & Svensson, 2015). First, an image source should be built only for sound and image sources and sources constructed opposite the reflective side of the polygon, while all the back sides should be discarded as reflective surfaces. Similarly, if the reflective surface is completely behind the previous reflector, it is not needed to create a new image source. All calculations can be developed independently from the listening position and thus, the resulting image source tree is valid for the entire space. The second verification concerns the visibility of an image source and requires information related to the location of the listener. For this, a specular reflection path is formed from

source to the listener using image sources. This path must hit all the reflective surfaces within their limits, and simultaneously, the path cannot intersect any other surface in the model, otherwise, the image source is not visible in the given listener position. The main problems of the technique are the exponential growth in the number of image sources and little manoeuvrability when higher-order reflections are simulated. However, the method efficiently calculates the early reflections and can be used in virtual reality systems that provide information for real time auralizations.

2.2.2.1.3 GA hybrid methods

The limitations of RT and ISM to estimate RIRs meant that new hybrid methods were investigated to the benefit of the advantages of both techniques. One of the first researchers to develop a hybrid algorithm was Vorländer (1989), who proposed the use of RT in order to find the visible image sources. It meant that once the specular reflection paths were identified, the image sources replaced the sound rays. According to the author, this approach provided more accurate impulse responses given by the exact reflections paths in the early part of the RIR, however, the general disadvantages of GA methods were not solved, and consequently diffraction and scattering from rough surfaces were ignored. The same principle was applied by Naylor (1993) in the hybrid room acoustic model implemented with the commercial code ODEON. In this case, a probabilistic approach was used to reduce the reflection density in the last part of the impulse in order to get a more efficient calculation of the reverberant tail. Another example is given by Lewers (1993), who implemented a triangular beam tracing model for the early part of the impulse and the radiosity approach for the late part. The author found that the lack of diffusion was the cause for the inaccurate reverberation time results. For this reason, Lewers applied the radiosity method to model diffuse reflections creating a network to characterize surfaces, receivers and sound paths in a room. Stephenson (1996), who combined the radiosity with quantized pyramidal beam tracing, also implemented the same technique. The hybrid model, in this case, used energy pyramidal beams that split into new ones, at each reflection or scattering. The number of pyramidal beams was quantized according to a reflection factor, the source position and the solid angles presented in the reflection path.

More recently, Siltanen et al (2012) tried to include the temporal behaviour of a diffuse reflection in a GA hybrid model by means of post-processing. Although the beam tracer did not reveal significant differences compared to the traditional beam tracing for the estimation of standard room acoustic parameters, according to the author informal listening tests showed clear differences when the RIR were convolved with anechoic material. However, further research is necessary to probe the advantages of including temporal spreading in GA.

2.2.2.1.3.1 Beam Tracing methods and Radiosity

The algorithms in Beam Tracing (BT) have been used to calculate the paths of early reflections, since it considers the image sources as part of the valid reflection path according to Laine, et al (2009). BT consists of tracing volumetric objects in a geometric model to determine possible specular reflection paths (Savioja & Svensson, 2015). However, there are two approaches to the technique, one related to ray tracing and the other closer to the image source method. The method related to ray tracing has a fundamental principle of expanding the discrete rays, to determine possible paths to volumetric objects that can be detected by specific receptors. The advantage of this method is that the propagation distances are found accurately. BT technique, as an optimizing method for the image source algorithm, aims to improve the performance of the technique, limiting the growth of the image sources (Savioja & Svensson, 2015). This is achieved by discarding the image sources which do not provide any valid reflection path. The first order image sources are generated as in the original algorithm, but a beam (cone or beam) is then formed for each image source at the edges of the surfaces that were used to create the image source. This means that the achieved cone tree has a minimum size, contrary to ray tracing where this tree begins with a large amount of cones.

In order to implement the algorithm when a listener moves, optimizations have been made that do not affect its accuracy (Laine, et al., 2009). The optimization consists of a set of planes which border the beam volume. The planes are defined by the image source and the edges of the polygon tree structure where the root corresponds to the sound source, and each subsequent node is a beam defined by the sound source and a polygon. To determine if a path segment intersects a polygon, it is compared to the side of the plane the starting point is located. When this test fails in one plane, the full path is considered invalid. This

optimization tends to give earlier negative results for invalid paths, which significantly accelerates calculation. The other optimization method groups the potential nodes of the reflection path in cubes, and by a test with a sphere (whose radius is determined by the current position of the listener) determines the invalid paths. On average, the optimization of is 40 times faster, reducing the calculation time by 50% without affecting the quality of the results, obtaining the same results as BT algorithms without optimization. Nevertheless, the technique only does the calculation for specular reflections (Laine, et al., 2009).

Another technique for modelling acoustic enclosures is the Radiosity method, which is based on surfaces and assumes ideally diffuse reflections (Savioja & Svensson, 2015). At first instance, the introduced equation and its descendants were used to determine the reverberation time of a space. Then, the method's foundations to determine the impulse response of an enclosure were introduced. The technique uses the shape factor between two surfaces, which is defined by their relative angles and the distance of mutual separation. It also introduces the concept of exchange factor, which is a cumulative factor for high order reflection paths, so the technique can address specular reflections of diffuse reflection. In the method, the sound energy radiated from a surface is independent of the angle, resulting in the reduction of the memory requirements. This makes the technique to be very suitable for real time applications.

The Radiosity simulation method can be extended to one which can handle diffuse reflection and not diffuse (Siltanen, et al., 2007). The principle of the improved Radiosity method is given by the rendering equation, which allows handling arbitrary materials represented by the Bidirectional Reflectance Distribution Function (BRDF). The algorithm has three phases, an initial shooting, in which the energy of each sound source is distributed to all visible patches (discretization geometry elements). Then the energy is propagated iteratively until the solution converges. A final stage consists of accumulating energy responses from all visible patches for each listener. Although the memory consumption of the algorithm is high, the simulation results indicate that the method is reliable for predicting the acoustics of enclosures (Siltanen, et al., 2007). Additionally it offers flexibility in modelling arbitrary reflections and the management of complex sound sources. This allows that the method can be extended to handling edge diffraction and transmission through materials.

2.2.2.2 Wave equation numerical methods

In this section, the numerical methods to estimate the sound field in a bounded environment are based on the wave equation theory, whose principle is to find a solution of the Helmholtz partial differential equation. This group of techniques can be divided in two domains: frequency and time.

2.2.2.2.1 Frequency domain numerical methods

The Finite Element Method (FEM) is a tool for the numerical solution of partial equations with boundary conditions. Some typical applications in acoustics go from the prediction of modal characteristics via structural, and coupled sound fields in enclosures (Aretz, 2009). This method gives an approximate solution for the distribution of sound pressure in a given spatial domain with boundary conditions. Such a solution is derived from discretization of the spatial domain in finite elements and limiting the possible solution space of the pressure function of each element. The function of pressure for each element is usually restricted to a polynomial approximation with a finite number of unknown variables. The finite element solution converges to the analytical solution with the element size decreased or increasing the order of the polynomial. The precision in acoustic simulations of enclosures by FEM is that the differential equations must adequately represent the physics of the problem. The finite element mesh should represent all relevant details in the geometry domain simulation, i.e., modelling all objects that are not small compared with the smallest wavelength considered. Similarly, the discretization must be sufficiently fine or the polynomial order high enough, taking as a rule a minimum of three elements per wavelength. Finally, it is crucial to have realistic boundary conditions on the surfaces of the simulation, which could be acoustic impedance, coupling conditions of structural fluid or source terms such as surface velocity or a given sound pressure. Finally, the final equation that describes the sound field within a body with arbitrary shape in terms of pressure is presented with the matrices of stiffness, mass and damping fluid and the normal velocity on the surface. This equation can be solved by two methods, direct and modal superposition (Kopuz & Lalor, 1995).

There is evidence of the use of FEM and BEM in acoustic enclosure predictions back in (1995), when Kopuz & Lalor analysed the sound field of a rectangular closed cavity, excited by the vibration of one of its walls. The rest of the

partitions in this study were treated as hard boundaries and the frequency range predicted was 40Hz-200Hz. The simulations were implemented with the commercial code SYSNOISE and the results obtained were almost identical for both numerical methods. Another example was given by the estimation of the resonance frequencies and mode shapes in a mining vehicle cabin done by Stanef et al (2004). In this case, the maximum frequency estimated in the software ANSYS was up to 300Hz. The results were verified against physical measurements and the FE model was used to predict the effect of active noise control. More recently, Aretz (2009) applied FEM to simulate the low frequency sound field in a recording studio. In order to evaluate the results obtained against the FE model and the real measurements in the room, free field acoustic measurements were realised. The authors found that the accuracy of room acoustics FE simulations relies on three main aspects: a realistic modelling of the geometry, the mesh discretization and realistic boundary conditions. For the last point, Aretz (2010) used a scale model to confirm which approach should be applied to estimate the acoustic impedance on the surfaces. The conclusion was that field incidence impedance captured the characteristics of both "locally" and "laterally" reacting porous absorbers in room acoustic FE simulations. In the FE recording studio model, the real part of the impedance corresponding to the random incidence absorption coefficient was used, assuming that all the boundaries were locally reacting. Very good agreement between simulation and measurement was achieved in this study up to the frequency of 400Hz, which verifies the potential of this technique to be part of an auralization system.

In the BEM only the discretization of the boundaries is required, which means that a BE model will always have a fewer number of nodes; however, the computational time necessary to solve the equations are shorter in the FEM. This is because the equations of the BEM are less structured than those of the FEM and, therefore cost-effectiveness is not as reduced as might be expected. This situation was first documented by Harari & Hughes (1992) in a study comparing the efficiency of both methods for time-harmonic acoustics problems. Likewise, Kopuz (1995) stated that total CPU time in the FEM was shorter than with the BEM for a single frequency. Moreover, according to the authors, complex boundary conditions were more easily treated with FEM. More recently, Fahline (2009) studied the difficulties of modelling boundary conditions in the BEM to solve interior acoustic problems. The author stated that in order to obtain more

accurate results, formulation of the admittance matrix using the reactive component is the best option, as long as the surface element discretization satisfies the standard six-element-per-wavelength rule.

Despite the disadvantages of the BEM, its application in enclosures continued in (1996), when Bai & Chang used this numerical method to study optimization techniques for active noise control. In this work, the author was able to model the sound field of a rectangular room and a vehicle cabin using a different number of sources and a number of acoustic impedance approaches. In (2006), Nagy et al implemented a modified numerical calculation and used the BEM as a reference to compare with the results of his proposed technique. The author modelled the interior noise caused by the vibration of the walls in a rectangular room and the results were compared with physical measurements, showing good agreement between the two numerical techniques.

To overcome the characteristic computational drawbacks of the BEM, in the last decades advanced algorithms have been developed with the intention of making this method more efficient to solve the Helmholtz equation in 3D domains. Two of those algorithms are the Regular Grid Method (RGM) and the Fast Multipole Method (FMM). These algorithms were applied in interior acoustics by Marburg & Schneider (2003), who used both approaches to implement a Fast Multipole Method with BEM (FBEM) with the purpose of predicting the sound field in a virtual concert hall of 1269m³. In this model, the source was assumed to be a monopole and the finest grid used had a maximum element size of 10 cm, which allowed the authors to analyse a highest frequency of 680Hz in a system with more than 100,000 degrees of freedom. According to Marburg & Schneider, the regular BEM and the FBEM generated the same results and the performance of iterative solvers is more efficient than a fast direct solver. More recently, the FBEM based on FMM along with the Adaptive Cross Approximation (ACA) has been applied in enclosures for passive (Mallardo, et al., 2012) and active noise control (Brancati & Aliabadi, 2012). In the first study, the complex geometry of an aircraft cabin was modelled with a fine mesh of 32,280 nodes in order to study the frequency range from 31.5Hz to 1000Hz. The impedance values were derived from the random incidence absorption coefficients and the simulation results were verified by acoustic measurements. The model was used to evaluate the reduction of noise in the aircraft cabin achieved with different upholstery materials and/or by changing the shape of the seats. In the second study, the

same model was applied to simulate a local active noise control called “control volume”, in order to reduce the typical jet cabin noise in the frequency range of 50-200Hz.

2.2.2.2.2 Time domain numerical methods

This section details the most widely used time domain numerical technique based on the wave equation theory implemented in acoustics, the Finite Difference in Time Domain (FDTD). The FDTD is an algorithm to solve iteratively the partial differential equations governing sound wave propagation. The principle is to discretise space and time to approximate the partial derivatives in either the second order wave equation or Euler’s conservation equations of mass and continuity. The first implementation of this numerical method in room acoustics was done by Botteldooren (1995), who was able to estimate the sound wave propagation for a virtual auditorium in the frequency bands of 63Hz and 125Hz. In this model, frequency independent impedance was applied on the boundaries, and the impulse responses resulting from independent octave band calculations were used to calculate the reverberation times in both frequency bands. It is important to point out how the author recognised the importance of combining this method with a GA technique such as RT, in order to get full frequency range estimation. Since this work, the academic community has focused its attention on four specific aspects of room acoustic simulations: the correct representation of frequency dependent boundary conditions, the source characterization, the reduction of geometric dispersion error and the increase of simulation efficiency.

The FDTD was developed by Yee to model electromagnetic systems, however due to the acoustic and electromagnetic waves sharing certain properties (Escolano, et al., 2004), an adaptation of this technique was developed to the acoustic discipline using the same mathematical principle (Barry, 2010). According to Bottledooren (1995) the technique is a numerical approach with great potential to apply in the sound propagation, specifically in room acoustics, since the calculations are performed directly in the time domain. The equations are discretized locally resulting in an explicit formulation, and the numerical formulation is itself conservative (Botteldooren, 1995). The main disadvantage of the FDTD method lies on the modelling of the frequency dependent impedance boundary conditions.

There are two main approaches to model frequency dependent boundary conditions in FDTD simulations. The most widely used is the implementation of digital filters. The second one is defined by the variation of density, sound speed and flow resistance. In the first approach, the mesh and the impedance boundary conditions can be implemented separately and joined using an interface based on a Wave Digital Filter (WDF). According to Escolano et al (2008), this method ensures stability in both elements because delay-free loops are avoided. Another advantage is that the digital filters can be designed with highly efficient structures allowing the modification of the coefficients and order without affecting the algorithm during the execution of the code. The importance of the digital filter approach lies on the possibility of using common acoustical data such as absorption coefficients, which can be transformed into impedance or reflection coefficients depending of the digital filter to be implemented (Jeong & Lam, 2010). In the second approach, Suzuki et al (2007) proposed to treat the boundary as two mediums with different characteristic impedances or using an extended version of the Rayleigh model. According to the authors, this approach is very easy to implement, by slightly modifying an existing FDTD code. However, they also recognised that this method requires a high computational effort in terms of processing power and memory resources.

2.2.2.2.2.1 The Rayleigh model

The Rayleigh model represents a porous material as a set of very fine channels. The model assumes that the channels have circular cross sections which are thin enough so that the movements of air inside them are ruled by viscous force (Vigran, 2008). These channels should be similar, parallel and equally spaced and pass through a skeleton material considered as completely rigid. It is assumed that the surface of this system is perpendicular to the axes of the channels. First it considers that the sound propagation in a single channel, assuming it is so narrow that the airflow profile is determined almost entirely by the viscosity of air and not by internal forces. This is always the case at low enough frequencies. Then the same lateral distribution of flow velocity prevails within the channel for constant flows, and similarly the flow resistance of the channel per unit length has almost the same value as the constant flow velocity (Kuttruff, 2000).

An implementation was carried out by the FDTD method, verifying the accuracy of the method in a pipe in one dimension (Suzuki, et al., 2007). The normal incidence absorption coefficient of a material with infinite & finite thickness, supported on rigid material was calculated using FDTD and compared to the exact solution. The porous material was represented by the Rayleigh model, modifying the equations in the area where the material was adding a flow resistance factor to these. Similarly, the normal incidence absorption coefficient of a material of finite thickness separated a distance (leaving an air layer) of a rigid material was compared. The numerical method in all cases produced very good results, matching a large extent the theoretical absorption coefficient. Additionally, the method can be extended to two or three dimensions, allowing the calculation of sound fields with porous materials of various sizes and condition. In order to establish non-uniform properties within the material, properties such as density, sound velocity and flow resistance with a random variation of up to 25% of the assigned value are established. This method allows examining cases in which the material has large distribution properties.

2.2.2.2.2.2 FDTD implementation

In room acoustics, there are two aspects to take into account for the implementation of FDTD sources: the source excitation and the source directivity. The most typically used source excitation function is the Gaussian pulse, which can be implemented using three different methods: hard source, soft source and transparent source. The last two methods introduce a more complex implementation process into the model, hence their application is avoided in room acoustic simulations (Jeong & Lam, 2012). In the hard source method, the source node value is fixed by the source excitation function introducing an expected scattering of the incident sound field and numerically artificial effects at low frequencies throughout the entire domain (Jeong & Lam, 2010). According to Jeong, the big ripples in the simulated impulse response can be removed by applying a proper truncation and high pass filter in a post-processing step, or implementing a Gaussian source limited in time by the arrival of the first reflection coming back to the source node (Jeong & Lam, 2012). To model specific source directivities in time domain numerical methods, Escolano & Lopez (2007) proposed the use of an array of monopoles to create a sound field given by a particular directivity diagram. In this method, the monopole array

located around the virtual source has a combination of amplitudes and phases with the purpose of simulating a desired sound field in a specific set of points.

Furthermore, Murphy et al (2014) conducted directivity measurements of first order through the use of two time-delayed omnidirectional point sources, weighted and summed implementing the method of finite difference time-domain (FDTD). In order to obtain a chosen directional wave front, the method relied on accurate wave interference. A 2000 points FDTD grid was used in order to define two monopole sources with opposite polarity in the centre of the rectilinear arrangement. To obtain a cardioid pattern, an appropriate fractional delay between the excitation sources was used. Different spacing between the two excitation points was used to carry out three tests, to ensure results were not influenced by the effective sampling density grid of the FDTD method. In order to define a directional source by locating a single central point in the grid, grid spacing in even numbers was used. An array of 180 receivers array was defined at a radial distance of 0.4 m from the source. This implementation, according to the authors, generated good approximations for the transparent and smooth sources of the method; however, for hard source excitation the desired directional pattern was not obtained (Murphy, et al., 2014). Additionally, the results are not frequency dependent, nevertheless, the method is limited at low frequencies by the excitation point separation and restricted at high frequencies by the dispersion error.

2.2.2.3 Hybrid models for the numerical modelling of acoustic propagation

In this section, there is a review of the literature regarding hybrid models including the numerical methods analysed in the previous sections.

2.2.2.3.1 FDTD/GA

The synthesis of room of impulse responses can be performed by various acoustic modelling techniques. Southern et al (2013) proposed a hybrid acoustic model that combines the 3-D FDTD method for modelling low frequency, the BT for low order reflections and the transfer method of acoustic radiation (ART) for the stage of late reflections. The hybrid model is based on the FDTD method implemented in the low frequency region. The cone tracing method is used in the high frequency region where the FDTD method requires large amounts of

memory. When the calculation time begins to be significantly higher in the BT method, due to higher order reflections, the response in the time domain is calculated with the ART method.

BT technique is used for the first part of the response to a specific order determined by the memory consumption of the cone tree. According to the authors, it is better to use the BT model as much as possible rather than the ART model because it is more physically correct when modelling specular reflected energy. The region where the BT & ART methods are mixed depends on the geometry and the diffusion coefficients of the surfaces. This mixing region is dependent on the path of reflection, that is, when the specular energy is negligible, thus the diffuse energy is significant, and BT's response does not contribute to the final answer, even if the path has already been calculated. Diffuse energy values are stored in each surface section (in which radiation does not vary) at each instant of time, and are used in the ART method as initial condition. The frequency responses of the BT and ART methods are added to each frequency band with a proper band pass filter, which allows the combination with the FDTD technique. The study compared the reverberation time with six responses measured by using the FDTD method and the BT-ART responses combined at 1 kHz, which, according to the authors, gave good concordance; however, the authors clarify that comparisons with data from actual measurements with listening tests are needed to complete the validity of the proposed hybrid model (Southern, et al., 2013).

2.2.2.3.2 FEM/GA

In order to obtain impulse responses in a wide bandwidth, models of acoustic propagation have been developed that combine simulation methods based on the wave equation FE and GA. These hybrid models are based on a crossover frequency, to which it will simulate the method based on the wave equation, and from which the geometrical acoustics begins (Aretz, et al., 2009). Taking as reference Schroeder frequency, the authors were able to combine the FEM and GA methods, which, according to preliminary listening test, showed no artificial audible artefact in the combined impulse responses. This type of formulation allows calculating impulse responses across the audible range even for small enclosures; however, satisfactory results are subject to the configuration of the simulation parameters of each technique. Aspects such as a good representation

of all the relevant details in geometry, mesh discretization finite enough to represent the sound field and the frequency resolution (discrete steps in frequency to solve the system of equation) are fundamental in the FEM. Similarly, parameters such as directivity of the source, absorption and diffusion coefficients, order of the image-source, number of rays, discrete time pattern for collecting rays and the size of the detection sphere at the receiver significantly influence the simulation results in the geometrical acoustics methods (Aretz, et al., 2009).

In the same way, Mahesh et al (2005) implemented the FEM method for developing a modal analysis and simulation in the frequency domain for indoors spaces. The modal analysis was performed starting from a rectangular enclosure in order to verify with the analytical solution, and then do it in an area with more complex geometry. Initially it was performed with rigid walls and then with absorbing surfaces. The analysis was developed up to 300 Hz due to the discretization of the mesh, and the domain was divided entirely in hexahedral elements. The practical feasibility of the FEM method in the simulation of enclosures and its combination with particles models in order to obtain the broadband response was investigated equally by Mahesh et al (2005). For this, transfer functions in the low frequency region using the FEM method and then using the method of image source in EASE software for two different locations were calculated, one in the shaded area and another directly opposite to the speaker. The image source calculation was done with the first 25 orders. The hybrid method between FEM and image source made possible to carry out an acoustic analysis of bandwidth for a given site. However, the precise knowledge of the acoustic properties of the enclosure is one of the limitations of the FEM method, due to the limited availability of values for the complex acoustic impedance of the walls (Mahesh, et al., 2005). Additionally, for the mid frequency range the large memory requirements along with the accuracy need special attention.

Furthermore, auralizations of car compartments using hybrid methods have been performed by Granieret et al (1995). Given the size of such enclosures, the limitations of the geometrical acoustics methods do not provide a solution at low frequencies. Therefore, binaural room impulse responses are calculated separately by the FE method at low frequency and the ISM at high frequency, and were subsequently superposed from a crossover frequency. The superposition

can be done in two ways, in the time domain, through meshes and by adding the impulse responses; and in the frequency domain using windows and adding the transfer functions. The approach in the frequency domain presented problems in the area near the crossover frequency, probably related to phase discontinuities in both calculations (Granier, et al., 1995). In this study, an analysis of the effect of diffraction by the acoustical prediction software EASE® was also conducted, which was not possible under Schroeder frequency in models of particle (Mahesh, et al., 2005). The authors concluded that the image source method can provide accurate results in a closed cavity provided there is no distribution of energy due to diffraction and a sufficient number of image sources have been considered.

2.2.3 Sound Reproduction

This section reviews the idea of spatial audio quality and the most relevant 3D sound reproduction systems for listening to an auralization. First, the concepts of Basic Audio Quality (BAQ) and Spatial Audio Quality Inventory (SAQI) are introduced. Second, the sound reproduction technologies to be described are classified in two groups: headphones and loudspeakers. Special attention is focused in the advantages and disadvantages of the techniques in terms of reproducing 3D sound in an auralization system.

2.2.3.1 Spatial audio quality

The Basic Audio Quality (BAQ) refers to the fidelity with which a signal is transmitted or reproduced by a system. Rozenn et al (2014) measured the impairments compared to a given reference. According to the authors, perceptual assessment of spatial audio systems can be based on singular listening qualities, such as the accuracy of the localization or the perception of coloration, in general criteria of accuracy such as perception plausibility and authenticity, or in detailed quality listening catalogues. In order to have all the vocabulary needed for the assessment, a Spatial Audio Quality Inventory (SAQI) was developed for the perceptual evaluation of spatial audio technologies used for the synthesis of acoustic environments. It consisted of 48 descriptors of listening qualities which are divided into 8 categories (timbre, pitch, geometry, enclosure, temporal behaviour, dynamics, singularities and general impressions) and should be considered as a description of the perceived differences from the

corresponding descriptor. Some SAQI attributes reflected a view of perception, being closed to temporal or spectral properties of the audio signal, while other attributes reflected a perspective representing high order psychological constructs, manners aspects, affective, of attitude or aesthetic (Lindau, et al., 2014).

The practical application of the SAQI started with a listening comparison of a stimulus test and a given or imagined reference. The subject had to indicate if any difference was perceived, if the answer was negative the test could be stopped, otherwise, the general perceived difference was rated by using a unipolar scale. Then, the subject indicated the temporal behaviour, the dependencies related to users or scene and/or assigned objects of reference to the perceived difference which was done by multiple-choice questions. The options were selected with regard to the interest of the investigation or the stimulus used. The procedure was repeated for all the selected attributes contained in the SAQI, randomly presented while the test stimuli were accessible for continuous comparison (Lindau, et al., 2014).

Rozenn et al (2014) proposed the method of Quality of Experience (QoE) to measure the experience of a subject against a given audio system. The QoE method addressed the following questions: "¿how can a listener describe his/her perception? ¿What are the objective characteristics (especially acoustic) and how do they correspond to perceptual dimensions?" Physical properties of the sound such as frequency content, the location of the sound source and the acoustic environment had an influence on the way it's perceived, these parameters were called *related physical attributes*. In addition, some other attributes concerned with the effect of the psychic or affective state of the listener were assessed by the following questions: "¿are the virtual sound sources plausible? ¿to what extend does the listener feel immersed in the virtual sound scene, what are your (s) emotion (s), etc.?" (Rozenn, et al., 2014).

The related physical attributes describe perceptual attributes that can be linked directly to a physical or mathematical property either of the sound source, the environment, or the sound reproduction system. These attributes can be timbre attributes, which are generally related to spectra-temporal characteristics of sound, localization of the source expressed in terms of azimuth and elevation angles, and distance. The Perceived width and Apparent Source Width (ASW)

corresponds to the spatial extent of an auditory event, expressed in terms of an angular coverage and depth. The attributes related to the enclosure that influences the perception of the quality of the room are the direct sound energy, the total energy of the reverberating sound, the reverberation decay time, spatial and temporal distribution of the early reflections, and the frequency balance of each criterion. The psychic and affective attributes indicate the results of further processing and analysis of the sound scene by the brain. The information of interest is how the psychic state of the listener is changed by sound. These effects are highly dependent on audio content and personal experience of the subject. The study proposes three potential attributes: naturalness, legibility and excitement. Naturalness is a comparison between an unknown reference (the original sound scene) and one known (binaural reproduction). Legibility refers to the ability to discriminate the different sound sources simultaneously, in order to focus on a specific component, and it is affected by the spatial and frequency separation. Finally, the last attribute corresponds to any emotion felt by the listener, whether positive or negative (Rozenn, et al., 2014).

According to the authors (Rozenn, et al., 2014), the evaluation of these perceptual attributes required adequate experimental protocols. From the experimental psychology, two main evaluation categories were distinguished, direct evaluation, which directly asks the subject to evaluate the attributes under study, and the indirect evaluation, where the perceptive evaluation of the subject is inferred. In the case of a direct evaluation, it is attempted to value each perceptive attribute separately, where the subject gives a score within an appropriate scale according to the attribute considered. Similarly, a direct evaluation of the attributes with a (auditory or visual) reference that is provided to the subject can be made, and the subject is required to compare it with the signal being evaluated. The subject's task is to score an attribute of an audio sample compared to the reference given. In case of the indirect evaluation, the subject is asked to perform a task in the context of binaural sound and QoE is inferred from its success. For example, the task of a listener may consist in describing the sound scene, which is to report the number, nature, and location of sound sources. The general intention of such experiments is to derive information about the naturalness and legibility of the sound scene from observations of the listener behaviour. Physiological measures such as heart rate, skin temperature, or the activity of the eye, can be recorded and linked to

the psychic state of the listener. This aims to observe cognitive, emotional or behavioural phenomena through the analysis of physiological responses of the listener. Also, images of magnetic resonance, electroencephalogram or magneto encephalograms are useful tools for observing brain activity. Although it is currently not possible to translate the maps of brain activity on what the subject think or feel, some information about their emotions can be inferred from knowledge of the neuronal activity and connections. The brain images appear as a promising tool to investigate spatial audio perception in general and binaural sound in particular, knowing that technological progress has made the electroencephalograms easier to measure with a simple handset (Rozenn, et al., 2014).

2.2.3.2 Headphones Systems

In virtual sound environments, the practicality of headphone systems relies on the reproduction of binaural signals, which can be achieved by means of measurements or simulations. The concept of binaural reproduction is based on two signals describing the sound pressures received at each of the eardrums in order to reproduce an audible experience. It means that when particular wave propagation reaches a human receiver, it is going to be affected by the spectral cues given by the head, pinnae and torso, also knows as a Human Related Transfer Function (HRTF). These binaural signals can be obtained using a recording dummy head with small microphones located at the ear positions. According to Møller (1992), the full spatial information of the HRTF can be captured placing the microphones at any point in the ear canals or a few millimetres outside, even if it is blocked. When a numerical simulation is intended to generate binaural signals, the computational process depends on the numerical method adopted in the transmission stage of the auralization system. For instance, if GA methods are used, sound rays detected by a receiver must be characterized with a single HRTF according to the particular arrival angles. On the other hand, when a wave equation numerical method is to be applied, a head and two points at ear positions must be modelled.

By measuring the sound pressure at the ear canal of the outer ear, the Head-Related Transfer Function (HRTF) can be obtained. The HRTF depends on the sound arrival direction, hence, a coordinate system must be used in order to obtain an accurate definition of the sound incidence. The azimuthal angle which

describes the direction in the horizontal plane and the elevation angle which represents the sound incidence for the upper and lower hemisphere (Vorländer, 2008). Vorländer (2008) stated that: "The HRTF database should cover all psychoacoustic effects of localization and dynamic changes in localization. Localization subjective tests showed that humans can discriminate between differences of 1° in azimuth in the frontal direction in the horizontal plane. In other directions, the angle intervals can be higher. By interpolation between the HRTF pairs available or by encoding the HRTF in a suitable functional basis, the size of the database can be reduced".

According to Kyriakakis (1998), although in principle it was possible to achieve adequate three dimensional sound field reproductions by using HRTF, it required accurate individual HRTF measurements of each listener. This fundamental requirement that derived from the inherent physiological and cognitive characteristics of the ear-brain interface made these systems impractical for widespread use. At that time, the research was focused on achieving good localization using synthetic HRTF (not individualized) derived through averaging, modelling, or based on the HRTF of subjects that have been determined as good localizators. The barriers that prevented a successful implementation of 3D audio system with HRTF were the large amount of data required to represent it accurately and the errors in the frequency responses and phase that raised from mismatches between non-individualized HRTF and measurement (Kyriakakis, 1998).

Although headphone systems are very convenient for reproducing binaural signals, these systems have drawbacks in terms of hearing sensation, lateralization of sources and complex equalization, which affects the realism in the virtual sound environments. The first aspect to take into account when listening via headphones is the unnatural occlusion of the ear. In this respect, it is important to keep in mind that individuals are not used to wearing transducers on the ears; in fact, wearing headphones reminds us that individuals are not immersed in a virtual sound environment, where an ideal situation would perceive sounds coming from all directions. Another problem is the "in head localization" of sources (see Figure 2.3). According to Toole (1970), the perception in axis between ears when wearing headphones is due to factors such as static pressure on the head, existence of independent paths for each ear, absence of body irradiation and unusual interaction of head and pinnae with the

sound field. Many efforts have been made to correct this problem. In (1996), Hartmann & Wittenberg did an experiment where they found the dependence of externalization on interaural phases of low frequency components and realistic spectral profiles in both ears. Consequently, Kim & Wonjae (2005) found that in order to obtain appropriate externalization and consistent distance perception of virtual sound sources, discrete binaural synthesis and individual equalization were required. The last is another relevant aspect to consider in terms of implementation complexity. As stated by Vorländer (2008), headphone equalization is a complicated procedure that considers radiation impedance into the ear canal, meaning that equalization model parameters depend on each individual anatomy. Although development of average artificial ears helped to deal with this situation, there are still uncertainties given by the mounting of headphones in real ears, where the inter-individual differences and the leakage are important issues to solve in these reproduction systems.

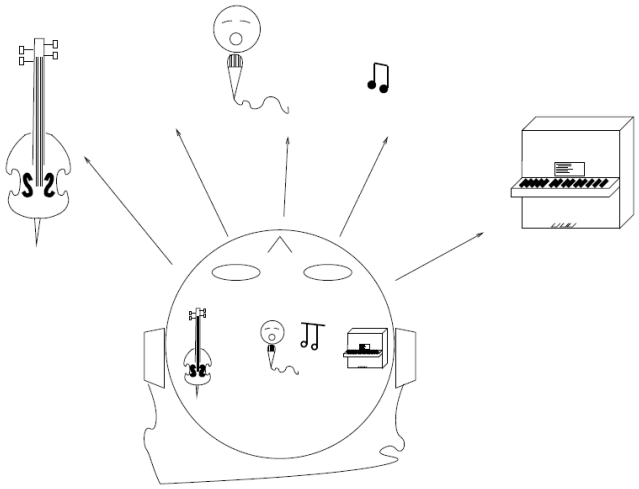


Figure 2.3: In head localization of sources in headphone systems reproduction (Liitola, 2006).

2.2.3.3 Loudspeaker Systems

This section describes 3D loudspeaker reproduction systems, which are classified according to operation principle into binaural and sound field. Special attention is focused on the advantages and disadvantages of the systems to reproduce 3D sound and the implementation in an auralization system.

2.2.3.3.1 Binaural Technology

The concept of loudspeaker systems using binaural technology relies on the reproduction of binaural signals obtained by a digital convolution process between dry audio messages recorded in the generation stage and Binaural Impulse Responses (BIR) estimated in the transmission stage. These binaural systems are loudspeaker arrangements designed to present a virtual acoustic environment in a sweet-spot, or a limited area where the listening experience is not affected by head movements. A common concern of binaural loudspeaker systems is how to deal with the effect of crosstalk. The filters applied in order to obtain Crosstalk Cancellation (CTC) use an inverse matrix, which design depends on the number and configuration of the loudspeakers and the operation principle of the binaural technology. The signal processing required to construct CTC filters have been widely investigated since the 1960s and it is not to be described in this research. An extensive study describing the interaural transfer function generic crosstalk canceler and the least square approximations in the frequency and time domain, can be appreciated in the analysis of designing parameters for CTC filters done by Lacouture (2010). In this work, the mentioned CTC methods were applied to more than two hundred different loudspeaker configurations, including two and four channel arrangements. The binaural systems to be described in this section are the Stereo Dipole, the Four-speaker system and Optimal Source Distribution (OPSODIS).

2.2.3.3.1.1 The Stereo Dipole

The stereo dipole is an arrangement of two closely spaced monopole transducers trying to control the sound field at the listener's ears. From a practical point of view, two loudspeakers, whit acoustic centres no more than 15-30cm apart, are placed in front of the listener with at most 10° angle span (Kirkeby & Nelson, 1998). Such an arrangement is capable of approximating the sound field reproduction generated by a combination of a point monopole source and a point dipole source. In comparison with a typical loudspeaker arrangement with an angle span of 60 degrees as seen by the listener, it has been demonstrated that the stereo dipole provides a larger sweet-spot in terms of sound field control and more robustness in respect to movement and misalignment of the listener's head (Kirkeby & Nelson, 1998) (Nelson, et al., 1997) (Takeuchi & Nelson, 2001). However, the same authors identified a limitation to implement efficient CTC

across all the frequency range. In fact, the need to generate more low-frequency energy to create a virtual image was identified. Likewise, Nelson & Rose (2005) theoretically verified that when the path length between two sources and one of field point is equal to one-half of an acoustic wavelength, CTC turns out to be more problematic. Despite the limitations of this technique, there is evidence of its use to reproduce virtual sound environments. Farina & Ugolotti (1999), who did a comparison with eight loudspeaker Ambisonic system for automotive purposes, give an example of the stereo dipole reproducing auralizations. In this study, the authors were able to implement both reproduction technologies in the same listening room, where the systems were used to blindly evaluate the sound field created by audio systems in cars. The capability of the stereo dipole 3D reproduction was demonstrated in this test, as listeners were not able to identify which system they were listening. However, as Farina & Ugolotti stated, the stereo dipole showed limitations to reproduce low frequencies and a very narrow sweet-spot at high frequencies. As a conclusion, the authors recognized the potential of the method stating that in case of reproducing a very different sound field, such as a concert hall, the stereo dipole might be a better option. Hence, Kwan & Yong (2008) applied a stereo dipole set up to evaluate the subjective preference regarding the sound field with and without diffusers on the sidewalls of a virtual concert hall. In this case, the BIR for both situations were obtained by means of acoustic measurements in a 1:10 scale model. It is important to note, that no other reproduction method was applied in this study; hence, it is not possible to evaluate the absolute performance of the binaural technique.

2.2.3.3.1.2 The Four-speakers system

In order to overcome the limitations of head misalignment, an expansion to the stereo dipole approach considering four speakers was suggested by Lentz & Behler (2004). This technique uses a symmetric arrangement of loudspeakers with dynamic CTC filters based on HRTF data, with the purpose of providing free rotation to the listener. The details of this reproduction method can be appreciated in Lentz (2007). According to the author, this arrangement provides eight combinations where a normal two channels approach can be approximated and a proper CTC can be applied for every listener orientation. An example of the application of this reproduction system is given by Krebber et al (2000), who verified the improvements in the perceived localization of the four-speaker

technique correlated with headphone systems. In this study, the subjective tests were implemented in three rooms with different absorption conditions and a car cabin. Lentz et al (2007), who applied this approach in a Virtual Reality (VR) system, gave another example. In this case, the four-speaker system was used to reproduce real time room acoustic auralizations with the possibility of having a moving listener. Although, this system provides a powerful tool for head movements and real time implementation, it is important to note that in terms of sound reproduction, this technique offers a similar sound field representation to a conventional two speaker approach.

2.2.3.3.1.3 The Optimal Source Distribution (OPSODIS)

In order to overcome the underperformance given by the CTC process when binaural sound signals are reproduced with loudspeakers, Takeuchi & Nelson (2002) proposed an Optimal Source Distribution that enables a lossless approach of inverse filters. The problem with the conventional CTC filters is the destructive interference at both ears reducing the amplitude of the original signals, as can be seen in Figure 2.4. In this figure the amplitude is represented by the length of the arrow and the phase is considered by the inclination. To deal with this interference, amplification is required in the inversion filters, which causes dynamic range loss and large errors around ill-conditioned frequencies. Likewise, the strong radiation of sound in directions other than the one of the receiver generates significant reflections reducing control performance, leading to fatigue of the transducers and loss of directional and spatial information. As stated by Takeuchi (2010), a stereo dipole system has a dynamic range loss of about 43dB for synthesized sound in comparison with a 16dB loss of a loudspeaker arrangement having an angle spanning of 180 degrees. However, the second requires more complex CTC, as can be appreciated in Figure 2.5. The figure indicates in the *x* axis the frequency, in the y axis the sound level loss related to the bit resolution and the curves the dynamic loss range of the synthesized sound. In order to take advantage of the benefits of the last two approaches, Takeuchi & Nelson developed a 3D sound reproduction technology applying a conceptual pair of monopole transducers, where the angle span is changing continuously as a function of frequency, as shown in Figure 2.6.

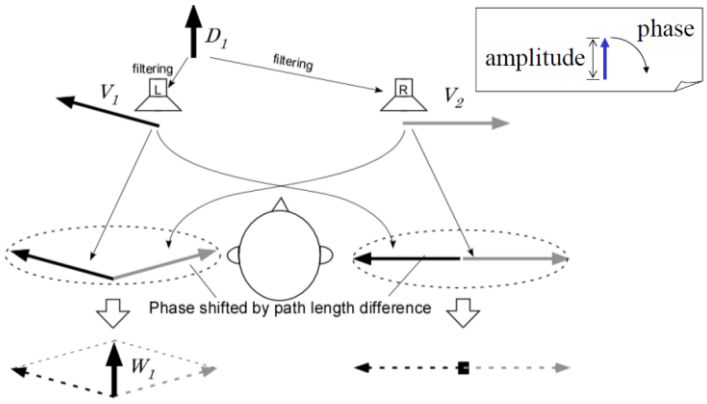


Figure 2.4: Schematic of the shortcomings in a conventional CTC process by implementing a rotating vector to show amplitude and phase changes (Takeuchi & Nelson, 2008).

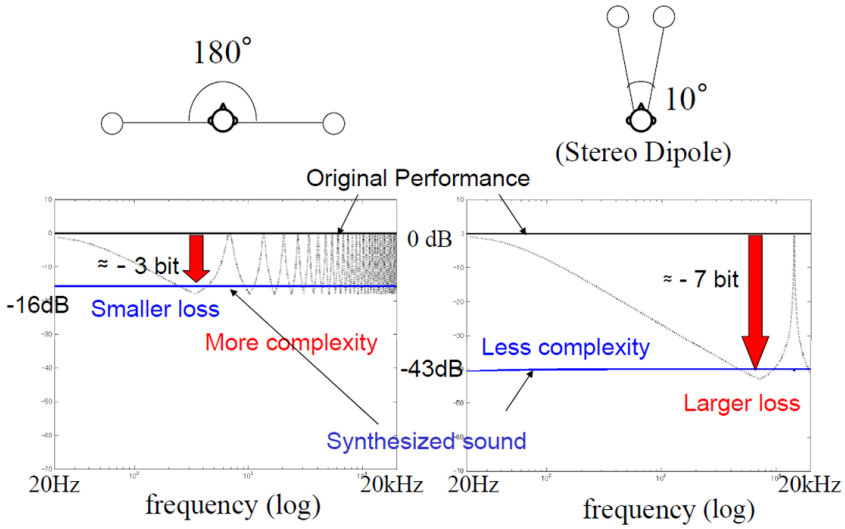


Figure 2.5: Dynamic loss range of the synthesized sound for a loudspeaker arrangement having an angle of 180 degrees and the stereo dipole (Takeuchi, 2010).

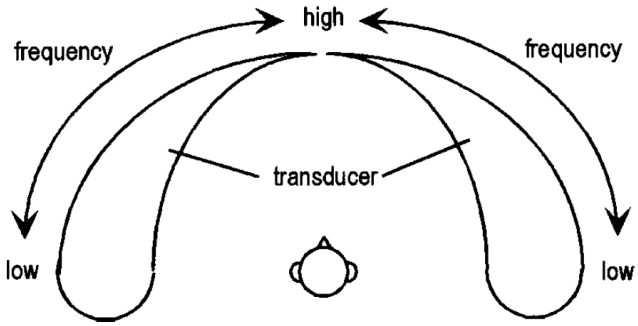


Figure 2.6: Principle of the OPSODIS system (Takeuchi & Nelson, 2002).

Another improvement of the OPSODIS system is given by the ninety degree phase modification on the crosstalk path in the inverse filter matrix, which guarantees that synthesized sound is always reproduced by constructive interference with no loss of dynamic range. Moreover, the radiation pattern becomes constant

over frequency and does not emanate excessive sound to the surrounding environment, which allows multiple listeners and robustness against spurious reflections (Takeuchi & Nelson, 2008). Although the OPSODIS system possesses multiple advantages in comparison with the stereo dipole, the use of this binaural technique to reproduce virtual sound environments has not been documented yet.

2.2.3.3.1.4 General limitations of Cross-Talk-Cancellation (CTC)

The purpose of a CTC network is to cancel the signals arriving from a contralateral path, so that the binaural signals reproduced on a speaker system reach the listener in the same way that it would be reproduced through headphones. Binaural signals must pass through a series of filters to ensure that each of these is equal to the signal that reaches the listener in each ear. To achieve this, it is necessary to determine the transfer functions describing the acoustic paths of the sources to the ears that will be contained in a matrix, known as the plant matrix. Basically the problem is to find the inverse of this matrix, which is generally singular and thus not invertible. Furthermore, when the reproduction system consists of more than two speakers, the equation system becomes over-determined and direct inversion is not possible. Thus, it is necessary to model the system so that it can obtain an approximation to be as close as possible to the required solution (Yesenia & Per , 2011). To obtain an efficient cancellation, each cancellation network element must be able to provide a significant enhancement at low frequencies. This is due to the difference between the HTRF direct path and the cross-talk path, which is very small at low frequencies. The closer the two sources are to the listener, it is easier to implement cross-talk cancellation network. The non-minimum phase behaviour of the electro-acoustic transducers at the extreme end of the frequency range makes it necessary to use a modelling delay in order to equalize the phase response and magnitude response (Kirkeby, et al., 199).

There are several methods to obtain an optimal inverse filter. The first is the Generic Crosstalk Canceller (GCC), which applies the exact definition of the inverse matrix. The filters are obtained by directly inverting the matrix composed of the section of minimum phase of the plant matrix transfer functions. This technique models the Interaural Transfer Functions (ITF) as the relationship between the ipsilateral and contralateral transfer functions'

minimum phase component. The section all-pass transfer functions is approximated to a frequency independent delay, assuming that the all-pass phase section is considered linear at low frequencies. This method is only applicable to two-channel arrangements, because it is based on a direct inverse matrix. The other two techniques are based on least squares approximations. These methods do not attempt to reverse the plant matrix directly but seek the best approach resulting in minimal errors. These techniques are the fast deconvolution method based on Fast Fourier Transform, and the calculation of the optimum filters in the time domain, using matrices containing digital FIR filters (Yesenia & Per , 2011).

In order to properly implement a crosstalk system, three angles of coverage, including 10 °, 60 ° and 120 ° have been objectively and subjectively compared by Bai & Lee (Bai & Lee, 2006). The Friedman test was applied to analyse the data of the subjective experiments, and the processed results indicated that the configuration of 120 ° performed well compared to the standard configuration of 60 °, and was better than 10 °. Arrangements with small separation angles produced a relatively large sweet spot because the displacement of the head only cause minimal changes in the arrival time differences between the two speakers. The arrangements with large separations appear to be more effective because the shadow of the head and the panning effect helps to provide a location to some degree. For large separation configurations the main problem was the stability of the sound image. Problems like bad conditioning, high gain, and low performance at low frequencies can arise from arrangements with very small angles of coverage, while there is a wider and useful frequency if an arrangement with a greater separation angle is used.

2.2.3.3.2 Sound Field Technology

This section details the techniques consisting of an array of loudspeakers reproducing a desired sound field around a sweet-spot. The methods to be described are Wave Field Synthesis (WFS) and Ambisonics.

2.2.3.3.2.1 Wave Field Synthesis (WFS)

WFS is a sound reproduction technique conceived by Berkhout (1988). It is based on wave decomposition analysis of the signals located in receiver positions according to a particular microphone array. These pressure points can be

considered as elementary sources generating new disturbances, where the main propagating wave is regarded to be the sum of all the secondary waves. According to Berkhout et al (1993), this microphone array is a configuration of source-oriented directional microphones, which has its repercussions in the transmission stage of an auralization system. The simulation of a sound field in a room using directional receivers is an easy task in GA, however, such implementation requires a very complex model if a wave equation numerical method is to be applied. In terms of reproduction, in order to reproduce a sound field within an enclosure of about 27m³, with frequency content up to 10 kHz, a two-dimensional array with 500 loudspeakers is required, (Vorländer, 2008). Nevertheless, this is not the only aspect to take into account, WFS is a technique that considers an anechoic reproduction room, consequently, an implementation in a normal environment would reduce significantly the quality of the sound reproduction. For this reason, adaptive processes have to be included in order to compensate the reduction of quality given by the room response (Gauthier & Berry, 2008) (Stefanakis, et al., 2010).

The Wave Field Synthesis (WFS) method is an approach to the reconstruction of the sound field. The algorithms are a type of Fourier transform between the space domains and the wave number, therefore; using an appropriate transformation, complex wave fields can be decomposed into wavelets such as planar, spherical or cylindrical waves. Wave decomposition is accomplished by analysing the signals in microphone arrays. According to the Huygens principle, the microphone positions where the sound pressure is recorded can be interpreted as elemental sources. In a situation of reproduction, the wave field is reconstructed sending waves from these points. If the discrete spatial sampling is high enough, any wave field can be reconstructed. For the reproduction of broadband signals, the conditions regarding size and distance between the speakers must be very strict in order to avoid masking at high frequencies (Vorländer, 2008). The usual formulation of the WFS method is based on two assumptions: the sources and listeners are located within the same horizontal plane, and the target sound field emanates from point sources with omnidirectional directivity characteristics. The first hypothesis allows deriving a feasible implementation based on linear arrays of loudspeakers in the horizontal plane. Using the latter case, the sound field can be extrapolated to any position in the space. The loudspeaker input signals (secondary source) are obtained

from a set of approximations of the Rayleigh integral considering omnidirectional secondary sources (Corteel, 2007).

The WFS method uses a large number of loudspeakers to create a virtual auditory scene on large listening areas. A loudspeaker system enclosing a listener can be treated in the wave equation as an inhomogeneous boundary condition. The solution of the homogeneous wave equation for a bounded region with respect to inhomogeneous boundary conditions is given by the Kirchhoff-Helmholtz integral. For an authentic reproduction of a sound field it is desirable to reproduce the wave field of a virtual source into a limited area as closely as possible. When a distribution of monopole and dipole sources in the listening area limits are carried by the directional gradient and wave field pressure of the virtual source respectively, then the reproduction can be accomplished if the bounded region is considered as the listening area. For a practical implementation, one of the two types of secondary sources that uses the Kirchhoff-Helmholtz integral must be discarded. Normally, the dipole sources are removed, since the monopole sources can be performed reasonably well with loudspeakers with closed cabinets (Spors, et al., 2008).

As a reproduction technique of holophonic sound, the WFS method directs the synthesis of the physical characteristics of the sound field within an extended listening area. This implementation allows the listener to wander around the installation and feel natural variations in the perception of various sound sources, contrary to the techniques based on the 'sweet spot'. The localization varies according to the relative position of the listener and the sound sources regardless of the speakers' position. The sense of presence is increased according to the coherent proprioceptive and auditory signals experienced by the listener that the virtual environment provides. Directivity can be incorporated as a tool to create or increase disparities in the listening area (Corteel, 2007).

The practical WFS implementations would consist of secondary sources (loudspeakers) that are located in discrete spatial locations, in order to obtain a continuous linear distribution. This spatial sampling of the continuous distribution can cause spatial aliasing in the wavefield reproduction. This can result in inaccuracies in the localization and coloration problems. Practical implementations of the secondary source distribution with a non-closed contour will always be of finite size. The WFS method assumes closed contours, infinitely

long linear or planar distributions of infinite size. The truncation of the distribution of the secondary source can lead to errors in the reproduction of wave field and are known as truncation errors. Two effects can be observed, first the wave field area correctly reproduced is limited by the finite aperture, and second, circular wave propagation of the external secondary sources is also presented (Spors, et al., 2008). Near field effects can also be presented for the sources located in the proximity of the loudspeaker array to which the approximation of a far field used for the derivation of the WFS filters is invalid, and the wavefront degradation formed as the loudspeakers are not ideal omnidirectional point sources (Corteel, 2007). Other inaccuracies in the WFS method occur when in the two-dimensional systems, points sources are used as secondary sources. This approach produces amplitude errors in the reproduced wave field. Also, the reproduction of moving virtual sources, which is performed using the stationary spherical wave model and changing its position over the time, present spatial aliasing errors and truncation more prominent than the stationary sources. Reproduction on a plane using only point sources as secondary sources will produce inaccuracies in the listeners that are not located in the plane of the loudspeakers (Spors, et al., 2008).

2.2.3.3.2.2 Ambisonics

The other well-known sound field technique is Ambisonics, which is a recording and 3D reproduction method based on the representation of the excitation of the sound field in terms of orthogonal basis functions, which are known as three-dimension harmonics (Frank, et al., 2015). Ambisonics is based on the representation of the sound field by means of spherical harmonic decomposition. These spherical harmonic components represent front-back (Y), up-down (Z), left-right (X) and non-directional (W) signals, called B-format. If a sound field is to be captured, a sound field microphone consisting of perfectly coincident figure-of-eight microphones pointing in X, Y and Z directions and an omnidirectional microphone W is required. This arrangement can be easily implemented in a room acoustics simulation using GA, taking into account that most commercial codes based on this theory already offer this option. However, the complexity of simulating directional receivers in wave equation numerical methods is a limitation to consider, if Ambisonic techniques is to be part of an auralization system. Another disadvantage of Ambisonic systems lies in a minimum number of speakers in order to reproduce correctly a sound field.

Moreover, the acoustic conditions of the listening room can reduce the performance of reproduction if strong reflections are reaching the receiver. An example regarding the application of Ambisonics to reproduce virtual sound environments was given in the stereo dipole section, where a comparison study implemented by Farina & Ugolotti (1999) was introduced. In this study, an array of eight loudspeakers showed how Ambisonics system has a wider effective frequency range with a larger sweet-spot, in comparison with a stereo dipole. On the other hand, in the same test the Ambisonics presented localization problems, and additionally, the listeners were not able to identify to which system they were listening.

In order to reproduce sound field auralizations, a method that combines three reproduction techniques based on speakers was proposed by Pelzer et al (2014). The sound field can consist of one or more sources and all of the source reflections on the walls of the virtual scene. The hybrid approach can take advantage of the individual strengths of each method of reproduction. Strong localization signals are necessary for the reproduction of the direct sound. The late diffuse sound field can be reproduced using immersive reproduction methods. For this, a hybrid system has been implemented which uses a common loudspeaker system to reproduce a crosstalk cancellation signal and Ambisonics simultaneously. The binaural signal ensures a high detail of the temporal and spectral characteristic of the direct sound and early reflections, while Ambisonics signal is used to produce a large diffuse sound field and envelopment. In order to mix the early and late part of the impulse response with different reproduction techniques, it must be ensured that the levels are precisely adjusted. For this, an equal distribution of virtual sources in a sphere was used. The listening enclosure with the speaker system installed is measured or simulated and the impulse responses are used for decoding the crosstalk cancellation signal, high order Ambisonics and vector base amplitude panning (Pelzer, et al., 2014).

Five reproduction methods were tested. Three pure implementations of crosstalk cancellation, fourth order Ambisonics amplitude panning, two hybrid variants using crosstalk cancellation or amplitude panning for the early reflections, and fourth order Ambisonics for the latter part. The transition moment from the early to the late part of the impulse response was defined by the mixing time. In typical cases, after three orders of reflections the sound field was expected to be mixed and diffused, then it could change from one method with strong

localization to another method with a great envelopment. Although the binaural crosstalk cancellation technique provided a precise and more homogeneous localization across different source positions being more suitable for the early part of the impulse response, the authors recommend for future investigations the designing of immersivity tests including different systems (Pelzer, et al., 2014).

2.2.3.3.2.3 Spatial Impulse Response Rendering Method

The Spatial Impulse Response Rendering method (SIRR) can be used to reproduce room acoustics with any multichannel speaker system. This method was designed to overcome the technical problems of conventional microphones to capture impulse responses using an analysis-synthesis approach that is perceptually motivated (Pulkki & Merimaa, 2005). SIRR processing method consists of the analysis of the direction and diffuse sound within frequency bands, followed by the synthesis that generates multi-channel impulse responses that can be adapted to an arbitrary speakers system. Although the technique is applicable to recordings in general, it is especially suitable for processing enclosures for convolution reverbs. The SIRR method assumes that the reconstruction of a sound field does not need to be identical to the original in order to closely reproduce the spatial impression of an existing place. Instead of rebuilding the sound field, the SIRR method aims to recreate time dependencies and frequency characteristics that are relevant to human perception, such as interaural time difference and interaural level differences (ITD and ILD respectively), monaural localization, interaural coherence (IC) and timbre.

The simplest approach to the problem is to analyse and synthesize the physical properties of the sound field to be transformed into binaural signals (Pulkki & Merimaa, 2005). More specifically, it is assumed that the direction of arrival of the sound will be transformed in ITD, ILD and monaural localization signals. The diffusion of sound would be transformed into signals of interaural coherence. The timbre depends on the monaural spectrum (time-dependent) together with the signals of ITD, ILD and IC. For a good perceptual quality of spatial reproduction, the arrival direction, diffusivity, and sound spectrum reproduced with the temporal and spectral resolution of human hearing is needed. When a room response has a good perceptual quality of the spatial reproduction, the

reproduced convolved sound with this response will also have a good perceptual quality of spatial reproduction. The implementation of the technique can be performed in several ways, one of these is using four audio channels recorded with four matching microphones (B-format), one omnidirectional and three bidirectional (figure 8) directed to three Cartesian orthogonal axes. In the analysis part of the method, the arrival direction and diffusivity can be estimated using sound intensity and energy from this B-format microphone. The sound intensity vector corresponds to the direction and magnitude of the net flow of sound energy. The arrival direction is an estimate of the opposite direction of the sound intensity vector, and the relationship between the sound energy density and the magnitude of the intensity vector is used to compute an estimation of the sound diffusivity. In the synthesis step, each frequency channel and time instant of the omnidirectional signal is reproduced as it was recorded, or diffused with crossfades between previous methods depending on the diffusivity analysed. The diffusivity of sound is created from a hybrid of two methods. The first is not to use any diffusion technique at low frequencies. With diffuse sound, the arrival directions of sound analysed behave stochastically. When sound is applied to such directions, it will spread around and different frequencies are panned in different directions. The second method uses phase randomization at high frequencies, which is done by creating continuous uncorrelated noise for each speaker, and setting the magnitude spectrum of each frequency component in each temporal window equal to the magnitude spectrum of diffuse energy divided among the speakers. Compared with conventional microphone techniques, the method is able to improve the directional quality of the reproduction (Pulkki & Merimaa, 2005).

Formal listening tests were carried out to evaluate the quality of SIRR method in order to determine how close the reproduction can be from the reference (Pulkki & Merimaa, 2005). The evaluation was done by creating virtual reality sound as natural as possible with a high number of speakers in an anechoic chamber, then that virtual reality was reproduced with the SIRR method and other techniques. The listeners reported how much differ the test sound differs from the reference using a single value according to the degradation scale of the ITU. The reference signal of acoustic virtual reality was created with the DIVA software, which models the direct sound and early reflections with the image source method and late reverberation statistically. The listeners were asked to hear three aspects of

the reproduction: sense of space, localization and coloration, and finally give an overall assessment based on these aspects between the test and reference sample. The test results were good for the SIRR technique when it was reproduced in large enclosures. With small enclosures the difference was audible, but in most cases the listener rated the difference as not-annoying. In the reference listening position, the samples were rated as high as the reference, therefore the technique produced almost transparent quality. In all scenarios, the SIRR system was judged as the best reproduction system.

2.3 Auralization systems to assess acoustic conditions in classrooms

This section reviews the literature in the following subjects, the academic impact of acoustic conditions in classrooms, the acoustic parameters to assess a classroom and the application of auralizations to assess acoustically a classroom.

2.3.1 Academic impact of acoustic conditions in classrooms

The worldwide interest in this topic aims to understand typical acoustic problems occurring in these kind of spaces, in order to find solutions to improve the teaching-learning process. According to Kumar (2009), a well-designed classroom takes into account acoustic parameters such as ambient noise, reverberation time and sound insulation in order to facilitate student listening, thereby improving learning experience. According to Sutherland & Lubman (2001) and Kristiansen et al (2013), there is sufficient evidence of the negative impact of background noise and reverberation on scholastic performance and professor's health, to indicate the importance of these two acoustic parameters when assessing acoustic conditions in a classroom.

Deficient acoustic conditions in classrooms can significantly affect student academic performance and teachers' well-being. In terms of scholar's yield, students exposed to noisy environments have more difficulty concentrating on cognitive tasks (Dockrell & Shield, 2006) (Ljung, et al., 2009) (Ali, 2013). According to all the authors, a classroom with poor intelligibility has a negative impact on students at the moment of executing tasks involving comprehension and memory, even interfering with the development of language in children.

Moreover, the use of mechanical ventilation in classrooms with minimum levels of insulation is enough to make the space unsuitable for academic tasks (Mydlarz, et al., 2013). This situation has also generated an occupational health problem for the teaching profession, since professors must raise their voice to overcome background noise in order to be heard by their students. According to Cantor (2013), the signal-to-noise ratio between the sound pressure levels radiated by a professor´s voice and the levels of background noise, must be greater than 5 dB for the brain to have the ability to distinguish both signals. This also has adverse physical and psychosocial effects on teachers´ health contributing to problems in the vocal system, fatigue, lack of motivation and sleepiness (Cardoso, et al., 2012) (Kristiansen, et al., 2013).

In order to enhance appropriate teaching and learning achievement, classrooms should meet established criteria of acoustic design in terms of background noise and reverberation times (Department of Education and Skills UK, 2004) (Acoustical Society of America, 2010), however, these conditions are more difficult to create in spaces already built (Trombetta Zannin, et al., 2009) (Trombetta Zannin & Reich Marcon, 2007). In these cases, the lack of an acoustic design at the moment of structuring the space makes the implementation of noise control measures more difficult. Consequently, in many cases, in order to improve the acoustic conditions of a classroom, an acoustic treatment to control reverberation is the only possible action. Regarding the above, there are studies indicating that annoying noise is usually perceived inside the classroom, which refers mainly to phenomenon of reverberation (Trombetta Zannin, et al., 2009). Moreover, according to Kristiansen (2013), professors exposed to high reverberation times are frequently less approachable to patients and students. The last means that a proper acoustic treatment might be the first step in order to improve teaching-learning activities in a classroom that is already built.

2.3.1.1 Acoustic parameters to assess a classroom

The main acoustic parameter to assess a classroom in terms of teaching-learning academic dynamic is given by speech intelligibility. Three modern measures have been used to determine the influence of interior acoustics and background noise on intelligibility: these are the ratio of useful sound to harmful sound, the percent Articulation Loss of Consonants (ALcons) and the Speech Transmission Index (STI). The first is defined as the relation between direct sound and early

reflections compared to noise and late reflections. Despite these variables being conceptually different, a strong relationship to each other has been found (Bradley, 1998). Moreover, Bradley found relationships between the assessment of intelligibility, using Fairbanks' Rhyme Test and some acoustic parameters derived from RIR in classrooms. These parameters were reverberation time, STI and the ratio of useful to harmful sound; the last two parameters were the most relevant to predict speech intelligibility and they had essentially the same accuracy. On the other hand, with variables as background noise and reverberation time, it was possible to estimate intelligibility with an accuracy slightly lower (Bradley, 1986).

There are other studies referring to the relationship between STI and subjective intelligibility test results, which have confirmed how sound quality is strongly related to background noise level and signal-to-noise ratio (Hodgson, 2002) (International Electrotechnical Commission, 2011). On the other hand, there are not many researchers studying the correlation between STI and subjective tests in the Spanish language. According to Sommerhoff's research (2007), there are two lists of Spanish words to evaluate subjective intelligibility, named after Miñana and Fuchs. Sommerhoff used them to study intelligibility with college students, finding in the second list similar results compared to the standardized English tests; although very different results were found in the first list. Sommerhoff´s contribution lies on the development of a list of logatoms CVC (Consonant - Vowel - Consonant) phonetically balanced in terms of the degree of difficulty, in order to assess intelligibility in Spanish (Rosas & Sommerhoff, 2008). Afterwards, the authors corroborated the relationship between STI and subjective tests by applying a list of words with CVC logatoms using a combination of Latin American Spanish. The test was conducted among university students in two conditions: a classroom with pink background noise and a reverberation chamber. The correlation between STI and subjective test results was different for each environment, obtaining for the same value of STI, dissimilar percentages of intelligibility (Sommerhoff & Rosas, 2011). Table 2.1 illustrates intelligibility classification according to ISO standard 9921 (2003) for CVC tests and the corresponding STI ranges estimated from correlation equations published by Sommerhoff.

Intelligibility measured by words on subjective tests, may not be suitable for evaluating speech transmission under certain conditions. It has been found that

assessment of intelligibility by words varies only between 90% and 100% over a range of signal-to-noise ratios of 4.5 to 14.5 dBA, within an adult population between 22 and 58 years old. Hence, this evidence suggests acceptable conditions for speech communication over a significant range of signal-to-noise ratio. However, in some conditions, excellent intelligibility is only possible with considerable additional effort from the listener. For this reason, a new measure called "listening difficulty" was proposed by Sato et al (2005). it is defined as the percentage of responses indicating some level of difficulty. In the signal-to-noise ratio mentioned above, this quantity differs from 95% to 5%, suggesting that to evaluate verbal communication both measures, intelligibility and listening difficulty, should be taken into account (Sato, et al., 2005). Moreover, it was found that listening difficulty evaluated in noisy and reverberant sound fields, is highly correlated with STI, regardless of the age of the adult listener. It is important to keep in mind that last results were obtained with a constant background noise (Sato, et al., 2012).

Table 2.1: Intelligibility classification ranges for CVC testing according to ISO standard 9921. STI ranges corresponding to the correlations found by Sommerhoff for Spanish language. Adapted from (Sommerhoff & Rosas, 2011).

	Excellent	Good	Fair	Poor	Bad
CVC	>81%	81% to 70%	70% to 53%	53% to 31%	<31%
STI (ISO)	>0.75	0.75 to 0.6	0.6 to 0.45	0.45 to 0.3	<0.3
STI (Noise)	>0.53	0.53 to 0.43	0.43 to 0.31	0.31 to 0.2	<0.2
STI (Reverberation)	>0.52	0.52 to 0.37	0.37 to 0.2	0.2 to 0.003	<0.003

The suitable voice level to maintain a very good intelligibility and low listening difficulty, regardless of background noise, depends on the reverberation time. It has been found that a level of 60 dBA is acceptable to a wide range of reverberation times (between 0 and 2 seconds) for both young and elderly listeners (Sato, et al., 2007). Likewise, Sato et al were able to find acceptable voice levels in terms of constant background noise with a broadband spectrum. Although intelligibility was maximized with a signal-to-noise ratio of at least 10 dB, this condition did not minimize the listening difficulty. It was found that for young adult listeners, the lowest suitable level in terms of intelligibility and listening difficulty was 60 dB, that in presence of background noise between 40 and 45 dB. However, for an elderly population this level was 65 dB with noise levels of 55 dB or less. It was concluded that for a noise between 40 and 50 dBA, the lowest voice level required was 65 dBA. If noise was around 50 and 55 dBA, voice level had to provide a signal-to-noise ratio of 15 dBA. For a noise level

between 60 and 70 dBA, signal-to-noise ratio should be 10 dBA. Furthermore, the maximum acceptable voice level was 80 dBA for noise levels between 40 and 55 dBA or 85 dBA if the noise was among 55 and 70 dBA (Sato, et al., 2011).

In regard to classrooms, different optimal values of reverberation times have been proposed. From a theoretical analysis, Bistafa & Bradley (2000) recommended for volumes between 100 and 500 cubic meters, reverberation times among 0.4 and 0.5 seconds in the 1 kHz octave band and the average of the octave bands from 500 to 4000 Hz. This is consistent with the recommendations for unoccupied classrooms of reverberation times between 0.6 - 0.7 seconds and less than 0.8 seconds (classrooms with fewer than 50 people), given by the standards ANSI/ASA S12.60 (2010) and Building Bulletin 93 (2004), respectively.

Another relevant subject of study, this time in university classrooms, has been the relationship between the subjective perception of room acoustics and the objective acoustic parameters used to evaluate them. The evaluation consisted of a questionnaire designed to find a global measure for a student´s subjective insight regarding the acoustic environment, which was called "Perceived Listening Ease" (PLE). Relationship between this measure and environmental factors such as lighting and temperature were found, suggesting a difficulty for students to separate their acoustic perceptions from the environment. Nevertheless, a significant relationship among PLE and acoustic parameters such as STI was found (Kennedy, et al., 2006).

2.3.2 Application of auralizations to assess acoustically a classroom

Auralizations have been used to subjectively assess acoustic conditions of rooms by evaluating parameters such as intelligibility and listening difficulty. There is evidence of the application of virtual sound environments in intelligibility studies since (1981), when Kleiner analysed the Gothenburg Town Theatre. According to the author, it was possible to assess speech intelligibility with simulated sound fields if the room echograms were foreseen reasonably accurately. The last indicates that Kleiner did sound reflections analysis by means of geometrical acoustics. He also indicated that there was a high correlation between the results obtained by means of direct listening, and the ones by listening to simulated

sound fields. There are other successful examples of intelligibility assessment using auralizations; Peng (2005) and Yang & Hodgson (2006) give some of them.

According to Peng (2005), for an auditorium already built intelligibility could be assessed using a list of words in four different ways. The first option consists of a person reading the words in a place using a certain level and speaking velocity. A second alternative involves recording the words in an anechoic chamber and reproducing them through the sound system of the auditorium. In the third option, the list is recorded in the auditorium to be reproduced to listeners in anechoic chamber. The last alternative consists of recording the words in an anechoic chamber, binaural impulse response measurements in the auditorium, and digital signal convolution processing using the outputs of last two and reproduction via headphones. The first three options not only represent the inconvenience of using the room under study, they also introduce additional disadvantages. If a person reads the words in the place, the success of the test lies on relevant factors as the diction, a proper speaking continuous level and an appropriate rhythm of reading. The use of the auditorium sound system might solve the vocal inconveniences; however, the test would depend on the reproduction system's design given by electroacoustic aspects such as distribution, coverage, quality of transducers and the acoustic power. On the other hand, the last alternative presents a virtual sound environment based on BIR measurements to obtain the source-receiver transfer function in the room as an approach used in the transmission stage of an auralization system.

In Peng´s research, intelligibility was assessed in three virtual rooms by means of auralizations based on geometrical acoustic numerical simulations to obtain BIR. In order to evaluate the quality of the auralizations, the simulated sound fields were compared to BIR measurements and direct listening in the room. Intelligibility results demonstrated that subjective tests based on auralization techniques presented a good agreement compared to tests applied in situ, which offers the possibility of subjectively assessing intelligibility of a non-constructed space. According to Peng, the use of auralizations would provide the possibilities of evaluating any position in the room and determining the impact of acoustic treatments during the building design stage.

Yang & Hodgson (2006) gave other examples on the application of auralizations for subjective tests. One of them was given by the assessment of intelligibility

in a virtual classroom, in order to investigate optimum reverberation times for verbal communication. The authors designed a subjective test to assess by wearing headphones, several acoustic conditions given by changes in voice levels, absorption surfaces and background noise levels. In the experiment two groups of people participated; a first group of 43 people (26 years old average) with normal hearing ability, and another group of 28 adults (46 years old average) with hearing impairment. Noise was integrated into the test in order to simulate an additional source in the classroom, modifying its position with respect to the voice source. For each acoustic condition, a list with 50 standardized words was used to assess intelligibility. The study concluded that when a listener was closer to the noise source rather than the voice source, the optimum reverberation time was zero. In other conditions, the optimal time varied from zero to values close to zero. The same authors compared virtual and real intelligibility values of two classrooms by means of acoustical simulations in CATT-Acoustic software and BIR measurements. For the subjective assessment, the same methodology previously applied was implemented in order to include the noise in the auralizations. The results determined that virtual intelligibility values were reliable if the room presented short reverberation times and low levels of noise (Yang & Hodgson, 2006).

2.4 Effects of noise on cognitive processes

Numerous studies have identified negative effects of noise exposure in the long-term; it affects health, specifically discomfort, sleep disturbances and daytime sleepiness, increased risk of hypertension and heart disease, and deterioration of cognitive performance in children (Basner, et al., 2014). In addition, although the noise is not related to serious psychological disorders, it can affect quality of life and well-being for children and adults (Clark & Stansfeld, 2007). Evidence for the effects of noise on cognitive performance in children is particularly strong: deficits of sustained visual and attention, spoken word perception and poor auditory discrimination, decreased memory in tasks with high demand of semantic material, reading ability impoverished and lower scores on standardized tests (Shield & Dockrell, 2003).

Nevertheless, not only chronic noise exposure affects cognitive performance. Some experimental studies have shown adverse effects due to noise exposure in adult groups. For example, in an office environment and using a low-

frequency noise, Hygge & Knez (2001) observed that an attention task was performed more quickly but with less accuracy in the presence of a higher noise level of 58 dBA compared to 38 dBA. The kind of noise determinates the type of effect on cognitive performance, as it was determined by Trimmel et al (2012). In this research, it was found that both irrelevant and intelligible background voice and aircraft noise negatively affected word learning in different ways. With the former, differences between the types of structure of the text were observed while they were not observed with the latter. Moreover, in an experimental office environment it was observed that a small difference in voice intelligibility listened as background noise presented effects on short-term memory, working memory and the subjective perception of the effect of noise, especially at greater distances from the noise source. In addition, greater affectation was observed in both cognitive performance and subjective perception in people with greater sensitivity to noise (Haapakangas, et al., 2014).

From a theoretical analysis, it can be argued that not only noise but also reverberation can be detrimental to cognitive performance. In intelligibility tests a spoken message should be processed beyond its recognition in presence of background noise and reverberation. In a condition of realistic communication, there is a need to extract information from a series of words or to maintain a continuous attention for long periods of time. However, the limited capacity of working memory means that more resources are needed for the phonological coding in speech, so that there will be fewer of these for processing. Thus, unfavourable reverberation conditions would reduce the cognitive resources of speech processing (Kjellberg, 2004).

This theory seems to be confirmed in part by Klatte et al (2010), who observed a negative effect of reverberation in classrooms on phonological processing in children with an average age of eight years old, as well as high subjective discomfort by the internal noise of the classroom and negative relationship with classmates and teachers. In another experimental study an effect of reverberation was not observed on speech perception as it was observed in the presence of background noise, for both children and adults, especially with background noise similar to the sounds generated in a typical classroom. Furthermore, in the same study an affectation was observed on the comprehension of complex oral instructions in children but not in adults (Klatte, et al., 2010).

Moreover, there are two studies supporting the theory mentioned by Kjellberg (2004). Ljung & Kjellberg (2009) studied the effect of reverberation time on the recall of words or sentences in adults, using a GA model of a classroom with the CATT-Acoustic software and implementing different conditions. With a long reverberation time, participants recalled fewer words; on the other hand, recalling sentences was less affected to the change in the reverberation time, however, longer response times were observed. Moreover, Ljung et al (2009) also observed a negative effect of both broadband background noise and reverberation time on adolescent participants on the recall of the contents of a speech.

2.5 Research Questions

In this research, two research questions have been formulated:

- The numerical simulations to estimate sound fields with the purpose of creating auralizations should consider the implementation of a hybrid model combining GA and FE. For the latter, a frequency dependent real impedance valued related to the diffuse field absorption coefficient is used to define acoustic impedance boundary conditions. What are the consequences in the sound field estimates if acoustic impedance values are defined according to GA material parameter databases?
- It is possible to create auralizations by means of binaural technology able to assess acoustical conditions of existing and virtual classrooms including the variables of background noise levels and reverberation time, in order to evaluate the impact of these variables on cognitive processes such as attention, memory and executive function, all in Spanish language?

Addressing the first research question, section 2.2.2.3 provides examples in the literature regarding the implementation of hybrid models considering the combination of wave equation numerical approaches with GA methods. These hybrid models are based on crossover frequency, from which the GA begins. According to authors as Aretz, et al. (2009), preliminary test showed no artificial audible artefact in the hybrid impulse responses. This means that if the numerical implementation meets the requirements in terms of realistically simulate the geometry, the acoustic source and the acoustic boundary

conditions, the sound field estimate should provide the sound transmission input to create a realistic auralization. From the requirements mentioned, the acoustic boundary conditions is the one providing more difficulty if a virtual environment is intended to be auralized. Aretz (2009) defined a real impedance valued, corresponding to the average absorption coefficient obtained from reverberation time measurements, an approach that can be used with materials in which waves can travel freely along a surface, such as sheets of glass, metal or plywood. This approach allows the use of GA material parameter databases. According to Aretz (2009), the representation of realistic boundary conditions in small rooms can be achieved in order to simulate the low frequency range. Nevertheless, it is not evident in the literature that a significant number of auralizations have been created by means of FEM-GA hybrid approach, in which realistic acoustic boundary conditions for the FEM have been modelled using GA material parameter databases.

It is important to take into consideration that an auralization involves a reproduction stage, which plays a significant role in the whole listening experience. As mentioned in section 2.1, auralizations created applying binaural technology are meant to be reproduced by headphones in most of the cases. It has been identified that the use of headphones systems presents issues such as in-head localization, back–front confusion and complex equalization. Other aspect related to the reproduction stage to take into account according to the literature revised in section 2.2.3.1, is the perceptual assessment of spatial audio systems. In this sense, singular listening qualities such as the accuracy of the localization is an important criteria of perception plausibility and authenticity. According to Rozenn et al (2014), there are physical properties of the sound such as frequency content, the location of the sound source and the acoustic environment that have an influence on the way it's perceived, these parameters are called *related physical attributes*. Addressing the second research question, it is relevant to consider that most of the studies regarding the application of auralizations to evaluate intelligibility, listening difficulty or the effects of background noise and reverberation, have been carried out by reproducing the auralizations via headphones, which could affect the general listening experience and hence, the results obtained.

3. Auralization system theoretical approach

The present chapter introduces the techniques and the theoretical foundations of the stages involved in the creation of the auralizations. First, in the generation stage, the recording technique used to get the source signal is described. Next, the transmission stage details the models implemented to estimate a sound propagation in a room for a singular source - receiver combination. This section finishes with the theory basics of the signal processing applied in the merged of both numerical approaches.

3.1 Generation stage

This section describes the recording technique applied to obtain a dry audio signal of three acoustic sources used in this research, bearing in mind the relation of the recording method with other stages in an auralization system. As it was mentioned previously in the literature review, an alternative to simulate low frequencies in small rooms is given by the application of a hybrid approach considering a wave equation numerical method instead of GA for this frequency range. Hence, it is pertinent that the dry audio signals to be convolved with the BIR had an important content of low frequency energy, in order to make more noticeable the advantages of FE to model the sound wave propagation in this particular frequency range. For that reason, the following three instruments were recorded: a saxhorn, a bass drum (Colombian percussion instrument) (see Figure 3.1) and a bass male voice.

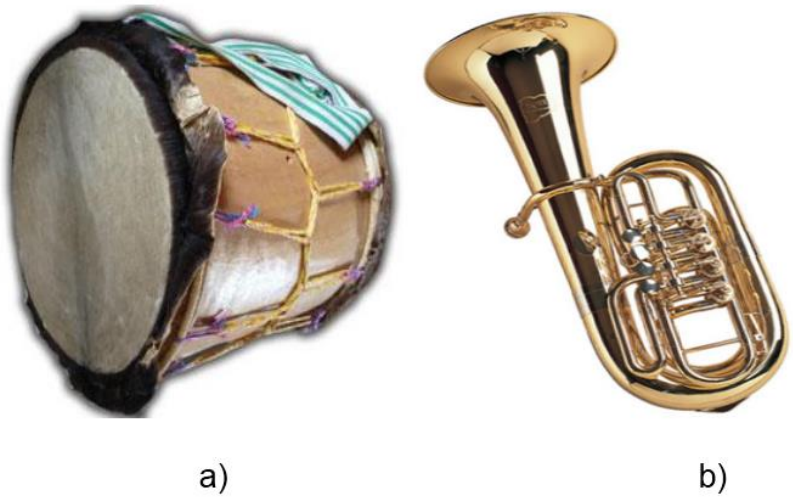


Figure 3.1: Instruments recorded in the generation stage. a) bass drum b) saxhorn.

The technique applied was the single channel recording at the point of main radiation of the instrument. This method provides an important advantage in the reproduction stage, since only one signal must be convolved with the BIR to obtain an auralization of a particular source-receiver combination. Two main aspects were taken into account in the recording of the signals, the analogue to digital conversion process and the acoustics of the room. For the first one, in order to have sound quality equivalent to CD reproduction, the sample rate was set to 44.1 kHz and the amplitude resolution was assigned a value of 16 bits. For the second one, although an anechoic chamber is in theory the proper place to record the acoustic signals, according to Buen (2008), a drier room compared to the spaces to be investigated would be sufficient to record the signals to auralize. In this research, a recording studio, located at the San Buenaventura University, was the place where audio messages were recorded. In order to corroborate Buen´s recommendation, acoustic measurements in the recording studio and the two rooms investigated (meeting room and classroom) were realized. The acoustic measurement procedures and results are discussed in the next chapter. As a reference, the mid reverberation time given by the average between the octave bands of 500 Hz and 1 kHz can be seen in Table 3.3.1.

Table 3.3.1: Mid reverberation of the rooms investigated given by the average between the octave bands of 500 Hz and 1 kHz.

Room	Mid reverberation time (s)
Recording studio	0.47
Meeting room	0.51
Classroom	2.6

3.1.1 Description of the instruments

The male voice recorded for the auralizations has a bass tessitura. The message recorded talks about auralizations. It explains the main goal of an auralization and a gives a brief history of it. The length of the signal is about 1 minute 21 seconds. It has energy content for the frequencies from 50 Hz to 10 kHz approximately. All the spectral plots of the signals recorded were analysed by means of a Power Spectral Density function implemented in MATLAB® software using the whole signal, a 1024 FFT size, 50% overlap and a Hamming window.

The spectral content of the recorded signal for the voice can be seen in Figure 3.2.

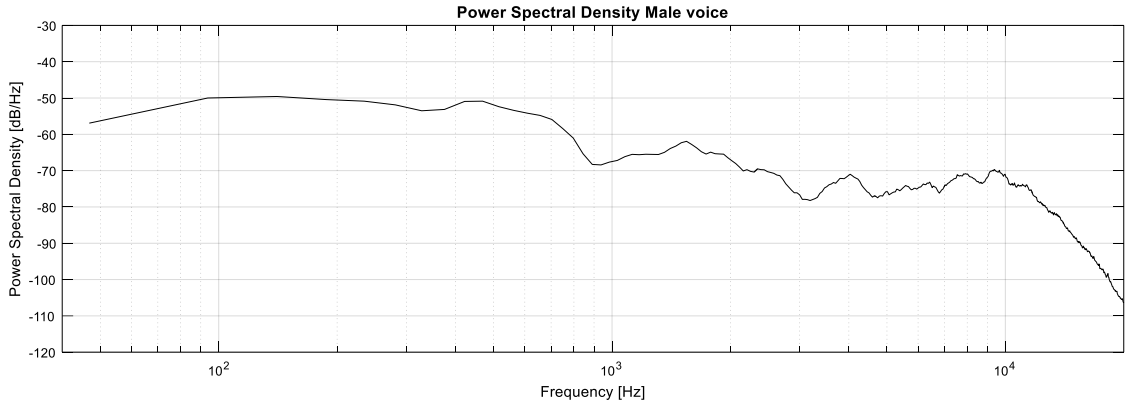


Figure 3.2: Power spectral density of the signal recorded for the male voice.

The saxhorn is a non-transporting baritone-voiced brass instrument with conical-bore, a controlling piston valve and tuned up in B^b. This instrument has a fundamental frequency range between C2 (65.4 Hz) through to B^b4 (466 Hz). It is commonly performed in marches and considered as a band instrument rather than an orchestra instrument. The saxhorn recorded for the auralization performed three pieces of melodies with a length of 1 minute 47 seconds. The recorded signals have energy content for the frequencies from 50 Hz up to 2 kHz approximately. The spectral content of the recorded signal for the saxhorn can be seen in Figure 3.3.

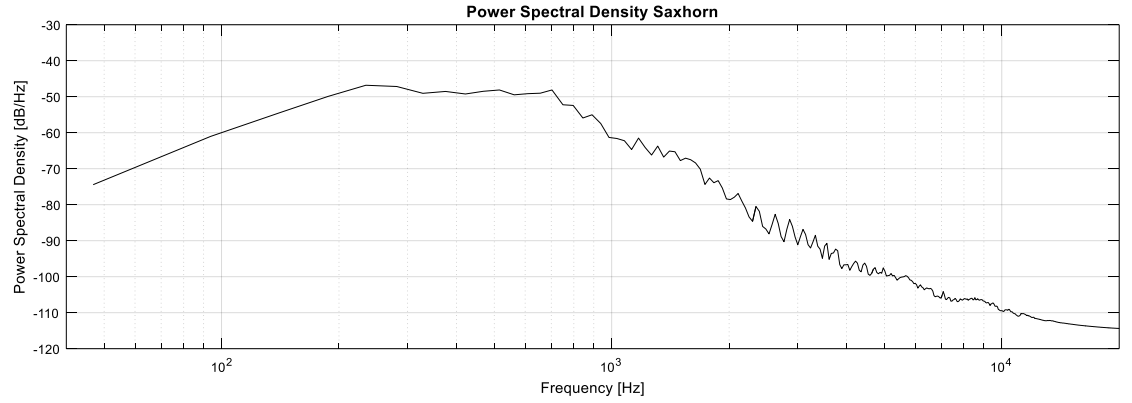


Figure 3.3: Power spectral density of the signal recorded for the saxhorn.

The Colombian bass drum is an instrument made of wood and leather patches (membrane) on both sides. Usually the bass drum is made of goatskin, cowhide or deerskin. The message recorded for the auralizations consisted of two basic rhythms of Colombian music, the base rhythm of “Cumbia” and a variation of

the base rhythm of “Son Corrido”. The signal recorded has a duration of 50 seconds. The recorded signal has energy content from 502 Hz up to 800 Hz approximately. The spectral content of the recorded signal for the bass drum can be seen in Figure 3.4.

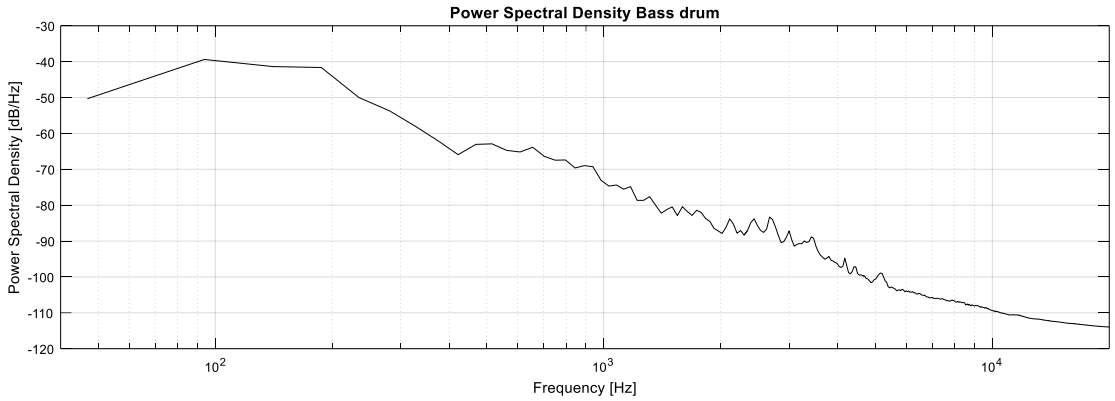


Figure 3.4: Power spectral density of the signal recorded for the bass drum.

3.2 Transmission stage

This section describes the models implemented for both numerical approaches to estimate RIR and BIR for a particular source-receiver combination. These models consider the definition of the following variables: a geometry, boundary conditions, an acoustic source and receivers. In this research, the GA models to be explained were implemented using CATT-Acoustic version 9. To implement the FE models, the commercial code COMSOL Multiphysics 4.3 was applied.

3.2.1 Analysis of wave propagation using FEM

In this section, the fundamental theory and the parameters involved in a FE model with the purpose of simulating sound wave propagation within a room are described. This method requires a discrete 3D model of the volume, a characterization of a source, the location of both ears for each receiver position and a definition of the boundary conditions. Bearing in mind the purpose of combining the results from both numerical approaches, the requirements in terms of input data to maintain similar conditions in both models are taken into account.

The wave propagation considering an enclosure of volume V bounded by a non-rigid surface S must satisfy the Helmholtz equation (Petyt & Jones, 2004):

$$\nabla^2 p + \left(\frac{\omega}{c_0}\right)^2 p = 0 \qquad [3.1],$$

Where ∇^2 is the Laplacian operator, p is the acoustic pressure, ω is the angular frequency and c_0 the speed of sound. Integrating the Helmhotz equation over the volume gives (Petyt & Jones, 2004):

$$\int_V \nabla^2 p dV + \omega^2 \int_V \frac{1}{c_0^{\,2}} p dV = 0 \qquad [3.2].$$

Taking into consideration the divergence theorem (Green´s theorem) stating that the integral of ∇p over the volume of the divergence is equal to the outward flux of ∇p from the closed volume, the first term of equation [3.2] gives (Petyt & Jones, 2004):

$$\int_V \nabla^2 p dV = \int_S \nabla^2 p \cdot \hat{n} dS - \int_V (\nabla \cdot \nabla p) dV \qquad [3.3],$$

where $\hat{n}$ is the outward-directed unit vector normal to the surface and ∇ is the gradient operator.

The next step consists of defining a suitable weak formulation weighting function *W* multiplying all terms, which results in (Petyt & Jones, 2004):

$$\int_S W \nabla p \cdot \hat{n} dS - \int_V (\nabla W \cdot \nabla p) dV + \omega^2 \int_V \frac{1}{c_0^{\,2}} W p dV = 0 \qquad [3.4],$$

Over the surface, the boundary condition can have the following options: acoustically rigid, vibrating with a harmonic normal velocity v or covered by a locally reacting material characterised by a specific acoustic impedance *Z*. The above indicates that the sound pressure field at the boundaries can have the following options (Petyt & Jones, 2004):

$$\nabla p \cdot \hat{n} = 0 \qquad [3.5],$$

$$\nabla p \cdot \hat{n} = -j\rho_0 \omega \hat{n} v \qquad [3.6],$$

$$\nabla p \cdot \hat{n} = -j\rho_0 \omega \frac{p}{z} \qquad [3.7],$$

where ρ_0 is the air density. Including the boundary conditions in equation [3.4] and integrating over the appropriate part of the surface gives (Petyt & Jones, 2004):

$$-j\omega \int_{S_1} W\rho_0 \hat{n} \cdot v dS = \int_V (\nabla W \cdot \nabla p) dV - \omega^2 \int_V \frac{1}{{c_0}^2} W p dV + j\omega \int_{S_2} W\rho_0 \frac{p}{z} dS \qquad [3.8],$$

where S_1 is part of the surface *S* vibrating with a harmonic normal velocity v and S_2 is part of the surface *S* covered by a material characterised by its specific acoustic impedance *Z*.

To derive a FE system of equations requires a discretization of the volume and boundary surfaces into finite elements. In this "meshing" process, the continuous acoustic pressure field is approximated by values of pressure at a finite number of discrete points called nodes, distributed throughout the entire room. At this point, the acoustic domain is divided into tetrahedral or hexahedral 3D fluid elements and triangular or quadrangular 2D sub-regions in the case of boundary surfaces. These subdivisions require a definition of shape functions in order to find the acoustic pressure at any point within each element, interpolating the known values approximated at the nodes of the particular sub-volume or sub-region. Afterward, a sum of nodal pressures and shape functions is defined as a global trial solution. Using the shape functions as the corresponding weighting functions and substituting the trial solution into the weak form of the equations, often referred to as a "Galerkin scheme", gives a set of linear equations of the following matrix notation form (Petyt & Jones, 2004):

$$\left[A(w)\right]_{n\times n}\left[p\right]_{n\times 1} = \left[f(w)\right]_{n\times 1} \qquad [3.9],$$

where, *n* is the number of nodes, matrix *p* contains the unknown nodal pressures p_i *(i=1,2...n)* and matrix *f(ω)* contains forcing terms arising from the excitation applied (Astley, 2010). The matrix *A* is defined as:

$$A = K + j\omega C - \omega^2 M \qquad [3.10],$$

where, *K* and *M* are constant matrices given by the shape functions defined for the interpolation of pressure and *C* is a frequency dependent matrix comprising the boundary conditions. The simplest case is given by a rectangular room with rigid walls assumed, which is the condition to obtain the room modes or natural frequencies. In this situation, the normal particle velocity vanishes at the surfaces and the linear equations of expression [3.9] reduce to the following expression (Petyt & Jones, 2004):

$$\left[K - \omega^2 M\right]\left[p\right] = 0 \qquad [3.11].$$

3.2.1.1.1 Mesh discretization and frequency resolution

In the generation of the mesh, the volume of the room and the wavelength of the frequency analysed define the number of degrees of freedom (DOF) to be solved in a model. This number can be estimated with the following expression:

$$DOF = \text{No. of points per } wavelength^3 \times \frac{volume}{\lambda^3} \qquad [3.12],$$

where λ is the maximum element size specified in the generation of the mesh. The above means that the application of this method in room acoustics is limited to the low frequency range in most of the cases, where the dimensions of the enclosure and the computational resources dictate the highest frequency to be solved.

The grade of detail in FE geometric room models is related to the mesh discretization. Different to what happens in GA, FE models are capable of simulating diffraction and interference phenomena. In order to include these physical effects it is necessary to include all the objects that are not small compared to the shortest wavelength analysed. To represent a single wavelength a minimum of 3 elements are required. However, it is not clear in the literature how many nodes per wavelength are necessary to represent a sound wave propagation in a 3D geometric model. Some authors consider "7-10 nodes per wavelength" a reasonable number as a rule of thumb (Astley, 2010). What is clear is the exponential growth of DOF in 3D meshes as the frequency of interest increases. For that reason, the simulations were divided in frequency groups,

where maximum element sizes of the meshes were determined according to the highest frequency of interest.

Bearing in mind FE is a frequency domain method where a system of equations solves a pressure field for one frequency at a time, frequency steps are to be defined by the user. This frequency resolution is related to the sample frequency used in the generation stage to record the audio signals to be auralized. This means that the impulse response length, which is related to the reverberation time, and the highest frequency to be estimated are the variables to determine these frequency steps. The product of a maximum reverberation time, expected in the room at any octave band, multiplying a hypothetic sampling frequency estimates the first variable. As rule of thumb, the frequency resolution would be given by the next expression:

$$\Delta f < f_s / L_{IR} \qquad [3.13],$$

where, f_s is the sample frequency of the audio signal recorded in the generation stage and L_{IR} is the apparent impulse response length, estimated by:

$$L_{IR} \approx T_{60\max} \times f_{SA} \qquad [3.14],$$

where, T_{60max} is the maximum reverberation time expected in the room at any octave band and f_{SA} is the apparent sampling frequency, given by the highest frequency to be simulated multiplied by a constant factor according to Nyquist criterion, in order to avoid aliasing. Although this criteria allows obtaining audible signals, if Δf is too large, there is a risk of ignoring the contribution of important frequency modes.

A room transfer function in the frequency range can be considered as the sum of contributions from many independent eigenmodes (Kuttruff, 2000). When the room transfer function is statistically dominated at the point at which the frequency spacing between modes becomes close defines the well-known "Schroeder frequency", which is related to the reverberation time T_{60} and the volume of the room by the following expression (Kuttruff, 2009):

$$f_{SH} = 2000\sqrt{T_{60}/V} \qquad [3.15].$$

According to Kuttruff (2000), the average spacing in a rectangular room between adjacent eigenmodes is a function of frequency, and can be expressed as (Kuttruff, 2009):

$$\left\langle df/dN_f \right\rangle = c^3/4\pi V f^2 \qquad [3.16],$$

where, c is the sound speed and f the frequency of interest. In order to guarantee a correct characterization of the room in FE simulations, where the predominant frequency contributions are taken into account, f can be given by f_{SH}. This implies that Δf should be assigned a value between the first criteria, which takes into account the apparent sample frequency, and the second criteria considering the modal density contribution as:

$$\frac{c^3}{4\pi V f_{SH}^{\ 2}} < \Delta f < f_s/L_{IR} \qquad [3.17],$$

where an ideal value would be closer to the left hand expression.

3.2.1.2 Impedance boundary conditions

Every time that a sound wave strikes a surface, a fraction of the acoustic energy is reflected, another is transmitted and another is absorbed (see Figure 3.5). Each of these fractions is identified by means of a coefficient, thereby, the absorbing capability of a material is called absorption coefficient, the rate of acoustic energy reflected is estimated by a reflection coefficient and the proportion of sound transmitted is determined by a transmission coefficient. Considering what is happening at any point of the boundary in terms of acoustic pressure and particle velocity, there is a quantity describing all the acoustic properties of a material mentioned previously. This quantity is called the acoustic impedance, which is defined by the next expression (Kuttruff, 2007):

$$Z = \frac{p}{u_n} \qquad [3.18],$$

where, p is the complex pressure at that point and u_n is the complex amplitude of the normal component of particle velocity at the same point. The acoustic

impedance is a frequency dependent complex value, which can be expressed as (Fahy, 2001):

$$Z = r_n + jx_n \qquad [3.19],$$

where, r_n denotes the real part called resistance, which is associated with "damping" or energy loss due to transmission or dissipation. The imaginary part represented by x_n is called reactance, which is related with the stiffness or mass of the material.

Bearing in mind the purpose of simulating a sound wave propagation inside an enclosure, it is important to describe the reflection and absorption phenomena associated to the surfaces enclosing a sound source. In this sense, assume a plane wave p_i incident upon a boundary with an angle θ in the *xy* plane, as shown in Figure 3.5 (*z* axis is perpendicular to sheet plane). The expression for the incident pressure on the surface is given by:

$$p_i(x,y,t) = \hat{p}e^{j(\omega t - k(x\cos\theta + y\sin\theta))} \qquad [3.20],$$

where, $\hat{p}$ is the complex acoustic pressure magnitude, ω is the angular frequency and k is the wavenumber.

The boundary reflects a portion of the incident pressure attenuated by a reflection factor $|R| < 1$ and with phase shift χ ; the resulting reflected pressure is expressed as:

$$p_r(x,y,t) = \hat{p}|R|e^{j(\omega t - k(-x\cos\theta + y\sin\theta)) + j\chi} \qquad [3.21].$$

Combining $|R|$ and $j\chi$ a reflection coefficient is obtained. This complex factor illustrates the changes in amplitude and phase taking place on the partial standing wave formed by incident and reflected waves (Kuttruff, 2007):

$$r = |R|e^{j\chi} \qquad [3.22].$$

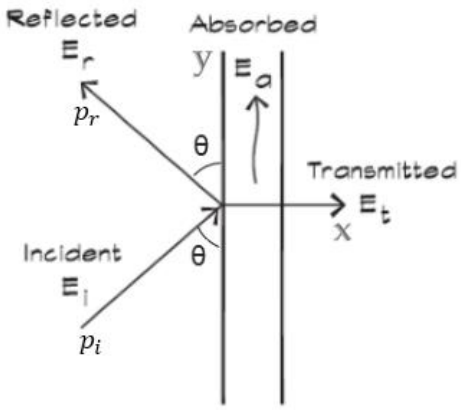


Figure 3.5: Distribution of the acoustic energy when an incident sound wave strikes a boundary (Long, 2014).

Considering the reflection coefficient, the reflected pressure p_r can be rewritten as:

$$p_r(x,y,t) = \hat{p}re^{j(\omega t - k(-x\cos\theta + y\sin\theta))} \qquad [3.23].$$

It is important to bear in mind that wall impedance is related to the complex reflection factor by the following expression (Kuttruff, 2009):

$$Z = \rho_0 c \frac{1+r}{1-r} \qquad [3.24].$$

In fact, the absorption coefficient is related to the acoustic impedance and can be rewritten in terms of resistance and reactance for normal incidence as (Kuttruff, 2009):

$$\alpha = \frac{4\mathrm{Re}\left(\frac{Z}{\rho_0 c}\right)}{\left|\frac{Z}{\rho_0 c}\right|^2 + 2\mathrm{Re}\left(\frac{Z}{\rho_0 c}\right) + 1} \qquad [3.25].$$

Beranek & Ver (1992), who state that when a uniform sound pressure field is created throughout the enclosure, give another approximation to this boundary condition expressed as:

$$Z = \frac{\left[C + j\omega\left(M - \frac{K}{\omega^2}\right)\right]}{A} \qquad [3.26],$$

where, *C* is a damping term, *M* is the mass, *K* is the stiffness and *A* is the surface area of the wall. According to the last expression, on a very stiff and reflective

wall without absorptive material or damping on it, the impedance can be approximated to:

$$Z \approx \frac{-jK}{A\omega} \qquad [3.27].$$

Likewise, when the frequency analysed is above the resonance frequency of the wall, the acoustic impedance is dominated by the mass of the partition and can be approximated with the expression:

$$Z \approx \frac{j\omega M}{A} \qquad [3.28].$$

The problem with Beranek's approach is given by the difficulty of finding stiffness and mass values that accurately approximate impedance conditions in a room. For this reason, in room acoustic simulations a simple approach of relating the acoustic impedance with an absorption coefficient (see expression 3.25) seems to be the most practical way to model the boundary conditions in a wave equation numerical method, especially if similar conditions with GA models are to be achieved. As it was stated in the literature, Aretz (2009), who successfully realised the acoustic impedance boundaries in a FEM model of a recording studio, gave an example of implementation. The contribution of Aretz lies on the use of field incidence absorption coefficient to find the resistance part of the acoustic impedance, which is associated with the energy loss by either dissipation or transmission. Taking the above into account, real and frequency dependent impedance values were approximated using the field incidence absorption coefficients applied in the GA models.

3.2.1.3 FE source characterization

In this thesis, when simulating sound wave propagation by means of combining two numerical approaches and where GA methods are used to predict sound radiation for mid and high frequencies, FEM is applied to estimate the sound pressure field in the low frequency range. Although real sound sources do not have an omnidirectional sound propagation pattern, most instruments and electroacoustic sources present a sound radiation pattern tending to be omnidirectional in lower frequency bands. This idea facilitates the application of a simple pulsating point source propagating sound in all directions. The sound

pressure field generated by this particular source within an acoustic domain satisfies the inhomogeneous wave equation given by:

$$\nabla^2 p - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} = -\rho_0 \frac{\partial q}{\partial t} \qquad [3.29],$$

where, c_0 is the sound velocity, ρ_0 is the air density, ∇^2 is the Laplacian operator and q is a term describing a volume velocity source, visualized as a small pulsating object injecting volume into an acoustic domain. The corresponding time domain solution is a pressure field of the form (Thompson & Nelson, 2015):

$$p(r,t) = \mathrm{Re}\left\{\frac{A}{r} e^{j(\omega t - kr)}\right\} \qquad [3.30],$$

where, r is the spherical radius or distance travelled by a sound propagation, ω is the angular frequency, k is the wavenumber and A is a constant dependent of the source volume velocity, expressed as:

$$A = \frac{jq\rho_0\omega}{4\pi} \qquad [3.31].$$

In FE simulations, a point source is fully characterized by defining the acoustic intensity or the acoustic power at 1 m distance in the free field. Keeping in mind the integration of numerical methods, all the frequencies were estimated with point sources radiating the same acoustic power, in order to maintain the same conditions in both models.

3.2.1.4 FE binaural receiver model

For source localization, the size of the head and position of the ears determine two main principles in binaural hearing, the Interaural Differences of Level (ILD) and Time (ITD). The first one is given by the diffraction effect, where the head generates an acoustic shadow at frequencies whose wavelengths are comparable to the head dimensions. The second one is associated to the time difference of the path lengths for both ears. According to Wightman & Kistler (1992), this time difference is frequency dependent, with larger values found at low frequencies. In terms of ILD, significant variations are presented at high frequencies and the variation of the angle of elevation is relevant at these frequencies due to pinna

cues. In this sense, Middlebrooks & Green (1990) stated that in the horizontal plane, localization is mostly based on ILD and ITD without pinna cues.

The function describing the binaural cues taking place when a sound wave reaches a person, receives the name of the Head Related Transfer Function (HRTF). In a strict sense, to characterize a binaural receiver in FE simulations that properly simulates a HRTF at low frequencies requires to model the head, torso and two receiver points at ear positions (Aretz, 2012). In this project, the torso was not included in the numerical simulation and a simple approach was used considering a cubic form instead of a head. The size of the cube was defined according to the average physical dimensions of a head established in the IEC/TS 60318-7:2011 standard (2011). This approach is based on the fact that the maximum frequency simulated in the FE model was around 600 Hz, whose wavelength is approximately 0.57 m at normal temperature and humidity conditions, which is more than twice the average size of a human head and almost four times the average separation between ears (International Electrotechnical Commission, 2011). In Figure 3.6 an example of the cubic form used in the FEM simulations can be appreciated. The arrow indicates the hypothetic direction at which the receiver is directing its head and hence, the sides of the cubic form where the point receivers must be placed.

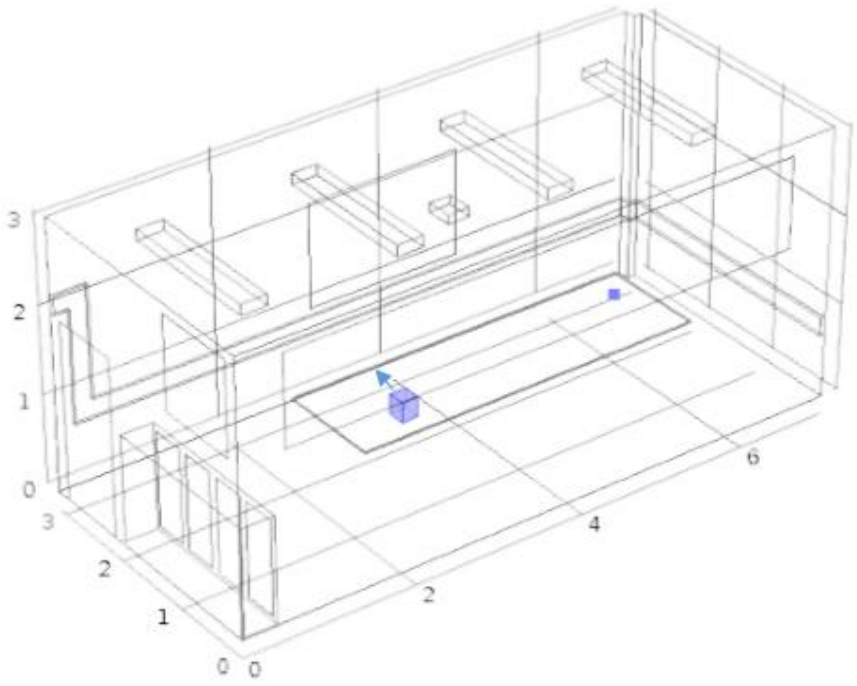


Figure 3.6: Binaural receiver model in FE simulations.

3.2.2 Analysis of wave propagation using GA

The methods based on GA theory analyse sound propagation by replacing the wave concept with the idea of a sound ray, just as a light ray in geometrical optics. The only differences with optical theory are the velocity of propagation and the decrease of the intensity, which is given by $1/r^2$, as in any spherical wave, where r denotes the distance from the origin. Another important fact in ray tracing theory is given by the reflection over the surfaces of the room. When a sound ray is incident upon a plane surface, there is a specular reflection where the reflection angle is equal to the incidence angle (see Figure 3.7). In order to estimate a sound pressure field generated by a source at a specific point of the room, the energy contributions of rays passing through a detection area (circular area or spherical volume defined at specific position) are added within set time intervals recording its direction and arrival time.

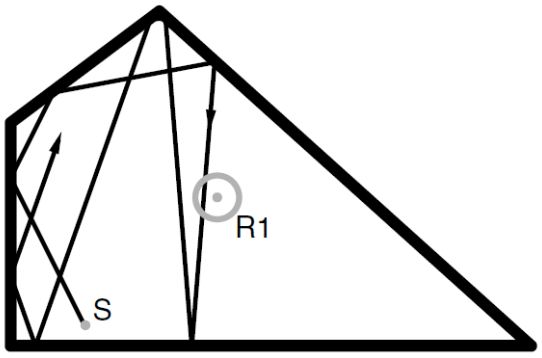


Figure 3.7: Ray tracing example of emitted ray from source S, entering the circular cross-section detection area of receiver R1 (D'Antonio, 2001).

Another way to estimate the reflection of a sound ray is given by the creation of a virtual image source placed in front of a reflection plane at its symmetrical position. The new source is separated by an equal distance from the reflecting boundary and is emanating with the same angle of the reflected ray (see Figure 2.2). In the Image Source Model (ISM), the surfaces are assumed to be perfectly planar, the effective power of an image source will depend on the absorption of the reflective plane and the scattering coefficient is neglected. Another difference with the basic ray tracing method lies in a simpler way to estimate the ray path times, obtained by calculating a three-dimensional vector length from each image source to the receiver (CATT, 2007).

In the ISM, the reflection is created by one imaginary source and is called a first order reflection. When a ray is reflected by two surfaces before reaching a

receiver, two imaginary sources are used to estimate the total path length, receiving the name of second order reflection, and so on. According to Kuttruff (2000), the construction of image sources for a given enclosure and source position does not have to be related to a specific ray path. In fact, each plane wall *N* can be associated with one first order image source, whose mirrored sources will lead to second order reflections as *N(N-1)*. The total number of images of order *i* can be calculated with the repetition of this procedure using the following expression (Kuttruff, 2009):

$$N(i) = N\frac{(N-1)^i - 1}{N-2} \qquad [3.32].$$

Another approach to deal with diffuse reflections in GA hybrid models is given by Dalenbäck (1996). In this case, a cone tracing algorithm allows surfaces with diffusion factor different to zero, generating both specular and diffuse reflections. In order to avoid an exponential growth, the assumption of a quadratic reflection density with time and a reflection order parameter are used. The application of this method can be appreciated in the software CATT-Acoustic, where Randomized Tail-corrected Cone-tracing (RTC) is given to provide more detailed calculations, capable of generating echograms that can be used for auralizations (see Figure 3.8) (CATT, 2007). The RTC is an algorithm that combines features of the standard ray tracing, the ISM and the specular cone-tracing. In this method, the direct sound, first order specular and diffuse reflections and second order specular reflections are estimated independently by the ISM. According to Catt's user manual, for each octave band a separate ray/cone-tracing takes place for the reason of the frequency-dependent diffuse reflection. Interested readers in GA prediction methods based on RTC algorithm are referred to the Dalenback's PhD thesis (Dalenbäck, 1995).

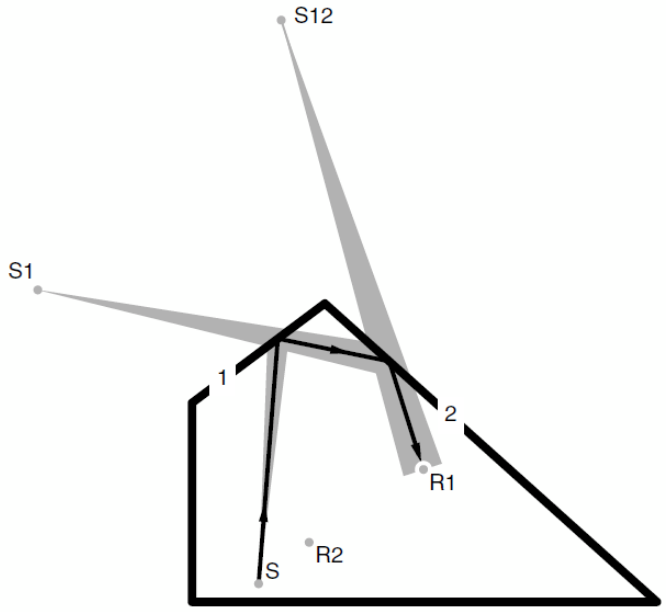


Figure 3.8: Second-order reflection using ISM and cone tracing (D'Antonio, 2001).

3.2.2.1 Prediction Method

The Randomized Tail corrected Cone-tracing (RTC) technique applied in this research is part of a full detailed calculation prediction method to estimate complete echograms and acoustical parameters like Early Decay Time and Reverberation Time. In order to simulate a BIR, this numerical technique includes an ISM for early specular reflections and a stochastic RT method for high-order reflections. This method requires a mathematical model to estimate sound propagation, generation of a 3D model describing the room geometry, acoustic properties of absorption and scattering on the boundaries and characterization of acoustic source and receiver positions.

The geometry in CATT is created in text language using input points with *x, y* and *z* coordinates to form planar surfaces. According to Catt's user manual, a maximum number of 99,999 planes are available with an aim of modelling any shape. The level of detail required in GA models to obtain accurate acoustical properties has been discussed by several authors, who agreed that high levels of geometric detail do not necessarily lead to better accuracy (see e.g. Smith, 2004). Considering that existing CAD (Computer Aided Design) tools facilitate the construction of high quality geometric room models, most of the GA commercial codes offer the possibility to import these CAD models, and CATT-Acoustic is not an exception. Bearing in mind the same geometric model must

be designed for FE simulations, room models were originally created in CAD language, which is compatible with both technologies.

This technique requires two input parameters for the calculations, a number of rays per octave-band and a truncation time. To estimate the first parameter, Smith (2004) carried out an exhaustive research quantifying the number of rays required to produce reliable results with different acoustic packages, including the software CATT-Acoustic. The conclusions of Smith's investigation showed that for a low number of rays most of the acoustic parameters calculated in CATT converged. Likewise, the author stated that the Auto number option in CATT provides the required number of rays to correctly estimate acoustical parameters (Smith, 2004). This option selects the larger number of the following two choices:

- A number corresponding to 1 ray per square meter arriving 80ms after the direct sound for all positions.
- A number corresponding to 1 ray for every 4 square meters at the longest hall dimension.

In this research, the Auto number option offered by CATT was doubled in order to guarantee reliable results in the sound propagation estimation of the rooms investigated.

The second input parameter in this prediction method is the ray truncation time. In order to obtain a well estimated reverberation time, it is recommended to set this time equal or higher to the maximum octave band reverberation time estimated by a classic Sabine or Eyring model.

3.2.2.2 Boundary conditions in GA

The next step consists of the acoustic characterization of room boundaries, in order to simulate the sound rays' reflections at the surfaces. In GA, the acoustic boundaries require defining two parameters: an absorption coefficient and a scattering coefficient.

3.2.2.2.1 Absorption Coefficient

In geometrical room acoustics, a reflection does not present a change in phase; hence, the reflection coefficient of expression [3.22] relies on the reflection magnitude. This indicates that every time that a sound wave strikes a surface,

the reflected energy is reduced by a percentage. This amount is estimated by the absorption coefficient of the material, which is a function of the reflection coefficient (Kuttruff, 2009):

$$\alpha = 1 - |r|^2 \qquad [3.33],$$

It is important to bear in mind that a specific surface can be struck several times for sound rays coming from different directions, thus the absorption coefficient must be defined in terms of a reflectivity contribution from every possible angle. This is the case when the absorption coefficient is measured in a reverberation chamber where a diffuse field is assumed. Taking this point into account, the absorption coefficient could be defined as "the ratio of intensity absorbed by the surface to the intensity incident to the surface" (Nelson, 1998) and its average in a room would be given by:

$$\overline{\alpha} = \frac{1}{S}\sum_i S_i \alpha_i \qquad [3.34],$$

where, *S* is the total surface area of the enclosure and $S_i \alpha_i$ are the area and absorption coefficient of individual surfaces in the room. This value is related to the reverberation time of the room by the following expression (Kuttruff, 2007):

$$T_{60} = \frac{0.161V}{S\overline{\alpha}} \qquad [3.35],$$

where, *V* is the volume, *S* is the total surface of the room and T_{60} is the reverberation defined as the time needed for the sound to disappear or decrease 60 dB after the sound source has been turned off.

3.2.2.2.2 Scattering Coefficient

A "Scattered Reflection" is applied when a high order reflection is incident upon a rough surface in a ray tracing algorithm. In this case, a percentage of the reflected sound energy is scattered with randomized directions when a sound ray strikes the surface. According to Vorländer & Mommertz (2000) the amount of energy randomized is defined by the scattering coefficient, symbolised with the letter δ, and defined "as the ratio of the non-specularly reflected sound energy to the total reflected energy". Assuming normalised incident energy of 1, the total reflected sound energy will be $1-\alpha$, the fraction of the sound energy that

is specularly reflected will be *(1-α)(1-δ)* and the component that is scattered reflected will be *(1-α)δ* (see Figure 3.9).

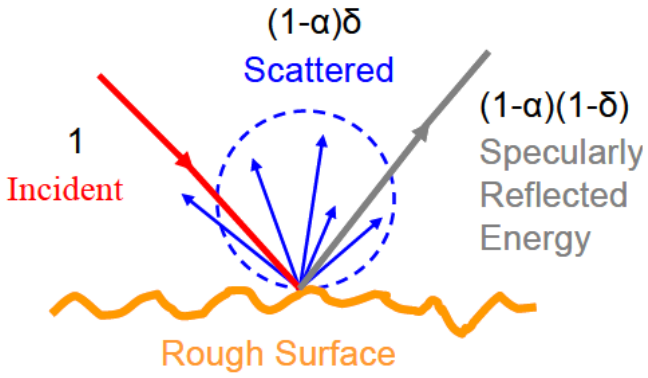


Figure 3.9: Scattering coefficient from a rough surface.

The importance of these kind of reflections, apart from including the effect of reflections on rough surfaces, is to find a way to include diffraction phenomenon given by the sound propagation interaction with obstacles, which is also ignored in basic ray tracing reflections. This can lead to unreliable results in a GA sound propagation prediction. Some authors state that a lack of scattering coefficients in geometrical room acoustics methods leads to severe overestimation of reverberation time and unnatural decay characteristics in binaural room impulse responses (Dalenbäck, 1995). Thus, a scattering coefficient is not only applied to irregular surfaces, but also to elements like tables, chairs or other kinds of furniture where the phenomenon of diffraction can be presented, according to its size and wavelength of the frequency analysed.

3.2.2.3 Source characterization in GA

In GA, the concept of a point source propagating sound rays in all directions is used. There are two requirements to characterize a source in a GA model: the sound radiation pattern and the broadband noise describing the sound pressure levels per octave band at 1 m distance. The first requirement is needed in the transmission phase to model a proper sound source directivity, in this case, the methods to obtain a sound source directivity are widely accepted and standardized (International Organization for Standardization, 2010). Bearing in mind the integration with FE, the second requirement was defined with a white noise, which exhibits a continuous sound pressure level in the entire spectrum.

The sound radiation pattern or the so-called directivity factor of a source describes the relationship, in free field conditions, of the acoustic intensity at a

given angle (θ, Ψ) with the acoustic intensity of the source assuming uniform radiation in all directions and far field propagation. The directivity factor is expressed as (Bies & Hansen, 2009):

$$D_\theta = \left(\frac{I_\theta}{\langle I \rangle}\right)_r \qquad [3.36],$$

where I_θ is the sound intensity at angle (θ, Ψ) and distance *r* from the source; and $\langle I \rangle$, is the average sound intensity over a spherical surface of radius *r*. The directivity factor of a source is defined in decibels as the directivity index (*DI*), as follows (Bies & Hansen, 2009):

$$DI = 10 \log_{10} D_{\theta ref} \qquad [3.37],$$

where θref indicates an angle of reference, which is given on axis for horizontal and vertical planes, in front of the source.

In order to measure the directivity characteristics of a source, ISO standard as the ISO 3744 (International Organization for Standardization, 2010) might be used. These standards specify methods to determine the sound power level of a source by means of sound pressure measurements, with the possibility to obtain directivity information. In this project, the loudspeakers data provided by the manufacturers was used in order to characterize the sound radiation pattern of the source. In this research, given that BIR measurements have been used as input data to create auralizations of the rooms investigated, the numerical BIR have been created simulating the directivity characteristics of the loudspeakers used in the acoustic measurements, in order to have similar conditions for the comparison between sets of auralizations. The reference of the loudspeakers used in this research can be seen in Chapter 4, which includes the acoustic measurement procedures and results. The loudspeaker directivity modelled in the software CATT-Acoustic can be appreciated in Chapter 5, which details the numerical implementation procedure of the rooms investigated.

3.2.2.4 GA binaural receiver model

The first step to characterize any receiver in GA is estimating an octave band echogram. In this representation, reflections are marked by perpendicular lines over a horizontal time axis according to their arrival times (see Figure 3.10). The height of each reflection is a function of the acoustic properties of the source

and the frequency-dependent characteristics of absorption and diffusion of the surfaces in the room. Hence, the highest level at the receiver position is given by the direct sound.

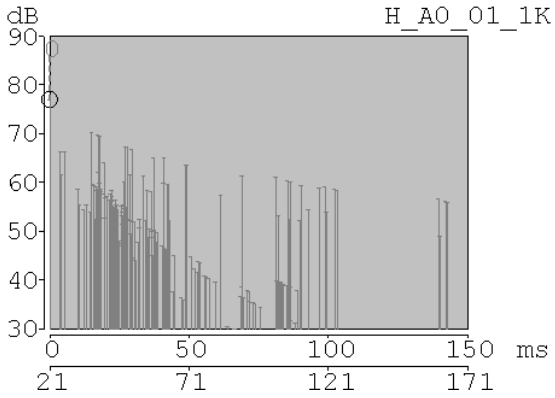


Figure 3.10: Echogram representation (CATT-Acoustic software).

At this stage, every single reflection contains information regarding its intensity, its arrival time and angle (D'Antonio, 2001). Taking into account that a binaural receiver is characterized by the corresponding HRTF, the following step is to convert the echogram reflections into an impulse response describing the responses at each ear. The Hilbert transform is used to provide a minimum phase construction of the magnitude information in the frequency domain. Then, an inverse fast Fourier transform (IFFT) is applied in order to obtain the impulse response of each reflection (see Figure 3.11). The size of the impulse response is given by the truncation time and the sample frequency. Afterward, the impulse responses obtained for each reflection are convolved with the left and right HRTF in order to get the BIR for that source-receiver combination.

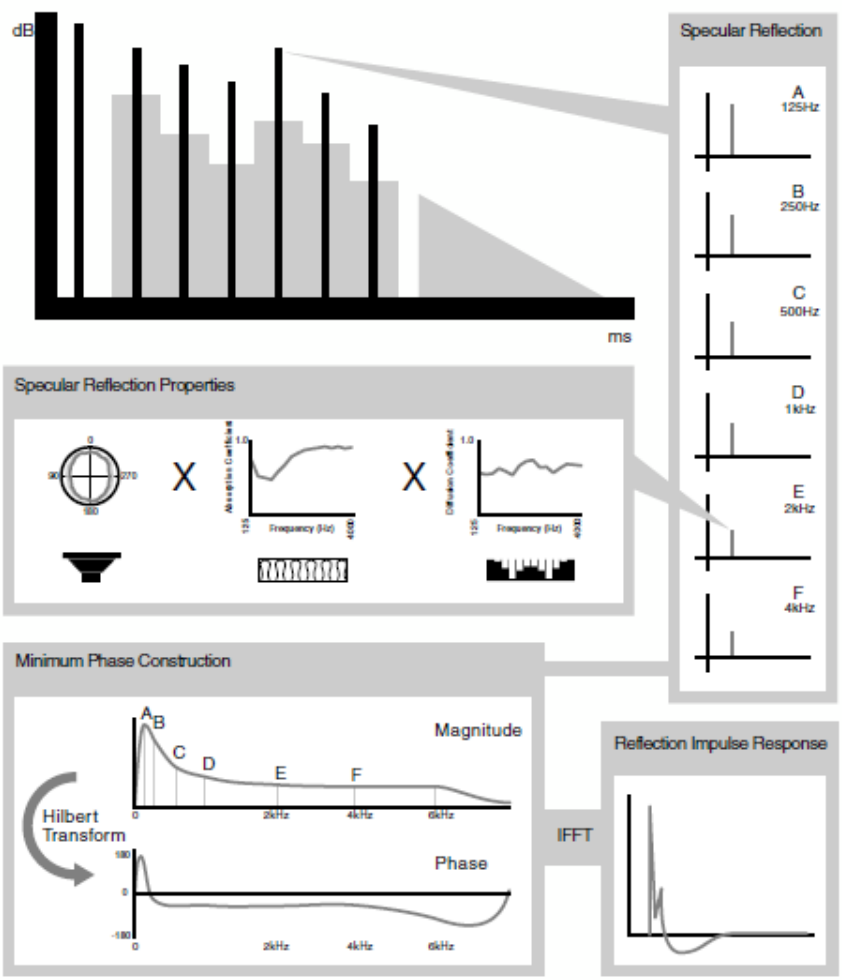


Figure 3.11: Transfer function construction to convert an echogram reflection into an impulse response (taken from D'Antonio (2001)).

3.2.3 Signal processing to create auralizations

In this section, the basics of the signal processing involved in the procedures applied to obtain the final auralizations are explained. In this research, three kinds of auralizations have been created, one by means of BIR measurements and the other two applying numerical simulations to estimate the room transfer path between source and binaural receiver (see sections 3.2.1 and 3.2.2). In all the cases, Linear Time Invariant (LTI) systems are assumed, since the pressure field at a given receiver position in a room, function of an impulse excitation signal at certain source position are to be estimated. The last is defined as the RIR, or the BIR, in the case when a binaural receiver is to be modelled. In the time domain, to create the auralizations binaural pressure fields at receiver positions can be calculated by the convolution operator using binaural room transfer functions as:

$$r_{l,r}(t) = \int_{-\infty}^{\infty} s(\tau) h_{l,r}(t-\tau) d\tau = s(t) * h_{l,r} \qquad [3.38],$$

where, $r_{l,r}$ is the time pressure response for left and right ear, $s(t)$ is the excitation signal and $h_{l,r}$ refers to the room transfer function including the HRTF for left and right ear.

Note that room transfer path estimation results are given in the time domain for GA and frequency domain for FE. In linear acoustics, any time dependent quantity can in theory be reconstructed as a sum of time-harmonic solutions, and vice-versa, by using the complex Fourier Transform pair of equations (Randall, 2008):

$$p(t) = \frac{1}{2\pi} \int_{-\omega}^{\omega} P(\omega) e^{j\omega t} d\omega \qquad [3.39],$$

$$P(\omega) = \int_{-t}^{t} p(t) e^{-j\omega t} dt \qquad [3.40],$$

where, ω is the angular frequency, $P(\omega)$ is the time harmonic acoustic pressure and $p(t)$ is the acoustic pressure in the time domain.

Bearing in mind that sampled signals are used, a numerically fast implementation of the Discrete Fourier Transform, is defined with the corresponding pair of equations (Proakis & Manolakis, 2007):

$$x[n] = \frac{1}{N} \sum_{K=0}^{N-1} X[k] e^{j 2\pi k n / N} \qquad [3.41],$$

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j 2\pi k n / N} \qquad [3.42],$$

where, N is the number of data in the sequence, and each $X[k]$ is a complex quantity encoding both amplitude and phase of a time-harmonic component of function $x[n]$.

In order to combine the results obtained with both numerical methods for each room transfer path, a filtering process had to be applied. Keeping in mind the limitations of GA to estimate the sound wave propagation at low frequencies, a high-pass filter for GA results and band-pass filter for FE values were implemented. In the FE method, it is important to bear in mind that frequency

domain results are obtained with an apparent sample frequency, according to the frequency steps defined by the user. In order to match the sample frequency of the GA simulation results it is therefore necessary, padding with zeros outside the frequency interval estimated with a number according to the GA BIR length. An Inverse Discrete Fourier Transform (IDFT) follows this procedure in order to obtain the FE BIR.

In linear systems, filtering can be described in the time domain by convolution and in the frequency domain, by its equivalent of multiplication. The pair of equations describing filtering in both domains are given by (Hammond & White, 2008):

$$y(t) = \int_{-\infty}^{\infty} x(t)h(t-\tau)d\tau = \int_{-\infty}^{\infty} x(t-\tau)h(t)d\tau = x(t) * h(t) \qquad \textit{[3.43],}$$

$$Y(f) = X(f).H(f) \qquad \textit{[3.44],}$$

where, $h(t)$ is the impulse response of the filter and $H(f)$ is the equivalent frequency response. In this project, recursive filters or the so-called Infinite Impulse Response (IIR) of order ten were applied.

In Figure 3.12, the signal processing involved to create the three types of auralizations used in this project is explained. The reference auralizations ("REF AURA") were created by means of BIR measurements, denoted as "M BIR", and a convolution process with anechoic material, represented as "ANC MAT". The same procedure was used to obtain the auralizations estimated by means of GA ("GA AURA"). The last group of auralizations was created applying FE to estimate the sound propagation at low frequencies and GA for the rest of the spectrum ("FE-GA AURA"). In this case, band-pass and high-pass filters had to be applied respectively before adding both simulation results. The next step after obtaining the final auralizations was given in the 3D reproduction system OPSODIS. Here, the binaural audio signals were filtered with the corresponding crosstalk cancellation filters, and then each audio signal was divided in three frequency bands. Finally, left and right signals were distributed correspondingly in three speakers according to the Optimal Source Distribution principle, in order to reproduce 3D sound at the listener position.

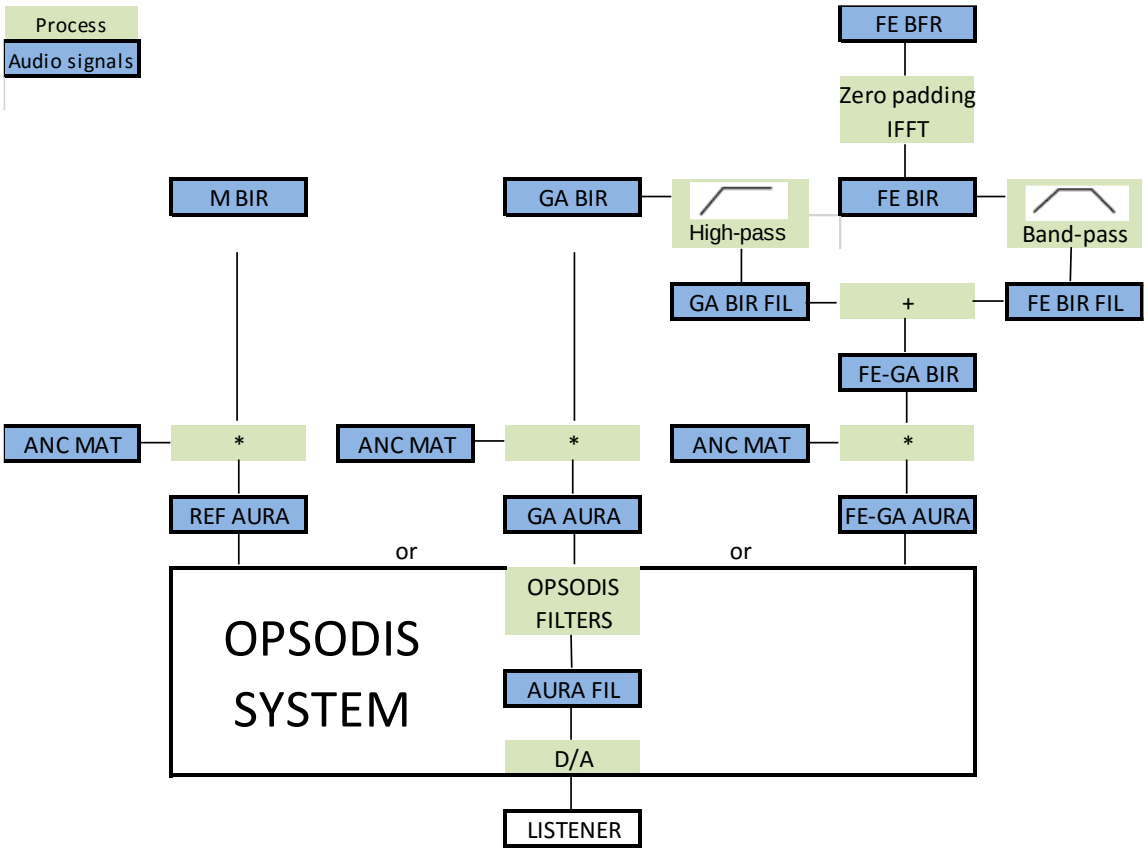


Figure 3.12: Process diagram to create the three kinds of auralizations researched in this project.

The -3dB cut-off frequency was defined taking into account the highest frequency estimated in FE simulations for each room investigated. Taking into account that the highest frequency estimated in FE simulations for the meeting room was around 700 Hz and for the classroom was approximately 500 Hz, two sets of filters were created. The band-pass filter responses applied for FE signals for both rooms investigated can be seen in Figure 3.13 and Figure 3.14. In order to provide an example of the use of the filters, in Figure 3.15 the phase response of a FE signal band-pass filtered can be seen. Likewise, in Figure 3.16 is possible to appreciate the phase response of a signal given by the combination of both numerical methods after the filtering process.

Figure 3.13 and Figure 3.14 show the magnitude and the phase responses of the filters applied to the FE results of both rooms investigated, this before the FE signals were combined with the GA results. The phase responses obtained in both cases indicate that there is a continuity in the frequency domain, which guarantee that the FE results had appropriate phase information before the combination with the GA results. Figure 3.15 shows the phase response of a

filtered FE signal, having as the input the meeting room results and the filter response. Figure 3.16 shows the phase response of a combination given by FE and GA results, which indicates that the outcome obtained by the hybrid approach had an appropriate frequency response in order to be used for auralizations.

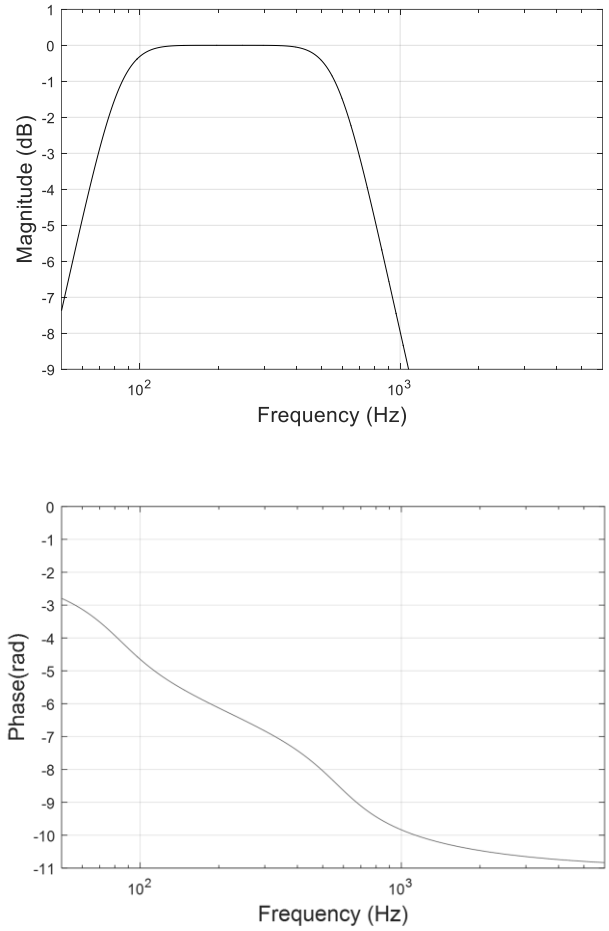
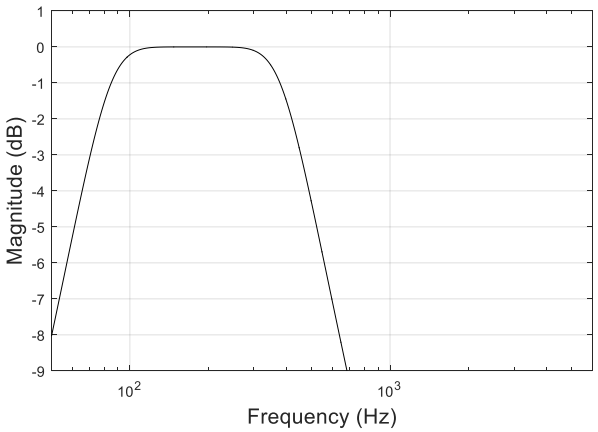


Figure 3.13: Magnitude and phase responses of the band-pass filter applied for the meeting room FE signals.



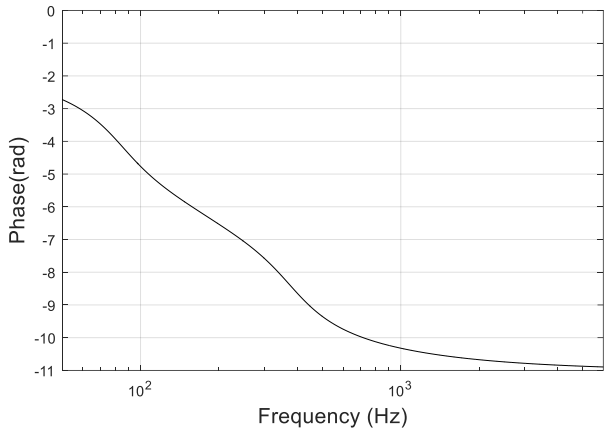


Figure 3.14: Magnitude and phase responses of the band-pass filter applied for the classroom FE signals.

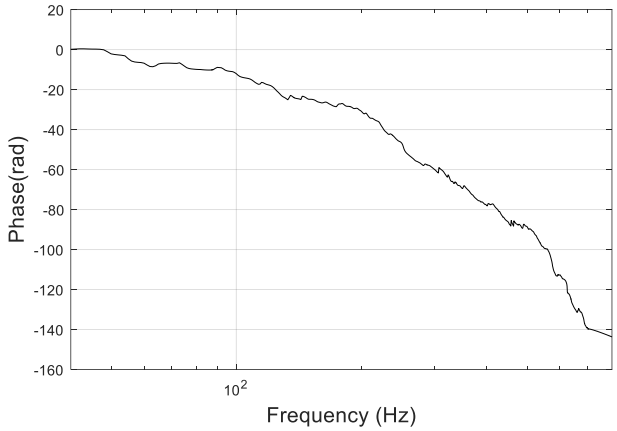


Figure 3.15: Example of phase response for a FE signal band-pass filtered.

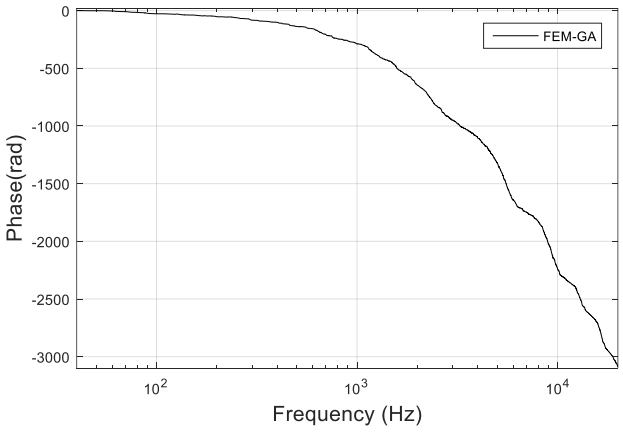


Figure 3.16: Example of phase response for a FEM-GA signal.

4. Acoustic measurements of the rooms investigated

This chapter presents the procedure and results of the acoustic measurements carried out in the rooms investigated in this thesis: the meeting room, the classroom and the recording studio. For all of them, RIR measurements were implemented in order to characterize the rooms, calculating the most relevant acoustic parameters according to the recommendations of ISO standard 3382 (2009). In addition, in the first two rooms BIR measurements were realised with the purpose of obtaining the room transfer functions for all source-receiver combinations, in order to create the reference auralizations used to evaluate current conditions of the rooms and for the comparison with the auralizations created by means of numerical simulations.

4.1 Measurement procedure of RIR and BIR

The measurements of room acoustic parameters were carried out applying the integrated impulse response method. The RIR measured were used to estimate the acoustic descriptors explained in the following chapter, according to ISO standard 3382 (2009). The measurements were implemented with the engineering degree of precision defined by the standard, consisting of six source-receiver combinations. The excitation signal used was a log sweep with a frequency range from 20 Hz to 20 kHz. In Figure 4.1, a block diagram explaining the methodology used in the RIR measurements can be seen.

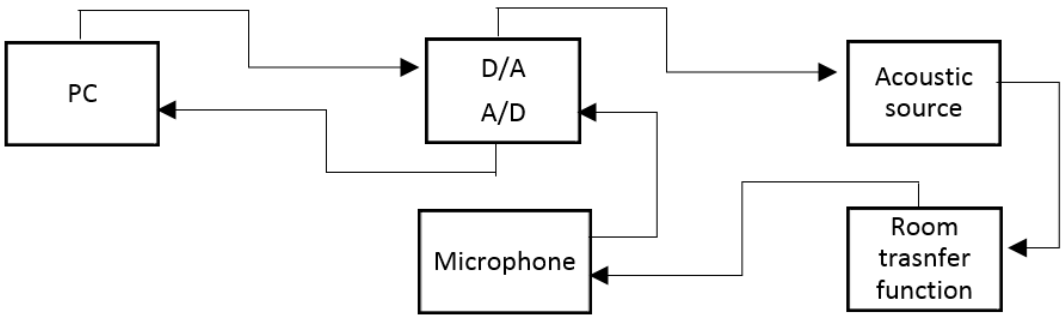


Figure 4.1: Block diagram sketch of room acoustic parameters measurements.

The BIR measurements were carried out using a dummy head, with the purpose of obtaining the binaural room transfer functions necessary to create the

reference auralizations. The measurement´s procedure is very similar to the one used for room acoustic parameters measurements, the only differences can be seen in the block diagram sketch of Figure 4.2.

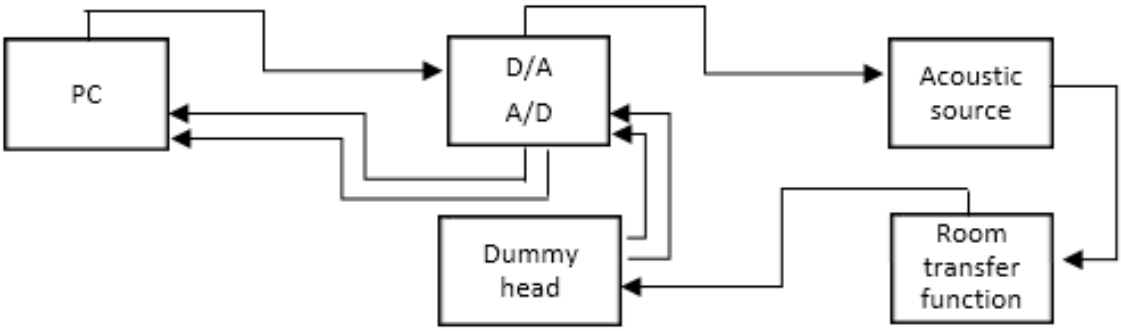


Figure 4.2: Block diagram sketch of BIR measurements.

4.2 Acoustic measurements of the Meeting Room

In this room, two types of measurements were carried out, BIR and RIR. The purpose of the first set of measurements was to obtain the room transfer path for all source-receiver combinations, in order to create the reference auralizations. The second set of measurements intended to characterize the room, calculating the most relevant acoustic parameters according to ISO standard 3382. The intention of the acoustic measurements was to provide reference values, in order to assess the numerical approaches used in this research to simulate sound propagation. The drawbacks of the procedure applied in the measurements were the loudspeaker used as sound source and the location of the loudspeaker on the table given by the reduced space available. According to the standard, the sound source shall be as close to omnidirectional as possible, especially if log sine sweeps are used to excite the room, where the requirements for directionality of the source should be fulfilled according to ISO 3382 (2009).

4.2.1 Meeting room description

The meeting room used in this research is the enclosure number 2011 located on the second floor in the ISVR building of University of Southampton. This room has a rectangular shape of 7.3 m long, 3.1 m width and 3 m height, with an approximate volume of 67 m³. The ceiling of the room is built with gypsum boards of 0.5 inches width and a heavy traffic carpet covers the floor. There are

two plasterboard walls of 2 inches width, one brick wall of 4 inches thickness and one concrete wall of approximately 9 inches, all of them painted. The wood door with an area about 1.7 m² and 1.5 inches thickness is located on the smallest plasterboard wall. There are six double-glazed windows of 3 mm thickness each one with approximately 4 cm space between them and surface area of about 1 m². One of them is placed on the same wall as the door and the other five are located in the biggest plasterboard wall. The furniture in the room is given by a big table in the middle of the room, one acrylic board of 2.2 m² located on the concrete wall and wood furniture with glass doors situated in a room corner. In Figure 4.3, the shape and dimensions of the room can be seen. Table 4.1 describes the material and area of the surfaces in the room.

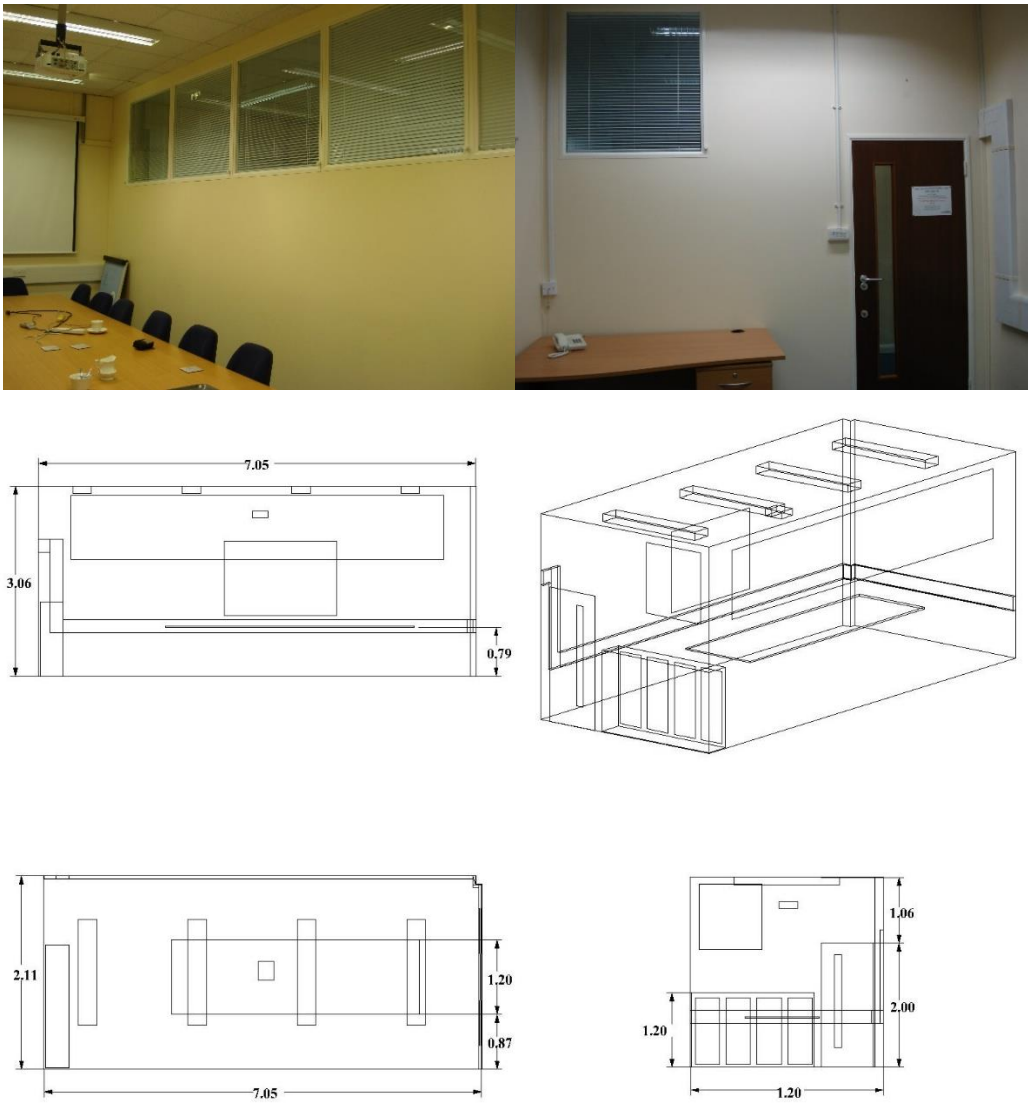


Figure 4.3: On top, photographs of the room. Next, drawings showing top, frontal, lateral and isometric views.

Table 4.1: Material and area of the surfaces in the meeting room.

Surface	Material	Area (m²)
Floor	Carpet	21,18
Door	Wood	1,65
Tables and Wood Furniture	Wood	11,90
Front Wall	Concrete	9,52
Back Wall (door and window)	Plaster	6,87
Left Wall	Brick	21,57
Right Wall (windows)	Plaster	15,57
Ceiling	Plaster	19,93
Lights	Metal	5,40
Door and Furniture windows	Glass	1,90
Windows	Doubled-Glass	7,00

4.2.2 Test report Meeting room measurements

Number and type of seats

There were no seats in the room during the measurements.

State of occupancy during measurements

The room was empty.

Condition of any variable equipment

The only variable equipment was the screen projector, which was up all the time.

Furniture

The furniture in the room consists of a big table, a furniture and one acrylic board of 2.2m² located on the concrete wall. The table and furniture are made of wood with approximate surface areas of 10m² and 4m² respectively. The furniture has small glazed doors.

Temperature and humidity

These variables were not measured and normal conditions of temperature and humidity during the measurements were assumed.

Equipment

- Sound source. 2-way active loudspeaker MACKIE SRM350v2
- Microphone. 1/2" Omni-directional microphone: B&K 4189 Sensitivity 50.5 mV/Pa serial number 2378983. Conditioning Amplifier: B&K 2694 Serial number 2165583.
- Dummy head NEUMANN.
- The Software TOTAL MIX was used to set the input and output levels and the application AEIRM to generate the sound signal and estimate the impulse responses.
- Laptop.

Sound signal used

The sound signal used was a Log Sine Sweep from 20Hz to 20 kHz.

Source and receiver positions

The loudspeaker was placed on the table, which has a height of 0.81m. The microphone height used was 1.2m.

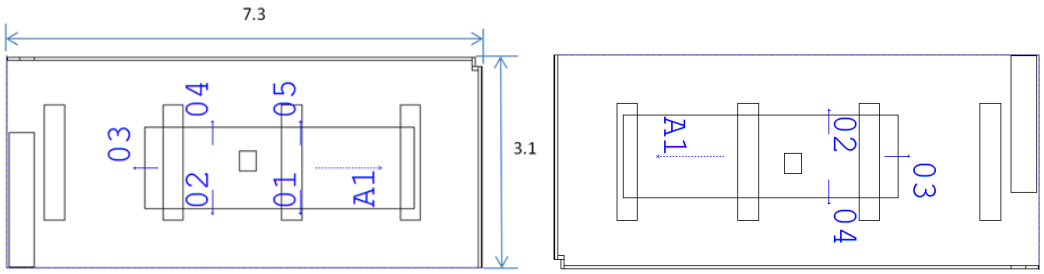


Figure 4.4: Source and receiver positions. Left, Binaural receiver positions and right, monaural receiver positions.

Date and measuring organization

The measurements were realised on July 2012 by the author.

4.2.3 Acoustic measurements results

In Table 4.2 and Figure 4.5 the reverberation time measurement results can be appreciated. The reverberation time estimated is the T_{20}, this concept along with other room acoustic parameters, analysed in order to objectively evaluate the numerical results, are explained in section 5.1.1.

Table 4.2: Reverberation Time in the meeting room and standard deviation for the spatial average, according to ISO 3382 (2009).

Frequency Band (Hz)	T_{20} [s]							
	Source Pos. 1						Spatial Average	Standard deviation
	Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5	Pos. 6		
125	0.50	0.56	0.46	0.53	0.40	0.45	0.48	±0.058
250	0.73	0.79	0.69	0.73	0.63	0.69	0.71	±0.054
500	0.56	0.57	0.55	0.55	0.55	0.58	0.56	±0.013
1000	0.45	0.51	0.43	0.47	0.45	0.46	0.46	±0.027
2000	0.45	0.43	0.44	0.45	0.44	0.45	0.44	±0.008
4000	0.48	0.49	0.48	0.51	0.50	0.49	0.49	±0.012

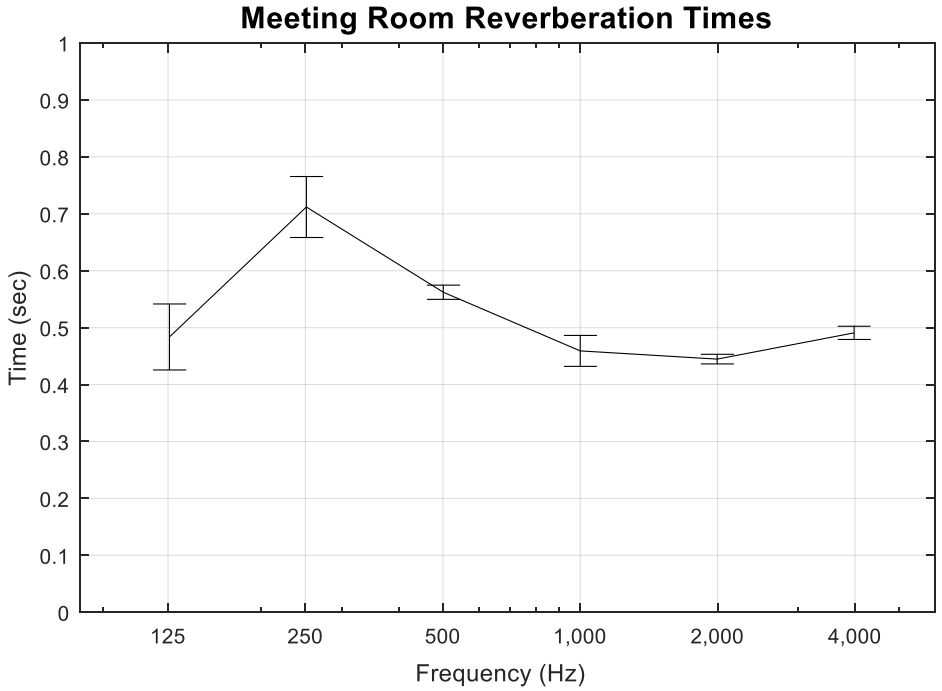


Figure 4.5: Meeting room Reverberation times estimated by means of acoustic measurements (T_{20}).

4.3 Acoustic measurements of the Classroom

In this room, three types of measurements were carried out RIR, BIR and background noise levels. In addition, a binaural background noise recording was realised for each binaural receiver position used. The purpose of the first set of measurements was to obtain the room transfer path for all source-receiver combinations, in order to create the reference auralizations. The second set of measurements intended to characterize the room, calculating the most relevant acoustic parameters according to ISO standard 3382 (2009). The intention of the

acoustic measurements was to provide reference values, in order to assess the numerical approaches used in this research to simulate sound propagation.

4.3.1 Classroom description

The selected classroom for this acoustic assessment procedure based on auralizations was the room named “Mini-auditorium 2”, located on the fourth floor of the engineering building at the University of San Buenaventura, in Medellin, Colombia. The room has a capacity of 40 people and a volume of about 135 m³. The enclosure has two painted concrete walls and two walls made of drywall. The ceiling is also made of drywall and the floor is made of tile. In Figure 4.6 the shape and dimensions of the room can be seen. Table 4.3 describes the material and area of the surfaces in the room.

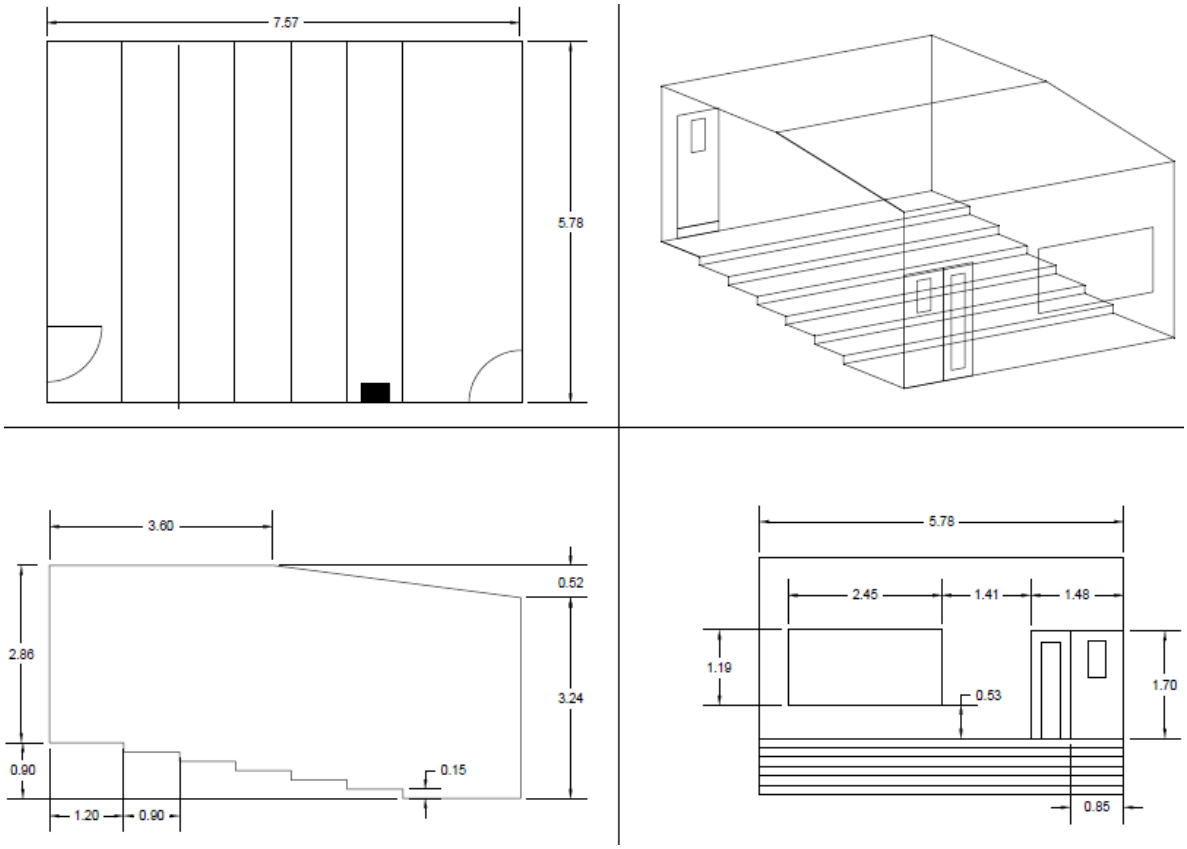


Figure 4.6: Classroom drawings. Top, frontal, lateral and isometric views.

Table 4.3: Area and material of the surfaces found in the classroom.

Surface	Material	Area (m²)
Floor	Tile	49.0
Doors	Wood	4.0
Board	Acrylic	2.9
Panel	Foam	0.5
Lights	Metal	5.0
Left and Back Walls	Plaster	37.0
Right and Front Walls	Concrete	37.2
Ceiling	Plaster	39.6
Windows (Doors)	Glass	0.8

4.3.2 Test report classroom measurements

Number and type of seats

There were no seats in the room during the measurements.

State of occupancy during measurements

The classroom was empty.

Condition of any variable equipment

There was no variable equipment in the classroom.

Sketch plan of the room

Sketch plan of the classroom including source and receiver positions used for BIR measurements.

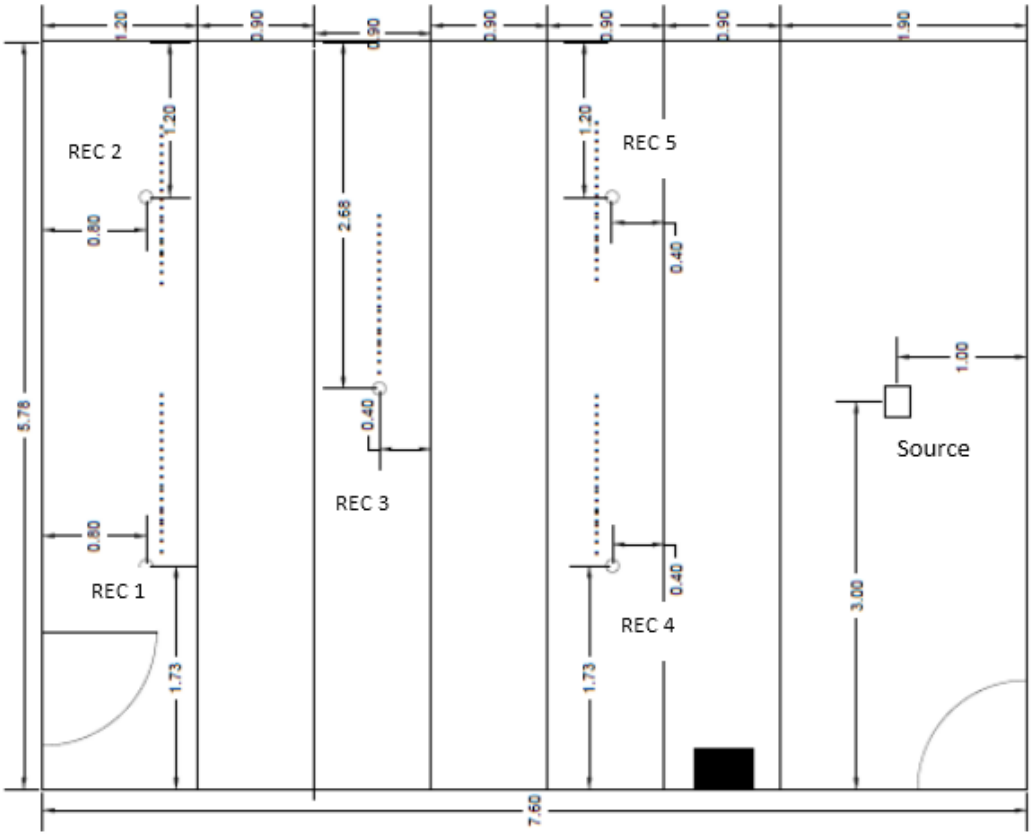


Figure 4.7: Sketch plan of the classroom, including source and receiver positions for BIR measurements.

Furniture

The only furniture present in the classroom during the measurements was a projector.

Temperature and humidity

These variables were not measured and normal conditions of temperature and humidity during the measurements were assumed.

Equipment

- Sound sources. For the BIR, the 2-way active loudspeaker JBL EON15 G2 and for the RIR, the dodecahedron 01dB OMNI12.
- Microphone. 1/2" Omni-directional microphone DBX.
- Dummy head of reference Cortex MK2B from the manufacturer 01dB.
- Sound level meter Cesva SC310sb, type I.
- Audio interface M-Audio MobilePre.
- Laptop.

Sound signal used

The sound signal used was a Log Sine Sweep from 20Hz to 20 kHz.

Source and receiver positions

The loudspeaker was placed at a height of 1.5m. The microphone height used was 1.2m. In Figure 4.8 the source and receiver positions for RIR measurements can be seen.

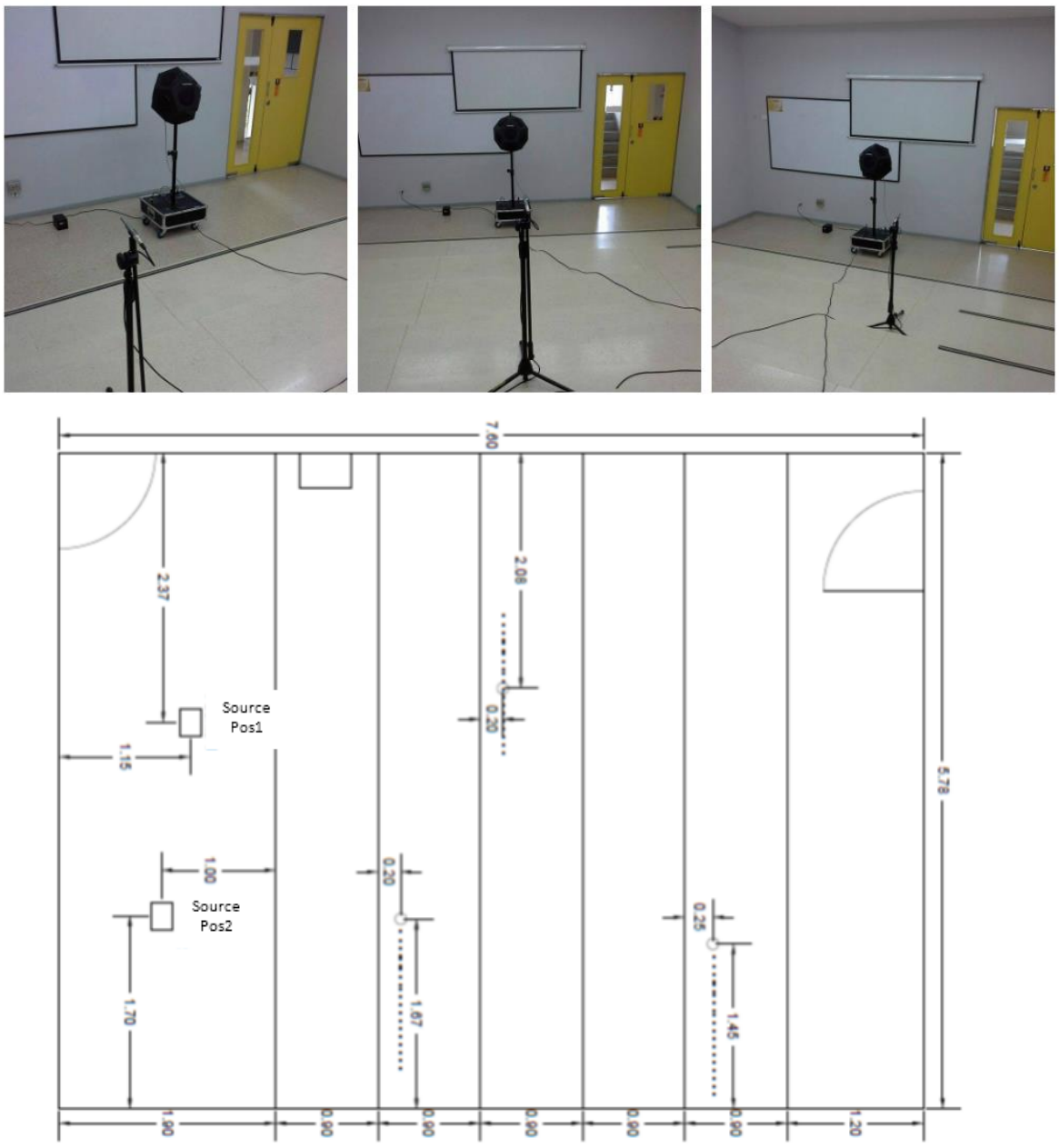


Figure 4.8: Photographs of the classroom showing the equipment used for the measurements and a sketch presenting the source and receiver positions for RIR measurements.

Date and measuring organization

The measurements were taken in April 2013 by Luis Tafur, PhD student at the ISVR and the San Buenaventura University undergraduate students: Daniel Urrego, Juan Camilo Rodríguez Villota and Anderson Naranjo Ruiz.

4.3.3 Acoustic measurements results

In this research, the room was considered "empty", with no furniture or persons inside the classroom during the sound field measurements and the corresponding simulations. In order to characterize the classroom, background noise, RIR and BIR measurements were taken according to ISO standard 3382 with engineering precision (2009). The sound pressure levels were measured for periods of thirty minutes, at three positions randomly distributed in the classroom, estimating the background noise levels as the spatially averaged energy of measurements (see Figure 4.9 and

Table 4.4Table 4.4). In order to compare the measurement results with the recommendations of the Building Bulletin 93 (2004) regarding reverberation time criteria, the measurements were carried out in the absence of furniture or chairs inside the classroom. During the data collection process, two different source locations and three microphone positions distributed randomly in the room were used (see Figure 4.8). The reverberation time estimated for the measurements was the T_{20}, given the Signal-to-Noise Ratio (SNR) obtained for each source-receiver combination (see Table 4.5). According to Table 4.5, the underlined values presented SNR values below 35 dB, which is the requirement to obtain proper T_{20} estimates (see section 5.1.1.1). In Figure 4.10 and Table 4.6, the results of T_{20} measurements by octave bands can be seen.

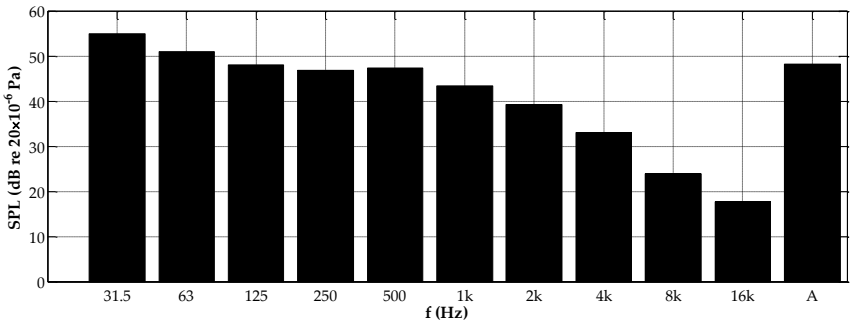


Figure 4.9: Background noise levels measured in the classroom.

Table 4.4: Background noise levels in the classroom (in dB re 20 x 10^{-6} Pa).

	Frequency [Hz]										LA_{eq}
	31.5	63	125	250	500	1k	2k	4k	8k	16k	
Pos. 1	57.2	50.1	50.0	47.7	48.0	42.8	38.7	33.4	24.4	17.7	48.6
Pos. 2	50.5	53.8	48.0	48.0	49.5	46.2	42.1	35.2	25.7	18.3	50.8
Pos. 3	55.7	48.2	47.0	44.2	43.4	40.0	36.0	29.4	20.9	16.7	45.1
Spatial Average	54.6	51.0	48.0	46.8	47.3	43.4	39.3	33.0	23.9	17.6	48.5

Table 4.5: SNR at each source-receiver combination obtained in the acoustic measurements of the classroom (SNR values below 35 dB underlined).

Frequency Band [Hz]	SNR [dB]						
	Source Position 1			Source Position 2			Spatially Averaged
	Rec. Pos. 1	Rec. Pos. 2	Rec. Pos. 3	Rec. Pos. 1	Rec. Pos. 2	Rec. Pos. 3	
31.5	8.70	25.70	24.80	25.60	20.60	21.60	21.17
63	34.10	33.10	23.50	30.20	31.00	27.30	29.87
125	40.00	46.10	37.30	41.90	44.30	42.20	41.97
250	47.40	49.80	43.60	47.90	48.50	47.30	47.42
500	44.20	42.50	40.10	43.20	42.40	43.70	42.68
1000	46.00	43.80	42.60	46.20	44.50	45.50	44.77
2000	40.30	39.10	39.60	42.60	41.10	42.70	40.90
4000	36.30	35.60	36.30	38.90	35.60	38.60	36.88
8000	31.00	29.30	30.20	32.00	27.80	31.10	30.23
16000	15.60	14.40	17.00	17.80	11.40	17.30	15.58

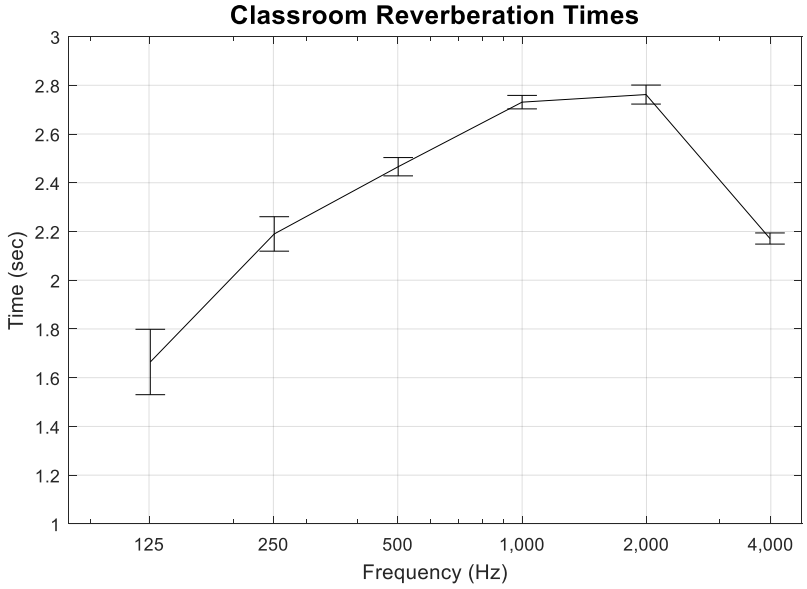


Figure 4.10: Classroom Reverberation times estimated by means of acoustic measurements (T_{20}).

Table 4.6: Reverberation Time in the classroom and standard deviation for the spatial average, according to ISO 3382 (2009).

Frequency Band (Hz)	T_{20} [s]							
	Source Pos. 1			Source Pos. 2			Spatial Average	Standard deviation
	Pos. 1	Pos. 2	Pos. 3	Pos. 1	Pos. 2	Pos. 3		
125	1.56	1.77	1.94	1.76	1.86	2.15	1.84	±0.198
250	2.32	2.19	2.32	2.25	2.34	2.2	2.27	±0.066
500	2.45	2.53	2.38	2.6	2.59	2.5	2.51	±0.084
1000	2.86	2.67	2.68	2.76	2.69	2.73	2.73	±0.071
2000	2.79	2.72	2.76	2.75	2.75	2.65	2.74	±0.048
4000	2.12	2.11	2.14	2.09	2.13	2.1	2.12	±0.019

4.4 Test Report Room Acoustic Parameters Measurements, "Recording Studio"

In this room, two types of measurements were carried out, RIR and background noise levels.

Statement

The procedure used to measure the room acoustic parameters was conformed with the standard ISO 3382 (2009).

Name and Place

The room acoustic parameters were measured in the "Recording Studio A", located on the basement of the engineering building at the University of San Buenaventura, in Medellin, Colombia.

Volume of the room

The volume of the recording studio is 60m³ approximately.

Number and type of seats

There were no seats in the room during the measurements.

State of occupancy during measurements

The room was empty.

Condition of any variable equipment

There was not variable equipment in the room.

Furniture

The only furniture presented in the room during the measurements were two small wood tables.

Temperature and humidity

These variables were not measured and normal conditions of temperature and humidity during the measurements were assumed.

Sketch plan of the room

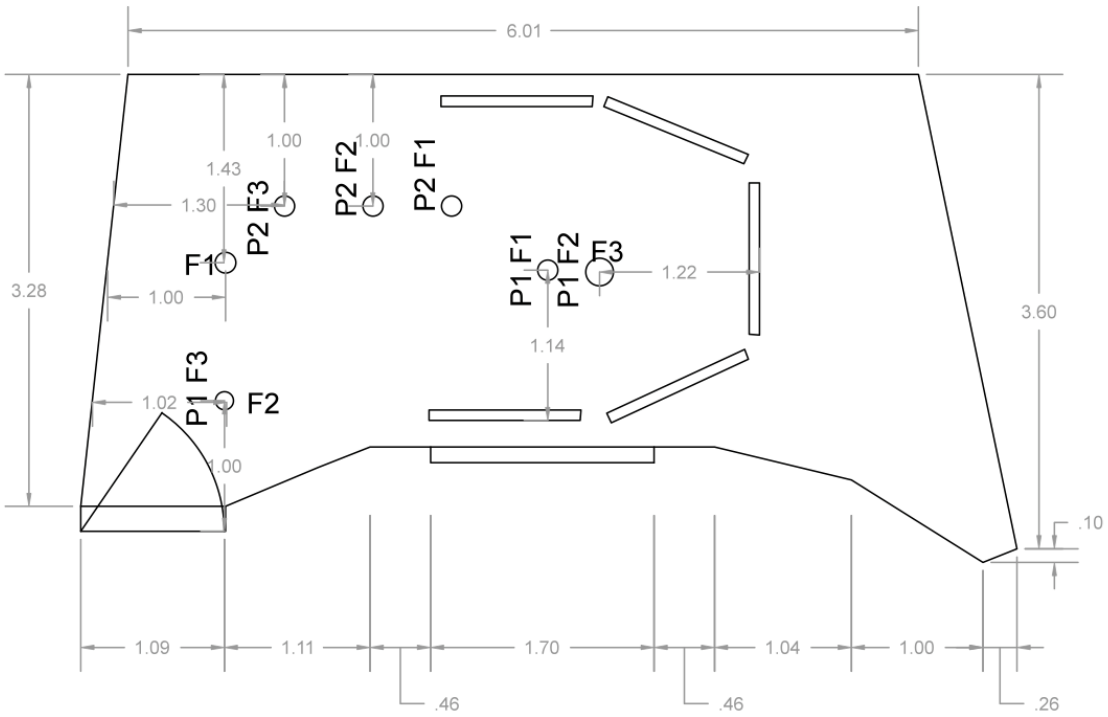


Figure 4.11: Sketch plan of the recording studio, including source positions (denoted as F) and receiver positions for RIR measurements (denoted as P).

Shape and materials of the room

The recording studio has an irregular shape in order to improve the diffusion acoustic characteristics. A description of the surfaces, materials (see Figure 4.12) and corresponding areas can be seen in the following table:

Table 4.7: Material and corresponding areas of the surfaces found in the recording studio.

Surface	Area (m^2)	Material
Floor	20.05	Clapboard
Window	2.89	Glass
Door	0.92	Metal
Absorbent panels	24.75	Fiberglass coated with coral cloth
Diffusers	23.70	Wood
Ventilation grill	0.66	Aluminium
Wood membranes	21.19	Wood
Groove	4.13	Metal
Wall	1.01	Concrete
Table panels	2.49	Wood
Absorption Panels	11.90	Mineral Rock Wool coated with coral cloth

Figure 4.12: Photographs of the recording studio showing the materials of the surfaces and the equipment used for the measurements.

Equipment

- Sound source, dodecahedron from manufacturer 01dB, reference OMNI12.
- Microphone. 1/2" Omni-directional microphone DBX.
- Sound level meter Cesva SC310sb, type I.
- Audio interface Focusrite Scarlett 2i2.
- Laptop.

Sound signal used

The sound signal used was a Sine Sweep from 20Hz to 20 kHz.

Source and receiver positions

The loudspeaker was placed at a height of 1.5m. The microphone height used was 1.2m (see Figure 4.11).

Date and measuring organization

The measurements were taken in September 2014 by Luis Tafur, PhD student at the ISVR and the San Buenaventura University undergraduate students: Jonathan Ochoa and Juan Camilo Rodríguez.

Measurements results

In the following tables the background noise levels and T_{20} measurements results can be seen:

Table 4.8: Background noise levels in the recording studio (in dB re 20 x 10^{-6} Pa).

	Frequency Band [Hz]										LA_{eq}
	16	31.5	63	125	250	500	1000	2000	4000	8000	
Position 1	44.5	46.0	39.4	26.3	25.2	20.3	18.6	20.1	20.8	27.4	28.7
Position 2	41.8	41.6	35.5	28.3	28.5	24.3	19.8	16.0	15.3	14.4	26.5
Position 3	44.4	39.9	36.0	24.4	27.2	23.3	19.5	19.5	15.2	14.2	26.3
Spatial Average	43.7	42.9	37.1	26.5	27.1	22.8	19.3	18.7	17.5	21.0	27.2

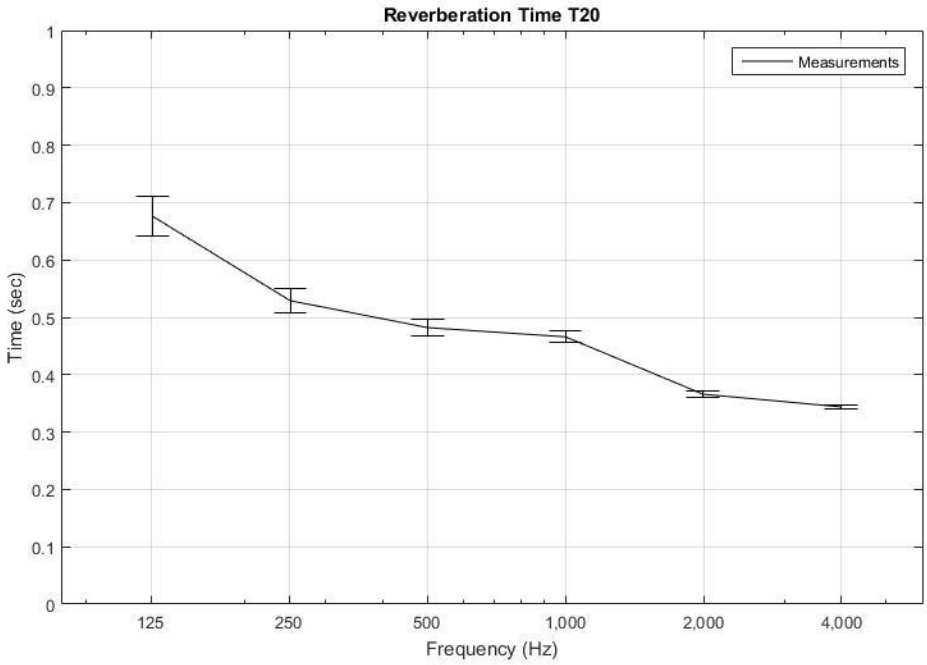


Figure 4.13: Reverberation Times in the recording studio estimated by means of acoustic measurements (T_{20}).

Table 4.9: Reverberation Time in the recording studio and standard deviation for the spatial average, according to ISO 3382.

Frequency Band (Hz)	T_{20} [s]												
	Source Pos. 1			Source Pos. 2			Source Pos. 3			Source Pos. 4			Spatial Average
	Pos. 1	Pos. 2	Pos. 3	Pos. 1	Pos. 2	Pos. 3	Pos. 1	Pos. 2	Pos. 3	Pos. 1	Pos. 2	Pos. 3	
63	0.705	1.077	0.856	0.978	1.010	1.159	0.983	1.487	0.695	0.884	0.838	1.337	1.001 ± 0.059
125	0.561	0.468	0.690	0.745	0.753	0.724	0.547	0.838	0.820	0.736	0.614	0.612	0.676 ± 0.034
250	0.544	0.577	0.513	0.472	0.544	0.518	0.493	0.592	0.456	0.536	0.559	0.548	0.529 ± 0.021
500	0.459	0.445	0.431	0.489	0.585	0.513	0.424	0.529	0.499	0.463	0.459	0.491	0.482 ± 0.014
1000	0.427	0.453	0.506	0.477	0.483	0.441	0.445	0.447	0.466	0.497	0.462	0.486	0.466 ± 0.010
2000	0.374	0.339	0.361	0.410	0.365	0.400	0.337	0.356	0.363	0.375	0.346	0.367	0.366 ± 0.006
4000	0.338	0.321	0.358	0.391	0.361	0.315	0.339	0.354	0.341	0.333	0.362	0.319	0.344 ± 0.004
8000	0.267	0.275	0.267	0.269	0.280	0.282	0.286	0.264	0.269	0.279	0.279	0.286	0.275 ± 0.003
16000	0.267	0.268	0.249	0.223	0.257	0.250	0.333	0.237	0.245	0.242	0.244	0.305	0.260 ± 0.002

5. Sound propagation numerical simulations of the rooms investigated

This chapter details the numerical modelling applied to obtain the RIR and BIR for the Meeting room and the Classroom, including an objective assessment of the results and a subjective evaluation of the meeting room auralizations. The construction of these models considers the methods described in Chapter 3 in which the required variables for GA and FE simulations were explained. The objective and subjective methods used to evaluate the numerical approaches applied in the transmission stage of an auralization are described. The first point is a comparison between simulations and acoustic measurements taking into account time domain responses, frequency responses and acoustic parameters obtained by the application of the ISO standard 3382 (2009). The second point is given by the parameters used to evaluate subjectively the application of the 3D system OPSODIS in the reproduction stage.

5.1 Objective evaluation

The present section describes the objective acoustic parameters taken into account to evaluate the accuracy of the numerical approaches used to estimate the sound wave propagation in a room. In order to have a reference, RIRs and BIR measurements were obtained in the rooms analysed. In this sense, all source-receiver combinations analysed in the rooms have a corresponding measured and simulated RIR. An initial evaluation is given by the basic comparison of time and frequency responses measured and simulated. In this regard, special attention is focused on the low frequency range, taking into account the frequency ranges modelled in the FEM and the limitations of GA reviewed in Chapters 2 and 3. In addition, RIRs and BIR were used to estimate the objective room acoustic parameters according to ISO 3382-1:2009 such as ´Reverberation Time´ (T_x), ´Clarity´ (C), ´Definition´ (D) and Interaural cross correlation.

5.1.1 Room acoustic parameters

This section describes the theoretical basics and the estimation procedures of the main acoustic parameters applied to evaluate or characterize a room. The integrated impulse response method according to ISO 3382 (2009) was the

procedure used to obtain some of the parameters mentioned above. Based on the standard procedures, it was decided to use a standard coverage, with six combinations of source-receiver positions. The RIR were used to estimate the Energy Decay Curve (*EDC*) for each octave band by a backward integration of the squared impulse response, applying the following expression (Mechel, 2009):

$$E(t) = \int_t^{\infty} p^2(\tau) d\tau = \int_{\infty}^{t} p^2(\tau) d(-\tau) \qquad [5.1],$$

where, P is the sound pressure, E is the energy decay curve as a function of time t. To express the energy decay in dB, the *EDC* was estimated using (International Organization for Standardization, 2009):

$$EDC = 10 \log_{10} E \qquad [5.2],$$

5.1.1.1 Reverberation Time *T*

When a diffuse field in an enclosure of volume V is assumed, the energy balance equation can be expressed as the rate of change of enclosure energy equals the power input, minus the rate of energy loss from enclosure (Thompson & Nelson, 2015):

$$\frac{d}{dt}\left[V\langle e(x)\rangle\right] = W - \langle I(x)\rangle S\overline{\alpha} \qquad [5.3],$$

Where, W is the sound power of the acoustic source, S is the surface area, $\overline{\alpha}$ is the average sound absorption coefficient, and $\langle e(x)\rangle$ and $\langle I(x)\rangle$ are the space averaged time values of energy density and intensity, which are defined in expressions [5.4] and [5.5], respectively (Thompson & Nelson, 2015):

$$\langle e(x)\rangle = \frac{\left\langle |p(x)|^2 \right\rangle}{2\rho_0 C_0} \qquad [5.4],$$

$$\langle I(x)\rangle = \frac{\left\langle |p(x)|^2 \right\rangle}{8\rho_0 C_0} \qquad [5.5],$$

with $\langle |P(x)|^2 \rangle$ as the space average modulus square pressure. Replacing expressions [5.4] and [5.5] into [5.3] and assuming that the source providing the input power W is switched off, the energy balance equation is reduced to:

$$-\frac{4V}{S\overline{\alpha}C_0}\frac{d}{dt}\left\langle\left|p(x)\right|^2\right\rangle=\left\langle\left|p(x)\right|^2\right\rangle \qquad \textit{[5.6],}$$

According to Nelson (1998), the solution of this equation is an exponential decay of the form:

$$\left\langle\left|p(x)\right|^2\right\rangle=\left\langle\left|p(x)\right|^2\right\rangle_0 e^{-\left(\frac{S\overline{\alpha}C_0t}{4V}\right)} \qquad \textit{[5.7],}$$

where $\langle|P(x)|^2\rangle_0$ is the initial value of the averaged square pressure, and t, is the reverberation time. Therefore, the time taken for the sound pressure level to decay 60dB is given by:

$$10\log_{10}\frac{\left\langle\left|p(x)\right|^2\right\rangle}{\left\langle\left|p(x)\right|^2\right\rangle_0}=10\log_{10}e^{-\left(\frac{S\overline{\alpha}C_0T_{60}}{4V}\right)}=-60dB \qquad \textit{[5.8].}$$

In order to find the reverberation time T_{60}, expression [5.8] is reduced to:

$$e^{-\left(\frac{S\overline{\alpha}C_0T_{60}}{4V}\right)}=10^{-6} \qquad \textit{[5.9].}$$

applying $L_n$on both sides, equation [5.9] takes the form of (Bies & Hansen, 2009):

$$S\overline{\alpha}=\frac{55.3V}{C_0T_{60}} \qquad \textit{[5.10].}$$

When normal conditions of temperature (~20˚ C) are assumed, the reverberation time equation expressed in equation [3.35] is obtained. Taking into account the air absorption (Kinsler, et al., 2000), expression [5.10] takes the following form (Kuttruff, 2009):

$$T_{60}=\frac{0.161V}{S\overline{\alpha}+4mV} \qquad \textit{[5.11],}$$

where, m is the air absorption constant, which depends on the humidity of air. Sabine´s equation provides suitable results only in rooms with an average absorption coefficient of less than 0.25. For more absorbent rooms, the Eyring-Norris equation may be used (Mechel, 2009):

$$T_{60} = \frac{0.049V}{-S\ln(1-\bar{\alpha})} \qquad [5.12].$$

Furthermore, the ISO standard 3382-2 (2008) establishes the procedures and requirements necessary to measure reverberation time in any type of enclosure. Since to obtain a range of 60 dB decay requires a very large dynamic range, two estimation ranges are defined in order to extrapolate a 60 dB decay curve. Hence, reverberation time estimation is given by the evaluation of the *EDC* in the ranges of -5dB to -25dB and -5dB to -35dB, in order to obtain the reverberation times T_{20} and T_{30} respectively. Taking into account that a 10 dB difference between the acoustic source signal and the background noise level is required in order to have an *EDC* not influenced by the latter, a signal-to-noise ratio of at least 35 dB is required to estimate T_{20}, and 45 dB, for the corresponding T_{30}.

5.1.1.2 Early Decay Time (*EDT*)

The *EDT* is derived from the slope of the *EDC* between 0dB and -10dB, below the initial level. The *EDT* should be calculated as the time required for 60 dB decay. According to the ISO standard 3382, *EDT* is the acoustic parameter that exhibits the best relation with the subjective perception of reverberation in a room.

5.1.1.3 Musical clarity C_{80}

The musical clarity C_{80} indicates the degree of separation between the different individual sound components of a musical composition. The C_{80} is defined as the ratio of the sound energy that reaches the listener during the first 80ms from the arrival of direct sound and the sound energy after the first 80ms, estimated from filtered RIRs for the frequency bands between from 125Hz to 4 kHz. It is expressed on a logarithmic scale and can be obtained from the following equation (International Organization for Standardization, 2009):

$$C_{80} = 10\log\frac{\int_0^{t_e} p^2(t)dt}{\int_{t_e}^{\infty} p^2(t)dt}\ dB \qquad [5.13],$$

where, *C* is the rate of early to late time, t_e is the early time limit of 80ms (C_{80}) and *p* is the sound pressure as function of time.

5.1.1.4 Definition D_{50}

D_{50} refers to the definition of speech intelligibility. It is the ratio of early to the total sound energy, where 50 indicates the first 50 ms. It is estimated from RIRs and can be obtained from the following equation (International Organization for Standardization, 2009):

$$D_{50} = \frac{\int_0^{0.05} p^2(t)dt}{\int_0^{\infty} p^2(t)dt} \qquad [5.14].$$

5.1.1.5 Interaural cross correlation (*IACC*)

Interaural cross correlation (*IACC*) is defined as the correlation between sounds reaching both ears at the same time, and is indicative of the degree of similarity between both signals. To estimate this acoustic indicator, it is necessary to have left and right signals from a BIR. If those sounds were equal, the IACC would be one (1), whereas if they were independent random signals, the *IACC* would tend to zero (0). The *IACC* coefficients are estimated applying the next equation (International Organization for Standardization, 2009):

$$IACC_{t1,t2} = \max\left|IACF_{t1,t2}\right| \; for -1ms < \tau < +1ms \qquad [5.15],$$

where *IACF* is the normalized interaural cross correlation function, defined in ISO standard 3382 (2009) as:

$$IACF_{t1,t2}(\tau) = \frac{\int_{t1}^{t2} p_l(t).p_r(t+\tau)dt}{\sqrt{\int_{t1}^{t2} {p_l}^2(t)dt \int_{t1}^{t2} {p_r}^2(t)dt}} \qquad [5.16],$$

where $P_l(t)$ and $P_r(t)$ are the impulse responses at the entrance of left and right ears respectively.

5.2 Numerical Simulations of the Meeting Room

The Meeting Room 2011 located in the TIZARD building of the University of Southampton was chosen as the first case of study. This room was selected for two reasons: the first is the simplicity of its geometry, which facilitates the 3D model construction for both numerical methods applied. The second reason was

its size. The idea of analysing a small room highlights the disadvantages of GA methods and the need of using a wave numerical method to predict sound wave propagation at low and mid frequencies. This section describes the analysis of the sound wave propagation estimations obtained by means of FE and GA numerical methods, the objective and subjective evaluation of the techniques implemented to create auralizations and a brief discussion of the results.

5.2.1 GA simulations of the Meeting room

In this research, the GA methods were implemented using CATT-Acoustic version 9. In order to create virtual sound environments, the hybrid technique "Randomized Tail-corrected Cone-tracing" (RTC), which includes ISM for low-order reflections and RT method for high-order reflections, has been applied. This method required the construction of a 3D model including the geometry, the source with its corresponding acoustic characteristics and the receiver positions. In the geometry, each planar face was assigned with its material acoustic properties of absorption and scattering. The source had included its acoustic intensity and directionality. Finally, the receiver required a definition of head direction with the purpose of applying the appropriate HRTF.

Table 5.1 describes the material, area of the surfaces in the room and the absorption coefficients, taken from CATT-Acoustic software library, used to estimate an analytical reverberation time according to the Eyring-Norris expression (see equation [5.13]). The reverberation times estimated by means of Eyring-Norris model can be seen in Figure 5.1.

Table 5.1: Absorption coefficients used on each surface of the meeting room used in both, Eyring-Norris and GA models.

Surface	Material	Area (m²)	Octave Band Centre Frequency (Hz)					
			125	250	500	1000	2000	4000
			Absorption Coefficients					
Floor	Carpet	21.18	0.08	0.24	0.57	0.69	0.71	0.73
Door	Wood	1.65	0.14	0.10	0.06	0.08	0.10	0.10
Tables and Wood Furniture	Wood	11.90	0.14	0.10	0.06	0.08	0.10	0.10
Front Wall	Concrete	9.52	0.01	0.01	0.01	0.02	0.02	0.02
Back Wall (door and window)	Plaster	6.87	0.15	0.10	0.06	0.04	0.04	0.05
Left Wall	Brick	21.57	0.02	0.02	0.03	0.03	0.04	0.05
Right Wall (windows)	Plaster	15.57	0.15	0.10	0.06	0.04	0.04	0.05
Ceiling	Plaster	19.93	0.20	0.15	0.10	0.08	0.02	0.01
Lights	Metal	5.40	0.02	0.02	0.03	0.03	0.04	0.05
Door and Furniture windows	Glass	1.90	0.35	0.25	0.18	0.12	0.07	0.04
Windows	Doubled-Glass	7.00	0.15	0.05	0.03	0.02	0.02	0.02

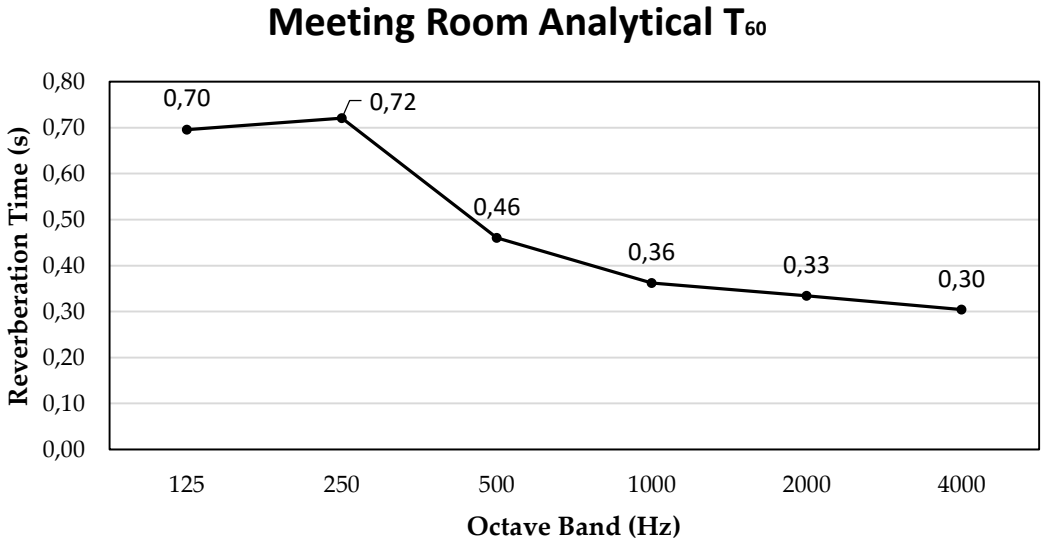


Figure 5.1: Reverberation times of the Meeting room estimated by means of analytical Eyring-Norris model.

5.2.1.1 The 3D model

As mentioned in Chapter 3, the 3D model was constructed in CAD language first, in order to import into each software (GA and FE), and hence, to guarantee

similar geometric conditions. After the importation process, a *GEO-file* was created by the CATT application. In this case, the meeting room was created using 313 corners and 229 planes. The windows, door and board were included in the *GEO-file* as subdivisions, as recommended by CATT-Acoustic user´s guide in order to create a "closed model", which increases the accuracy of sound wave propagation predictions. This is achieved by keeping the sound rays inside the model, which is guaranteed when there are no duplication of surfaces. The final 3D model created in the CATT software can be seen in Figure 5.2.

5.2.1.2 Absorption and scattering coefficients

The next step consisted of defining material acoustic properties for absorption and scattering. The absorption coefficients used on each surface of the CATT model were taken from the library provided by the software (see Table 5.1). The auralization system proposed in this project attempts to create a virtual sound environment from scratch, with information of acoustic properties available in libraries. Regarding acoustic scattering, all of the flat surfaces were assigned with a minimum scattering coefficient of 10%. For standalone objects such as table, furniture, projector and lights, a frequency dependent scattering coefficient was applied using the *Automatic edge diffusion* option provided by the software. This function applies significant diffusion to a surface if its size is small compared to the sound wavelength (CATT, 2007). At this point, it was possible to check the volume of the room and the analytical reverberation times provided by the software, in order to have an idea of the accuracy of the model in terms of shape and acoustic energy absorption/diffusion distribution. Figure 5.3 illustrates the global Eyring reverberation time estimates obtained with CATT-Acoustic, in comparison to the analytical estimates and measured values found previously (see 4.2.3).

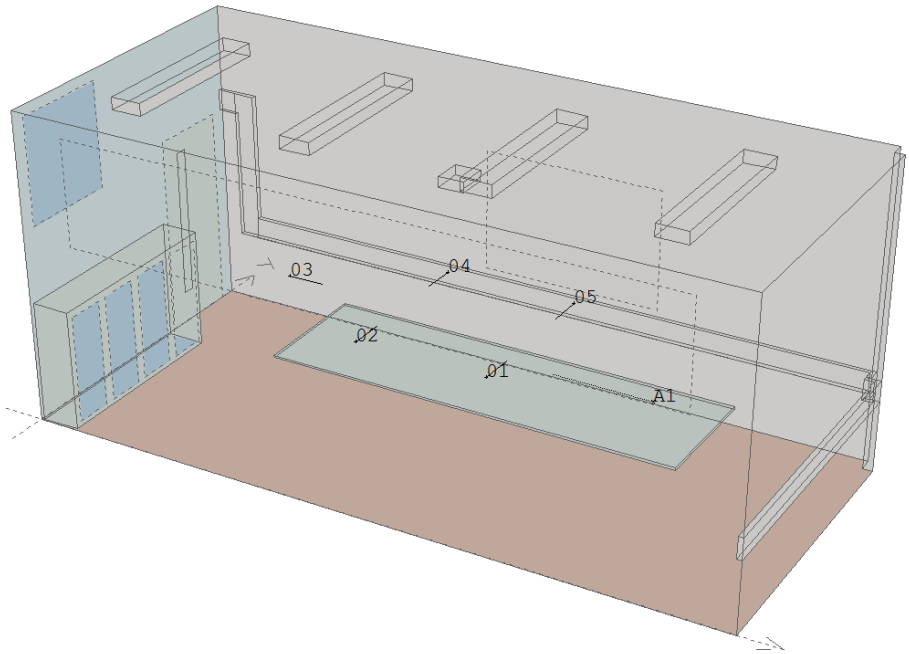


Figure 5.2: GA model created in CATT-Acoustics after importing 3D model in CAD language. A1 denotes the acoustic source location and the consecutive numbers from 01 to 05, the receiver positions simulated.

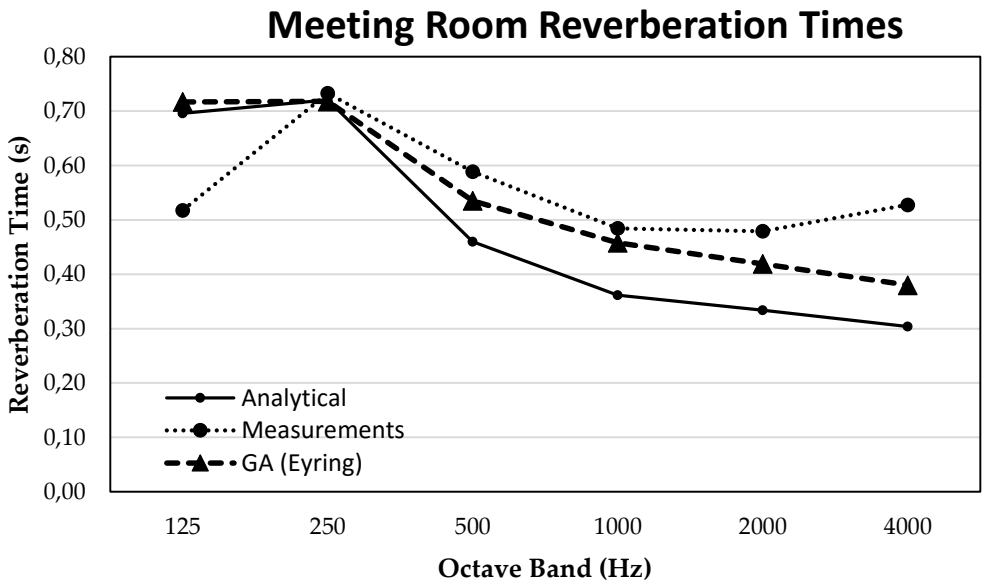


Figure 5.3: Meeting room Reverberation times estimated by means of acoustic measurements (T_{20}), Eyring-Norris equation and GA analytical Eyring model provided by software CATT-Acoustic.

5.2.1.3 Source and receivers in Meeting room GA simulations

In order to model a sound source in GA simulations, the following information was required: acoustic centre location, source orientation, sound level pressures defined at 1 m distance for all octave band frequencies and directivity data. For

the first two, the corresponding Cartesian coordinates were introduced matching the same location and source orientation used in the acoustic measurements. As a source signal, white noise at 94 dB was used in order to facilitate the implementation of a similar acoustic source in a frequency domain wave equation numerical method as FE. The directivity information provided by the manufacturer was introduced using the source directivity module available in CATT software. In this case, a *SD0* format was implemented interpolating horizontal and vertical polar measurements every 15˚. In Figure 5.4, directivity patterns plots of the sound source modelled in CATT are shown.

Five receiver positions around the table were considered in this study (see Figure 5.2). To obtain the RIR at those positions, just the Cartesian coordinate definition was necessary. On the other hand, for the BIR, orientation information had to be defined, in order to apply the procedure indicated in Chapter 3 to find left and right responses.

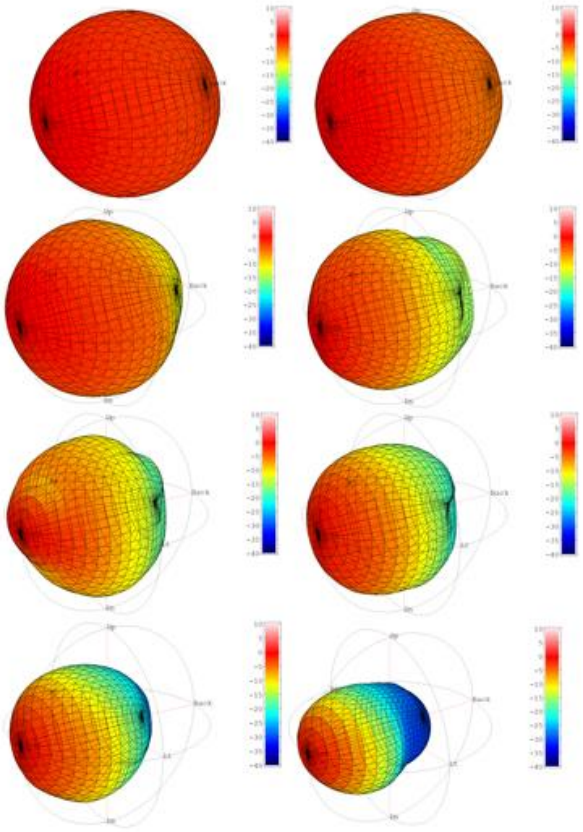


Figure 5.4: Directivity patterns plots for octave bands from 125 Hz to 16 kHz of the MACKIE loudspeaker modelled in CATT.

5.2.2 FE simulations

In this study, FE simulations were implemented using the software COMSOL 4.3. In this case, a time-harmonic simulation up to 700Hz with 1Hz frequency steps (see Chapter 3) was applied for the same source-receiver combinations applied in GA simulations and acoustic measurements. In order to have simulation results to combine with the ones obtained by the GA method, the same parameters defined in the CATT software were used in the FE model. The creation of this model required the construction of a 3D geometry with its corresponding meshing process, the specification of boundary conditions, the definition of a monopole source and the characterization of a binaural receiver.

5.2.2.1 The Geometry and generation of the mesh

In order to generate the geometry, the original CAD model was imported into COMSOL software. It is important to note that geometry construction in FE operates with the same structure as CAD, hence, there was no need to execute an additional procedure. In Figure 5.5, the 3D model created in the COMSOL software can be seen.

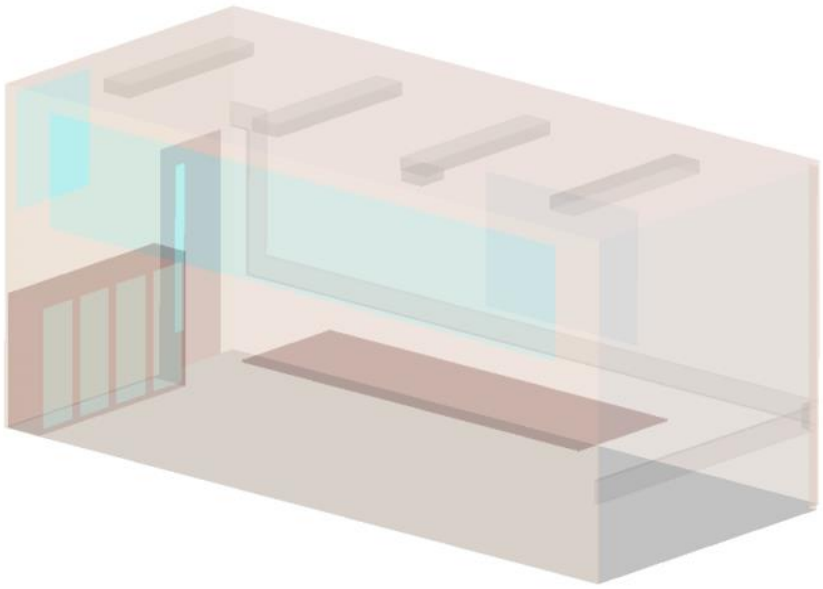


Figure 5.5: FE model created in COMSOL after importing the 3D model in CAD language.

In the generation of the mesh, the number of DOF to be solved in the model was defined according to the room volume and wavelength of the frequency analysed, as stated in the expression [3.12]. Hence, according to [3.12], to estimate a frequency of 700 Hz in this model, a system with around 1,000,000

DOF had to be solved. In a normal computer, the computation of such a number of DOF could take an unreasonable time for just one frequency. For this reason, groups of frequencies with different mesh resolutions divided the simulations (see Figure 5.6), varying the maximum element size according to the maximum frequency estimated, as is shown in Table 5.2. Another parameter defined in the model was the algebraic linear system, which solved the matrix equation resulting from the spatial discretization. For this model the MUMPS solver was applied, which was a method capable of dealing with symmetric and non-symmetric matrices. The application of this procedure allowed running FE simulations on a desktop PC.

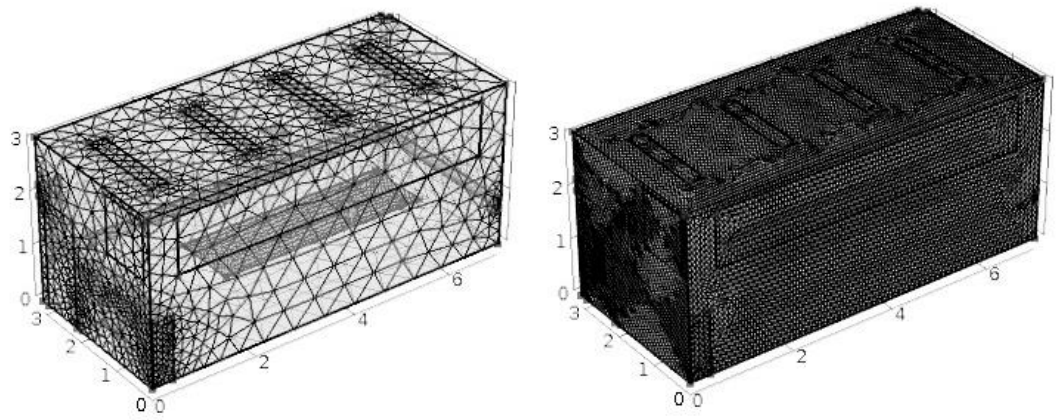


Figure 5.6: The coarsest and the finest mesh resolutions implemented in COMSOL.

5.2.2.2 FE Boundary Conditions

As mentioned in Chapter 3, in FE room acoustic simulations a simple approach relating acoustic impedance with absorption coefficients seems to be the most practical way to model boundary conditions. In this research, the approach given by Aretz (2009) was used (see equation [3.25]), stating that impedance boundary conditions can be defined using a field incidence absorption coefficient to find the resistance part of impedance, which is associated with energy loss by either dissipation or transmission. Taking the above into account, real and frequency dependent impedance values were approximated using the field incidence absorption coefficients applied in the GA model, as can be seen in Table 5.3.

Table 5.2: Maximum frequency estimated, maximum element size, DOF, average time estimation per frequency and approximate number of points per wavelength for each simulation ran in COMSOL.

Max. Frequency Estimated (Hz)	Max. Element Size (m)	DOF	Average time estimation per frequency (min)	Number of points per wavelength
50	0.686	122.728	0.3	10.0
100	0.343	207.869	2	10.0
150	0.229	326.760	4	10.0
200	0.172	486.905	9	10.0
250	0.140	558.951	12	9.8
300	0.122	802.695	18	9.4
350				8.8
375				8.2
400				7.7
425				7.2
450				6.8
475				6.5
500				6.2
525	0.111	922.648	26	5.9
550				5.6
575				5.4
600				5.1
625				4.9
650				4.7
675				4.6
700				4.4

Table 5.3: Real acoustic impedance estimated from absorption coefficients used in GA model, implemented in COMSOL.

Surface	Material	Area (m²)	Octave Band Centre Frequency (Hz)			
			125	250	500	1000
			Real acoustic impedance (Pa.s/m)			
Floor	Carpet	21.18	20250.4	6095.9	2003.2	1458.7
Door	Wood	1.65	10985.2	15707.2	26758.6	19923.2
Tables and Wood Furniture	Wood	11.90	10985.2	15707.2	26758.6	19923.2
Front Wall	Concrete	9.52	165180.9	165180.9	165180.9	82256.9
Back Wall (door and window)	Plaster	6.87	10191.1	15694.0	26758.6	40627.2
Left Wall	Brick	21.57	82173.8	82173.8	54578.0	54504.1
Right Wall (windows)	Plaster	15.57	10191.1	15694.0	26758.6	40627.2
Ceiling	Plaster	19.93	7430.8	10191.1	6699.8	4669.7
Lights	Metal	5.40	82173.8	82173.8	54578.0	54504.1
Door and Furniture windows	Glass	1.90	3857.7	5759.3	8357.9	12977.2
Windows	Doubled-Glass	7.00	10161.7	31844.1	54356.9	82090.9

5.2.2.3 Source and receivers in the FE approach

Source and receivers were simulated in FE applying the same acoustic conditions and positions used in the GA model. The procedure to model an omnidirectional source in FE simulations was explained in Chapter 3. The source was characterised as a monopole point radiating uniformly 1Pa of acoustic pressure at 1m distance, in a frequency independent spherical propagation. In this sense is important to take into consideration that FE simulations go up to 700 Hz, being useful up to approximately 600 Hz according to the frequency response of the band-pass filter explained in Chapter 3. This implies that differences between GA and FE source directivities can be appreciated only at the octave bands of 125, 250 and 500 Hz. It is important to take into consideration that directivity information used in the GA simulations was given according to the loudspeaker used in the measurements, taking the values from the datasheet provided by the manufacturer. The dissimilarities in terms of DI (see equation [3.37]) can be appreciated in Table 5.4. In order to estimate the RIR, the same five receiver positions around the table were defined. To obtain the corresponding BIR, receivers were determined by following the procedure described in Chapter 3. In this case, a cube and two receiver points at corresponding ear positions simulated the HRTF.

Table 5.4: DI applied in the GA and FE source simulations for the octave bands of 125, 250 and 500 Hz.

Numerical approach	Directivity Index DI (dB) Frequency Octave Bands		
	125 Hz	250 Hz	500 Hz
FE	0	0	0
GA*	2	3.1	5

*Obtained from manufacturer datasheet, according to loudspeaker used in the measurements.

Regarding the implications of DI differences (see Table 5.4) between acoustic sources in the numerical simulations implemented, the results of a workshop on room acoustics comparative measurements have been taken into consideration (Adrian James Acoustics Limited, 2004). In this exercise, a number of room acoustic measurements were carried out in a large auditorium and a small room using a range of different measurement systems, sound sources and microphones. The sound sources comparison included Omni-directional

loudspeakers, directional loudspeakers and impulsive sources. The data analysed from those measurements was given by the Reverberation Time (T_{20} and T_{30}), the EDT and the D_{50}. The conclusions of the measurement exercise in the small room, which size is comparable to the Meeting Room investigated in this thesis, indicate that the effect of changing the directionality of the source did not affect significantly the measured reverberation time, neither the EDT. The D_{50} values were not considered in the small room for this exercise.

5.2.3 Meeting Room objective results

In this section, objective results are presented comparing the numerical approaches used to create auralizations. For all cases, RIR and BIR measurements results were taken as the reference or ideal condition. In the objective assessment, a comparison between measurements and simulations considered time and frequency responses and room acoustic parameters estimated for both conditions. The objective evaluation of the simulations includes a comparison with reference-measured results of frequency and time responses, and room acoustic parameter estimations. First, the room transfer function responses are presented for particular source-receiver combinations. Second, the results of room acoustic parameters according to ISO 3382-1:2009, such as Reverberation Time, Early Decay Time (*EDT*), Clarity, Definition and Inter-Aural Cross Correlation (*IACC*) are described. This section finishes with the results of natural frequencies of the room calculated analytically and numerically in order to underline the capability of a wave equation method.

5.2.3.1 Time domain room transfer function results of the Meeting Room

This section presents measurements and simulation results of RIR and BIR for a particular source-receiver combination. In order to facilitate the comparison of the numerical approaches implemented, the impulse responses were filtered in two different frequency ranges. A frequency range from 80 Hz to 600 Hz is used to visualize the impulse responses obtained with the FE method. In order to see the impulse responses simulated by both numerical approaches a wide frequency range from 80 Hz to 20 kHz was used. In order to illustrate the RIR results, three of the six source-receiver position combinations are described.

The following figures (from Figure 5.7 to Figure 5.12) present the RIRs obtained for receiver positions number two, three and four (see Figure 5.2).

In Figure 5.7, it can be seen that the time of arrival of direct sound in the impulse responses coincide largely with the highest value in magnitude in the modelling with FEM in sample 384 (0.0087 seconds) and in the measured RIR in sample 349 (0.0079 seconds). In the simulated FEM RIR, at sample 653 (0.0148 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 6.1 ms. As for the measured RIR, the peak which could correspond to the first reflection is in sample 627 (0.0142 seconds), which would mean an ITDG of 6.3 ms. In both cases the peaks do not match in magnitude. Likewise, a peak with greater level is observed in the RIR obtained by FEM that even exceeds the peak of direct sound, which does not occur in the measured RIR.

In Figure 5.8, the measured RIR peak with the greatest magnitude is in sample 294 (0.0067 seconds) while in the RIRs modelled by GA and FEM-GA same peak is located in sample 325 (0.0074 seconds). A second peak is presented, which could correspond to the first reflection in sample 385 (0.0087 seconds) in the measured RIR, whereby an ITDG equal to 2.0 ms would be obtained. Similarly, in the RIRs obtained by GA and FEM-GA this peak is in sample 430 (0.0098 seconds) that would give a result of ITDG of 2.4 ms. The peaks corresponding to the first reflection are quite similar in magnitude in all RIRs. The following early reflections are similar in magnitude, except from the peaks around sample 1018 (0.0231 seconds) in which the magnitude is greater in the simulated RIRs with respect to the measured case. In the case of the RIR obtained by FEM-GA, the peak of the early reflections both in magnitude and time as well as the reverberant tail does not differ significantly from the GA RIR.

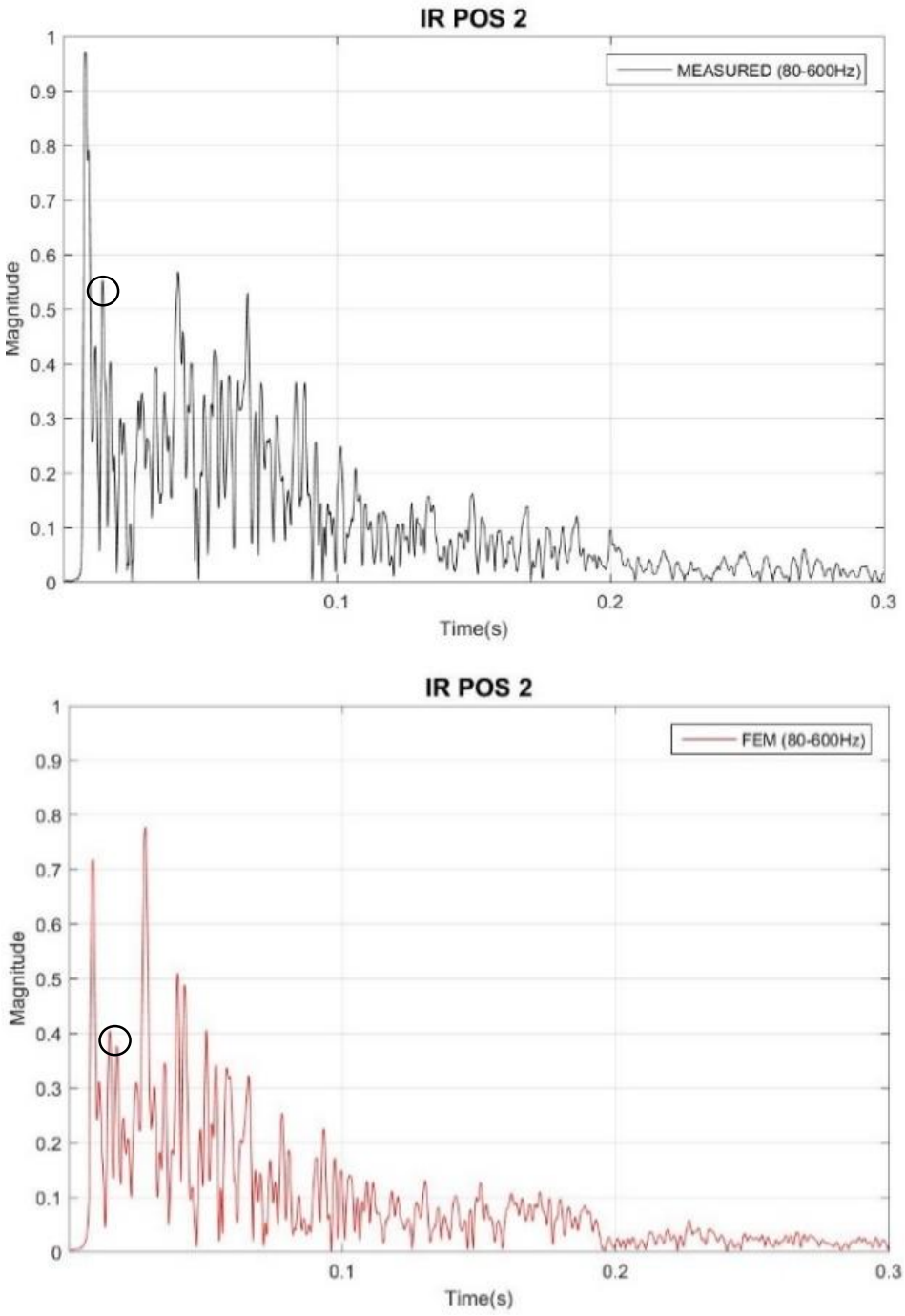


Figure 5.7: Top, the measured RIR at position 2. Bottom, the estimated FEM RIR for the same position.

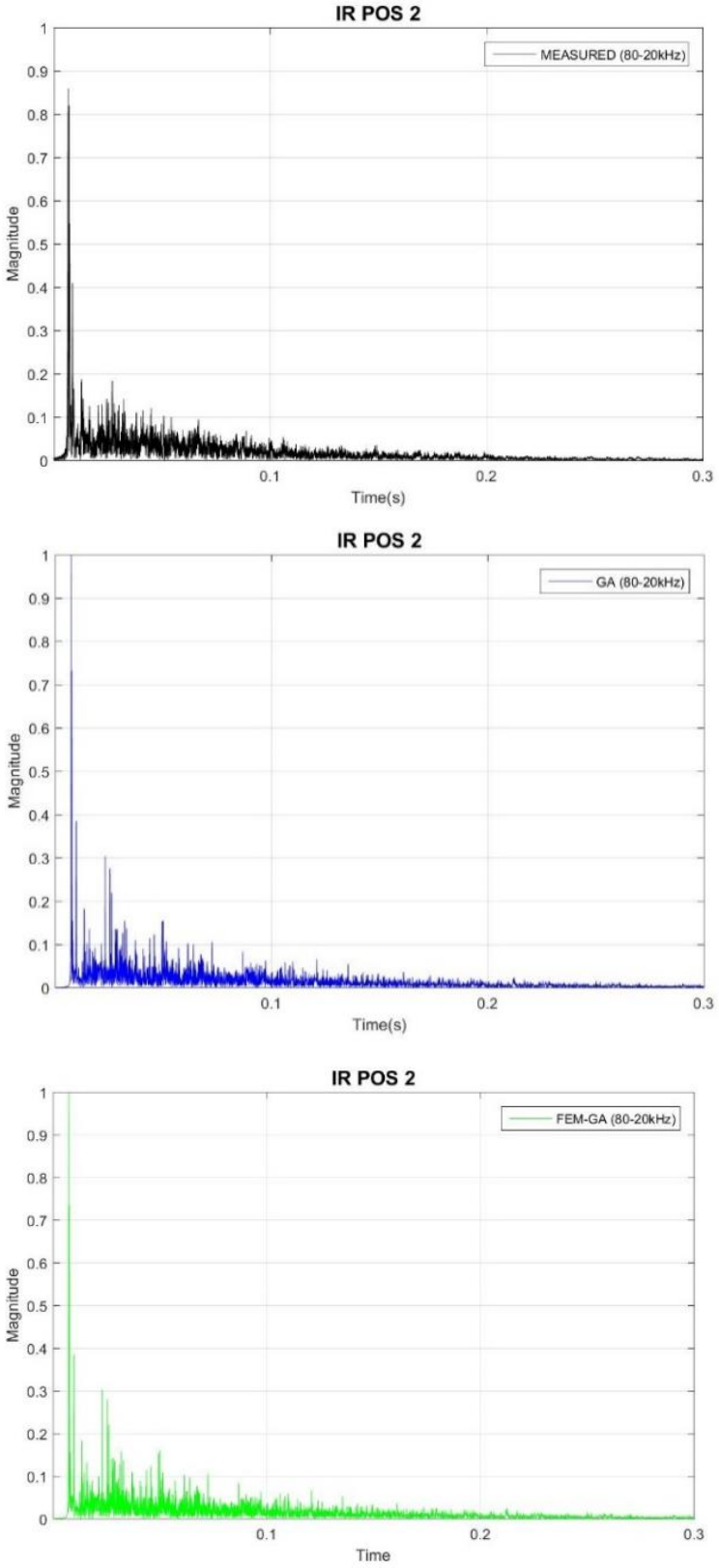


Figure 5.8: RIRs obtained at position 2. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

In Figure 5.9, the time of arrival of the direct sound in both RIRs is the same, but the magnitudes of the first peak in the RIRs do not match, being in the measurement around 0.9 and in the FEM case by approximately 0.6. The peak corresponding to the first reflection coincides in time but not in magnitude. By determining the ITDG in both responses, it is obtained that in the measurement the time corresponds to 6.3 ms, while in the FEM case it is 6.1 ms. The following peaks of the early reflections are seen with great similarity across both RIRs, but with a larger amplitude in the measured case than that of the FEM case.

In Figure 5.10, the direct sound of all the responses arrives at the same instant of time, approximately at sample 472 (0.0107 s). The ITDG in the measured RIR is 2.9 ms, while in the numerical RIRs is 2.6 ms. Some peaks of the early reflections coincide in time, but not in magnitude, with the RIRs simulated by both approaches having greater amplitude. As in the previous position (see Figure 5.8), the RIR simulated by FEM-GA does not present a significant difference compared to the one obtained by GA, in terms of magnitude and arrival time of the early reflections, or the reverberation tail.

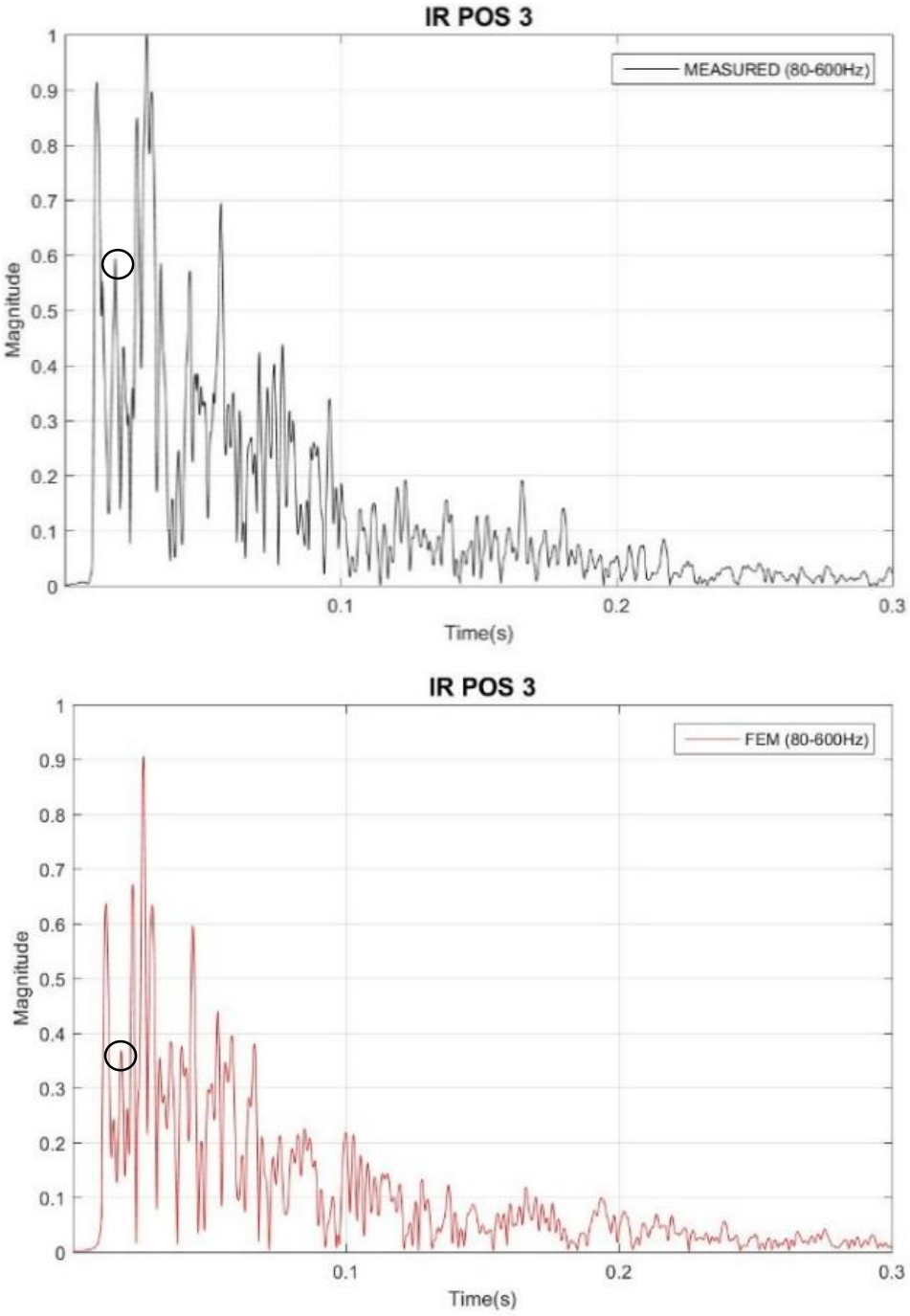


Figure 5.9: Top, the measured RIR at position 3. Bottom, the estimated FEM RIR for the same position.

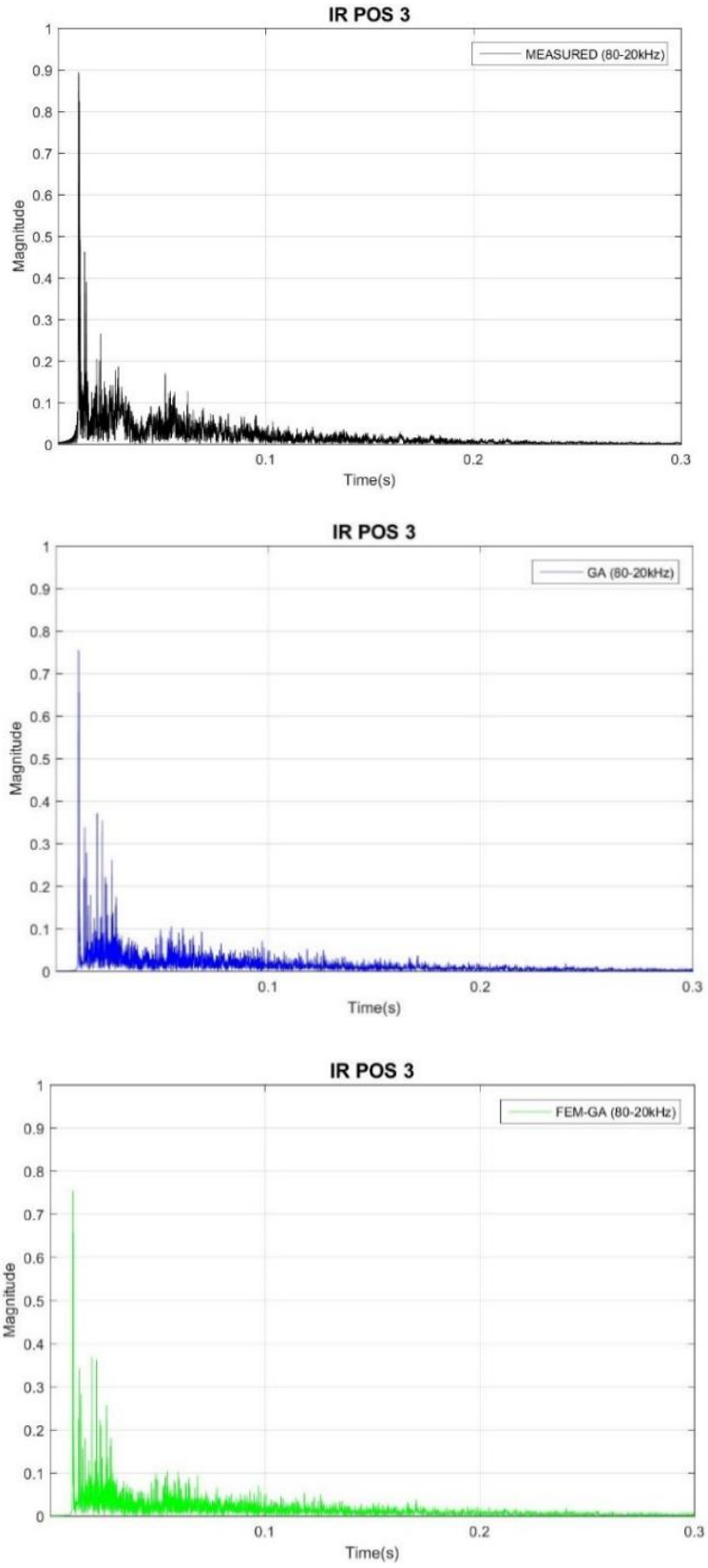


Figure 5.10: RIRs obtained at position 3. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

In Figure 5.11, it can be seen how the time of arrival of direct sound in the RIRs coincide largely, however, the magnitude of this first peak do not match, being in the measurement over 0.9 and in the FEM case close to 0.6. In the FEM RIR, at sample 629 (0.0143 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 5.2 ms. As for the measured RIR, the peak which could correspond to the first reflection is in sample 627 (0.0142 seconds), which would mean an ITDG of 6.0 ms. In both cases the peaks do not match in magnitude.

In Figure 5.12 the measured RIR peak with the greatest magnitude is in sample 309 (0.007 seconds) while in the RIRs modelled by GA and FEM-GA same peak is located in sample 342 (0.0078 seconds). A second peak is presented in sample 425 (0.0096 seconds), which could correspond to the first reflection in the measured RIR, whereby an ITDG equal to 1.8 ms would be obtained. Similarly, in the RIRs obtained by GA and FEM-GA this peak is in sample 469 (0.0106 seconds) that would give an ITDG of 2.8 ms. The peaks corresponding to the first reflection are quite similar in magnitude in all RIRs. The simulated FEM-GA RIR does not present a notorious difference compared to the one obtained by GA, in terms of magnitude and arrival time of the early reflections, neither the reverberation tail.

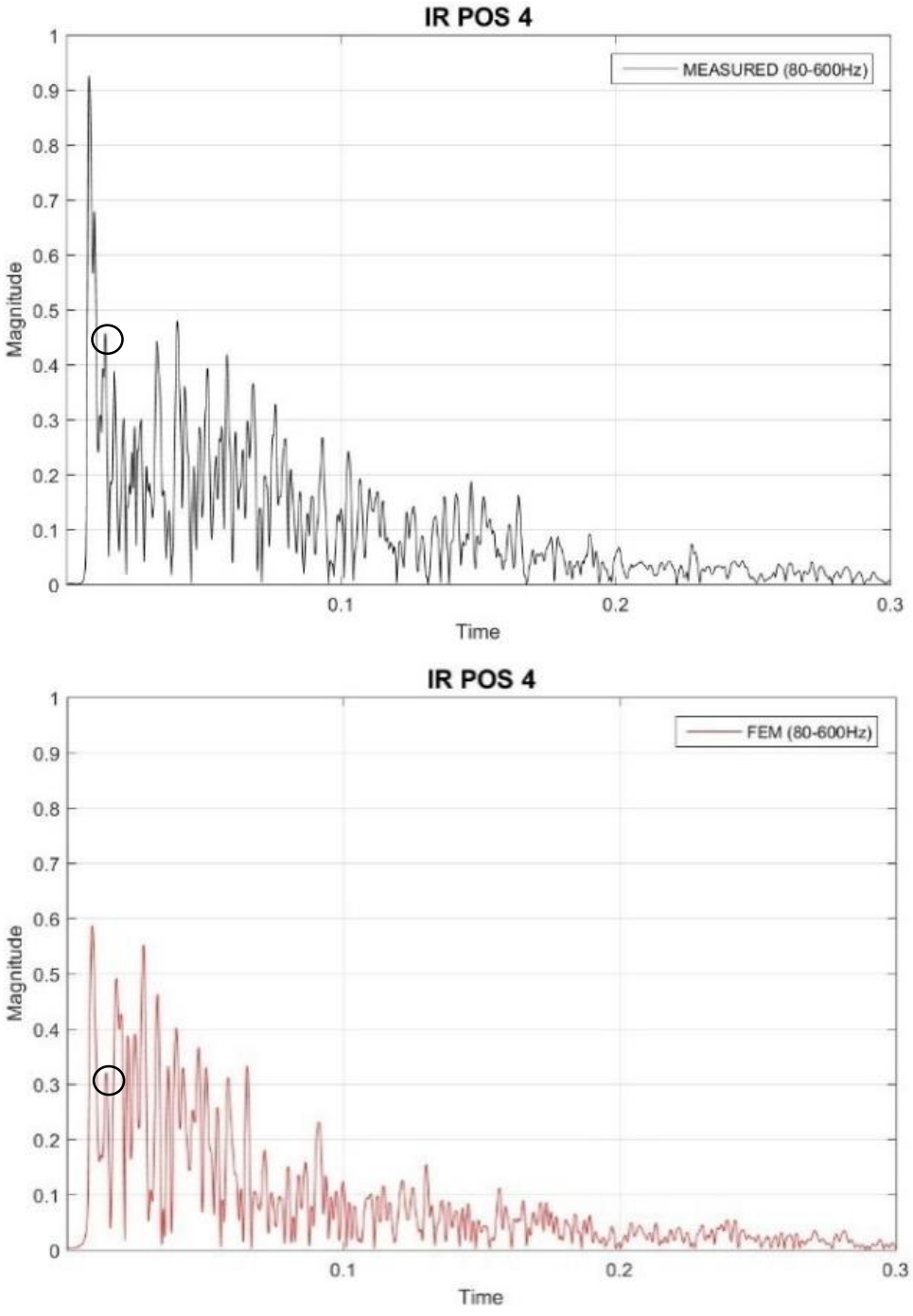


Figure 5.11: Top, the measured RIR at position 4. Bottom, the estimated FEM RIR for the same position.

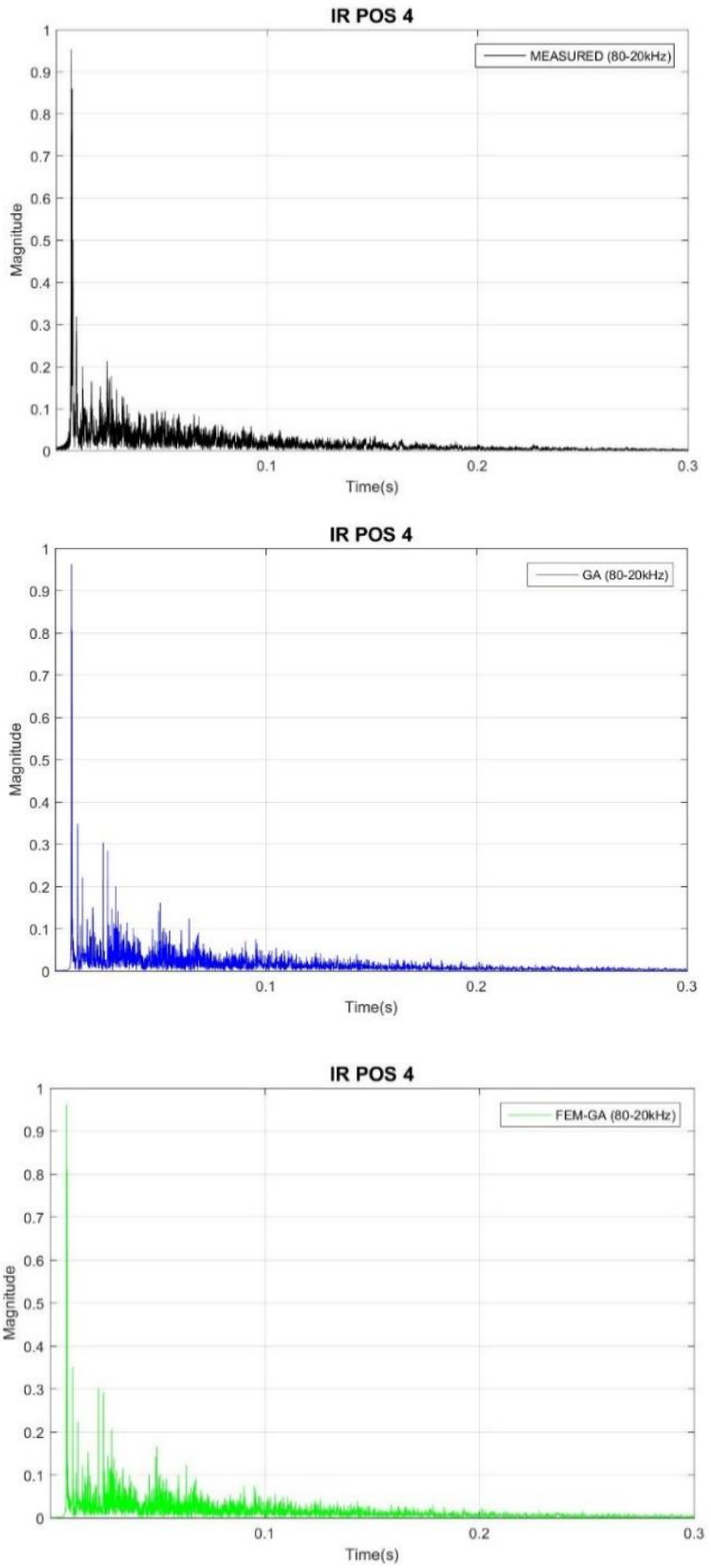


Figure 5.12: RIRs obtained at position 4. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

From Figure 5.13 to Figure 5.18 the BIRs obtained by means of acoustic measurements and numerical simulations are presented for the same positions.

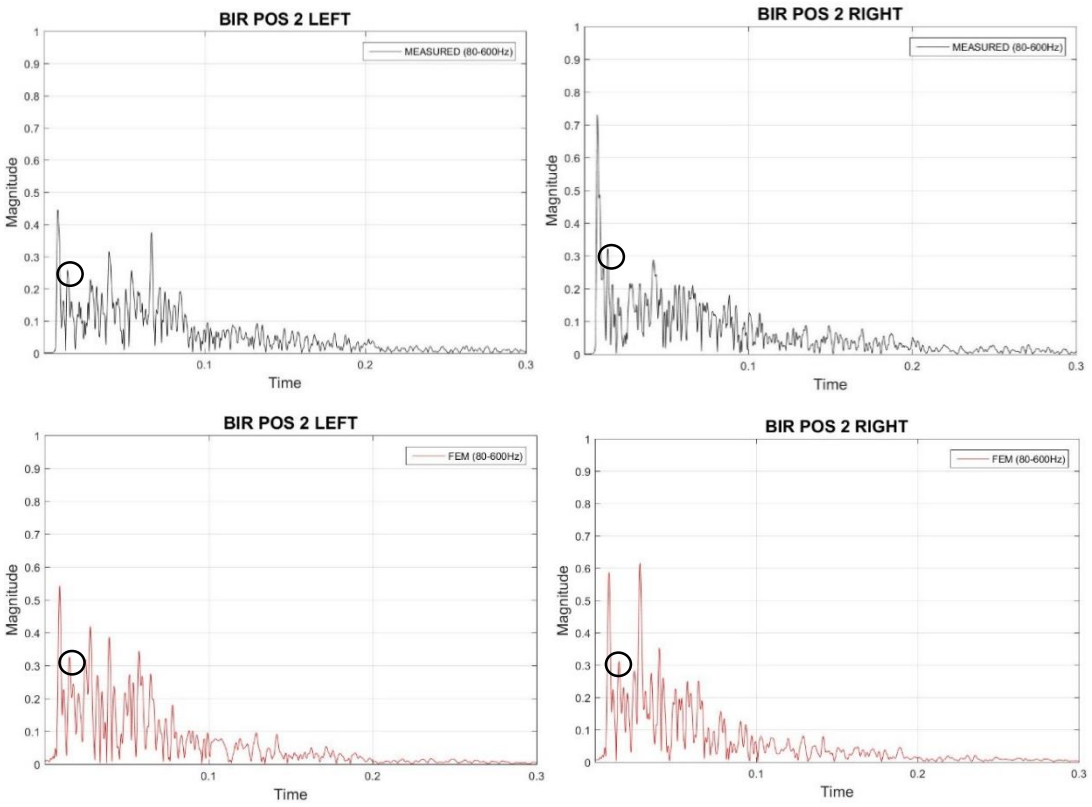


Figure 5.13: Top, the measured BIR at position 2. Bottom, the estimated FEM BIR for the same position.

In Figure 5.13, it can be observed that the time of arrival of the direct sound in the measured and simulated BIR coincide largely. In the left responses the first peak is located at sample 386 (0.0088 seconds) for the measured BIR and at sample 409 (0.0093 seconds) for the simulated BIR. In the right responses the first peak is located at sample 346 (0.0078 seconds) for the measured BIR and at sample 384 (0.0087 seconds) for the simulated BIR. In the left responses the magnitude of the first peak is similar, and in the case of the right it differs slightly. All the responses show that the magnitude of the first reflection maintains a strong concordance in all the responses, being around 0.3. The ITDG that appears on the left measured BIR is 5.8 ms and on the left simulated BIR is 6.1 ms. Similarly, the ITDG that appears on the right measured BIR is 6.5 ms, whereas in the simulated BIR is 6.1 ms.

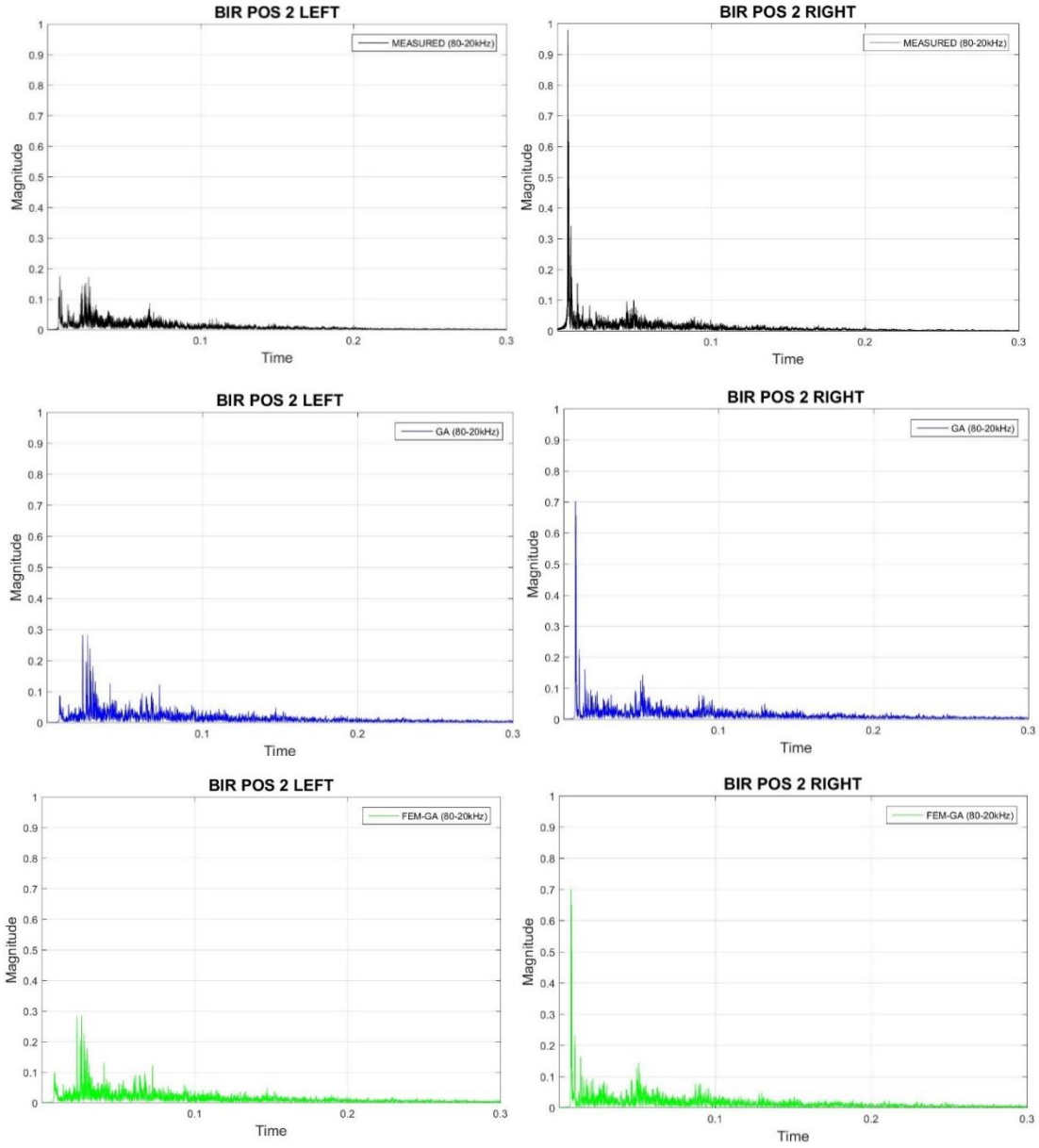


Figure 5.14: Measured and simulated BIR at position 2. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

In Figure 5.14, it can be seen that the time of arrival of the direct sound arrives at the same time for the left BIRs, approximately at sample 359 (0.0081 seconds), and coincide largely for right BIRs, being in sample 301 (0.0068 seconds) for the right measured BIR and sample 337 (0.0076 seconds) for the BIRs obtained by means of GA and FEM-GA. Regarding the early reflections in the left responses, although the peaks coincide in time, they differ in magnitude. In the left measured BIR, at sample 585 (0.0133 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 5.3 ms. As for the left BIRs simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 613 (0.0139 seconds),

which would mean an ITDG of 5.7 ms. In the right BIR case, a second peak is found at sample 573 (0.013 seconds) in the measured response, resulting in an ITDG of 6.2 ms. As for the simulated BIRs, a second peak is found at sample 605 (0.0137 seconds), which corresponds to an ITDG of 6.1ms.

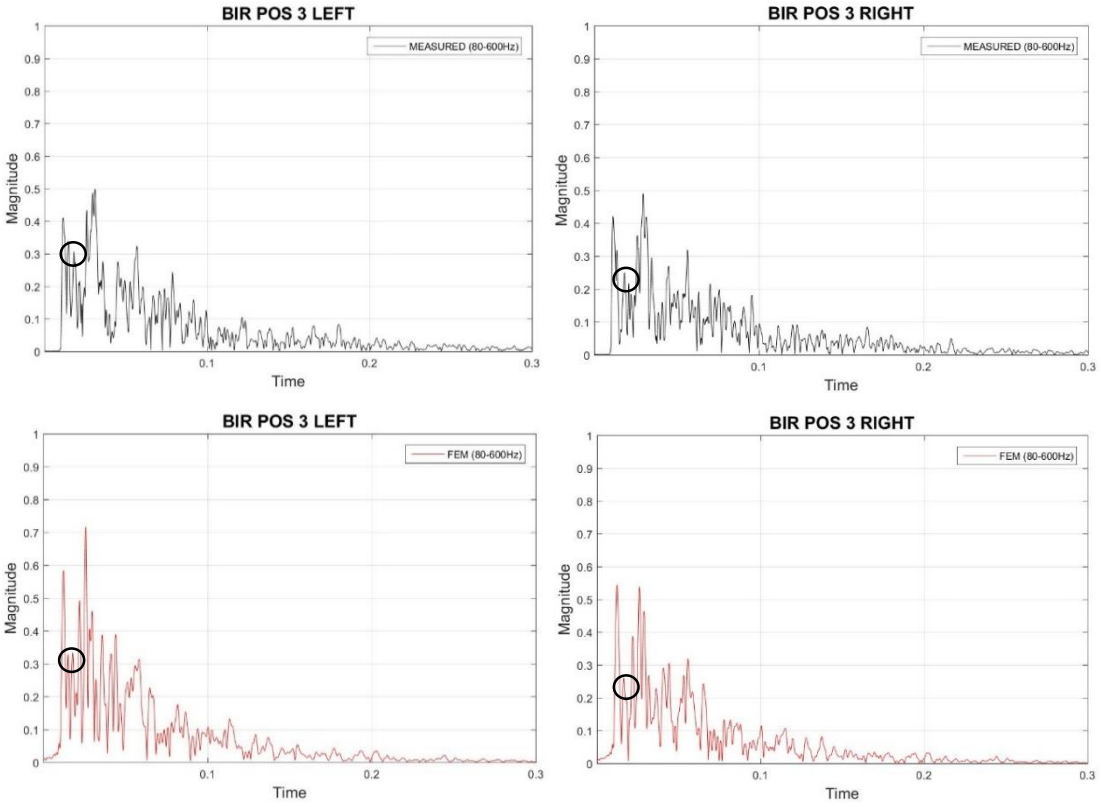


Figure 5.15: Top, the measured BIR at position 3. Bottom, the estimated FEM BIR for the same position.

In Figure 5.15, it can be seen that the time of arrival of the direct sound in the measured and simulated BIR coincide largely. In these, the peaks corresponding to the first reflection coincides in time, but not in magnitude. In the left responses the first peak is located at sample 499 (0.0113 seconds) for the measured BIR and at sample 545 (0.0124 seconds) for the simulated BIR. In the right responses the first peak is located at sample 510 (0.0116 seconds) for the measured BIR and at sample 542 (0.0123 seconds) for the simulated BIR. The early reflection peaks that follow are more similar for the right BIR case that the left BIR case, as in the latter case, the FEM simulation results in peaks with greater amplitude. The ITDG that appears on the left measured BIR is 3.3 ms and on the left simulated BIR is 2.6 ms. Similarly, the ITDG that appears on the right measured BIR is 6.8 ms, whereas in the simulated BIRs is 4.0 ms.

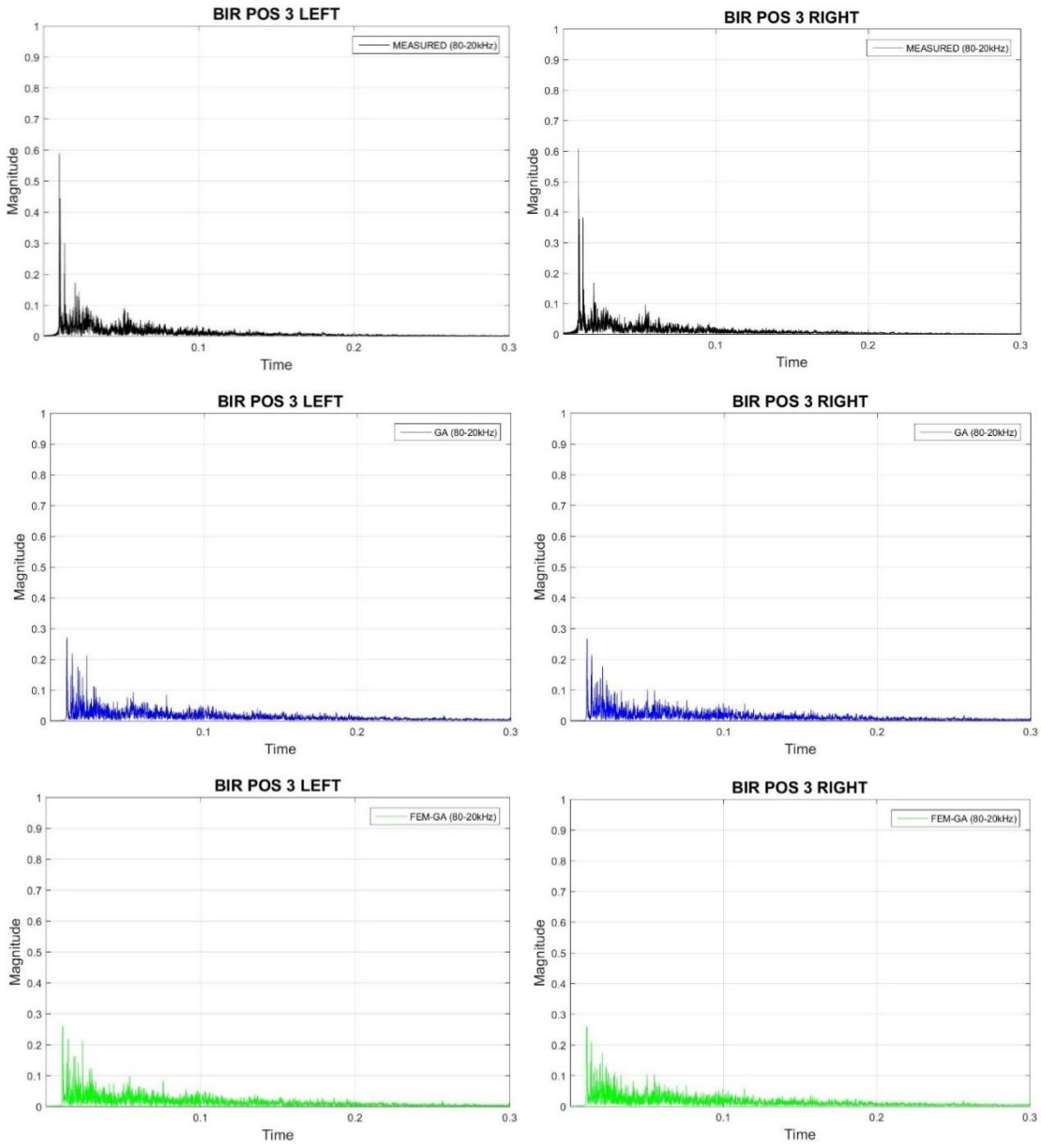


Figure 5.16: Measured and simulated BIR at position 3. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

Figure 5.16 compares measured BIR against GA and FEM-GA simulated BIRs for position 3. It can be seen that the time of arrival of the direct sound in the measured and simulated BIRs coincide largely, although the magnitude of the first peaks differ in the simulated cases, being higher in the measured responses. In the left responses the first peak is located at sample 451 (0.0102 seconds) for the measured BIR and at sample 486 (0.011 seconds) for the simulated BIRs obtained by GA and FEM-GA. In the right responses the first peak is located at sample 447 (0.0101 seconds) for the measured BIR and at sample 486 (0.011 seconds) for the simulated BIRs. In the left measured BIR, at sample 600 (0.0136 seconds) a second peak is found, which could correspond to the

first reflection, therefore the ITDG in this case would be 3.4 ms. As for the left BIR simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 635 (0.0144 seconds), which would mean again an ITDG of 3.4 ms. In the right BIRs case, a second peak is found at sample 576 (0.0131 seconds) in the measured response, resulting in an ITDG of 3.0 ms. As for the simulated BIRs, a second peak is found at sample 613 (0.0139 seconds), which corresponds to an ITDG of 2.9 ms.

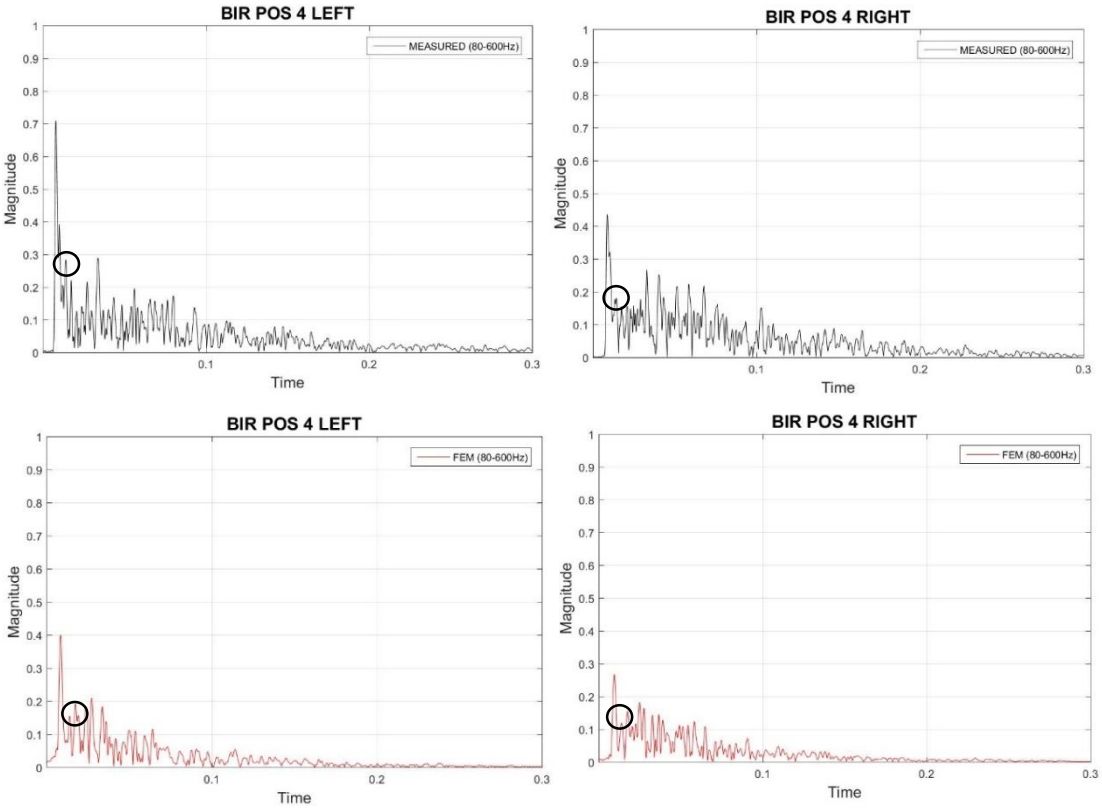


Figure 5.17: Top, the measured BIR at position 4. Bottom, the estimated FEM BIR for the same position.

Figure 5.17 compares measured BIR against FEM simulated BIRs for position 4. It can be seen that the time of arrival of the direct sound in the measured and simulated BIR coincide largely. In the left responses the first peak is located at sample 349 (0.0079 seconds) for the measured BIR and at sample 374 (0.0085 seconds) for the simulated BIR. In the right responses the first peak is located at sample 387 (0.0088 seconds) for the measured BIR and at sample 427 (0.0097 seconds) for the simulated BIR. The ITDG that appears on the left measured BIR is 6.2 ms and on the left simulated BIR is 5.2 ms. Similarly, the ITDG that appears on the right measured BIR is 5.7 ms, whereas in the simulated BIRs is 3.8 ms.

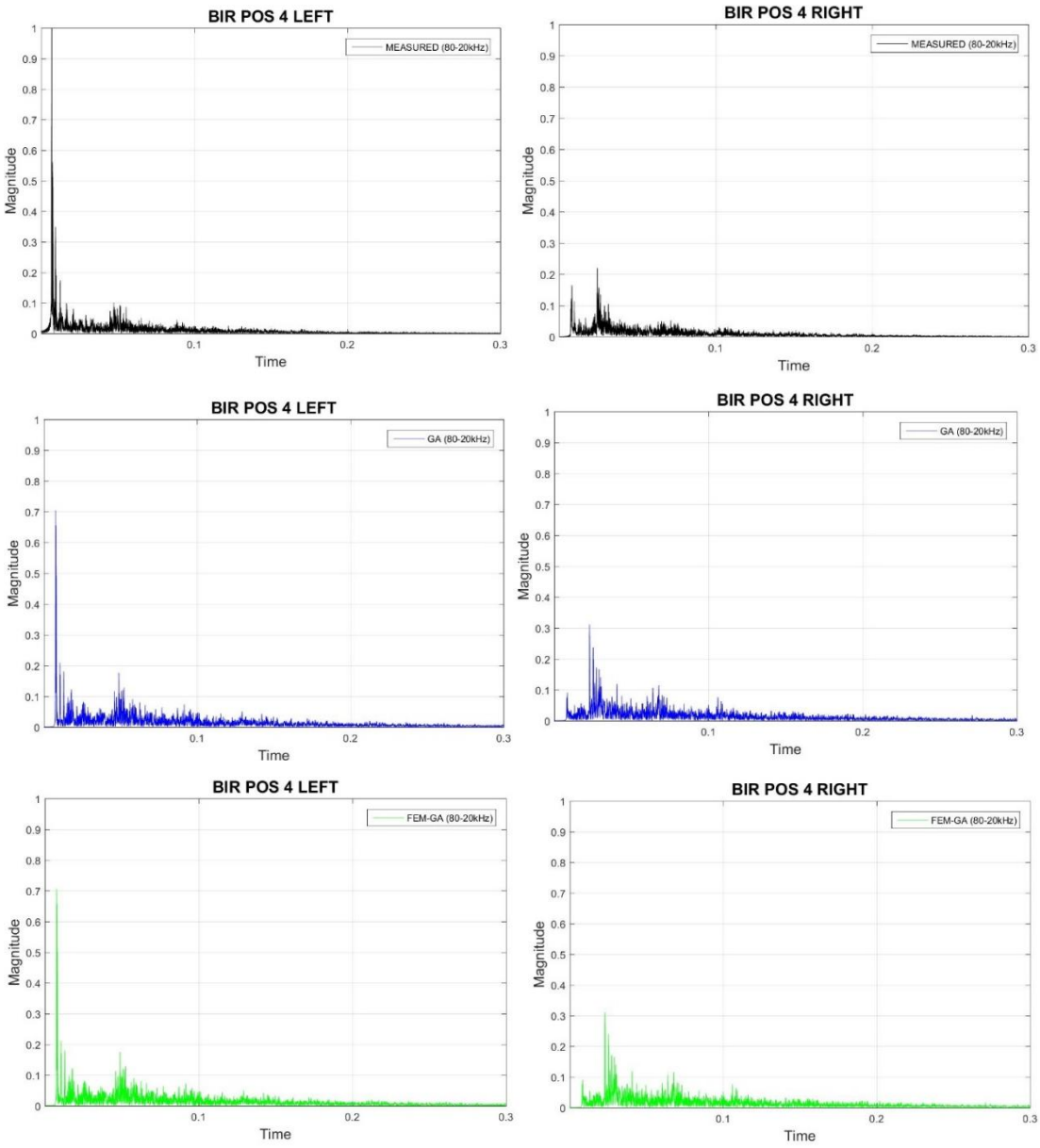


Figure 5.18: Measured and simulated BIR at position 4. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

In Figure 5.18, it can be seen that the time of arrival of the direct sound in the measured and simulated BIRs coincide largely. In the left responses the first peak is located at sample 302 (0.0068 seconds) for the measured BIR and at sample 337 (0.0076 seconds) for the simulated BIRs. In the right responses the first peak is located at sample 353 (0.008 seconds) for the measured BIR and at sample 375 (0.0085 seconds) for the simulated BIRs. In the left measured BIR, at sample 414 (0.0094 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 2.6 ms. As for the left BIR simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 466 (0.0106 seconds), which would mean an ITDG of

3.0 ms. In the right BIRs case, a second peak is found at sample 433 (0.0098 seconds) in the measured response, resulting in an ITDG of 1.8 ms. As for the simulated BIRs, a second peak is found at sample 481 (0.0109 seconds), which corresponds to an ITDG of 2.4 ms.

5.2.3.2 Frequency domain room transfer function results of the Meeting Room

In this section, the frequency responses simulated are compared against measured frequency responses. From Figure 5.19 to Figure 5.21 the RFRs are presented for positions 2, 3 and 4 (see Figure 5.2). The RFRs figures were obtained by applying a FFT to the normalised RIRs, having as the reference the measured RIRs. The RFRs graphs are split up in two frequency ranges for clarity. First, a lower frequency range from 80 Hz to 700 Hz is used to visualise the simulated GA and FEM responses against the measured RFR. Second, a frequency range starting at 700 Hz until the upper limit of the 4 kHz octave band is used in order to compare GA RFR and measured response.

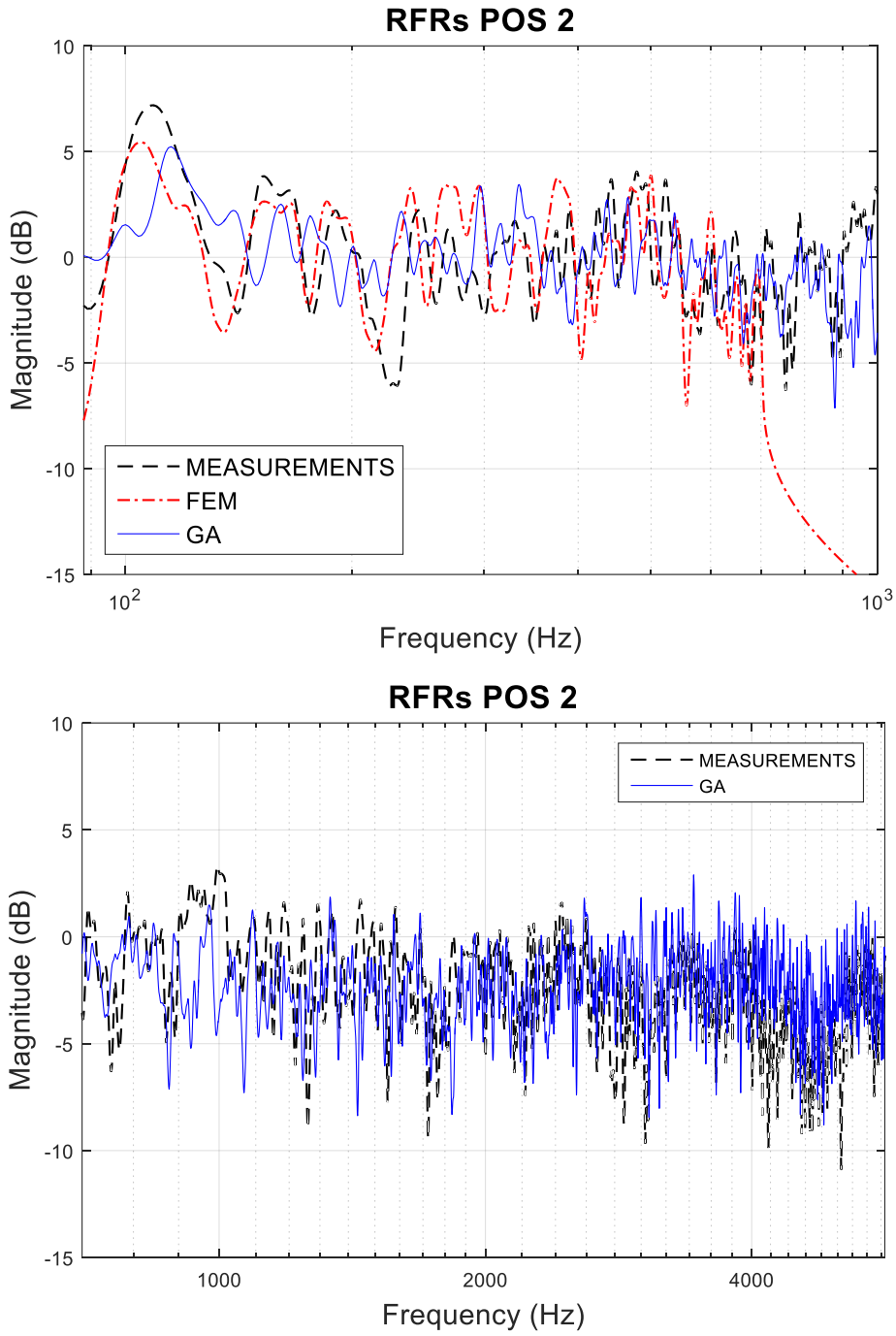


Figure 5.19: RFRs obtained for position 2. Top, measured and simulated by GA and FEM RFRs up to 700 Hz. Bottom, measured and GA RFRs from 700 Hz up to 5.6 kHz.

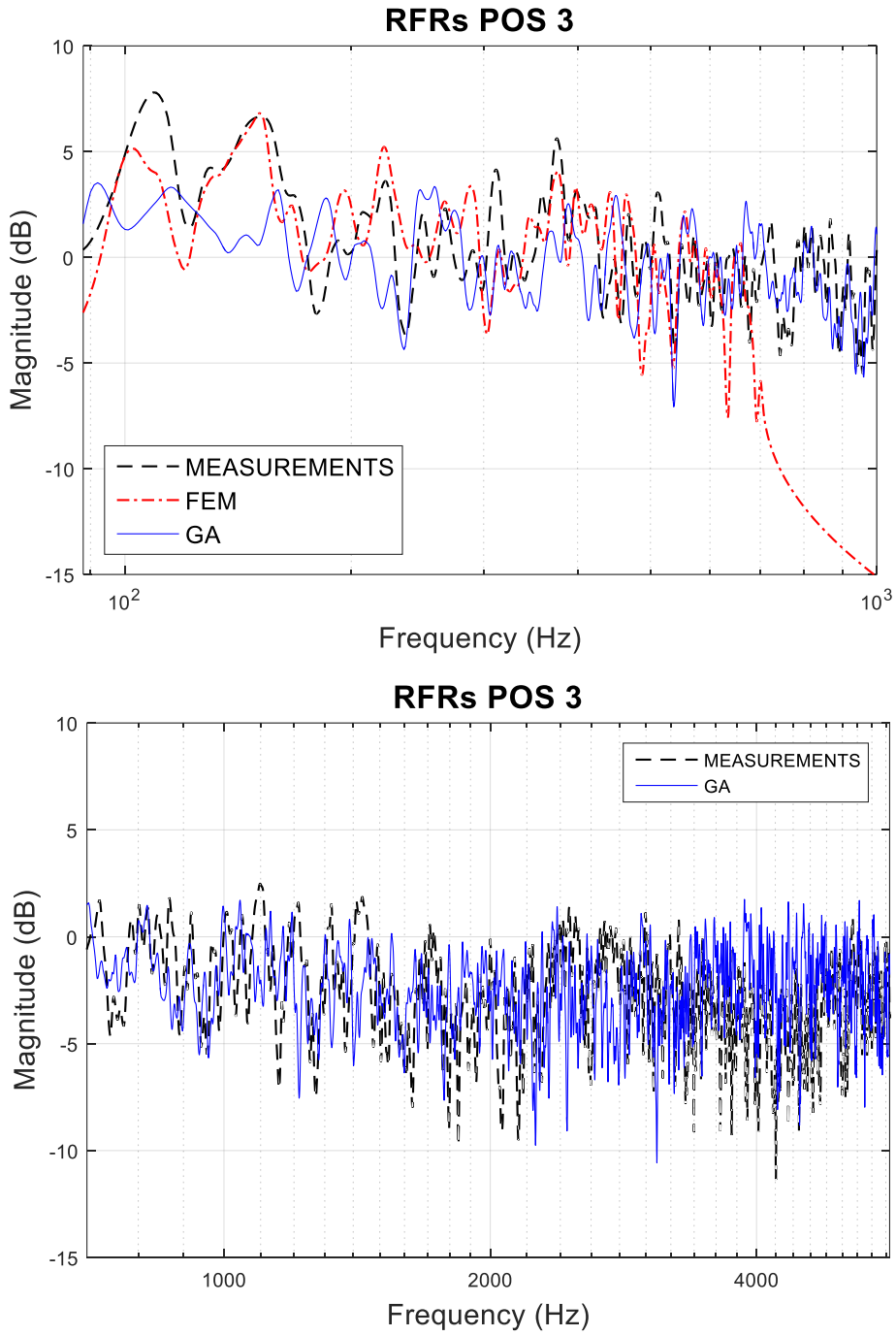


Figure 5.20: RFRs obtained for position 3. Top, measured and simulated by GA and FEM RFRs up to 700 Hz. Bottom, measured and GA RFRs from 700 Hz up to 5.6 kHz.

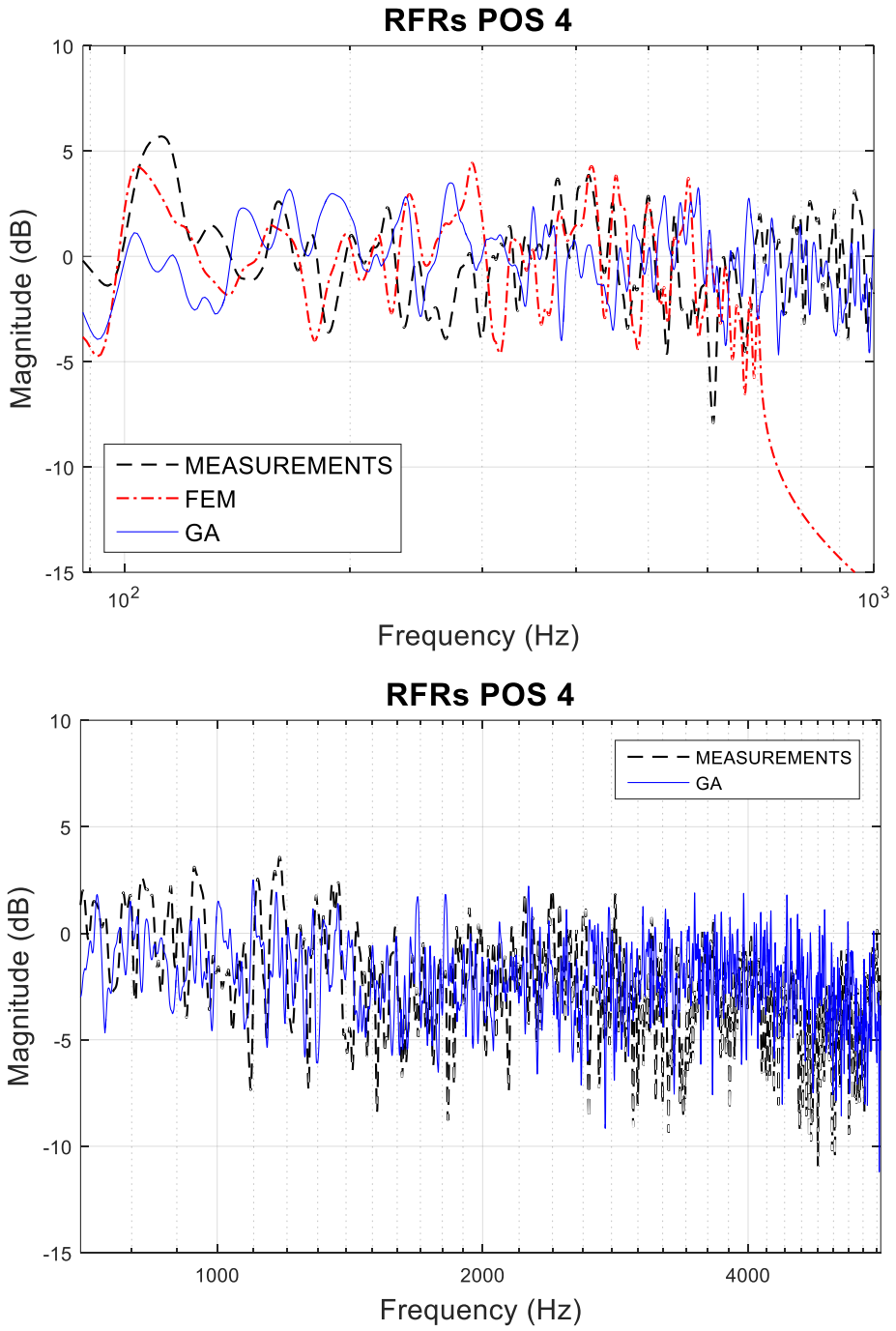


Figure 5.21: RFRs obtained for position 4. Top, measured and simulated by GA and FEM RFRs up to 700 Hz. Bottom, measured and GA RFRs from 700 Hz up to 5.6 kHz.

5.2.3.3 Room acoustic parameters results of the Meeting Room

In this section, the spatially averaged room acoustic parameters results for as T_{20} (see Figure 5.22) and EDT (see Figure 5.23), and the D_{50} (see Figure 5.24), C_{80} (see Figure 5.25) and *IACC* (see Figure 5.26) are presented according to ISO standard 3382 (2009). The last parameter was estimated from the measured and simulated BIR.

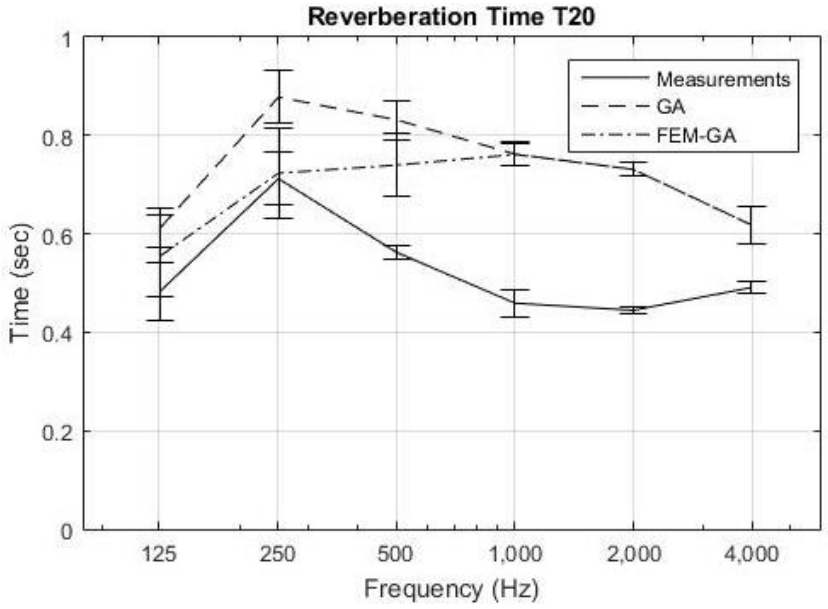


Figure 5.22: Spatially averaged T_{20} results of the Meeting Room calculated, applying the integrated impulse response method according to ISO 3382, from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

In Figure 5.22 T_{20} spatially averaged over the six source-receiver position combinations obtained by means of measurements, GA and FEM-GA hybrid approaches can be seen. It is noted that results obtained by FEM-GA have a better agreement than GA with respect to the measured T_{20}, in the octave bands of interest 125, 250 Hz and 500 Hz. It can be appreciated how reverberation times estimated by GA are overestimated in comparison to measurements for all the frequencies.

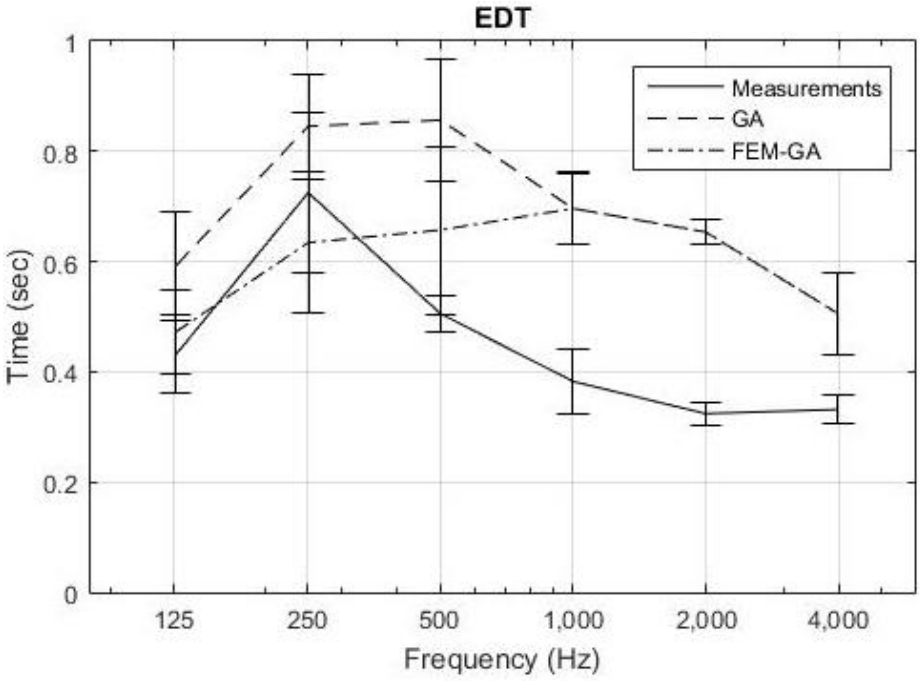


Figure 5.23: Spatially averaged *EDT* results of the Meeting Room calculated, applying the integrated impulse response method according to ISO 3382, from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

Figure 5.23 presents EDT results spatially averaged over the six source-receiver position combinations obtained by means of measurements, GA and FEM-GA. As for T_{20}, EDT obtained by FEM-GA have more similar results to measured values for the frequencies of interest. In this case, in the octave band of 250 Hz the EDT estimated by FEM-GA had a lower value than the measured EDT. The EDT results obtained by GA present a similar situation than the one observed in results for T_{20}.

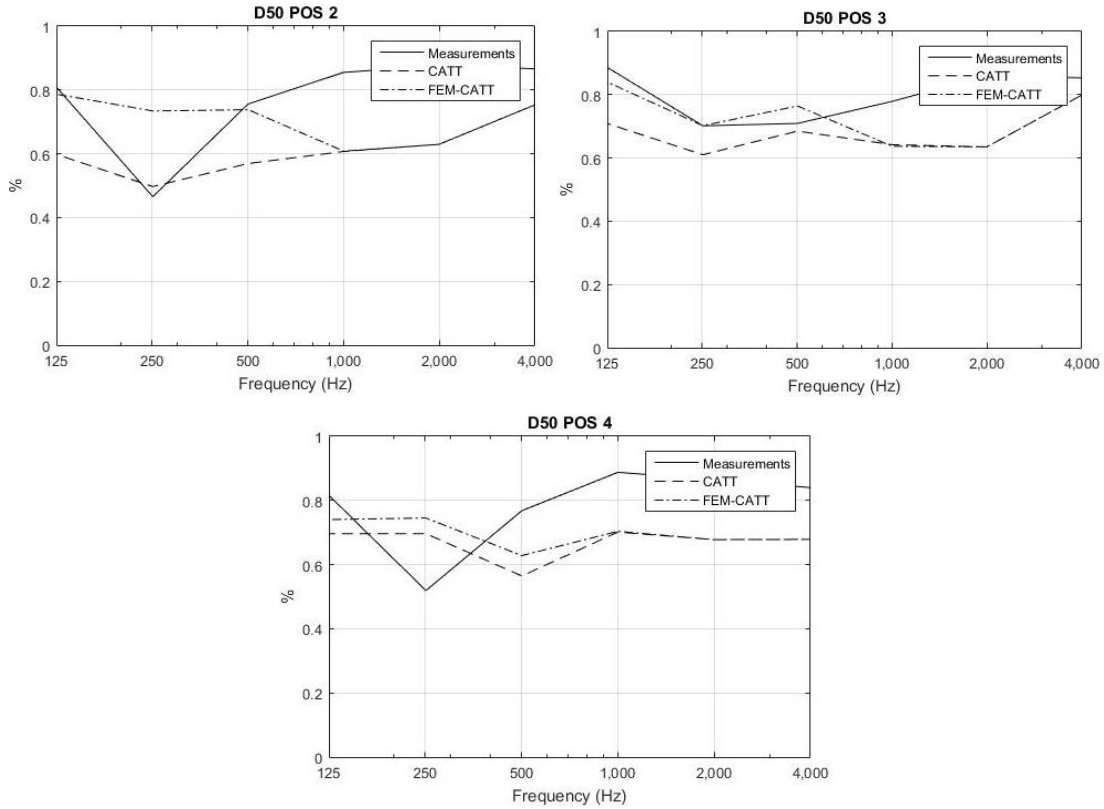


Figure 5.24: D_{50} results of the Meeting Room for positions 2, 3 and 4 calculated from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

Figure 5.24 shows the D_{50} results obtained from each RIR measured and estimated by GA and FEM-GA, at three different positions in the Meeting Room. In general, results obtained by FEM-GA are closer to D_{50} values measured in the octave bands of 125 Hz and 500 Hz, however, it is important to mention that the measured D_{50} for 250 Hz at position 2 presents an unusual dip compared with 125 Hz and 500 Hz. In position 2 the values obtained by GA and FEM-GA differ significantly between them. In position 3, numerical results are very similar to measured values to a greater extent in the frequency bands of interest (125 Hz, 250 Hz and 500 Hz). For the octave bands of 1 kHz, 2 kHz and 4 kHz GA D_{50} results are overestimated with respect to measurement results.

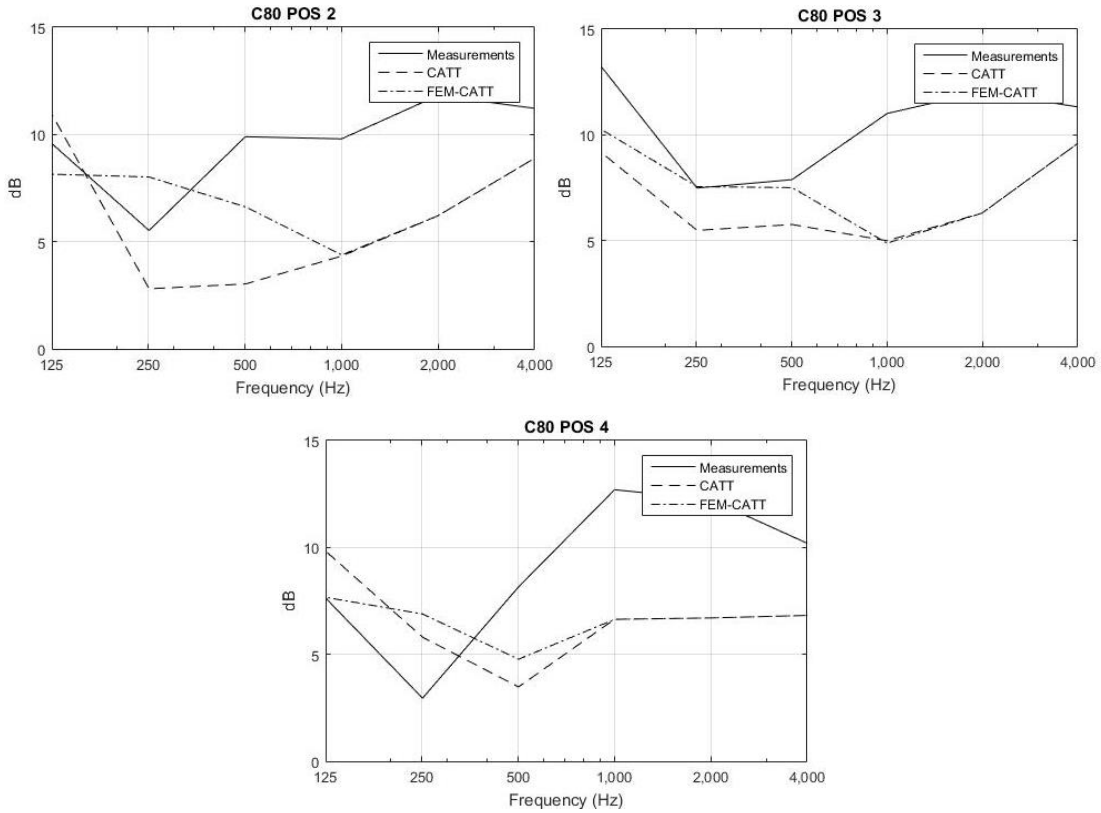


Figure 5.25: C_{80} results of the Meeting Room for positions 2, 3 and 4 calculated from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

In Figure 5.25 can be appreciated that for all positions, FEM-GA C_{80} results are in general closer to measured C_{80} values in the frequencies of 125 Hz, 250 Hz and 500 Hz, with the exception of the octave band of 250 Hz at position 4. As mentioned in D_{50} results, measured C_{80} for 250 Hz at position 2 and 4 presents an unusual dip compared with 125 Hz and 500 Hz. In these frequencies, position 3 gives results that demonstrate good agreement between measurements and the FEM-GA approach. GA results of C_{80} are underestimated with respect to measured values for the octave bands of 500 Hz, 1 kHz, 2 kHz and 4 kHz.

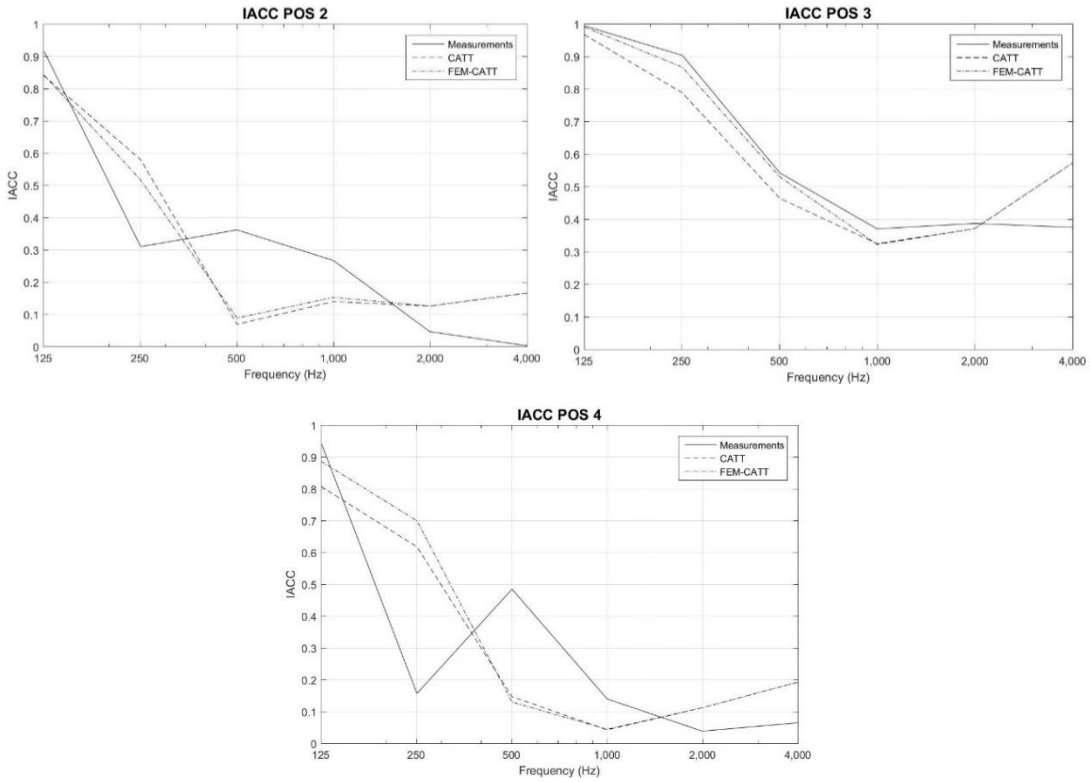


Figure 5.26: *IACC* results of the Meeting Room for positions 2, 3 and 4 calculated from BIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

In Figure 5.26 can be appreciated how the numerical approaches present similar trends of IACC results between them for all positions. Once again, measured IACC for 250 Hz at position 2 and 4 presents an unusual dip compared with 125 Hz and 500 Hz. Position 3 presents the most similar results between measurements and numerical approaches. Positions 2 and 4 exhibit significant differences in IACC results for the octave bands of 250 Hz and 500 Hz. In this case, GA results of IACC are underestimated with respect to measured values for the octave bands of 500 Hz, 1 kHz and 4 kHz, being overestimated in the octave band of 2 kHz.

5.2.3.4 Natural frequencies analytical solution of the Meeting Room

The natural frequencies of the room were estimated below the Schroeder frequency (see equation [3.15]), using the following expression (Kuttruff, 2007):

$$f_{n_x n_y n_z} = \frac{c_0}{2\pi}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2} \qquad [5.17],$$

where, c_0 is the sound speed, L_x, L_y and L_z are the dimensions of the room and n_x, n_y and n_z are consecutive integer numbers denoting the axial, tangential and oblique modes. According to Kuttruff (2000), the number of eigenfrequencies from 1 Hz to a determined upper frequency limit in small rooms can be estimated with expression (Kuttruff, 2009):

$$N_f = \frac{4\pi}{3}V\left(\frac{f}{c}\right)^3 + \frac{\pi}{4}S\left(\frac{f}{c}\right)^2 + \frac{L}{8}\frac{f}{c} \qquad [5.18],$$

where, $L = 4(L_x + L_y + L_z)$ is the sum of all edge lengths in the rectangular room and $S = 2(L_xL_y + L_xL_z + L_yL_z)$ is the area of all walls (Kuttruff, 2000). According to expression [5.18], the number of eigenfrequencies increases cubically with frequency. In Figure 5.27, the 123 eigenfrequencies below the Schroeder frequency of 224 Hz for the Meeting Room are grouped according to frequency bands, with a 10Hz bandwidth.

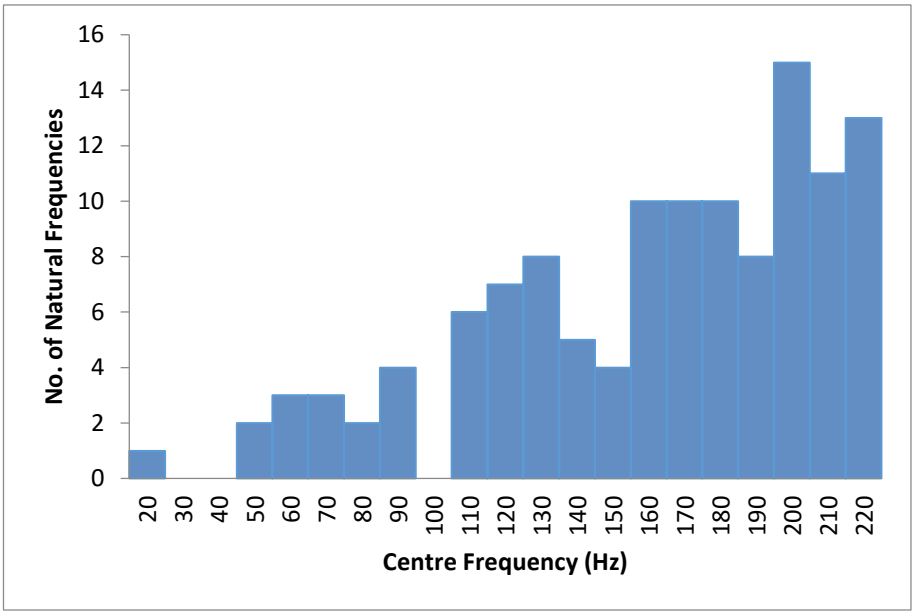


Figure 5.27: Histogram of natural frequencies in the Meeting Room below Schroeder frequency.

In order to see the accuracy of FEM to estimate room modes, the first 32 eigenfrequencies were calculated in COMSOL and the percentage of error was estimated by the following expression:

$$Error\% = (f_{ana} - f_{fem}) / f_{ana} \qquad [5.19],$$

where, f_{ana} is the natural frequency calculated analytically and f_{fem} is the eigenfrequency found with FEM. The natural frequencies calculated by both methods, the mode shapes and the percentage of error can be seen in Table 5.5.

Table 5.5: Mode shapes, analytical and numerical natural frequencies and the correspondent percentage of error.

Mode shapes			**Eigenfrequencies (Hz)**			**Mode shapes**			**Eigenfrequencies (Hz)**		
nz	ny	nx	Analytical	FEM	%Error	nz	ny	nx	Analytical	FEM	%Error
0	0	1	23.5	24.3	-3.55%	0	2	0	107.2	110.3	-2.90%
0	0	2	47	48.7	-3.55%	0	1	4	108.2	111.8	-3.39%
0	1	0	53.6	55.1	-2.89%	0	2	1	109.7	112.1	-2.15%
1	0	0	57.2	56	1.96%	1	0	4	110	112.3	-2.09%
0	1	1	58.5	60.3	-3.00%	2	0	0	114.3	112.9	1.21%
1	0	1	61.8	61.1	1.15%	2	0	1	116.7	114.7	1.73%
0	0	3	70.5	73	-3.55%	0	2	2	117	120.5	-3.00%
0	1	2	71.3	73.5	-3.18%	0	0	5	117.5	121.6	-3.55%
1	0	2	74	74.2	-0.30%	1	2	0	**121.5**	122.2	-0.59%
1	1	0	78.4	78.6	-0.34%	1	1	4	**122.4**	123.7	-1.11%
1	1	1	81.8	82.3	-0.61%	2	0	2	**123.6**	124.9	-1.07%
0	1	3	88.5	91.5	-3.31%	1	2	1	**123.7**	125.1	-1.11%
1	0	3	**90.7**	92	-1.40%	2	1	0	126.3	126.1	0.14%
1	1	2	**91.4**	92.5	-1.20%	0	2	3	128.3	127.3	0.78%
0	0	4	94	97.3	-3.55%	2	1	1	128.4	132.3	-2.97%
1	1	3	105.4	107.3	-1.79%	0	1	5	129.1	132.9	-2.97%

A graphical representation of the modes 0-0-3, 1-1-0 and 1-1-1 can be appreciated in Figure 5.28:

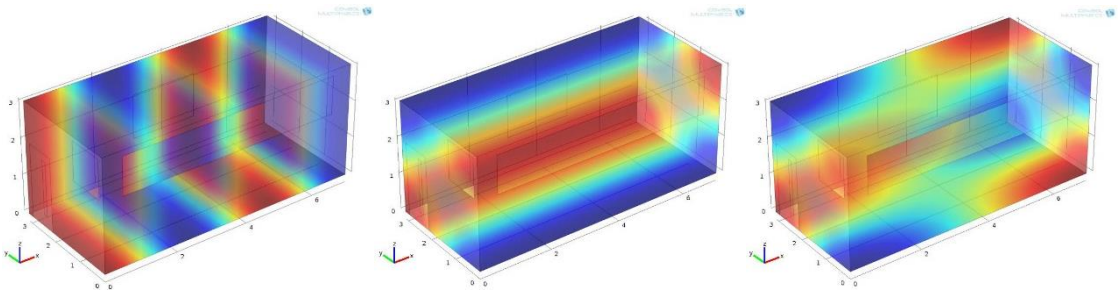


Figure 5.28: Graphic representation in COMSOL of axial mode 0-0-3, tangential mode 1-1-0 and oblique mode 1-1-1.

According to the eigenfrequencies presented in Figure 5.27, there is an important number of natural frequencies over 200 Hz, which indicates that there are places in the room that have maximums and minimums of pressure in the

octave band of 250 Hz. This situation could affect the measurements results and might be the explanation of the unexpected values of measured D_{50}, C_{80} and IACC obtained in this octave band. The importance of the eigenfrequencies calculated numerically lies on the possibility of predicting these frequencies with a high degree of precision (see Table 5.5), appreciating the places in the room where these modes present maximums and minimums of pressure (see Figure 5.28).

5.3 Numerical Simulations of the Classroom

The "Mini-auditorium 2", located on the fourth floor of the engineering building at the University of San Buenaventura, in Medellin, Colombia was chosen as the second case of study. This room was selected two reasons: the differences with respect to the meeting room in terms of size and surface materials and the simplicity of its geometry, which facilitates the 3D model construction for both numerical methods applied. In terms of size, this room doubles the volume of the meeting room. Regarding the material of the surfaces found in the classroom, practically all of them are hard reflective surfaces, in which waves can travel freely along a surface. This simplifies the definition of the boundary conditions in both numerical methods and provides useful information, regarding the application of a real valued frequency dependent impedance related to the absorption coefficient obtained from a GA material parameter database. This section describes the analysis of the sound wave propagation estimations obtained by means of FE and GA numerical methods and the objective evaluation of the techniques implemented to create auralizations.

5.3.1 GA simulations of the Classroom

In this section, the steps applied to create the GA model of the classroom with existing conditions are described. The first step consisted of the estimation of the reverberation time by means of a Sabine model, considering air absorption. The selection of materials and coefficients of absorption and scattering applied in the model can be seen in Table 5.6 and Table 5.7. A comparison between estimated reverberation time and measured (T_{20}) reverberation time results is presented in Figure 5.29. The next step involved the creation of the GA model applying the theoretical methods explained in Chapter 3 and the same procedure used to generate the GA Meeting Room model.

Table 5.6: Absorption coefficients used on each surface of the classroom for both, Sabine and GA models.

Surface	Material	Area (m^2)	Octave Band Centre Frequency (Hz)					
			125	250	500	1000	2000	4000
			Absorption Coefficients					
Floor	Tile	48.99	0.01	0.01	0.02	0.02	0.02	0.02
Doors	Wood	4.06	0.14	0.10	0.06	0.08	0.10	0.10
Windows	Glass	0.84	0.35	0.25	0.18	0.12	0.07	0.04
Board	Acrylic	2.91	0.02	0.02	0.03	0.03	0.04	0.05
Panel	Foam	0.50	0.08	0.22	0.55	0.70	0.85	0.75
Air conditioning	Plastic	1.38	0.02	0.02	0.03	0.03	0.04	0.05
Video projector	Plastic	0.30	0.02	0.02	0.03	0.03	0.04	0.05
Lights	Metal	5.04	0.02	0.02	0.03	0.03	0.04	0.05
Right wall	Plaster	22.41	0.12	0.10	0.08	0.06	0.06	0.06
Left wall	Painted concrete	24.43	0.01	0.01	0.01	0.02	0.02	0.02
Back wall	Plaster	14.54	0.12	0.10	0.08	0.06	0.06	0.06
Front wall	Painted concrete	12.75	0.01	0.01	0.01	0.02	0.02	0.02
Ceiling	Plaster	39.63	0.12	0.10	0.08	0.06	0.06	0.06

Table 5.7: Scattering coefficients used on each surface of the classroom for the GA model.

Surface	Material	Area (m^2)	Octave Band Centre Frequency (Hz)					
			125	250	500	1000	2000	4000
			Scattering Coefficients					
Floor	Tile	22.56	10	10	10	10	10	10
Floor (staircase area)	Tile	26.43	80	60	40	20	10	10
Doors	Wood	4.06	10	10	10	10	10	10
Windows	Glass	0.84	10	10	10	10	10	10
Board	Acrylic	2.91	10	10	10	10	10	10
Panel	Foam	0.50	10	10	10	10	10	10
Air conditioning	Plastic	1.38	98	98	79	39	20	10
Video projector	Plastic	0.30	98	98	98	80	40	20
Lights	Metal	5.04	80	60	40	40	40	40
Right wall	Plaster	22.41	10	10	10	10	10	10
Left wall	Painted concrete	24.43	10	10	10	10	10	10
Back wall	Plaster	14.54	10	10	10	10	10	10
Front wall	Painted concrete	12.75	10	10	10	10	10	10
Ceiling	Plaster	39.63	10	10	10	10	10	10

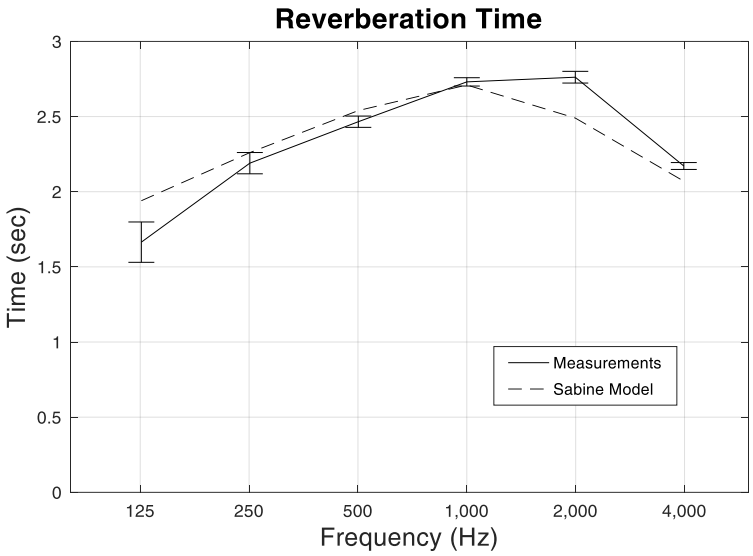


Figure 5.29 Classroom Reverberation times estimated by means of Sabine model and spatially averaged measurements.

5.3.1.1 Source and receivers in Classroom GA simulations

The sound source and receivers were modelled following the same procedure described in section 5.2.1.3. As a source signal, white noise at 94 dB was used in order to facilitate the implementation of a similar acoustic source using the FEM approach. The directivity information provided by the manufacturer was used in the source directivity module available in CATT software. In this case, a *SD0* format was implemented interpolating horizontal and vertical polar measurements every 15˚. In Figure 5.30, directivity pattern plots of the sound source modelled in CATT are shown.

Five receiver positions were considered in this study (see Chapter 4). To obtain the RIR at those positions, just the Cartesian coordinate definition was necessary. On the other hand, for the BIR the direction had to be defined, in order to apply the procedure indicated in Chapter 3 to find left and right responses.

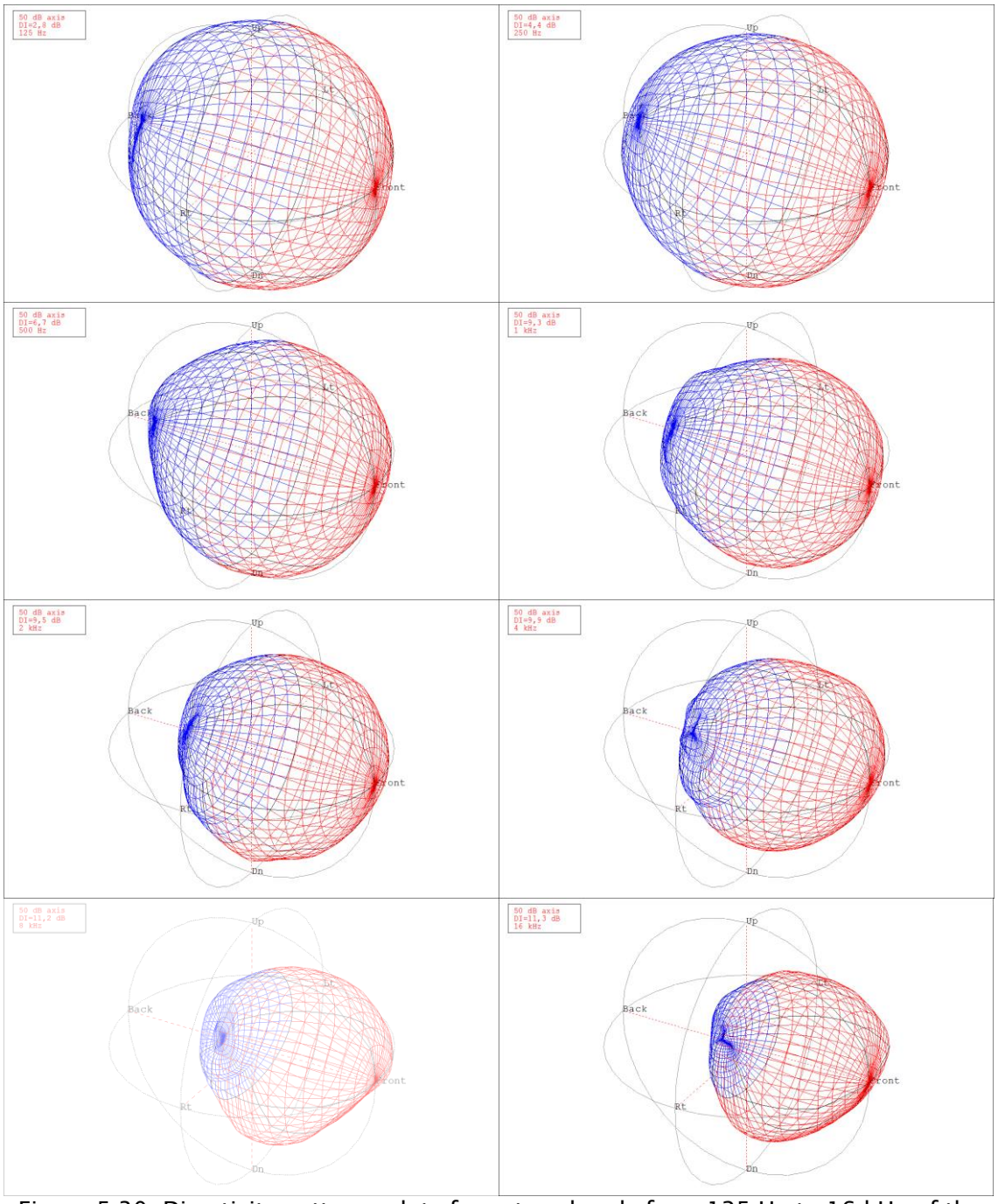


Figure 5.30: Directivity patterns plots for octave bands from 125 Hz to 16 kHz of the MACKIE loudspeaker modelled in CATT.

5.3.1.2 The 3D model

The development of the 3D model took into consideration the geometry of the room, the Schroeder frequency, the Reverberation Time estimation and the acoustic characteristics of the source. The volume of the classroom corroborated the precision of the GA model geometry, obtaining an analytical value of 135 m³ in comparison with the numerical value of 142 m³. In terms of Schroeder frequency, the analytical estimation gave 279 Hz and the GA method 275 Hz. A

Reverberation Time comparison taking into account an analytical estimation and spatially averaged measurements and numerical values results can be seen in Figure 5.31. The final 3D model created in the CATT software can be seen in Figure 5.32.

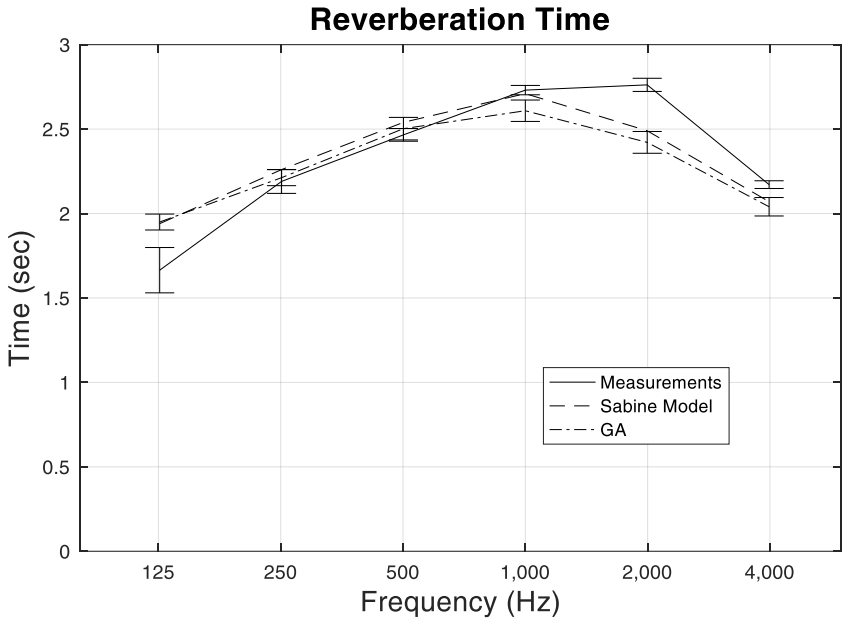


Figure 5.31: Classroom Reverberation times estimated by means of Sabine model and spatially averaged measurements and GA numerical approach.

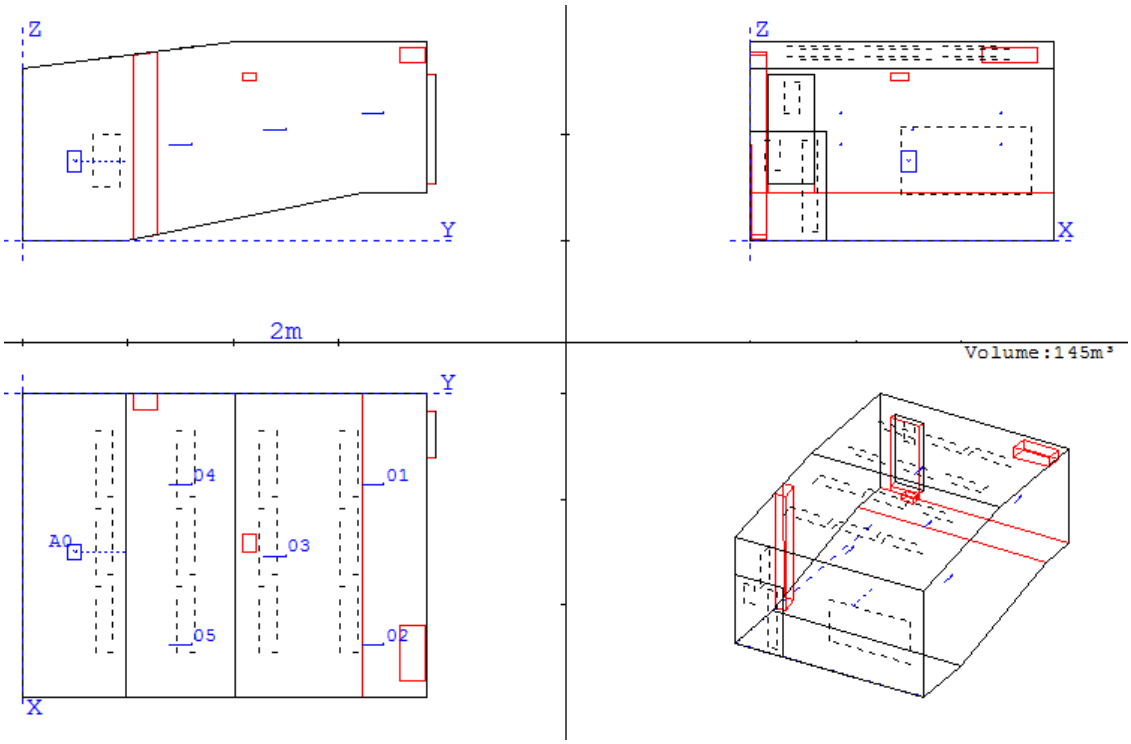


Figure 5.32: Top, frontal, lateral and isometric views of the classroom modelled in CATT-Acoustic software.

5.3.2 FE simulations

In this study, FE simulations were implemented using the software COMSOL 4.3. In this case, a time-harmonic simulation up to 500Hz with 1Hz frequency steps (see section Chapter 3) was applied for the same source-receiver combinations applied in GA simulations and acoustic measurements. In order to have simulation results to combine with the ones obtained by the GA method, the same parameters defined in the CATT software were used in the FE model. The creation of this model required the construction of a 3D geometry with its corresponding meshing process, the specification of boundary conditions, the definition of a monopole source and the characterization of a binaural receiver.

5.3.2.1 The Geometry and generation of the mesh

In order to generate the geometry, the original CAD model was imported into the COMSOL software. It is important to note that geometry construction in FE operates with the same structure as CAD, hence, there was no need to execute an additional procedure. In Figure 5.33, the 3D model created in the COMSOL software can be seen.

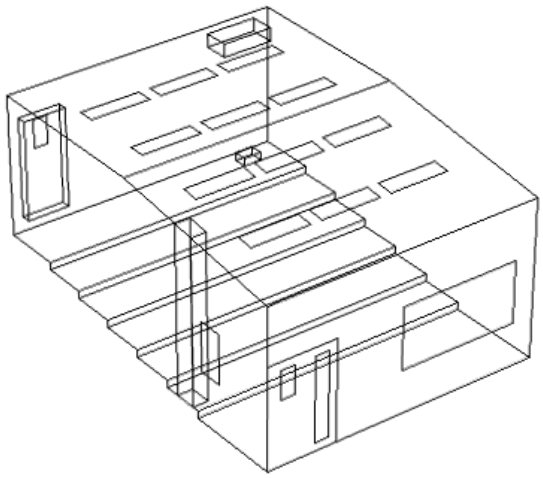


Figure 5.33: FE model created in COMSOL after importing the 3D model in CAD language.

In the generation of the mesh, the number of DOF to be solved in the model was defined according to the room volume and wavelength of the frequency analysed, as stated in the expression [3.8]. According to that, to estimate a frequency of 500 Hz in this model, a system with around 1,000,000 DOF had to be solved. In a normal computer, the computation of such a number of DOF could take an unreasonable time for just one frequency. For this reason, groups of frequencies with different mesh resolutions divided the simulations (Figure 5.34), varying the maximum element size according to the maximum frequency estimated, as is shown in Table 5.2. Another parameter defined in the model was the algebraic linear system, which solved the matrix equation resulting from the spatial discretization. For this model the MUMPS solver was applied, which was a method capable of dealing with symmetric and non-symmetric matrices. The application of this procedure allowed running FE simulations in a desktop PC.

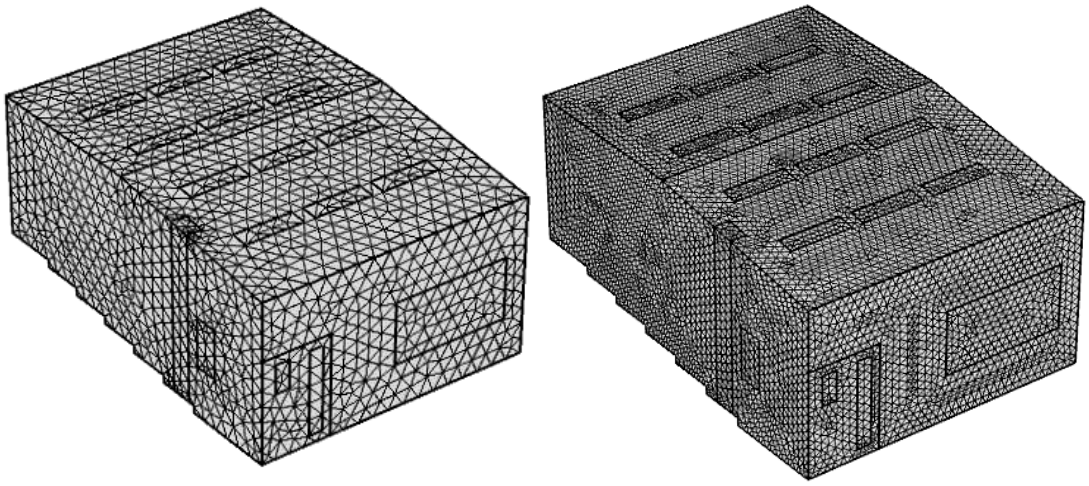


Figure 5.34: The coarsest and the finest mesh resolutions implemented in COMSOL.

5.3.2.2 FE Boundary Conditions

As mentioned in Chapter 3, in FE room acoustic simulations a simple approach relating acoustic impedance with absorption coefficients seems to be the most practical way to model boundary conditions. In this research, the approach given by Aretz (2009) was used (see equation [3.25]), stating that impedance boundary conditions can be defined using a field incidence absorption coefficient to find the resistance part of impedance, which is associated with energy loss by either dissipation or transmission. Taking the above into account, real and frequency dependent impedance values were approximated using the field incidence absorption coefficients applied in the GA model, as can be seen in Table 5.8.

Table 5.8: Real acoustic impedance estimated from absorption coefficients used in GA model, implemented in COMSOL.

Surface	Material	Area (m^2)	Octave Band Centre Frequency (Hz)			
			125	250	500	1000
			Real acoustic impedance (Pa.s/m)			
Floor	Carpet	21.18	20250.4	6095.9	2003.2	1458.7
Door	Wood	1.65	10985.2	15707.2	26758.6	19923.2
Tables and Wood Furniture	Wood	11.90	10985.2	15707.2	26758.6	19923.2
Front Wall	Concrete	9.52	165180.9	165180.9	165180.9	82256.9
Back Wall (door and window)	Plaster	6.87	10191.1	15694.0	26758.6	40627.2
Left Wall	Brick	21.57	82173.8	82173.8	54578.0	54504.1
Right Wall (windows)	Plaster	15.57	10191.1	15694.0	26758.6	40627.2
Ceiling	Plaster	19.93	7430.8	10191.1	6699.8	4669.7
Lights	Metal	5.40	82173.8	82173.8	54578.0	54504.1
Door and Furniture windows	Glass	1.90	3857.7	5759.3	8357.9	12977.2
Windows	Doubled-Glass	7.00	10161.7	31844.1	54356.9	82090.9

5.3.2.3 Source and receivers in the FE approach

Source and receivers were simulated in FE applying the same acoustic conditions and positions used in the GA model. The procedure to model an omnidirectional source in FE simulations was explained in Chapter 3. The source was characterised as a monopole point radiating uniformly 1Pa of acoustic pressure at 1m distance, in a frequency independent spherical propagation. In this sense is important to take into consideration that FE simulations go up to 500 Hz. This implies that differences between GA and FE source directivities should be considered only at the octave bands of 125, 250 and 500 Hz. In this sense, it is important to take into consideration that directivity information used in the GA simulations was given according to the loudspeaker used in the measurements, taking the values from the datasheet provided by the manufacturer. The dissimilarities in terms of DI can be appreciated in Table 5.9. In order to estimate the RIR, the same five receiver positions were defined. To obtain the corresponding BIR, receivers were determined by following the procedure described in Chapter 3. In this case, a cube and two receiver points at corresponding ear positions gave a coarse approximation to the HRTF.

As mentioned in section 5.2.2.3, the implications of DI differences (see Table 5.9) between acoustic sources in the numerical simulations implemented are

considered taking into consideration the results of an analysis of comparative measurements in a workshop on room acoustics (Adrian James Acoustics Limited, 2004). In this exercise a number of room acoustic measurements were carried out in a large room, which size is larger than the Classroom investigated in this thesis. The conclusions of the measurement exercise in the large room indicate that the effect of changing the directionality of the source did not affect the measured reverberation time, neither the EDT. A different situation was given by the measured D_{50} in the large room, which results evidenced a significant dependence on the directionality and on the orientation of the source.

Table 5.9: DI applied in the GA and FE source simulations for the octave bands of 125, 250 and 500 Hz.

Numerical approach	Directivity Index DI (dB) Frequency Octave Bands		
	125 Hz	250 Hz	500 Hz
FE	0	0	0
GA	4.5	6.4	8.3

5.3.3 Classroom objective results

In this section, objective results are presented comparing the numerical approaches used to create the auralizations. For all cases, RIR and BIR results based on the measurements were taken as the reference or ideal condition. In the objective assessment, a comparison between measurements and simulations considered time and frequency responses and room acoustic parameters estimated for both conditions. The objective evaluation of the simulations includes a comparison with reference-measured results of frequency and time responses, and room acoustic parameter estimations. First, the room transfer function responses are presented for particular source-receiver combinations. Second, the results of room acoustic parameters according to ISO 3382-1:2009, such as Reverberation Time, Early Decay Time (*EDT*), Clarity, Definition and Inter-Aural Cross Correlation (*IACC*) are described. This section finalizes with a discussion of the objective results obtained for both rooms, in order to quantify the accuracy of both numerical studied in this thesis to estimate acoustic parameters.

5.3.3.1 Time domain room transfer functions results of the Classroom

This section presents measurements and simulation results of RIR and BIR for a particular source-receiver combination. In order to facilitate the comparison of the numerical approaches implemented, the impulse responses were filtered in two different frequency ranges. A frequency range from 80 Hz to 400 Hz is used to visualize the impulse responses obtained with FE. In order to see the impulse responses simulated by both numerical approaches, a wide frequency range from 80 Hz to 20 kHz was used. In order to illustrate the RIR results, three of the five source-receiver position combinations are described. The following figures (from Figure 5.35 to **¡Error! No se encuentra el origen de la referencia.**) present the RIRs obtained for receiver positions number one, two and three (see section 4.3.2).

In Figure 5.35, it can be seen that the time of arrival of direct sound in the RIRs coincide largely with the highest value in magnitude in the modelling with FEM in sample 905 (0.0205 seconds) and in the measured RIR in sample 912 (0.0207 seconds). The RIR simulated by FEM, in sample 1128 (0.0256 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 5.0 ms. As for the measured RIR, the peak which could correspond to the first reflection is in sample 1050 (0.0238 seconds), which would mean an ITDG of 3.2 ms. In both cases the peaks do not match in magnitude. Moreover, a peak with greater level is observed in the measured RIR that even exceeds the peak of direct sound, which does not occur in the RIR simulated by FEM.

In Figure 5.36, the measured RIR peak with the greatest magnitude is in sample 791 (0.0179 seconds) while in the RIRs simulated by GA and FEM-GA this peak is located in sample 779 (0.0177 seconds). A second peak is presented, which could correspond to the first reflection in sample 910 (0.0206 seconds) in the measured RIR, whereby an ITDG equal to 2.7 ms would be obtained. Similarly, in the numerical RIRs this peak is in sample 871 (0.0198 seconds) that would give a result of ITDG of 2.1 ms. The following early reflections are similar in magnitude, except in some peaks around the sample 2932 (0.0665 seconds) in which the magnitude is greater in the simulated RIRs with regard to the measured RIR.

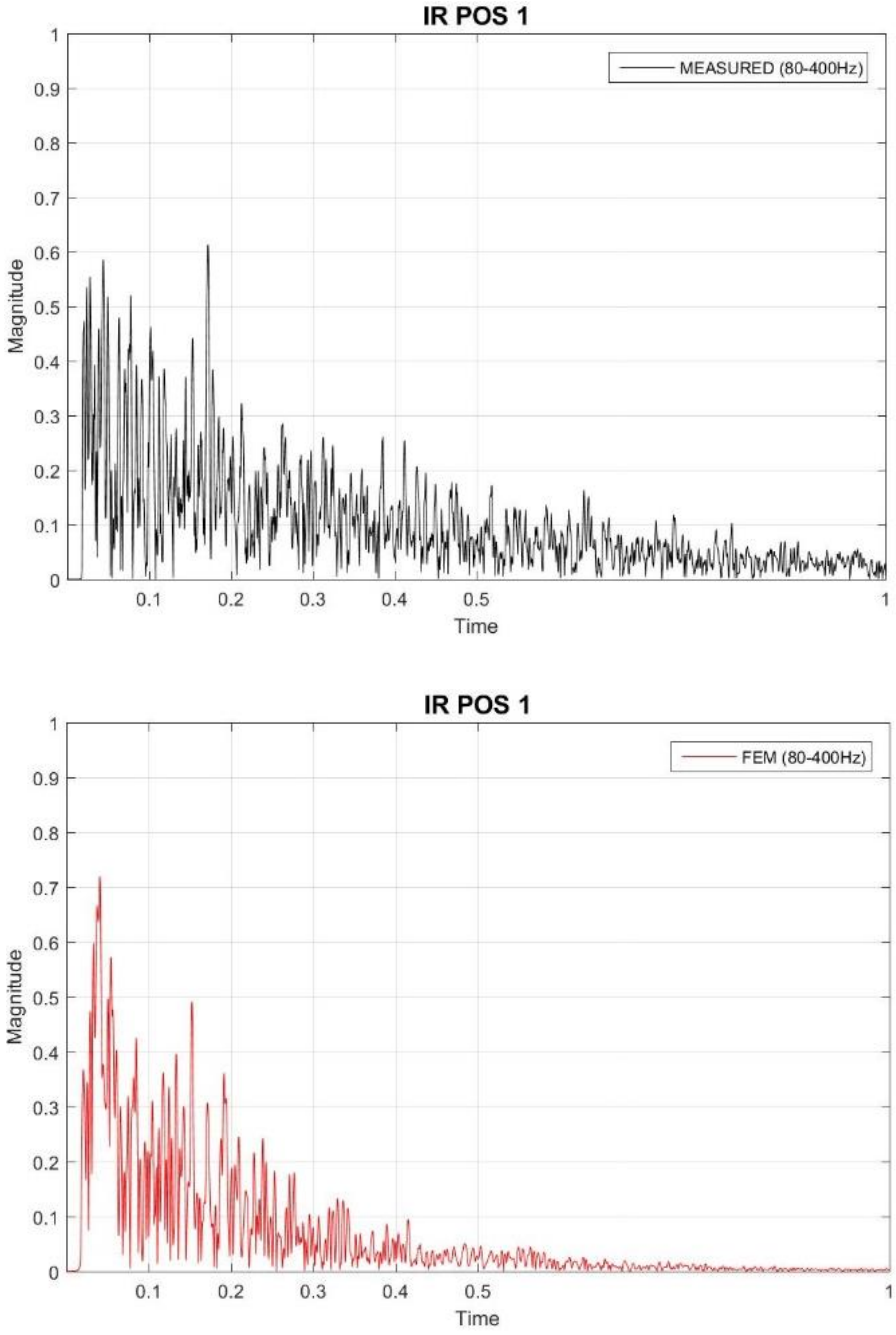


Figure 5.35: Top, the measured RIR at position 1. Bottom, the estimated FEM RIR for the same position.

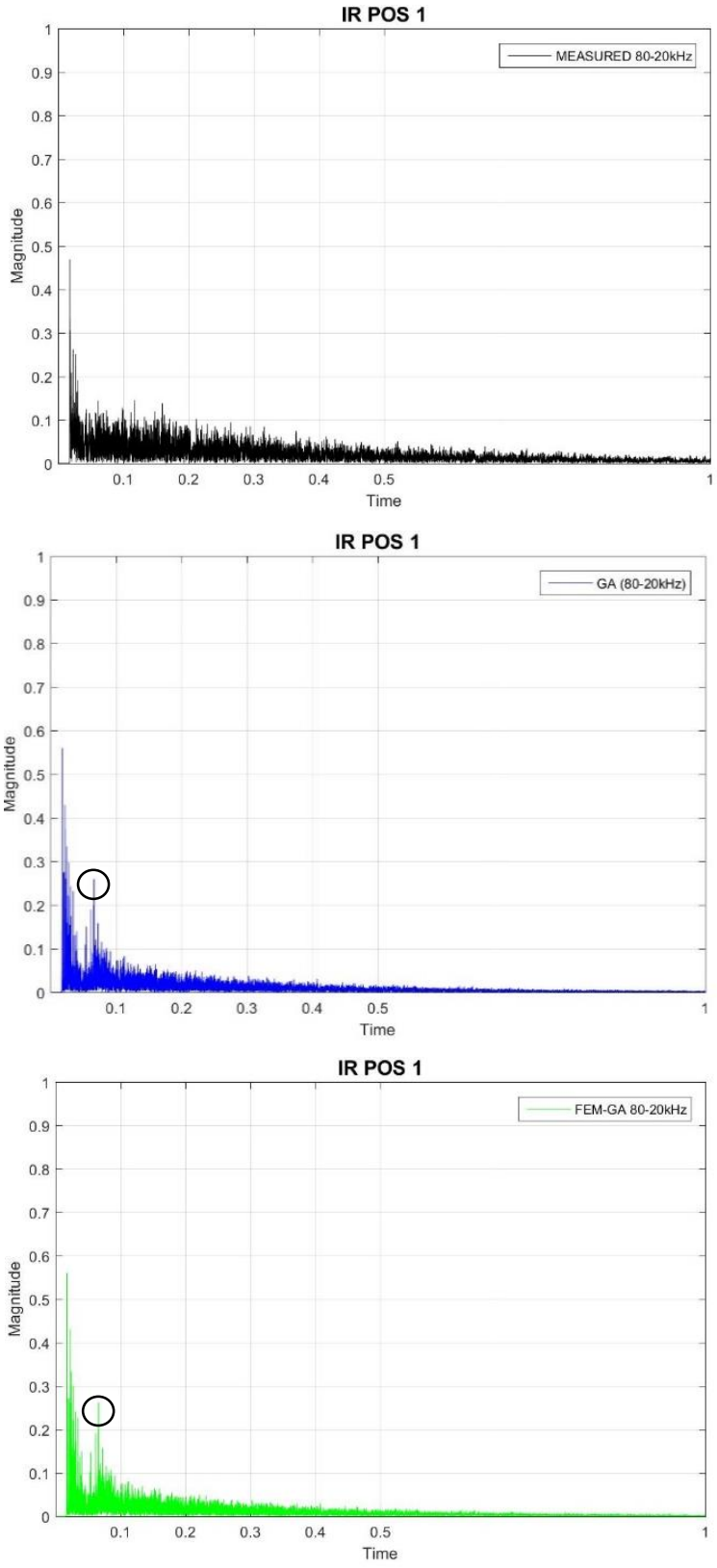


Figure 5.36: RIRs obtained at position 1. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

Figure 5.37 compares measured RIR against simulated FEM RIR for position 2. It can be seen that the arrival of the direct sound in the simulated RIR is at sample 980 (0.0222 seconds) and in the measured RIR at sample 1084 (0.0246 seconds). The magnitudes of the first peaks in the RIRs are very similar, around 0.55 in the one simulated by FEM and approximately 0.5 in the measured RIR. By determining the ITDG in both responses, it is obtained that in the measured response it corresponds to 3.5 ms, whereas in the simulated case it is 5.3 ms.

Figure 5.38 compares measured RIR against numerical RIRs obtained by the numerical approaches of GA and FEM-GA, for position 2. It can be seen that the time of arrival of the direct sound is quite similar in all responses, with the first peak in sample 813 (0.0184 seconds) in the measured RIR, and in sample 795 (0.0180 seconds) for the simulated RIRs. Nevertheless, the magnitude of the peaks differs significantly, being higher in the numerical RIRs. The ITDG in the measured RIR is 2.5 ms, whereas in the simulated RIRs is 2.0 ms. The following peaks of the early reflections are seen with great similarity across all RIRs, however, they differ in magnitude, having greater level the peaks in the numerical RIRs around sample 2918 (0.0662 seconds).

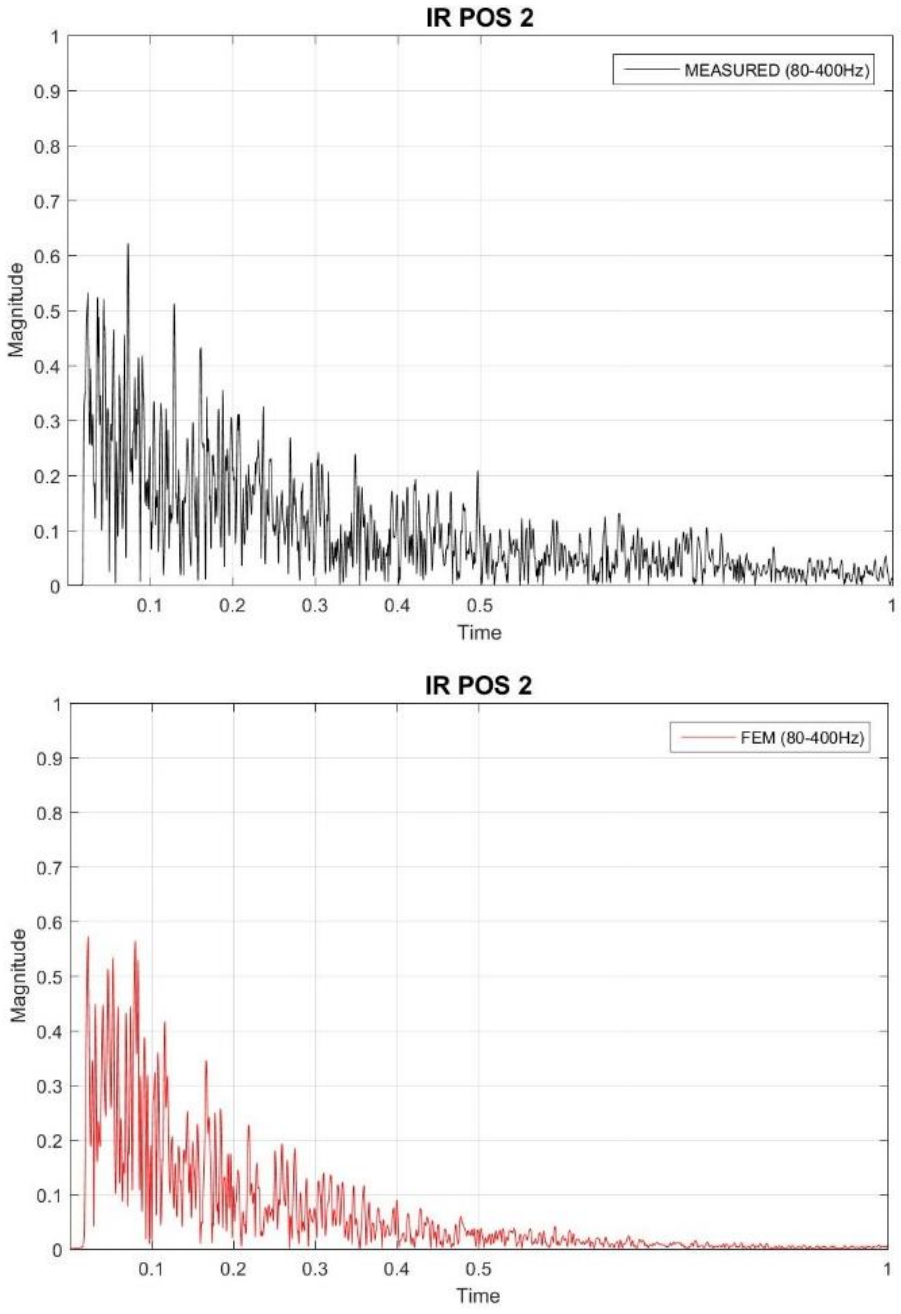


Figure 5.37: Top, the measured RIR at position 2. Bottom, the estimated FEM RIR for the same position.

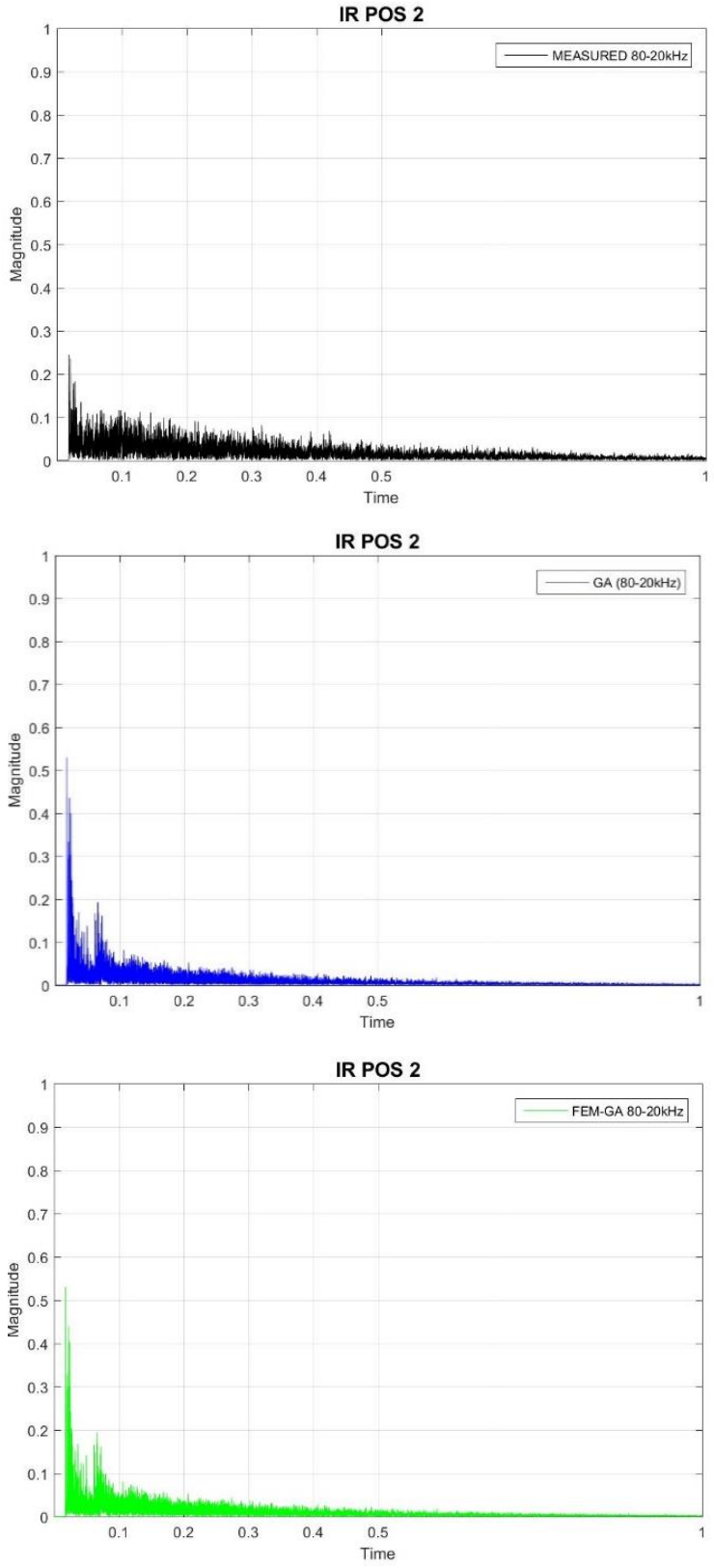


Figure 5.38: RIRs obtained at position 2. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

Figure 5.39 compares measured RIR against simulated FEM RIR for position 3. It can be seen that the time of arrival of the direct sound is quite similar in both responses, having the first peak in the measured case at sample 448 (0.0102 seconds) and in the simulated RIR at sample 455 (0.0103 seconds). The ITDG in the measured case is around 8.5 ms while in the simulated response is about 7.7 ms. The early reflections differ in magnitude, being higher in the FEM RIR.

Figure 5.40 compares measured RIR against simulated RIRs obtained by the numerical approaches of GA and FEM-GA, for position 3. The first peak is found in sample 376 (0.0085 seconds) in the measured RIR, while in the numerical RIRs it is found about sample 364 (0.0083 seconds), although differing significantly in magnitude. A second peak is presented in sample 544 (0.0123 seconds), which could correspond to the first reflection in the measured RIR, whereby an ITDG equal to 3.8 ms would be obtained. Similarly, in the RIRs obtained by GA and FEM-GA this peak is in sample 546 (0.0124 seconds) that would give an ITDG of 4.1 ms.

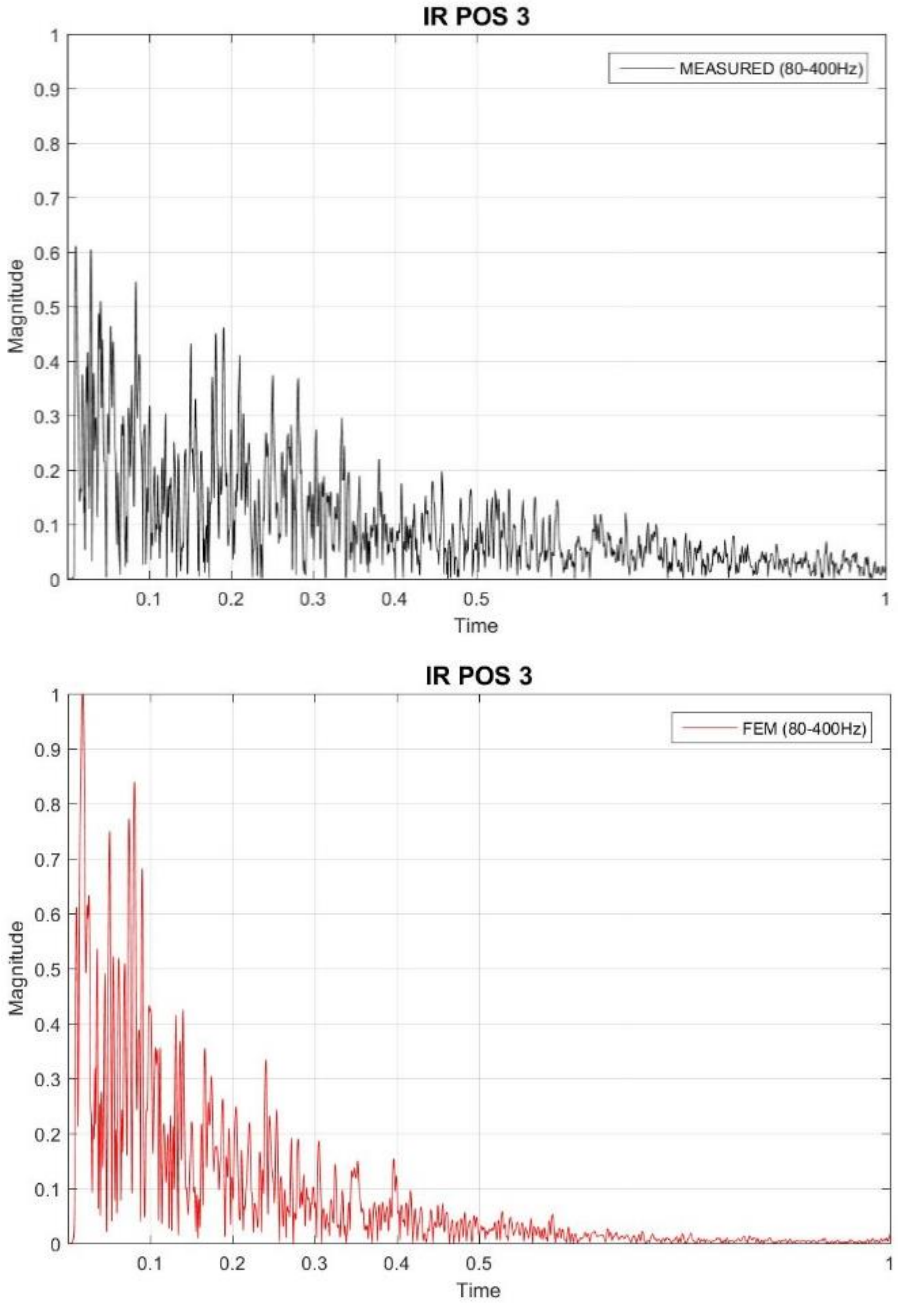


Figure 5.39: Top, the measured RIR at position 3. Bottom, the estimated FEM RIR for the same position.

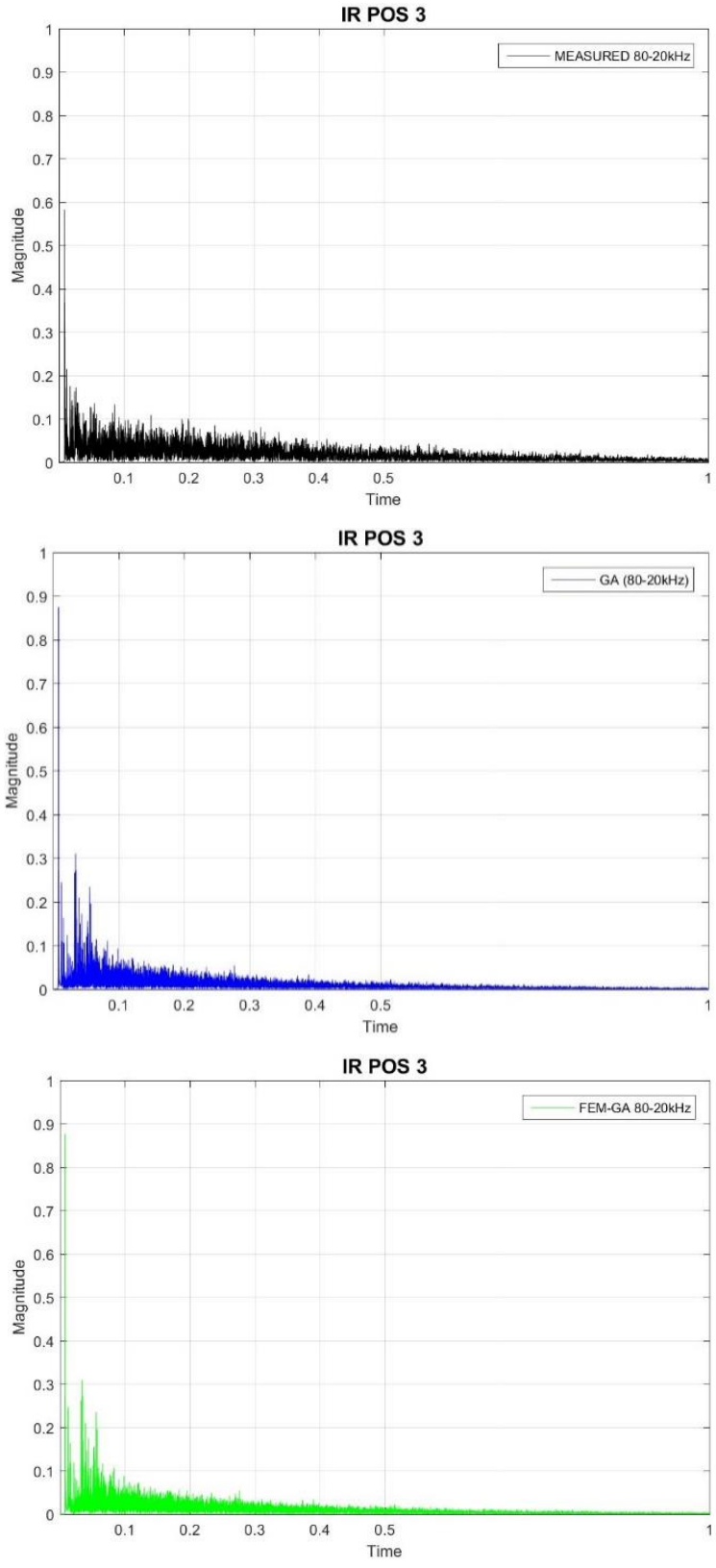


Figure 5.40: RIRs obtained at position 3. Top, the measured RIR. Middle, the estimated GA RIR. Bottom, the estimated FEM-GA RIR.

From Figure 5.41 to Figure 5.46**¡Error! No se encuentra el origen de la referencia.** the BIRs obtained by means of acoustic measurements and numerical simulations are presented for the same positions.

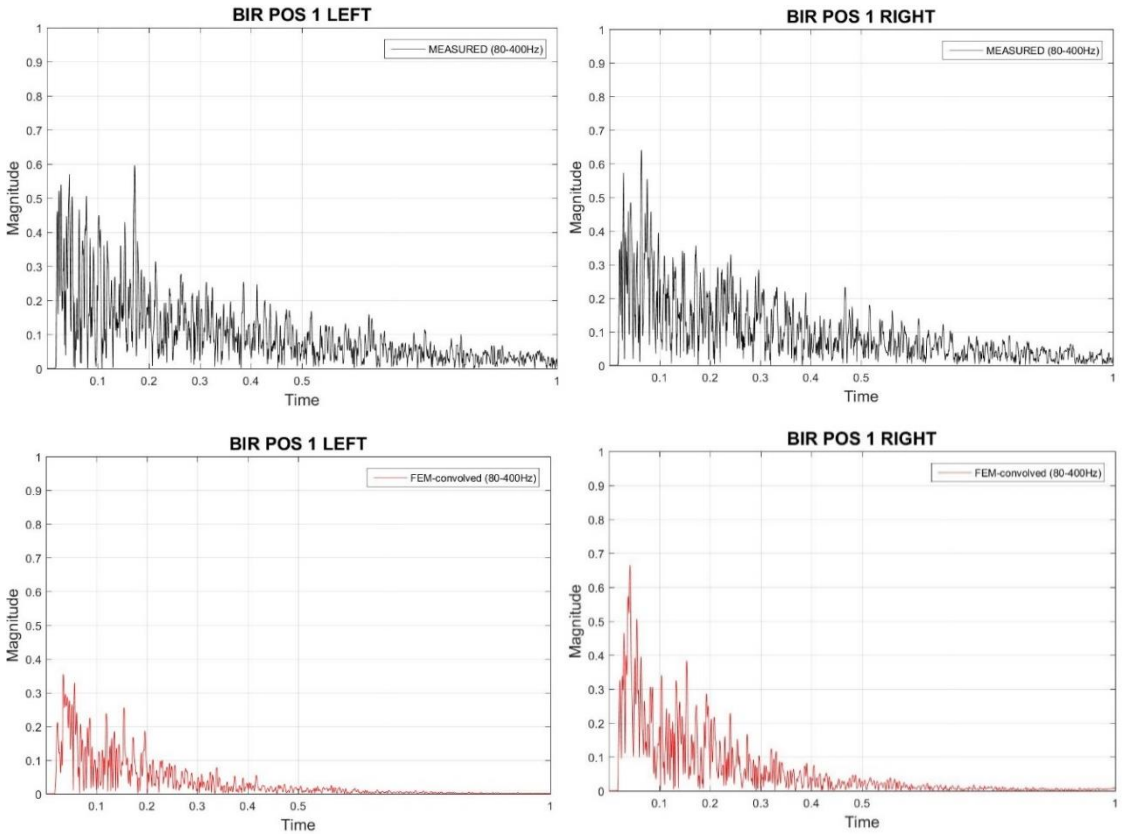


Figure 5.41: Top, the measured BIR at position 1. Bottom, the estimated FEM BIR for the same position.

In Figure 5.41, it can be observed that the time of arrival of the direct sound in the measured and simulated BIR coincide largely. In the left responses the first peak is located at sample 912 (0.0207 seconds) for the measured BIR and at sample 996 (0.0205 seconds) for the simulated BIR. In the right responses the first peak is located at sample 906 (0.0205 seconds) for the measured BIR and at sample 964 (0.0219 seconds) for the simulated BIR. The ITDG that appears on the left measured BIR is 3.1 ms and on the left simulated BIR is 2.9 ms. Similarly, the ITDG that appears on the right measured BIR is 3.1 ms, whereas in the simulated BIR is 3.5 ms.

Figure 5.42 compares measured BIR against GA and FEM-GA simulated BIRs for position 1. It can be seen that the direct sound coincide largely for measured and numerical approaches. In the left BIRs, the first peak is in sample 791 (0.0179 seconds) for the measured case and sample 786 (0.0178 seconds) for

the numerical responses. In the left measured BIR, at sample 910 (0.0206 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 2.7 ms. As for the left BIRs simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 907 (0.0206 seconds), which would mean an ITDG of 2.8 ms. In the right BIR case, a second peak is found at sample 914 (0.0207 seconds) in the measured response, resulting in an ITDG of 2.7 ms. As for the simulated BIRs, a second peak is found at sample 880 (0.02 seconds), which corresponds to an ITDG of 2.1ms.

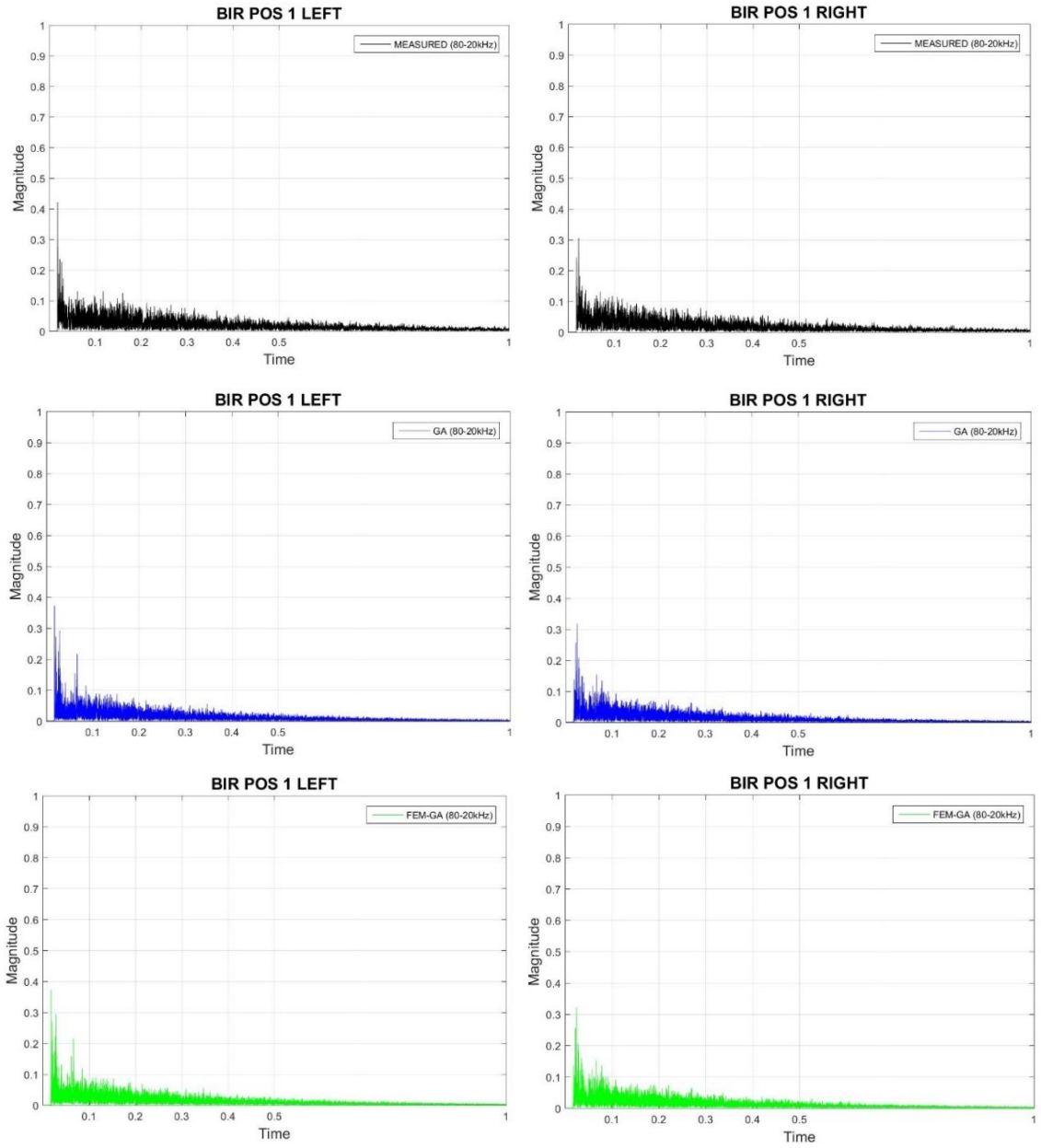


Figure 5.42: Measured and simulated BIR at position 1. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

In Figure 5.43, it can be seen that the time of arrival of the direct sound in the measured and simulated BIRs coincide largely. In the left responses the first peak is located at sample 1082 (0.0245 seconds) for the measured BIR and at sample 968 (0.022 seconds) for the simulated BIR. In the right responses the first peak is located at sample 1044 (0.0237 seconds) for the measured BIR and at sample 985 (0.0223 seconds) for the simulated BIR. The ITDG that appears on the left measured BIR is 3.5 ms and on the left simulated BIR is 6.0 ms. Similarly, the ITDG that appears on the right measured BIR is 5.3 ms, whereas in the simulated BIRs is 5.5 ms.

Figure 5.44 compares measured BIR against GA and FEM-GA simulated BIRs for position 2. It can be seen that the time of arrival of the direct sound in the measured and simulated BIRs coincide largely. In the left responses the first peak is located at sample 813 (0.0184 seconds) for the measured BIR and at sample 805 (0.0183 seconds) for the simulated BIRs obtained by GA and FEM-GA. In the right responses the first peak is located at sample 807 (0.0183 seconds) for the measured BIR and at sample 802 (0.0182 seconds) for the simulated BIRs. In the left measured BIR, at sample 925 (0.021 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 2.6 ms. As for the left BIR simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 907 (0.0206 seconds), which would mean again an ITDG of 2.3 ms. In the right BIRs case, a second peak is found at sample 930 (0.0211 seconds) in the measured response, resulting in an ITDG of 2.8 ms. As for the simulated BIRs, a second peak is found at sample 921 (0.0209 seconds), which corresponds to an ITDG of 2.7 ms.

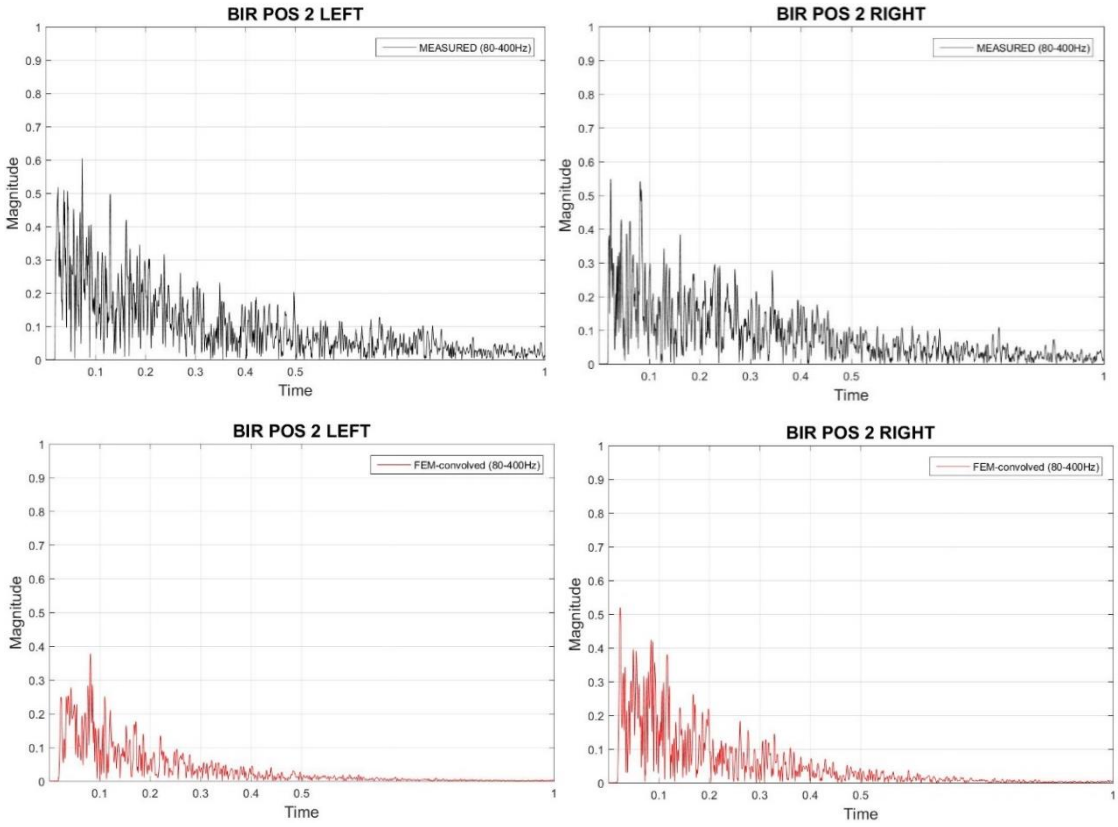


Figure 5.43: Top, the measured BIR at position 2. Bottom, the estimated FEM BIR for the same position.

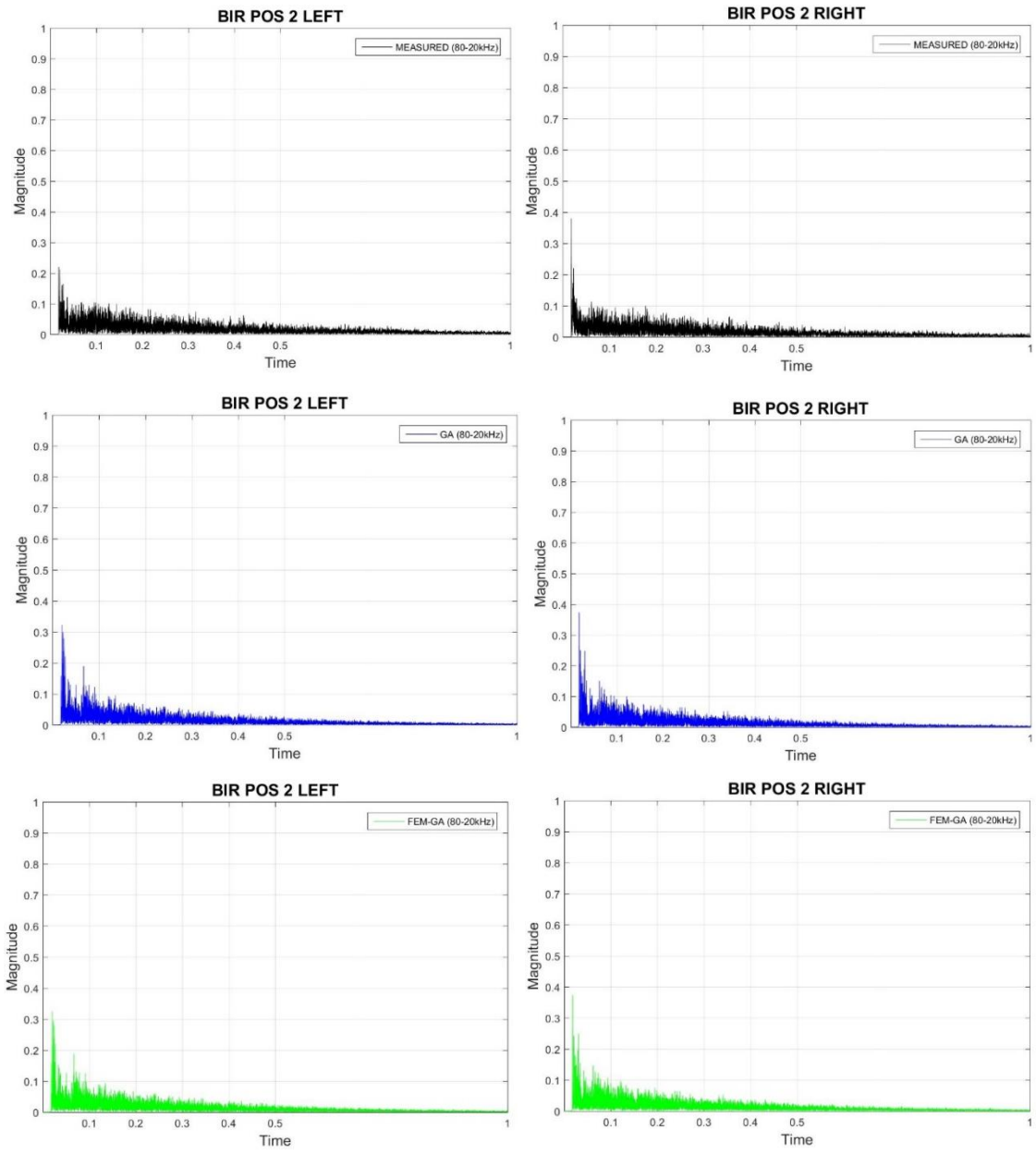


Figure 5.44: Measured and simulated BIR at position 2. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

Figure 5.45 compares measured BIR against FEM simulated BIRs for position 3. It can be seen that the time of arrival of the direct sound in the measured and simulated BIR coincide largely. In the left responses the first peak is located at sample 492 (0.0112 seconds) for the measured BIR and at sample 459 (0.0104 seconds) for the simulated BIR. In the right responses the first peak is located at sample 465 (0.0105 seconds) for the measured BIR and at sample 454 (0.0103 seconds) for the simulated BIR. The ITDG that appears on the left measured BIR is 4.9 ms and on the left simulated BIR is 3.4 ms. Similarly, the ITDG that appears on the right measured BIR is 7.8 ms, whereas in the simulated BIRs is 6.3 ms.

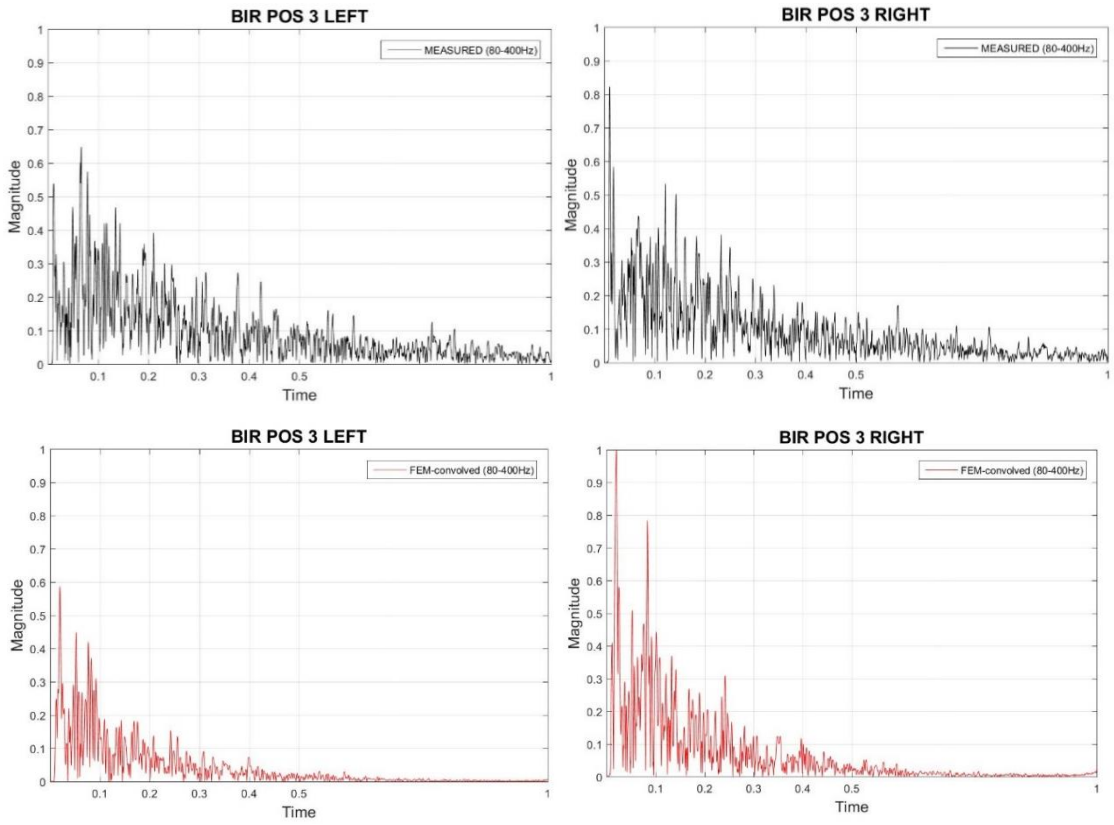


Figure 5.45: Top, the measured BIR at position 3. Bottom, the estimated FEM BIR for the same position.

Figure 5.46 compares measured BIR against GA and FEM-GA simulated BIRs for position 3. It can be seen that the time of arrival of the direct sound in the measured and simulated BIRs coincide largely. In the left responses the first peak is located at sample 399 (0.009 seconds) for the measured BIR and at sample 380 (0.0086 seconds) for the simulated BIRs obtained by GA and FEM-GA. In the right responses the first peak is located at sample 378 (0.0086 seconds) for the measured BIR and at sample 369 (0.0084 seconds) for the simulated BIRs. In the left measured BIR, at sample 618 (0.014 seconds) a second peak is found, which could correspond to the first reflection, therefore the ITDG in this case would be 5.0 ms. As for the left BIR simulated by GA and FEM-GA, the peak which could correspond to the first reflection is in sample 601 (0.0136 seconds), which would mean again an ITDG of 5.0 ms. In the right BIRs case, a second peak is found at sample 564 (0.0128 seconds) in the measured response, resulting in an ITDG of 4.2 ms. As for the simulated BIRs, a second peak is found at sample 548 (0.0124 seconds), which corresponds to an ITDG of 4.0 ms.

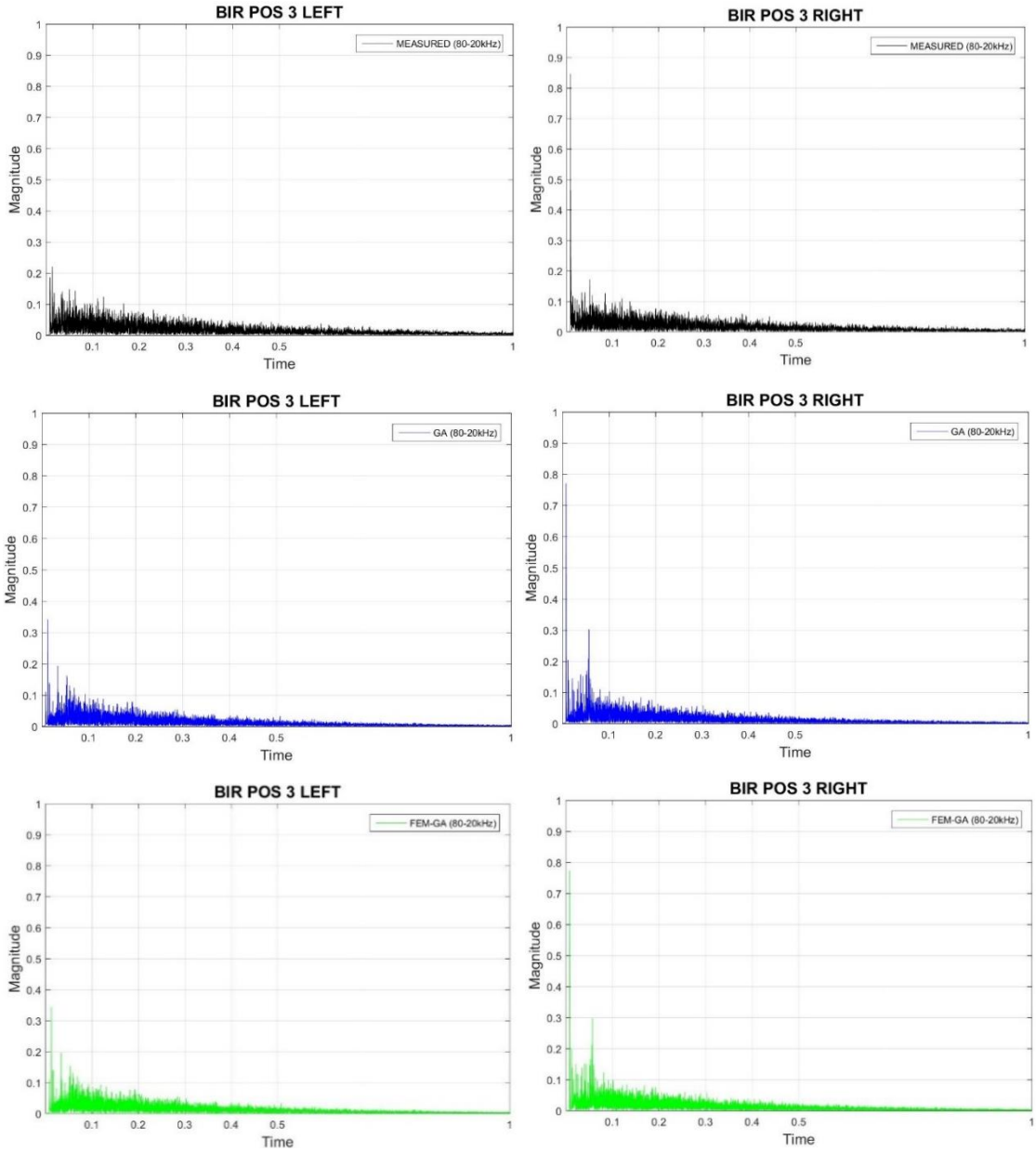


Figure 5.46: Measured and simulated BIR at position 3. Top, the measured BIR. Middle, the estimated GA BIR. Bottom, the simulated FEM-GA BIR.

5.3.3.2 Frequency domain room transfer functions results of the Classroom

In this section, the frequency responses simulated are compared against measured frequency responses. From Figure 5.47 to Figure 5.49 the RFR are presented for positions 1, 2 and 3 (see section 4.3.2). The RFRs figures were obtained by applying a FFT to the normalised RIRs, having as the reference the measured RIRs. The RFRs graphs are split up in two frequency ranges for clarity. First, a lower frequency range from 80 Hz to 500 Hz is used to visualise the simulated GA and FEM responses against the measured RFR. Second, a frequency

range starting at 500 Hz until the upper limit of the 4 kHz octave band is used in order to compare GA RFR and measured response.

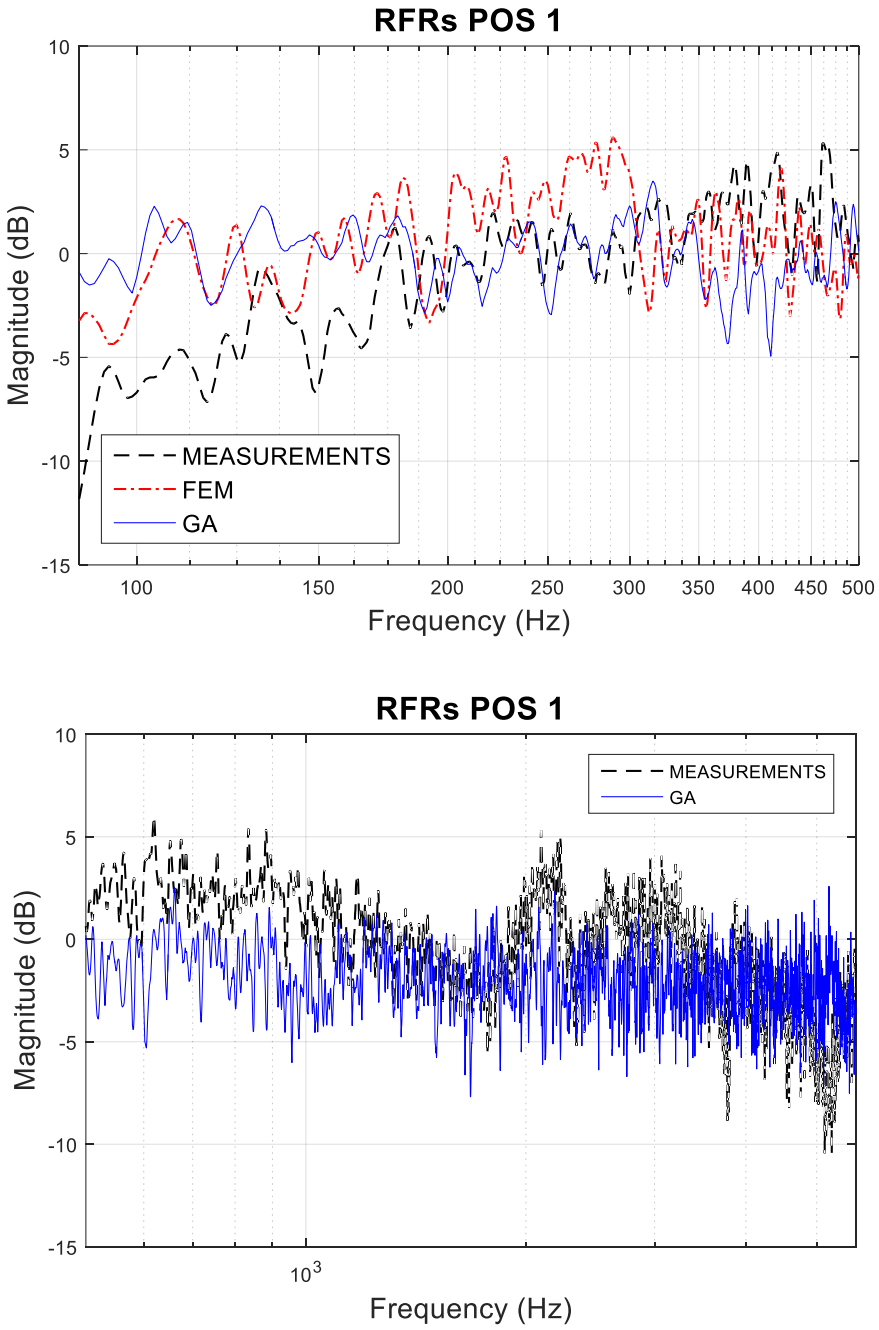


Figure 5.47: RFRs obtained for position 1. Top, measured and simulated by GA and FEM RFRs up to 500 Hz. Bottom, measured and GA RFRs from 500 Hz up to 5.6 kHz.

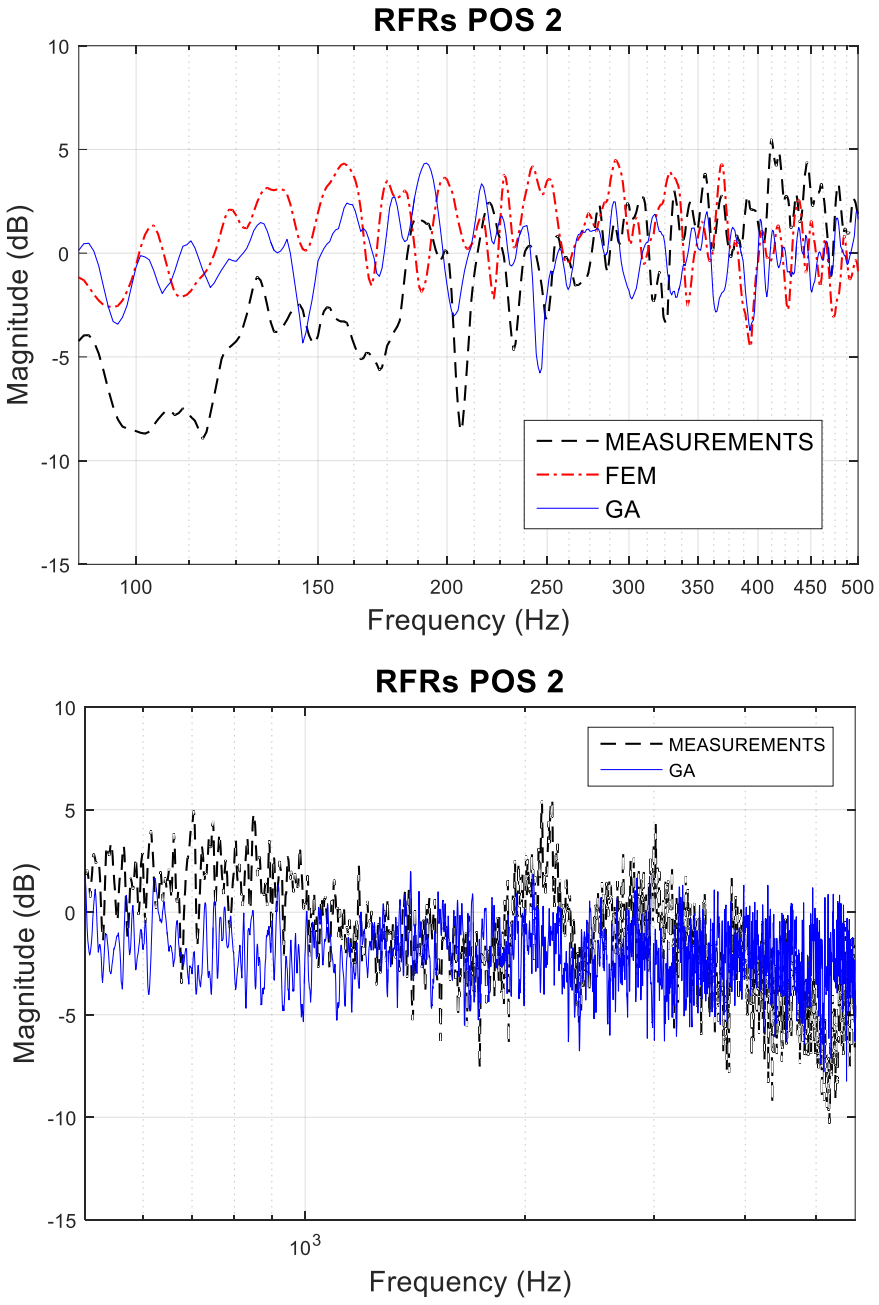


Figure 5.48: RFRs obtained for position 2. Top, measured and simulated by GA and FEM RFRs up to 500 Hz. Bottom, measured and GA RFRs from 500 Hz up to 5.6 kHz.

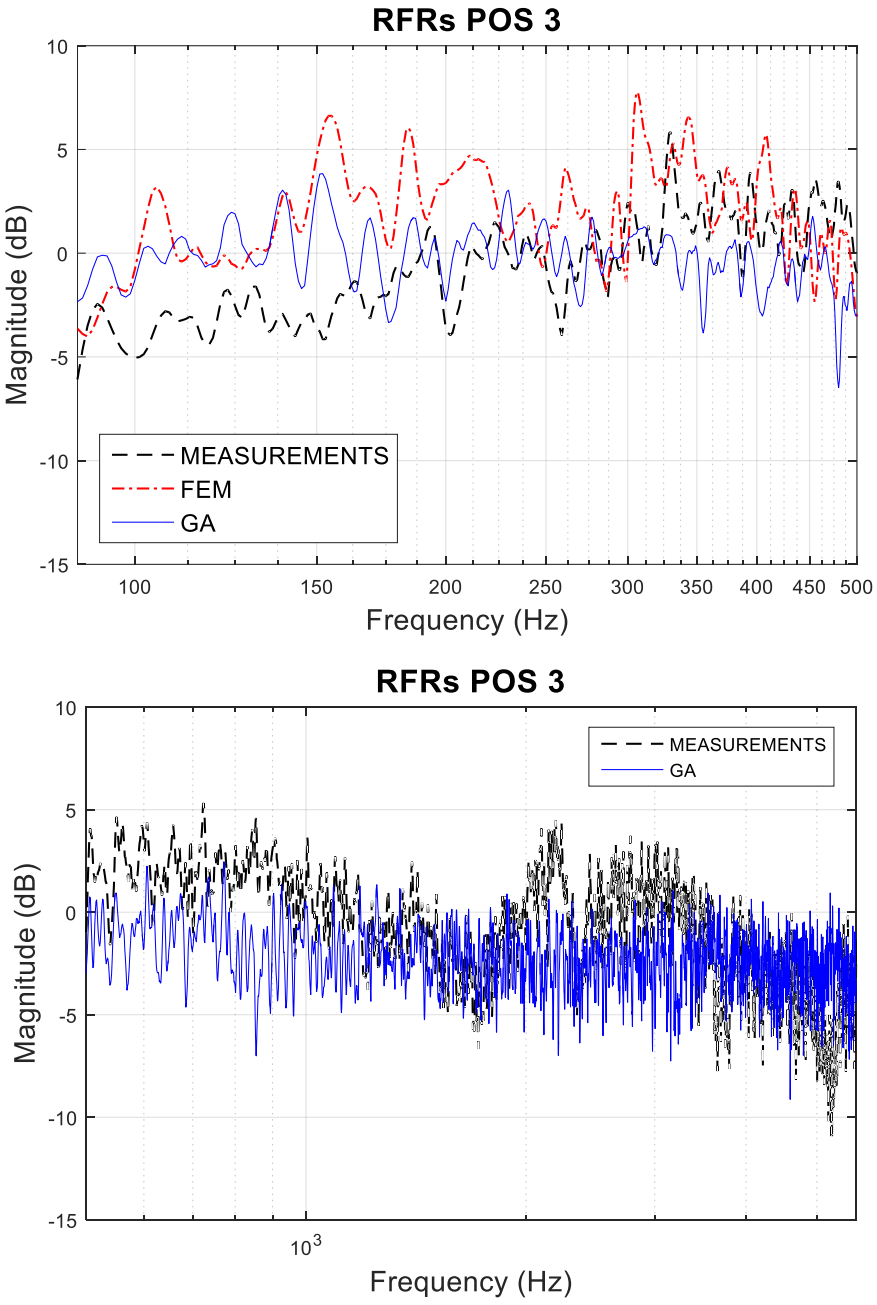


Figure 5.49: RFRs obtained for position 3. Top, measured and simulated by GA and FEM RFRs up to 500 Hz. Bottom, measured and GA RFRs from 500 Hz up to 5.6 kHz.

5.3.3.3 Room acoustic parameters results of the Classroom

In this section, the spatially averaged room acoustic parameters results such as T_{20} (see Figure 5.50), EDT (see Figure 5.51), D_{50} (see Figure 5.52), C_{80} (see Figure 5.53) and *IACC* (see Figure 5.54) are presented according to ISO standard 3382 (2009). The last parameter was estimated from the measured and simulated BIR.

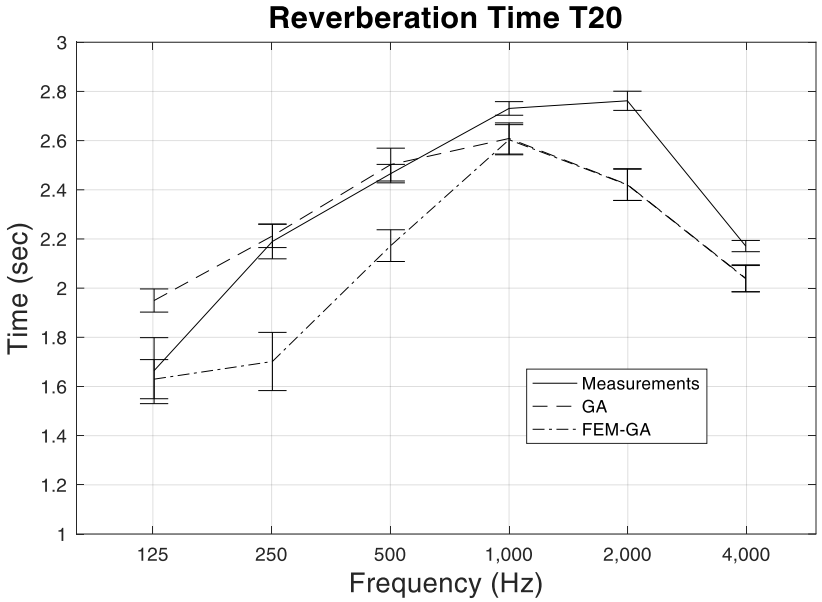


Figure 5.50: Spatially averaged T_{20} results of the Classroom calculated, applying the integrated impulse response method according to ISO 3382, from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

In Figure 5.50 T_{20} spatially averaged over the six source-receiver position combinations obtained by means of measurements, GA and FEM-GA hybrid approach can be seen. It is noted that results obtained by GA have a better agreement than FEM-GA with respect to the measured T_{20}, in the octave bands of 250 Hz and 500 Hz. It can be appreciated how T_{20} estimated by FEM-GA are underestimated in comparison to measurements for all the frequencies.

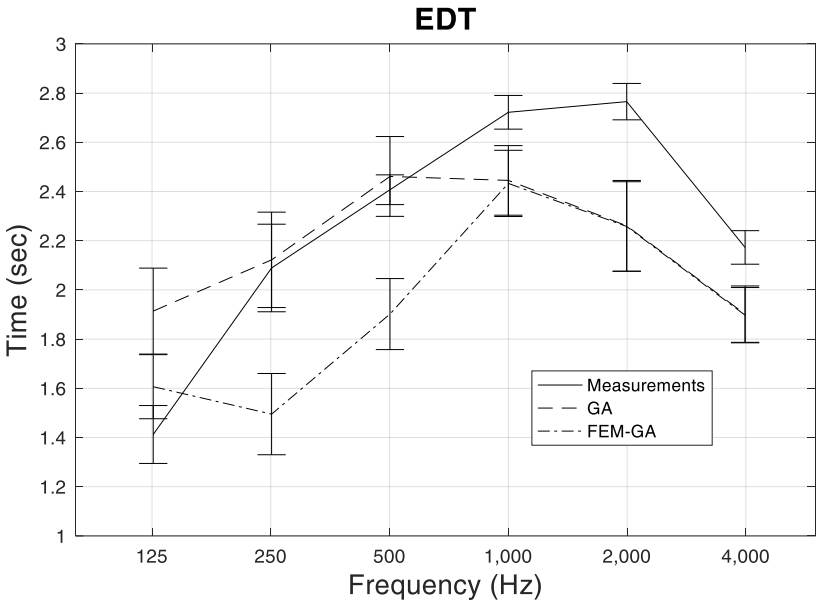


Figure 5.51: Spatially averaged *EDT* results of the Classroom calculated, applying the integrated impulse response method according to ISO 3382, from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

Figure 5.51 presents EDT spatially averaged over the six source-receiver position combinations obtained by means of measurements, GA and FEM-GA. It is noted that results obtained by GA have a better agreement than FEM-GA with respect to the measured EDT, in the octave bands of 250 Hz and 500 Hz. It can be appreciated how EDT estimated by FEM-GA are underestimated in comparison to measurements for all the frequencies but 125 Hz.

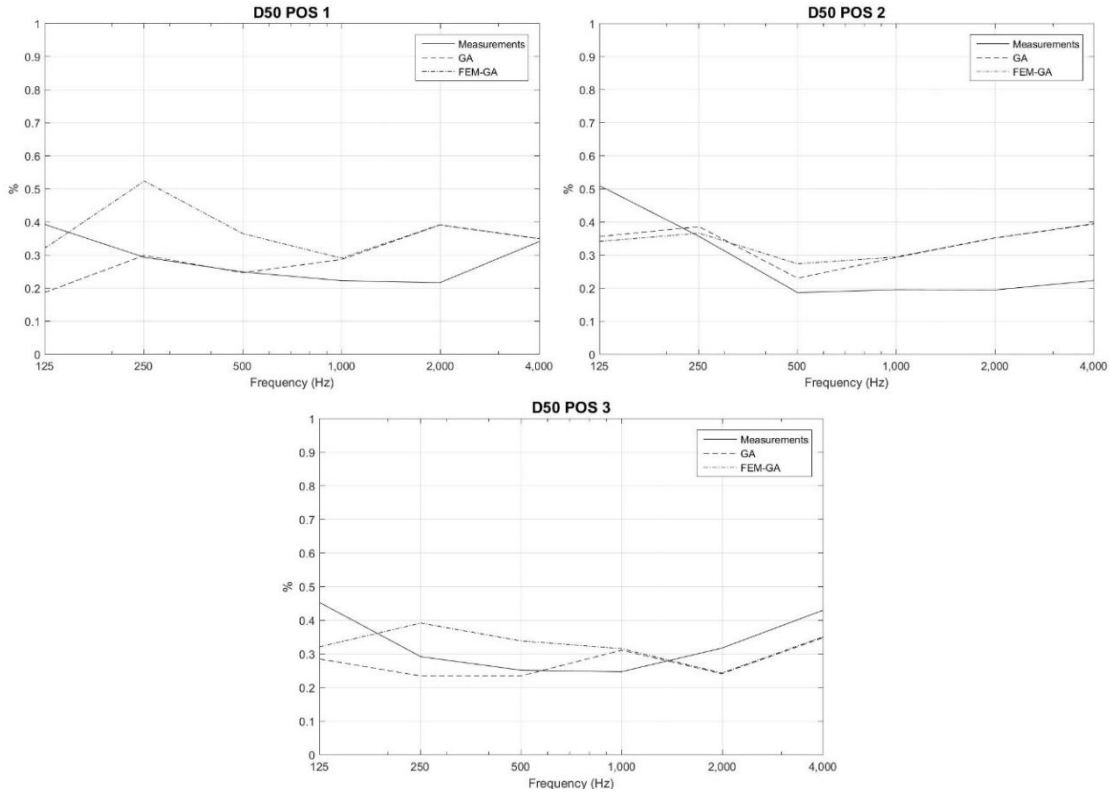


Figure 5.52: D_{50} estimates results by means of measurements, GA simulations and the numerical approach combination of FEM-GA, for positions 1, 2 and 3. D_{50} results of the Classroom for positions 1, 2 and 3 calculated from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

Figure 5.52 shows the D_{50} results obtained from each measured RIR and estimated by GA and FEM-GA, at three different positions in the Classroom. In general, results obtained by GA are closer to D_{50} values measured in the octave bands of 250 Hz and 500 Hz. In position 1 the values obtained by GA and FEM-GA differ significantly between them. In position 2, numerical results are very similar to measured values to a greater extent in the frequency bands of 250 Hz and 500 Hz. For the octave bands of 1 kHz, 2 kHz and 4 kHz GA D_{50} results are overestimated with respect to measurement results in positions 1 and 2.

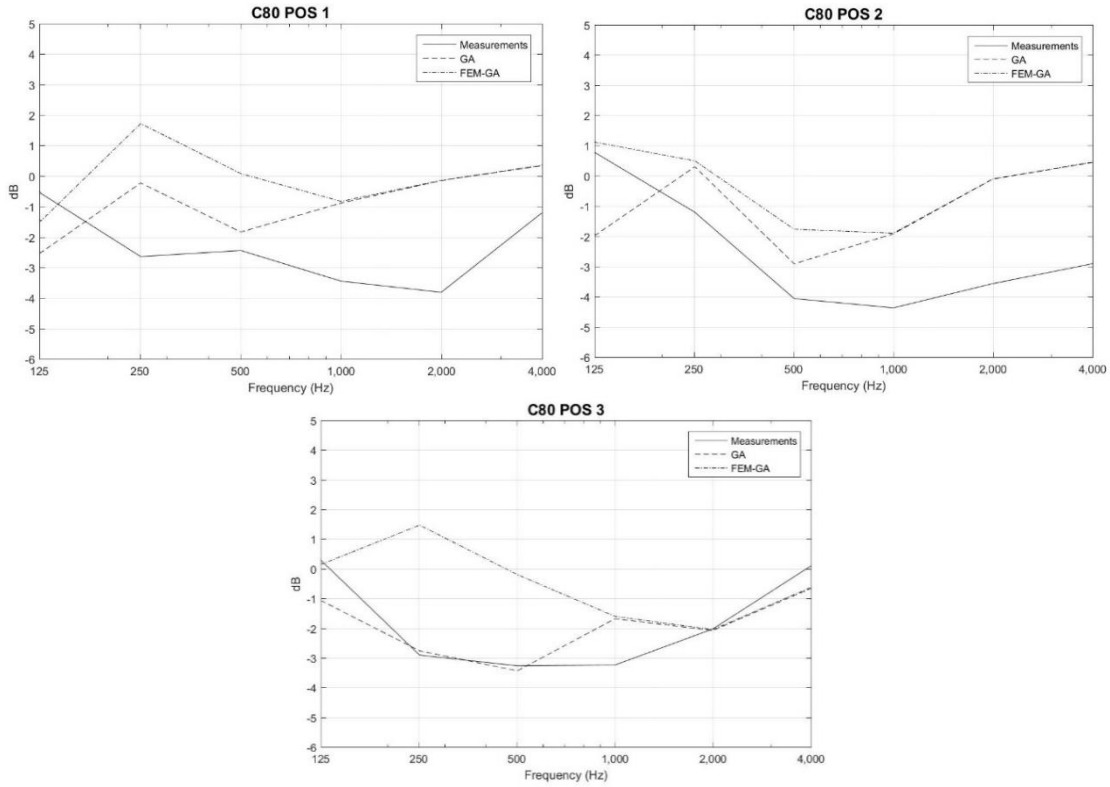


Figure 5.53: C_{80} results of the Classroom for positions 1, 2 and 3 calculated from RIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

Figure 5.53 shows the C_{80} results obtained from each measured RIR and estimated by GA and FEM-GA, at three different positions in the Classroom. It can be appreciated that for all positions, GA C_{80} results are in general closer to measured C_{80} values in the frequencies of 250 Hz and 500 Hz. In these frequencies, position 3 gives results that demonstrate good agreement between measurements and GA approach. GA results of C_{80} are overestimated with respect to measured values for the octave bands of 1 kHz, 2 kHz and 4 kHz.

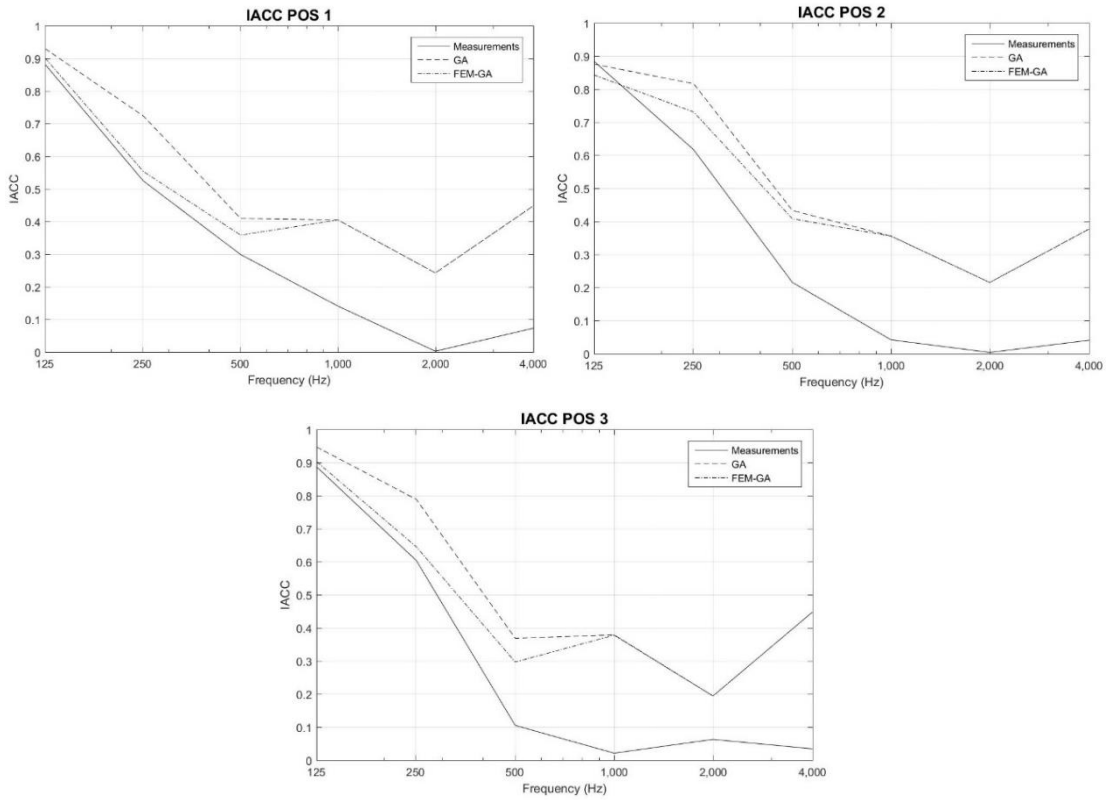


Figure 5.54: *IACC* results of the Classroom for positions 1, 2 and 3 calculated from BIRs obtained by measurements, GA simulations and the numerical approach combination of FEM-GA.

In Figure 5.54 can be appreciated how the numerical approaches present similar trends of IACC results between them for all positions. Position 1 presents the most similar results between measurements and numerical approaches. Position 2 and 4 exhibit significant differences for IACC results in the octave band of 500 Hz. In this case, numerical results of IACC are overestimated with respect to measured values for all the octave bands. It can be appreciated that for all positions, GA IACC results are in general closer to measured IACC values in all the frequencies.

5.4 Discussion of objective results

In this section, the objective results obtained from the simulations of the rooms investigated in this thesis are discussed in order to quantify the accuracy of the numerical approaches implemented. For this, the three groups of objective results are discussed independently: time domain room transfer functions, frequency domain room transfer functions and room acoustic parameters. In all

cases, the results of the acoustic measurements are taken as the reference or ideal condition.

5.4.1 Discussion of time domain room transfer function results

In order to analyse the measured and simulated time domain room transfer functions of sections 5.2.3.1 and 5.3.3.1, the ITDG values resulting from the RIRs are considered. A percentage of error is calculated according to expression [5.19], taking as the reference value the measured ITDG result of the corresponding RIR. In Table 5.10 the measured and simulated ITDG results obtained for both rooms investigated can be seen. According to the % of error of the table, in the Meeting Room the RIRs estimated applying the FEM exhibit ITDG values with a less percentage of error than the ITDG values obtained from the RIRs simulated by means of GA. On the other hand, in the Classroom the highest percentage of error is presented in the ITDG values obtained from the RIRs estimated by means of the FEM.

Table 5.10: ITDG results obtained from measured and numerical RIRs of the rooms investigated, taking into account the frequency ranges estimated in the FEM and % of error according to the measured reference.

Room	Receiver Position	ITDG (ms)				%Error	
		RIR (Filtered 80-600Hz Meeting Room and 80-400Hz Classroom)		RIR (Filtered 80-20kHz)			
		FEM	Measurements	GA	Measurements	FEM	GA
Meeting Room	Pos. 2	6.1	6.3	2.4	2.0	3%	20%
	Pos. 3	6.1	6.3	2.6	2.9	3%	10%
	Pos. 4	5.2	6.0	2.8	1.8	13%	56%
Classroom	Pos. 1	5.0	3.2	2.1	2.7	56%	22%
	Pos. 2	5.3	3.5	2.0	2.5	51%	20%
	Pos. 3	7.7	8.5	4.1	3.8	9%	8%

5.4.2 Discussion of frequency domain room transfer function results

In order to analyse the measured and simulated frequency domain room transfer functions of sections 5.2.3.2 and 5.3.3.2, the coherence function has been taken as a measure of similarity between responses. This function is an absolute measure of how well two signals are linearly related. To compute and plot the coherence, the magnitude squared coherence estimate in MATLAB® software has

been used. This estimate is a function of frequency with values between 0 and 1, indicating how well one signal corresponds to the other at each frequency. In order to evaluate the estimated RFRs, each numerical approach has been applied a coherence function having as a reference signal the measured RFR at the corresponding position. This means that for a particular position, two plots are related, one realising a coherence function between the FEM approach and the measured reference and a second figure realising the GA response with respect to the measured RFR. In all cases, the magnitude squared coherence estimate was applied using the following input arguments: a Hamming window, a 50% of sections overlap and a FFT length of 512.

Figure 5.55 to Figure 5.57 present the magnitude square coherence estimate for the RFRs of the Meeting Room (see section 5.2.3.2). The plots on the left side relate the coherence estimates between the FEM RFRs and the measured responses. The plots on the right side are the coherence estimates between the GA RFRs and the measured responses. It can be seen in these figures that, in general, the FEM approach presents a better correlation with the measured reference in two out of the three positions analysed in the Meeting Room, in positions 2 and 4.

Figure 5.58 to Figure 5.60 present the magnitude square coherence estimate for the RFRs of the Classroom (see section 5.3.3.1). As mentioned previously, the plots on the left side relate the coherence estimates between the FEM RFRs and the measured responses and the plots on the right side are the coherence estimates between the GA RFRs and the measured responses. It can be seen in these figures that the GA approach presents a better correlation with the measured reference in two out of the three positions analysed in the Classroom, in positions 1 and 3.

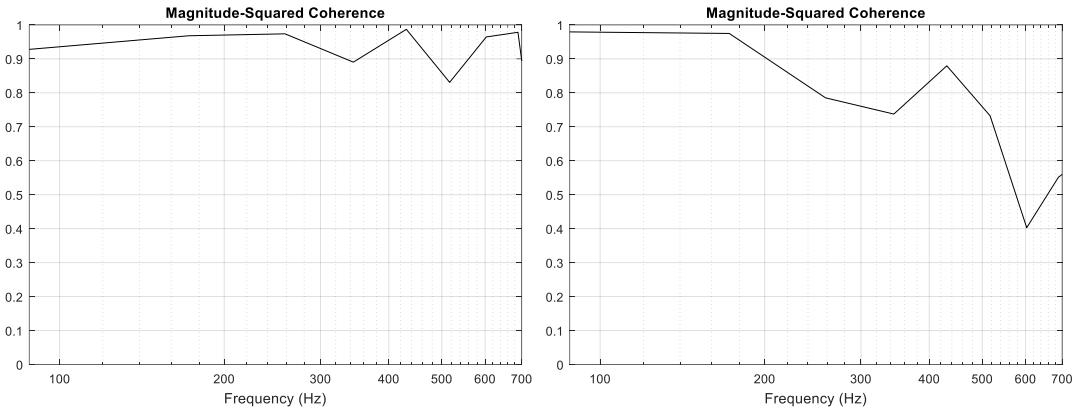


Figure 5.55: Magnitude-Squared Coherence plots for position 2 in the Meeting Room. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

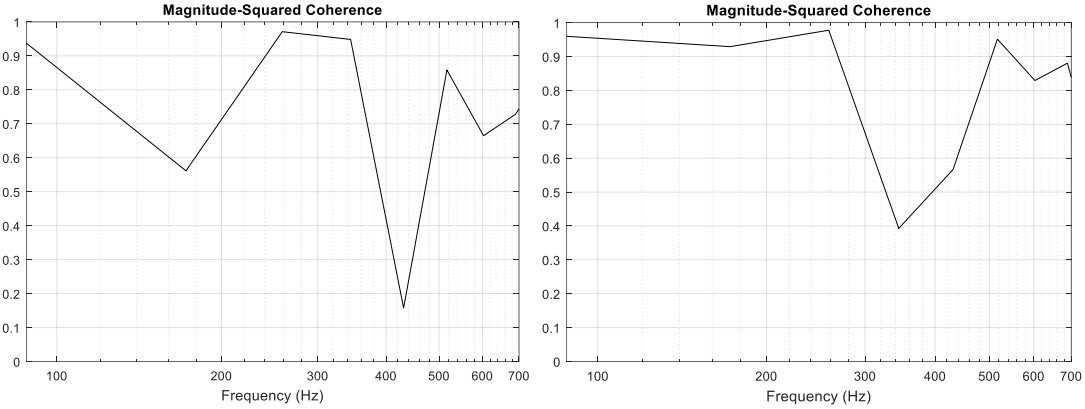


Figure 5.56: Magnitude-Squared Coherence plots for position 3 in the Meeting Room. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

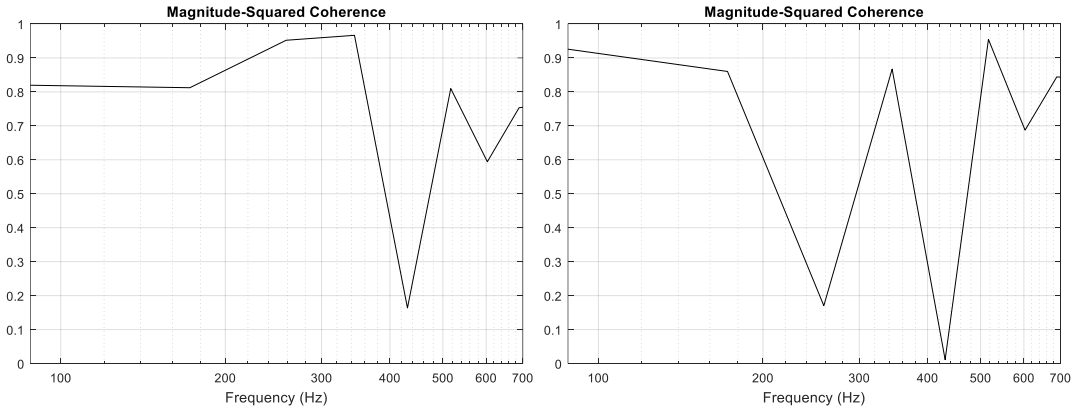


Figure 5.57: Magnitude-Squared Coherence plots for position 4 in the Meeting Room. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

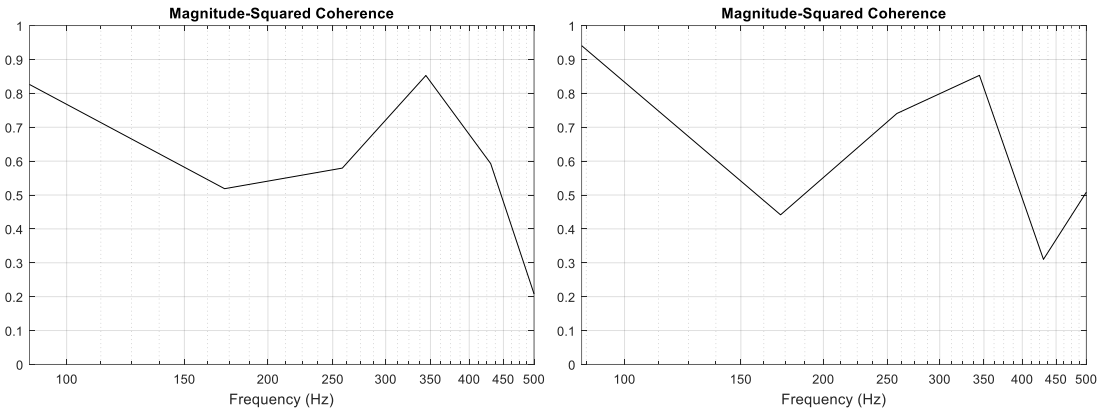


Figure 5.58: Magnitude-Squared Coherence plots for position 1 in the Classroom. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

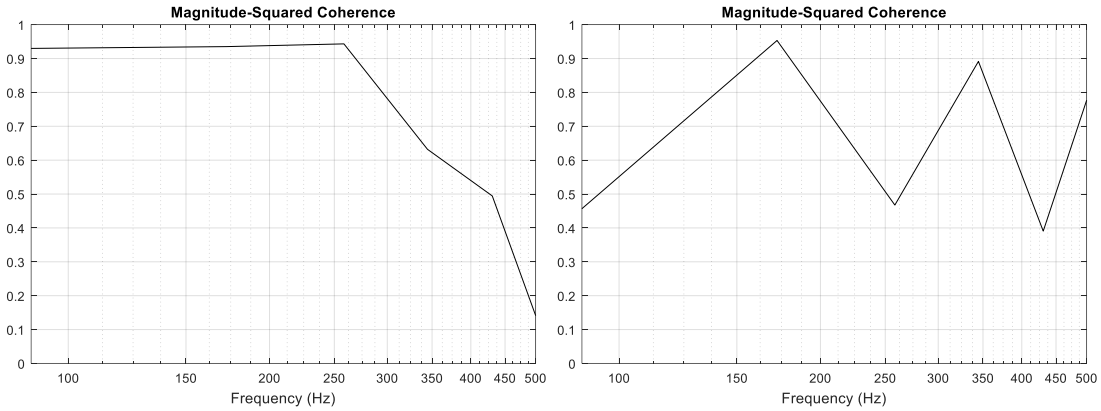


Figure 5.59: Magnitude-Squared Coherence plots for position 2 in the Classroom. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

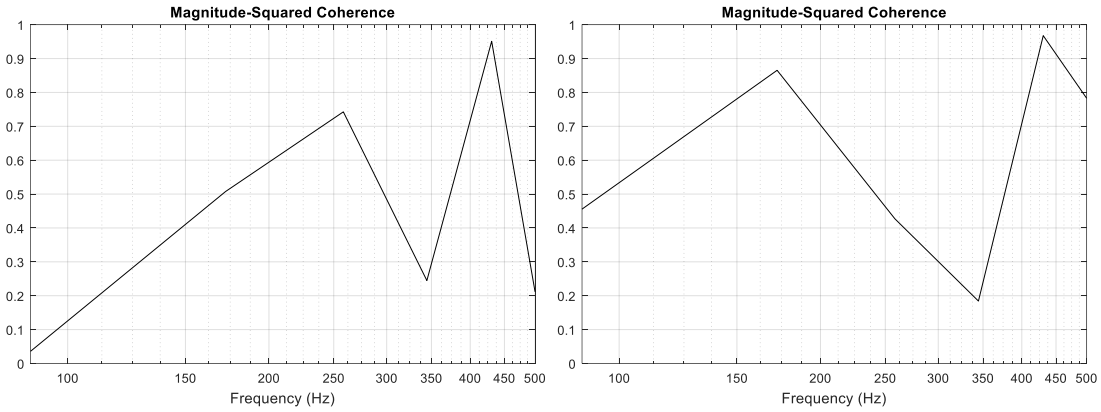


Figure 5.60: Magnitude-Squared Coherence plots for position 3 in the Classroom. Left, the estimate for FEM RFR and measured RFR signals. Right, the estimate for GA RFR and measured RFR signals.

5.4.3 Discussion of room acoustic parameter results

In order to quantify the accuracy of both numerical approaches to estimate acoustic parameters, three statistical descriptors were estimated (see Table 5.11and Table 5.12). First, the Squared Error was calculated for each indicator with the intention of evaluating differences with frequency (see equation [5.20]). After that, an Average Squared Error was estimated in order to have a statistical indicator to determine which numerical approach was more appropriate to estimate each acoustic indicator. The last statistical indicator estimated was the Standard Deviation Squared Error, with the purpose of corroborating the decision regarding the Average Squared Error:

$$SE_i\% = \left(\frac{M_i - S_i}{M_i} \right)^2 \qquad [5.20],$$

where for each acoustic indicator, M_i denotes the measured reference value at the octave band i and S_i represents the simulated result.

Table 5.11 and Table 5.12 present for both rooms investigated the squared error by frequency band, the average and standard deviation of the squared error for each acoustic indicator analysed in this thesis, with all the statistical descriptors taking into account the numerical approaches of GA and FEM-GA. According to Table 5.11, in the Meeting Room all the acoustic indicators gave results that best compares with the measured reference when the FEM was applied. On the other hand, the results in Table 5.12 indicate that in the Classroom the acoustic indicators such as EDT, T_{20} and D_{50} presented a higher average squared error when the FEM-GA approach was applied. It is important to note that this situation is not presented for C_{80} and IACC given the unexpected squared error values obtained in the GA approach for C_{80} in the octave band of 125 Hz and the IACC result in the octave band of 500 Hz.

Table 5.11: Squared error by frequency band, average and standard deviation of squared error for each acoustic indicator and numerical approach, according to the measured references for the Meeting Room.

Acoustic Indicator	Numerical approach	Squared Error %			Average Squared Error (%)	Standard Deviation Squared Error
		Frequency Band of 125 Hz	Frequency Band of 250 Hz	Frequency Band of 500 Hz		
EDT [s]:	GA	0.1396	0.0276	0.4820	0.2164	0.23672
	FEM-GA	0.0092	0.0156	0.0902	0.0383	0.04505
T20 [s]:	GA	0.0708	0.0539	0.2290	0.1179	0.09660
	FEM-GA	0.0218	0.0002	0.0993	0.0405	0.05210
C80 [dB]:	GA	0.0118	0.0991	0.2841	0.1317	0.13904
	FEM-GA	0.0057	0.1934	0.0006	0.0666	0.10988
D50 [%]:	GA	0.0061	0.0009	0.0271	0.0114	0.01384
	FEM-GA	0.0002	0.0002	0.0002	0.0002	0.00005
IACC [%]:	GA	0.0370	0.1056	0.1385	0.0937	0.05178
	FEM-GA	0.0001	0.0037	0.0135	0.0058	0.00696

Table 5.12: Squared error by frequency band, average and standard deviation of squared error for each acoustic indicator and numerical approach, according to the measured references for the Classroom.

Acoustic Indicator	Numerical approach	Squared Error %			Average Squared Error (%)	Standard Deviation Squared Error
		Frequency Band of 125 Hz	Frequency Band of 250 Hz	Frequency Band of 500 Hz		
EDT [s]:	GA	0.0290	0.0418	0.0327	0.0345	0.00661
	FEM-GA	0.0003	0.1933	0.1347	0.1094	0.09893
T20 [s]:	GA	0.0007	0.0310	0.0270	0.0196	0.01649
	FEM-GA	0.0346	0.1342	0.0753	0.0814	0.05006
C80 [dB]:	GA	**5.1600**	0.2876	0.0604	1.8360	2.88093
	FEM-GA	1.5891	2.0511	0.6090	1.4164	0.73641
D50 [%]:	GA	**0.1638**	0.0019	0.0171	0.0609	0.08939
	FEM-GA	0.0543	0.2282	0.1697	0.1507	0.08848
IACC [%]:	GA	0.0020	0.1156	**1.2282**	0.4486	0.67754
	FEM-GA	0.0001	0.0141	0.8578	0.2906	0.49121

5.5 Subjective evaluation

In order to evaluate the performance of the numerical methods applied to estimate the sound wave propagation in the transmission stage and the use of OPSODIS 3D reproduction system, a subjective test to evaluate the virtual sound environments was implemented. The room used for this test was the “Meeting Room” located on first floor of ISVR building. The parameters evaluated were:

localization of the source, reverberation or sense of space, warmth and brightness. The meaning of these parameters are explained in the following section (see section 5.5.1). All the variables were assessed against a reference value given by the auralizations created by means of BIR measurements.

5.5.1 Subjective test design

Three groups of auralizations (BIR measurements, GA and FEM-GA) were created for three "dry" audio samples: human voice, percussion instrument and wind instrument. Likewise, for each group of auralizations, three different receiver positions were auralized in order to have information for spatial evaluation. The methodology of the survey consisted of a pairwise comparison of samples A and B, where the first was the reference auralization created by means of BIR measurements and the second was given by the virtual sound environment generated numerically, either with GA or by the combination of FEM-GA. The participants were asked to rate the parameters mentioned above of sample B, with respect to reference sample A (see Appendix B: "Subjective test to assess virtual sound environments"). During the test, the participants did not know that sample A was the reference and B was the stimuli to be evaluated. It is important to note at this point that the test was designed so that each participant had to rate all the sources, parameters and positions. The signals were reproduced through an OPSODIS system positioned inside a recording studio at 2m distance of the listener, who was free to play, stop or repeat samples using a tablet with an application designed for the experiment in Pure Data language (see Appendix C: "Block diagrams of the subjective test in Pure Data"). The same comparison base on a unipolar scale was applied for each attribute, as it is recommended by Lindau et al (2014). As can be seen in Table 5.13, five indicates "not different" and one "completely different", a five-grade assessment scale as it is recommended by the ITU Radiocommunication Assembly in the document "Subjective assessment of sound quality" [ITU-R BS.562-3] (1990).

Table 5.13: Subjective assessment scale.

RATING	ASSESSMENT
Not different	5,0
Slightly not different	4,0
Slightly different	3,0
Rather different	2,0
Completely different	1,0

The definitions of the parameters evaluated are as follows:

- Localization: attribute associated to a subjective perception of the direction indicating the origin of sound and the relative position of the source.
- Sense of space: Similar to reverberation, this parameter refers to a subjective permanence of reflected sound in the enclosure. In other words, it indicates a subjective size impression of the room in acoustic terms.
- Warmth: attribute denoting a subjective perception of loudness at low frequencies of the corresponding source.
- Brightness: parameter indicating a subjective perception of loudness at high frequencies of the corresponding source.

According to the Spatial Audio Quality Inventory (SAQI) developed for the perceptual evaluation of spatial audio technologies, the parameters evaluated are categorized in the following groups: geometry, room and timbre. Localization for instance, is a parameter of the geometry descriptors group indicating the direction of a sound source including the listening quality descriptors of horizontal plane, the vertical planes and the perceived distance. Sense of space is a parameter included in the room descriptors group related to the listening quality descriptor of level of reverberation. The last two parameters are in the timbre descriptors group including the listening quality descriptors related to the timbral change at high-frequency and low-frequency (Lindau, et al., 2014).

The subjective test consisted of two stages, adaptation and evaluation. During the adaptation process, participants were allowed to become familiar with the

devices, the test environment, to clarify the meaning of the attributes of study and to get familiar with the assessment scale. At this point, the participants had the opportunity to listen to the instruments recorded in “dry” conditions and to experience examples of auralizations for different source-receiver combinations. The estimated time for this stage was between 2 and 5 minutes in order to avoid fatigue before the following phase of evaluation. The next step involved the pairwise comparison of auralizations, where the participants were asked to read the instructions and to register their answers in the test form (see Appendix B: “Subjective test to assess virtual sound environments”). The maximum time available for the evaluation stage was 20 minutes. Figure 5.61 shows the visual interface provided with the tablet, in order for each participant to manipulate the audio files at their own will.

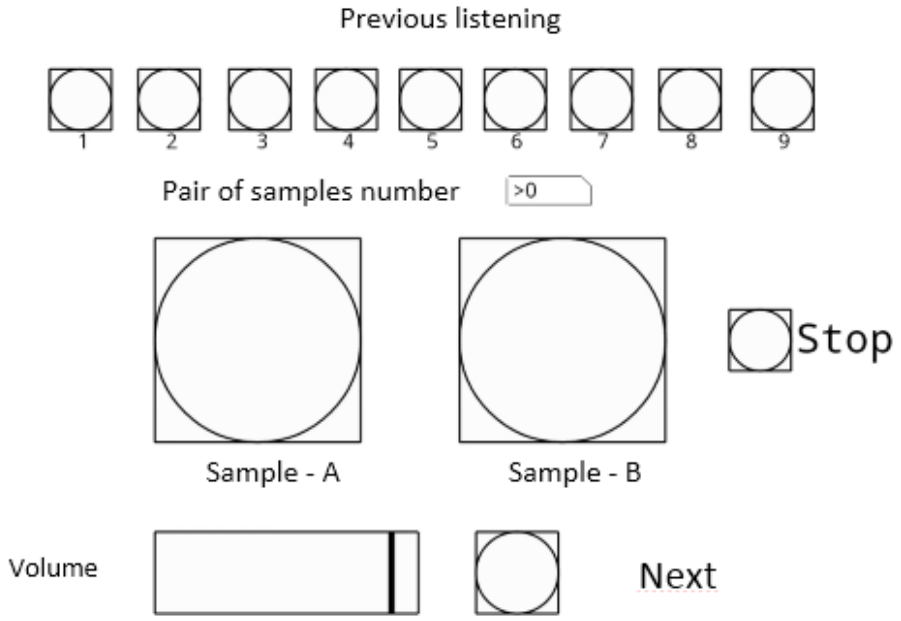


Figure 5.61: User interface for controlling audio files, running on a Tablet.

5.5.2 Basic statistical considerations

In this section, the basics of the statistics applied in the definition of a sample size and the analysis of the subjective tests results are described. The population defined for this study considered students from the sixth semester onwards of the Sound Engineering programme at the University of San Buenaventura, in Medellin Colombia. According to the University data, there were 103 students enrolled in these semesters at the moment the test was applied. It was suggested that the participants had normal audition and experience in subjective listening

tests. In order to get a representative sample, a simple random sampling method was chosen.

Sample size was determined taking into account three aspects:

1) The estimated confidence level.
2) The permissible error.
3) The finite character of the population (less than 100,000 people).

A general formula to estimate a sample size is given by the following expression (Wayne, 2009):

$$n = \frac{Z^2 * P * Q * N}{E^2 (N-1) + Z^2 * P * Q} \qquad [5.21],$$

where, n is sample size, N is the population size, *P* and *Q* are probabilities indicating when the phenomenon occurs (positive and negative variability), Z is the standardized variable of normal distribution according to the chosen confidence level and E is the permissible error defined by the head of the study.

When values of *P* and *Q* are unknown or when the survey covers different aspects, in which these values are unequal, it is convenient to take the most appropriate case that is given by the maximum size of sample, which occurs for *P = Q = 0.5*. The confidence level is a function of the significance level α, therefore, for a confidence level of 90%:

$$1-\alpha = 90\% \qquad [5.22].$$

The last corresponds to a standardized variable of normal distribution $Z = 1.645$. Taking a maximum error of 10% ($E = 0.1$) and assuming maximum variance for a known population, $N = 103$, sample size estimation applying expression [5.21] would be given by:

$$n = \frac{Z^2 * P * Q * N}{E^2 (N-1) + Z^2 * P * Q} = \frac{1.645^2 * 0.5 * 0.5 * 103}{0.1^2 (103-1) + 1.645^2 * 0.5 * 0.5} = 41.07 \qquad [5.23],$$

hence, in this study a simple random sampling of 40 students was taken into account.

5.5.2.1 Measures of central tendency

The central tendency measures are unique values aiming to describe a data set by identifying a centre position. The mean, median and mode are all valid measures of central tendency:

- The mean or arithmetic average results from the sum of all the data divided by the number of data. A disadvantage of this measure is given by the fact of being particularly susceptible to the influence of outliers, or unusual values compared with the rest of data. Furthermore, as the data distribution is less symmetrical, the mean is less representative of the central tendency. In these cases, it is preferable to use the median.
- The median is the intermediate value of a set of data organized from the lowest to the highest. This measure is less affected by outliers or by the possible asymmetry of the data.
- The mode is the most common value in a data set and the one that is repeated the most. It is commonly used in nominal variables. A drawback of this measure is given when more than one data share this characteristic.

In the case when a data distribution is normal, mean and median are equal and both can be used as measures of central tendency. However, if a data has a more asymmetric distribution, there will be more difference between mean and median, where median is the most preferable measure to use.

5.5.2.2 Dispersion and asymmetric coefficients

The dispersion in a data set is a measure used to describe the variability in a statistic sample. The main measure of dispersion is the standard deviation s, which represents the average distance of data with respect to the mean:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \qquad [5.24],$$

where, x_i is each one of the values in the set, $\bar{x}$ is the mean and n is the number of data points. A coefficient of asymmetry is able to measure the symmetry of a data set and determines the best way to describe a central tendency. A good example for this coefficient is the Skewness operator.

In statistics, the Skewness is a measure that describes asymmetry from a normal distribution. In other words, this coefficient quantifies the extent at which a sample differs from a normal distribution. When one random variable is to be analysed, the Skewness is considered the third standardized moment and the formula to estimate it is given by the following expression:

$$skewness = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)s^3} \qquad [5.25].$$

5.5.2.3 Box-and-Whisker Plot

The Box-and-Whisker Plot is an exploratory procedure used to create a plot designed to analyse a data sample summarizing five statistic measures such as a minimum, a lower quartile, a median, an upper quartile and a maximum. Moreover, this technique also indicates the presence of outliers. These measures can be used to identify statistic features such as centre, dispersion and asymmetry. The box encloses the interquartile range, when the inferior border represents the first quartile and the superior limit represents the third quartile. This interval is covered by the 50% of the data sorted from smallest to largest. A line inside the box represents the second quartile, or median. If sample mean is requested, a plus sign is placed at the corresponding location. The whiskers characterize the largest and smallest data values considering a range given by 1.5 times the interquartile range, which are drawn from the edges of the box. Any data value beyond that limit is considered an outlier, and is represented by point symbols. Any point which value exceeds more than three times the interquartile range is called *far outside point,* and is indicated with a plus sign.

5.5.3 Subjective tests results and discussion

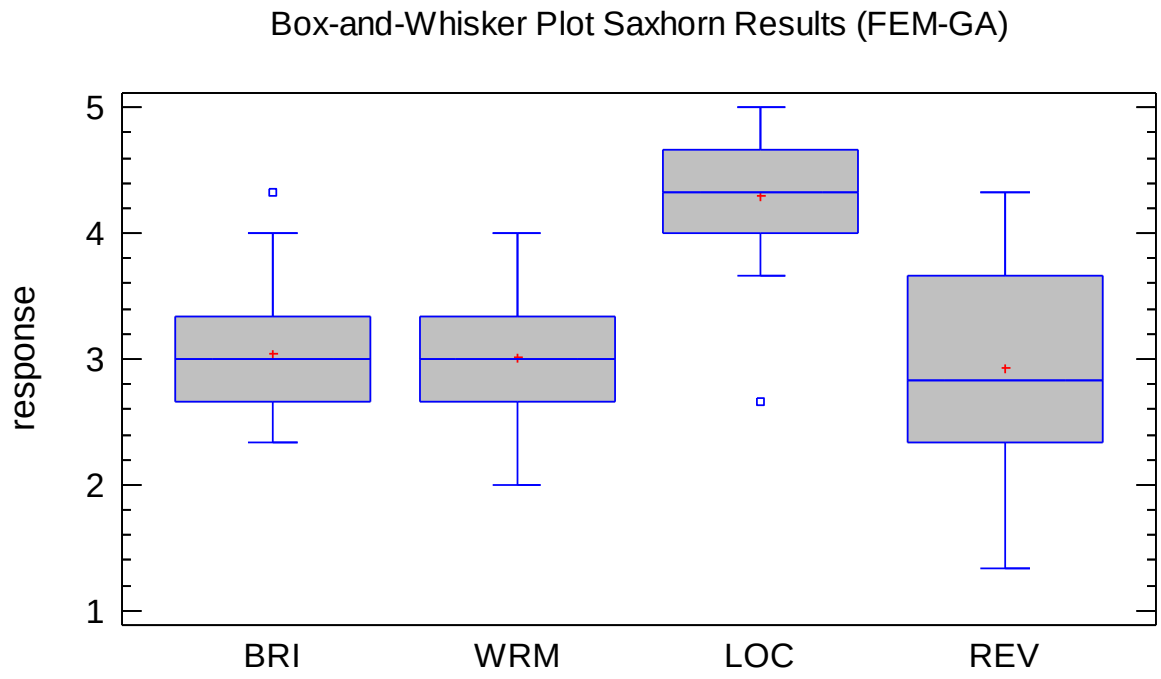
In this section, the results and respective discussion obtained by the application of the subjective tests are presented (see section 0). According to the test design, the subjective evaluation included the rating of four parameters for both numerical approaches used (localization=LOC, sense of space=REV, warmth=WRM and brightness=BRI), assessing three different acoustic sources (saxhorn, bass drum and male voice), at three specific positions in the room (see section *5.5.1*). All the results are indicated in graphical and tabular forms by means of box-and-whisker plots and tables including dispersion and asymmetric coefficients such as standard deviation, coefficient of variation, the Skewness

and the Kurtosis operator. This section is divided in two parts, beginning with spatially averaged values calculated by considering results for all specific positions. On the second part, the subjective tests results include a comparison of statistics obtained at each position auralized in the room. The plots and tabular information were generated using the software STATGRAPHICS Centurion XV Version 15.2.11 Portable©.

5.5.3.1 Spatially averaged subjective results

This section presents the spatially averaged estimates obtained for each numerical approach and instrument. The statistical indicators were calculated by averaging the results of the three positions auralized in the room. First, the box-and-whisker plots (see Figure 5.62 to Figure 5.64) and the tabular statistical information (see Table 5.14 to Table 5.19) are exposed for each instrument, followed by a discussion of spatially averaged results.

5.5.3.1.1 Spatially averaged saxhorn results



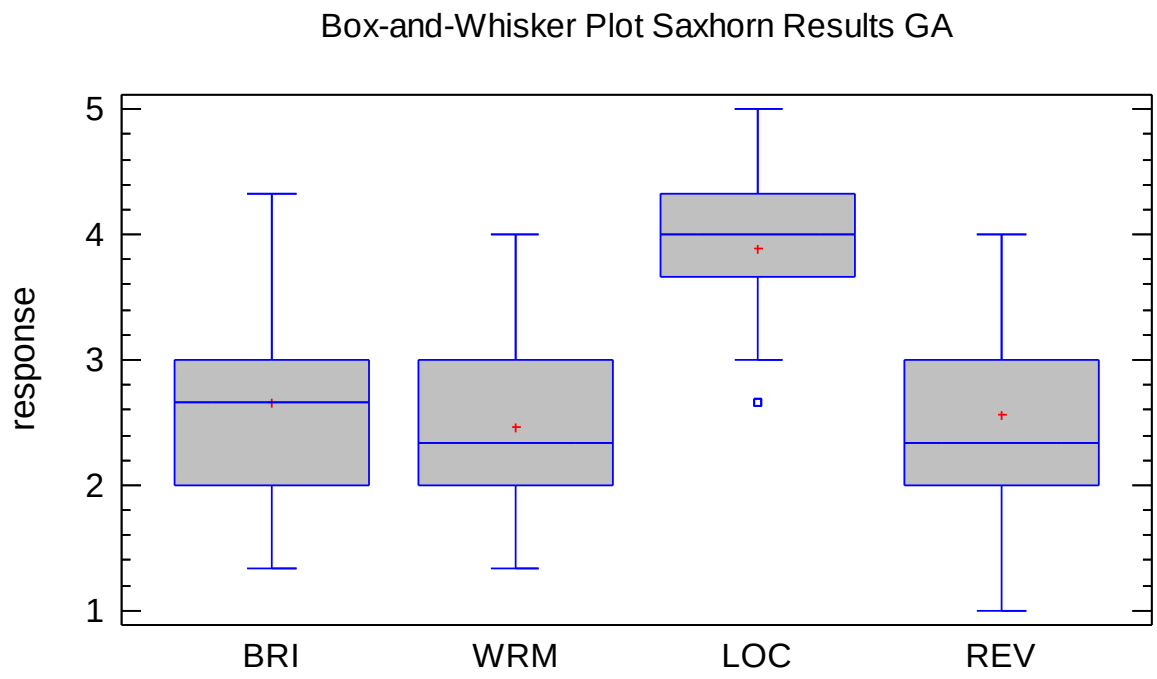


Figure 5.62: Spatially averaged estimates of the four parameters evaluated for the saxhorn on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

Table 5.14: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the saxhorn, using the hybrid approach FEM-GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	3.04	3.02	4.30	2.93
Median	3.00	3.00	4.33	2.83
Mode	3.00	3.00	4.67	3.67
Standard deviation	0.61	0.52	0.59	0.82
Coeff. of variation	20.23%	17.14%	13.80%	28.06%
Standard error	0.14	0.12	0.14	0.19
Minimum	2.33	2.00	2.67	1.33
Maximum	4.33	4.00	5.00	4.33
Lower quartile	2.67	2.67	4.00	2.33
Upper quartile	3.33	3.33	4.67	3.67
Skewness	0.77	-0.10	-1.15	-0.19
Kurtosis	-0.24	-0.24	1.97	-0.91

Table 5.15: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the saxhorn, using the numerical approach GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	2.67	2.46	3.89	2.56
Median	2.67	2.33	4.00	2.33
Mode	2.67	2.00	4.33	2.33
Standard deviation	0.74	0.84	0.75	0.85
Coeff. of variation	27.79%	34.16%	19.28%	33.18%
Standard error	0.17	0.20	0.18	0.20
Minimum	1.33	1.33	2.67	1.00
Maximum	4.33	4.00	5.00	4.00
Lower quartile	2.00	2.00	3.67	2.00
Upper quartile	3.00	3.00	4.33	3.00
Skewness	0.33	0.62	-0.40	0.06
Kurtosis	0.26	-0.57	-0.71	-0.69

5.5.3.1.2 Spatially averaged bass drum results

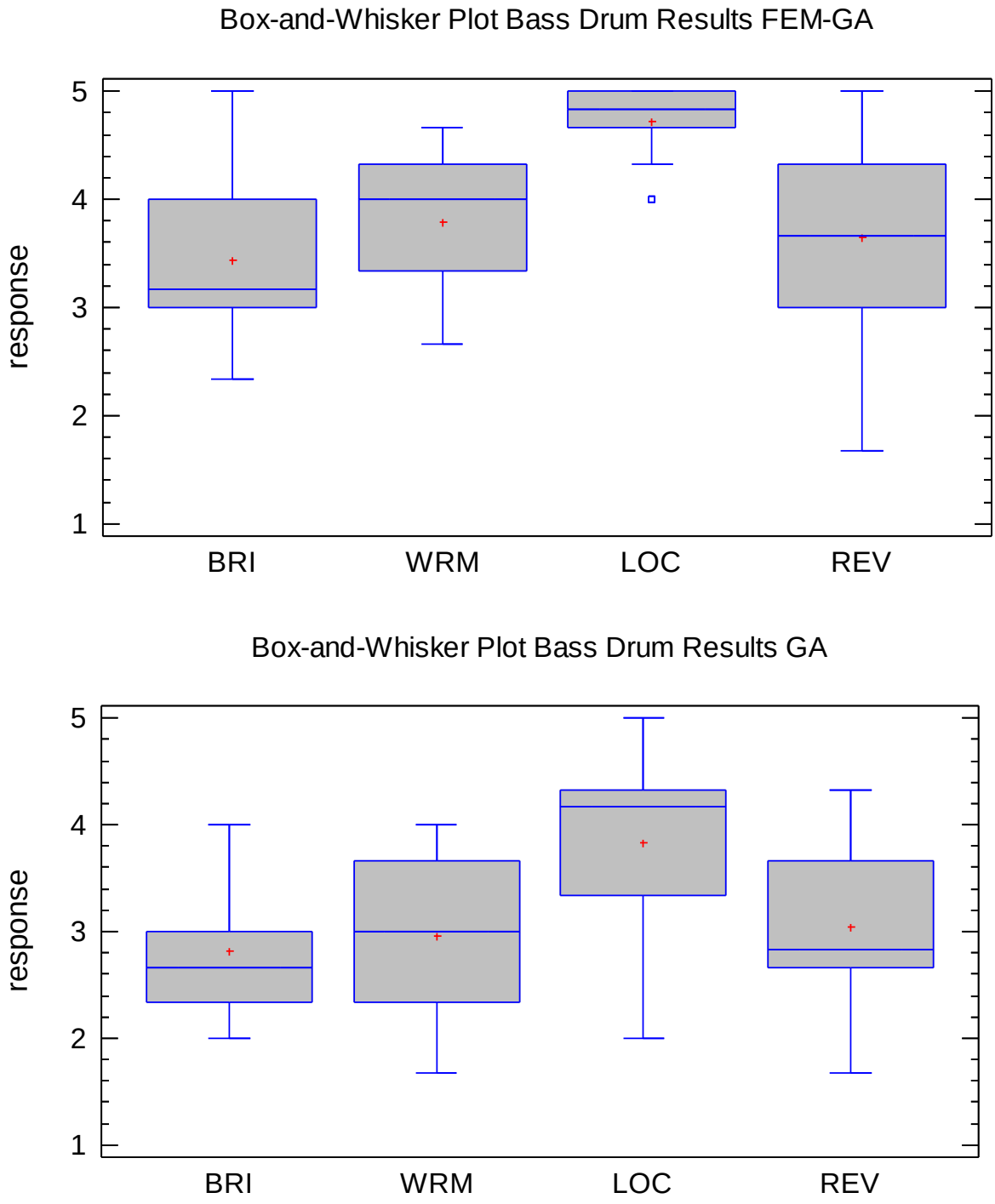


Figure 5.63: Spatially averaged estimates of the four parameters evaluated for the bass drum on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

Table 5.16: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the bass drum, using the hybrid approach FEM-GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	3.44	3.80	4.72	3.65
Median	3.17	4.00	4.83	3.67
Mode	3.00	4.00	5.00	
Standard deviation	0.79	0.65	0.35	0.96
Coeff. of variation	23.00%	17.07%	7.36%	26.31%
Standard error	0.19	0.15	0.08	0.23
Minimum	2.33	2.67	4.00	1.67
Maximum	5.00	4.67	5.00	5.00
Lower quartile	3.00	3.33	4.67	3.00
Upper quartile	4.00	4.33	5.00	4.33
Skewness	0.65	-0.51	-1.07	-0.30
Kurtosis	-0.17	-0.93	0.08	-0.52

Table 5.17: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the bass drum, using the numerical approach GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	2.81	2.96	3.83	3.04
Median	2.67	3.00	4.17	2.83
Mode		3.67	4.33	2.67
Standard deviation	0.57	0.74	0.87	0.86
Coeff. of variation	20.40%	24.97%	22.57%	28.39%
Standard error	0.14	0.17	0.20	0.20
Minimum	2.00	1.67	2.00	1.67
Maximum	4.00	4.00	5.00	4.33
Lower quartile	2.33	2.33	3.33	2.67
Upper quartile	3.00	3.67	4.33	3.67
Skewness	0.53	-0.50	-0.81	0.18
Kurtosis	-0.37	-0.98	-0.26	-1.07

5.5.3.1.3 Spatially averaged male voice results

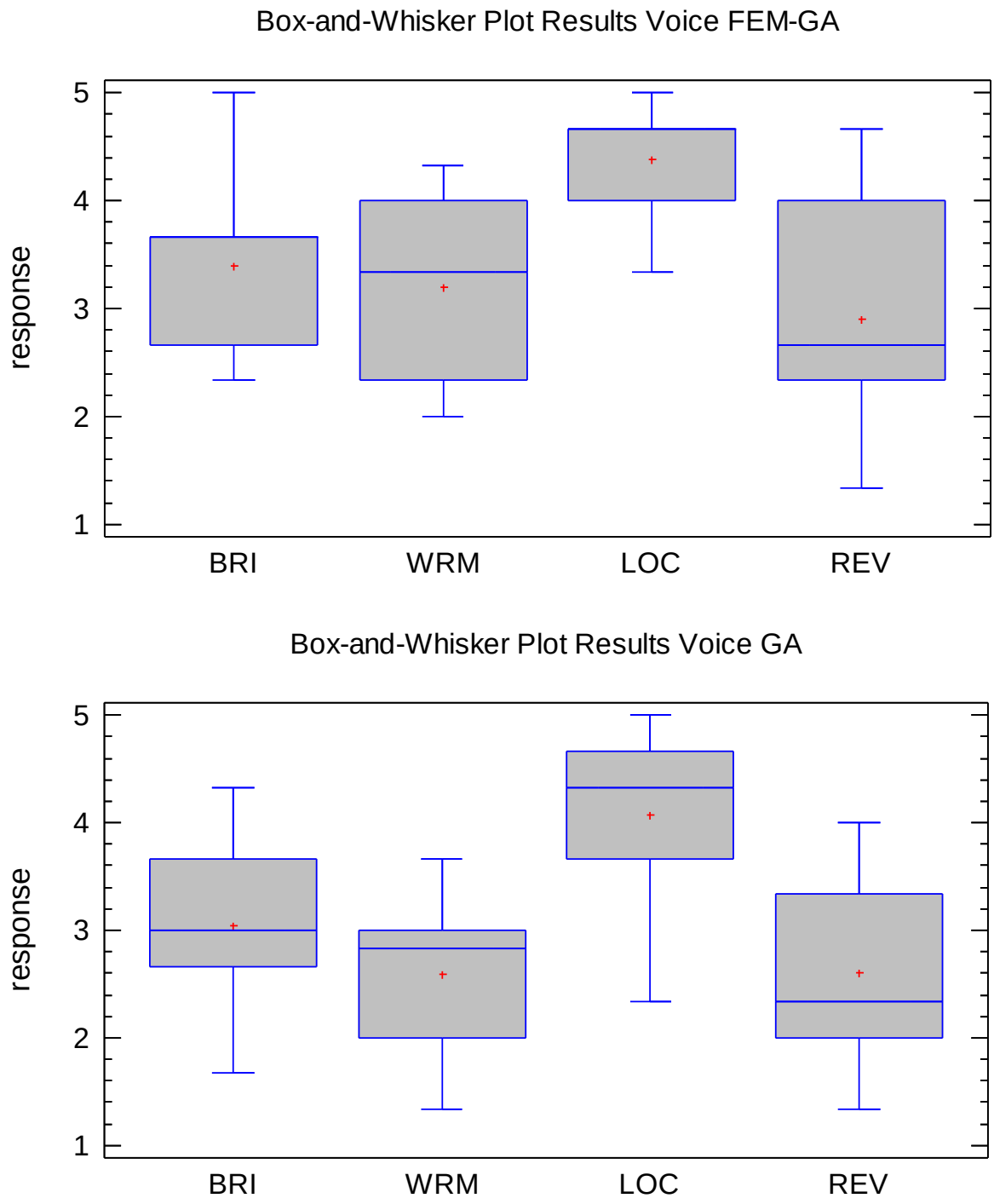


Figure 5.64: Spatially averaged estimates of the four parameters evaluated for the male voice on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

Table 5.18: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the male voice, using the hybrid approach FEM-GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	3.39	3.20	4.39	2.91
Median	3.67	3.33	4.67	2.67
Mode	3.67	4.00	4.67	2.33
Standard deviation	0.68	0.83	0.50	0.96
Coeff. of variation	20.03%	25.77%	11.43%	32.86%
Standard error	0.16	0.19	0.12	0.23
Minimum	2.33	2.00	3.33	1.33
Maximum	5.00	4.33	5.00	4.67
Lower quartile	2.67	2.33	4.00	2.33
Upper quartile	3.67	4.00	4.67	4.00
Skewness	0.26	-0.10	-0.67	0.42
Kurtosis	0.49	-1.60	-0.55	-0.93

Table 5.19: Summary of exploratory spatially averaged statistics obtained on the four parameters evaluated for the male voice, using the numerical approach GA.

Statistical indicator	BRI	WRM	LOC	REV
Count	40	40	40	40
Average	3.04	2.59	4.07	2.61
Median	3.00	2.83	4.33	2.33
Mode	3.00	3.00		2.00
Standard deviation	0.74	0.72	0.87	0.82
Coeff. of variation	24.37%	27.74%	21.47%	31.35%
Standard error	0.17	0.17	0.21	0.19
Minimum	1.67	1.33	2.33	1.33
Maximum	4.33	3.67	5.00	4.00
Lower quartile	2.67	2.00	3.67	2.00
Upper quartile	3.67	3.00	4.67	3.33
Skewness	-0.12	-0.35	-0.73	0.48
Kurtosis	-0.54	-0.74	-0.63	-0.92

5.5.3.1.4 Discussion of spatially averaged subjective results

Figure 5.62 to Figure 5.64 clearly illustrate how every single source obtained higher scores in all the parameters evaluated, when the auralizations simulated by means of the hybrid approach (FEM - GA) were listened. It is observed from the figures how the localization of the source was the parameter subjects judged to be the most similar with respect to measured reference auralizations. This situation manifests itself in the BIR renderings provided in this chapter, which

show a suitable agreement in the arrival of direct sound for both numerical approaches. In the evaluation of this parameter, the bass drum is the acoustic source with the best scores meaning that it was very difficult to distinguish between measurement and simulation. It is important to note for that instrument how the hybrid approach exhibited the best responses in all the parameters but brightness, which had comparable results in the male voice auralizations.

According to Table 5.14 and Table 5.15, exploratory spatially averaged statistics for the saxhorn auralizations indicate that warmth was the worst assessed parameter, with an average value of (2.4) in the auralizations created by means of GA simulations. Another aspect to note is given by the dispersion difference between both numerical approaches. Statistical indicators such as Standard deviation, Coefficient of variation and Standard error present higher values in GA simulations for all the parameters evaluated. Moreover, the Skewness estimate of (0.1) for warmth in the hybrid approach represents a normal distribution, evidencing a consistent subjective assessment.

In the bass drum auralizations case, a comparison of Table 5.16 and Table 5.17 point out brightness and warmth in GA as the parameters with the lowest rate, with (2.8) and (2.9) as the correspondent averages values. In this case, the assessment of localization in the hybrid approach presents the lowest dispersion taking into account the Standard deviation, Coefficient of variation and Standard error results. Considering the last three statistical indicators, the reverberation was the most dispersed parameter in both numerical approaches, although Skewness values evidence a normal distribution of data.

Regarding the male voice auralizations, a comparison of statistical results in Table 5.18 and Table 5.19 indicate once again, that warmth has the lowest rate in GA simulations and localization in hybrid method is the least dispersed parameter. In contrast, this numerical approach presents the poorest evaluation in terms of dispersion for reverberation parameter.

5.5.3.2 All positions subjective results

This section presents all position estimates obtained for each numerical approach and instrument. The statistical indicators were calculated on each position auralized in the room. First, the box-and-whisker plots (see Figure 5.65, Figure 5.66 and Figure 5.67) and the tabular statistical information (see from

Table 5.20 to Table 5.25) are presented for each instrument, followed by a discussion of specific positions results.

5.5.3.2.1 All positions saxhorn results

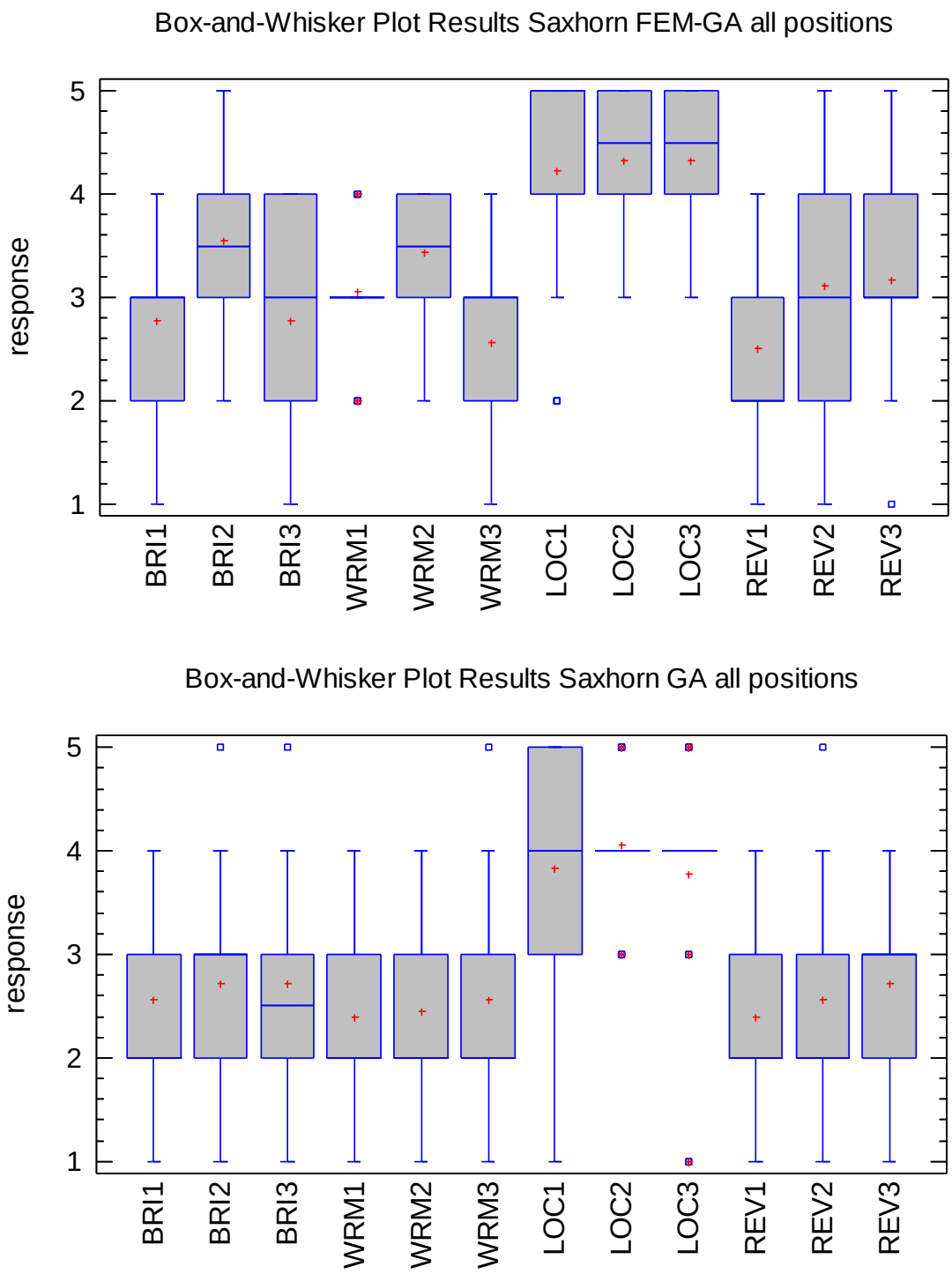


Figure 5.65: Estimates at each position of the four parameters evaluated for the saxhorn on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

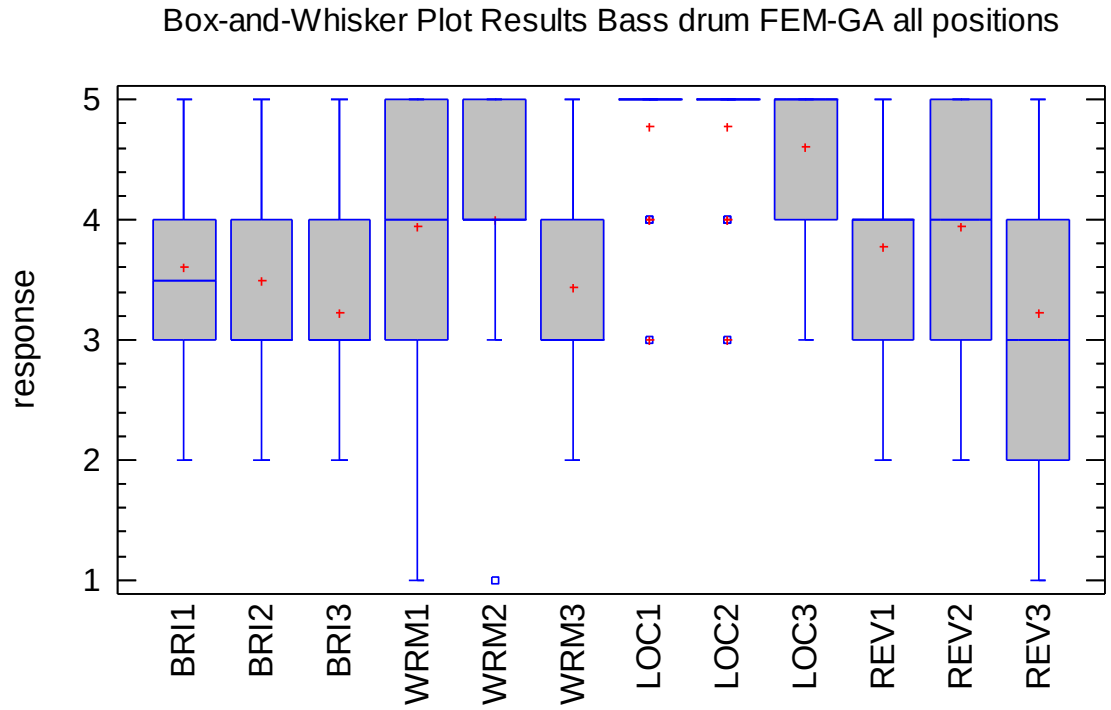
Table 5.20: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the saxhorn, using the hybrid approach FEM-GA.

Statistical Indicator	BRI1	BRI2	BRI3	WRM 1	WRM 2	WRM 3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	2.78	3.56	2.78	3.06	3.44	2.56	4.22	4.33	4.33	2.50	3.11	3.17
Median	3.00	3.50	3.00	3.00	3.50	3.00	5.00	4.50	4.50	2.00	3.00	3.00
Mode	3.00	3.00	2.00	3.00	4.00	3.00	5.00	5.00	5.00	2.00		3.00
Standard deviation	0.81	0.78	0.94	0.64	0.62	0.86	1.17	0.77	0.77	1.04	1.02	0.99
Coeff. of variation	29.10%	22.05%	33.94%	20.92%	17.88%	33.48%	27.62%	17.70%	17.70%	41.73%	32.87%	31.11%
Standard error	0.19	0.18	0.22	0.15	0.15	0.20	0.27	0.18	0.18	0.25	0.24	0.23
Minimum	1.00	2.00	1.00	2.00	2.00	1.00	2.00	3.00	3.00	1.00	1.00	1.00
Maximum	4.00	5.00	4.00	4.00	4.00	4.00	5.00	5.00	5.00	4.00	5.00	5.00
Lower quartile	2.00	3.00	2.00	3.00	3.00	2.00	4.00	4.00	4.00	2.00	2.00	3.00
Upper quartile	3.00	4.00	4.00	3.00	4.00	3.00	5.00	5.00	5.00	3.00	4.00	4.00
Skewness	-0.30	0.21	0.02	-0.04	-0.62	-0.19	-1.24	-0.68	-0.68	0.17	-0.24	-0.37
Kurtosis	0.02	-0.15	-1.10	-0.14	-0.39	-0.28	-0.01	-0.87	-0.87	-1.06	-0.36	0.11

Table 5.21: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the saxhorn, using the numerical approach of GA.

Statistical Indicator	BRI1	BRI2	BRI3	WRM 1	WRM 2	WRM 3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	2.56	2.72	2.72	2.39	2.44	2.56	3.83	4.06	3.78	2.39	2.56	2.72
Median	2.00	3.00	2.50	2.00	2.00	2.00	4.00	4.00	4.00	2.00	2.00	3.00
Mode	2.00		2.00	2.00	2.00	2.00		4.00	4.00	2.00	2.00	3.00
Standard deviation	0.86	0.96	1.02	0.85	0.98	1.10	1.15	0.64	1.17	0.98	1.04	0.83
Coeff. of variation	33.48 %	35.20 %	37.39 %	35.57 %	40.24 %	42.91 %	30.01 %	15.76 %	30.86 %	40.96 %	40.76 %	30.36 %
Standard error	0.20	0.23	0.24	0.20	0.23	0.26	0.27	0.15	0.27	0.23	0.25	0.19
Minimum	1.00	1.00	1.00	1.00	1.00	1.00	1.00	3.00	1.00	1.00	1.00	1.00
Maximum	4.00	5.00	5.00	4.00	4.00	5.00	5.00	5.00	5.00	4.00	5.00	4.00
Lower quartile	2.00	2.00	2.00	2.00	2.00	2.00	3.00	4.00	4.00	2.00	2.00	2.00
Upper quartile	3.00	3.00	3.00	3.00	3.00	3.00	5.00	4.00	4.00	3.00	3.00	3.00
Skewness	0.44	0.63	0.63	0.39	0.17	0.74	-0.94	-0.04	-1.52	0.35	0.71	-0.11
Kurtosis	-0.49	0.66	-0.03	-0.11	-0.80	-0.05	0.64	-0.14	2.23	-0.68	0.45	-0.29

5.5.3.2.2 All positions bass drum results



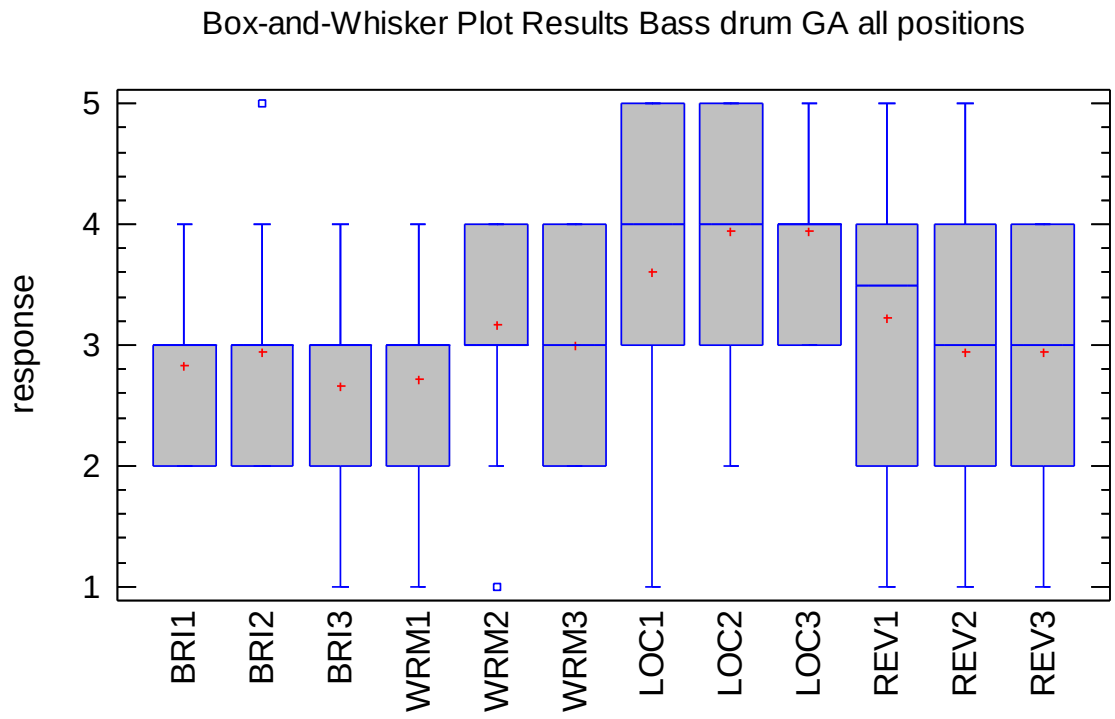


Figure 5.66: Estimates at each position of the four parameters evaluated for the bass drum on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

Table 5.22: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the bass drum, using the hybrid approach FEM-GA.

Statistical indicator	BRI1	BRI2	BRI3	WRM1	WRM2	WRM3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	3.61	3.50	3.22	3.94	4.00	3.44	4.78	4.78	4.61	3.78	3.94	3.22
Median	3.50	3.00	3.00	4.00	4.00	3.00	5.00	5.00	5.00	4.00	4.00	3.00
Mode	3.00	3.00	3.00	5.00	4.00	3.00	5.00	5.00	5.00	4.00	5.00	
Standard deviation	0.85	0.92	0.94	1.11	0.97	0.86	0.55	0.55	0.61	0.94	1.16	1.26
Coeff. of variation	23.53 %	26.39 %	29.26 %	28.14 %	24.25 %	24.84 %	11.48 %	11.48 %	13.18 %	24.96 %	29.45 %	39.19 %
Standard error	0.20	0.22	0.22	0.26	0.23	0.20	0.13	0.13	0.14	0.22	0.27	0.30
Minimum	2.00	2.00	2.00	1.00	1.00	2.00	3.00	3.00	3.00	2.00	2.00	1.00
Maximum	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Lower quartile	3.00	3.00	3.00	3.00	4.00	3.00	5.00	5.00	4.00	3.00	3.00	2.00
Upper quartile	4.00	4.00	4.00	5.00	5.00	4.00	5.00	5.00	5.00	4.00	5.00	4.00
Skewness	0.26	0.25	0.45	-1.04	-1.74	0.19	-2.57	-2.57	-1.36	-0.45	-0.64	-0.27
Kurtosis	-0.53	-0.60	-0.39	1.24	4.71	-0.28	6.36	6.36	1.13	-0.39	-1.07	-0.74

Table 5.23: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the bass drum, using the numerical approach of GA.

Statistical indicator	BRI1	BRI2	BRI3	WRM1	WRM2	WRM3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	2.83	2.94	2.67	2.72	3.17	3.00	3.61	3.94	3.94	3.22	2.94	2.94
Median	3.00	3.00	3.00	3.00	3.00	3.00	4.00	4.00	4.00	3.50	3.00	3.00
Mode	3.00	3.00	3.00	2.00	4.00			5.00	4.00	4.00		4.00
Standard deviation	0.71	0.87	0.77	0.89	0.92	0.84	1.20	1.21	0.73	1.17	1.11	1.06
Coeff. of variation	24.96%	29.64%	28.76%	32.87%	29.16%	28.01%	33.09%	30.71%	18.39%	36.19%	37.70%	35.85%
Standard error	0.17	0.21	0.18	0.21	0.22	0.20	0.28	0.29	0.17	0.27	0.26	0.25
Minimum	2.00	2.00	1.00	1.00	1.00	2.00	1.00	2.00	3.00	1.00	1.00	1.00
Maximum	4.00	5.00	4.00	4.00	4.00	4.00	5.00	5.00	5.00	5.00	5.00	4.00
Lower quartile	2.00	2.00	2.00	2.00	3.00	2.00	3.00	3.00	3.00	2.00	2.00	2.00
Upper quartile	3.00	3.00	3.00	3.00	4.00	4.00	5.00	5.00	4.00	4.00	4.00	4.00
Skewness	0.25	0.71	-0.20	0.07	-0.87	0.00	-0.53	-0.78	0.09	-0.23	0.41	-0.55
Kurtosis	-0.78	0.20	0.10	-0.81	0.01	-1.59	-0.35	-0.97	-0.90	-0.98	-0.38	-0.88

5.5.3.2.3 All positions male voice results

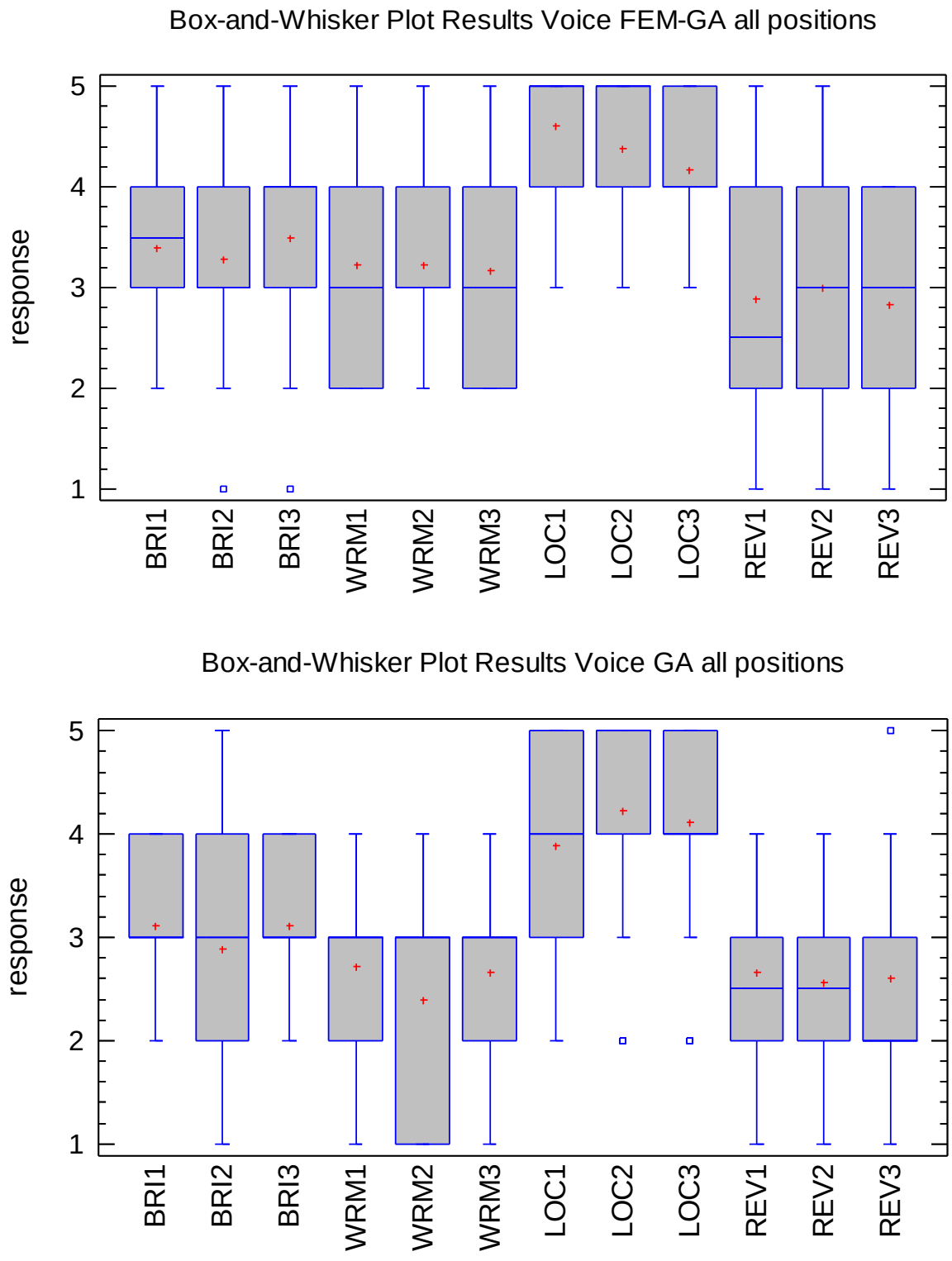


Figure 5.67: Estimates at each position of the four parameters evaluated for the male voice on each numerical technique (on top the hybrid approach FEM-GA and on bottom, GA).

Table 5.24: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the male voice, using the hybrid approach FEM-GA.

Statistical indicator	BRI1	BRI2	BRI3	WRM 1	WRM 2	WRM 3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	3.39	3.28	3.50	3.22	3.22	3.17	4.61	4.39	4.17	2.89	3.00	2.83
Median	3.50	3.00	4.00	3.00	3.00	3.00	5.00	5.00	4.00	2.50	3.00	3.00
Mode	4.00		4.00	4.00	3.00	4.00	5.00	5.00	4.00	2.00	2.00	
Standard deviation	0.85	0.96	1.04	0.94	0.88	0.99	0.70	0.78	0.71	1.13	1.24	0.92
Coeff. of variation	25.08%	29.24%	29.81%	29.26%	27.25%	31.11%	15.13%	17.72%	16.97%	39.18%	41.22%	32.60%
Standard error	0.20	0.23	0.25	0.22	0.21	0.23	0.16	0.18	0.17	0.27	0.29	0.22
Minimum	2.00	1.00	1.00	2.00	2.00	2.00	3.00	3.00	3.00	1.00	1.00	1.00
Maximum	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	4.00
Lower quartile	3.00	3.00	3.00	2.00	3.00	2.00	4.00	4.00	4.00	2.00	2.00	2.00
Upper quartile	4.00	4.00	4.00	4.00	4.00	4.00	5.00	5.00	5.00	4.00	4.00	4.00
Skewness	-0.26	-0.63	-0.87	-0.02	0.10	0.05	-1.61	-0.85	-0.25	0.24	0.42	-0.14
Kurtosis	-0.53	0.66	0.64	-1.10	-0.64	-1.32	1.40	-0.71	-0.78	-1.23	-0.96	-0.91

Table 5.25: Summary of exploratory statistics obtained at each position on the four parameters evaluated for the male voice, using the numerical approach of GA.

Statistical indicator	BRI1	BRI2	BRI3	WRM1	WRM2	WRM3	LOC1	LOC2	LOC3	REV1	REV2	REV3
Count	40	40	40	40	40	40	40	40	40	40	40	40
Average	3.11	2.89	3.11	2.72	2.39	2.67	3.89	4.22	4.11	2.67	2.56	2.61
Median	3.00	3.00	3.00	3.00	3.00	3.00	4.00	5.00	4.00	2.50	2.50	2.00
Mode	3.00	3.00	3.00	3.00	3.00	3.00		5.00	4.00	2.00	2.00	2.00
Standard deviation	0.76	1.18	0.68	0.96	1.04	0.77	1.02	1.06	0.96	0.91	1.04	0.98
Coeff. of variation	24.38%	40.94%	21.74%	35.20%	43.41%	28.76%	26.30%	25.11%	23.43%	34.03%	40.76%	37.48%
Standard error	0.18	0.28	0.16	0.23	0.24	0.18	0.24	0.25	0.23	0.21	0.25	0.23
Minimum	2.00	1.00	2.00	1.00	1.00	1.00	2.00	2.00	2.00	1.00	1.00	1.00
Maximum	4.00	5.00	4.00	4.00	4.00	4.00	5.00	5.00	5.00	4.00	4.00	5.00
Lower quartile	3.00	2.00	3.00	2.00	1.00	2.00	3.00	4.00	4.00	2.00	2.00	2.00
Upper quartile	4.00	4.00	4.00	3.00	3.00	3.00	5.00	5.00	5.00	3.00	3.00	3.00
Skewness	-0.19	0.24	-0.13	-0.27	-0.20	-0.20	-0.50	-1.16	-1.13	0.24	0.01	0.92
Kurtosis	-1.12	-0.49	-0.53	-0.66	-1.20	0.10	-0.77	0.16	0.83	-0.89	-1.07	0.79

5.5.3.2.4 Discussion of all positions subjective results

Figure 5.65 illustrates the subjective test results of the saxhorn instrument auralizations for receiver positions 1, 3 and 4 created by means of both numerical approaches (see Figure 5.2). It is important to note that the localization parameter was assessed with higher scores in every receiver position simulated in the meeting room with the hybrid approach. In the case of brightness and reverberation, according to the Box-and-Whisker plot positions 2 and 3 presented better results with the same numerical technique. Warmth parameter responses point out that positions 1 and 2 exhibited a more similar behaviour to measured reference auralizations, than GA simulations. One final aspect to note is given by a tendency in GA auralizations of assessing the differences with respect to measured reference, for all the parameters but localization, between slightly and rather different (responses among (2) and (3)).

According to Table 5.20 and Table 5.21, exploratory statistical results for the saxhorn confirm that all the parameters were assessed as having more similarity with respect to the measured reference, in the hybrid approach. For instance, the results in this numerical technique indicate that brightness tended to a value of (3.5) at receiver position 2, warmth presented a tendency to (3.5) at the same location, reverberation tended to (3) at position 3 and in the case of localization, the trend indicated (4.5) at receiver positions 2 and 3. For the hybrid approach, the lowest rate is given by the reverberation parameter with an average of (2.5) at location 1. In the case of GA auralizations, the highest tendency scores are obtained at receiver position 3 for reverberation (2.7) and warmth (2.5) and location 2 for the parameters of localization (4.0) and brightness (2.7). The lowest subjective assessment in this numerical approach is given at position 1 on reverberation (2.3).

The exploratory statistical results for the bass drum in section 5.5.3.2.2 indicate that the auralizations created by means of the hybrid approach were assessed as being more similar with respect to reference auralizations than GA simulations, for all the parameters and receiver positions (see Figure 5.66). The observation of Table 5.22 and Table 5.23 corroborates this situation, providing tendency results for the hybrid approach sorted from the highest to the lowest as follows: localization (4.8), warmth (4.0), reverberation (3.8) and brightness (3.5), the first two in position 2 and the last two in receiver location 1. This numerical technique has presented the lowest rate for the parameter of reverberation (3.2) at position 3. The GA subjective results exhibits the best tendencies at position 2 for the parameters of localization (4.0), warmth (3.1) and brightness (3.0), and location 1 for reverberation (3.3). In this case, the parameter exposing the lowest score was the brightness (2.7) at position 3.

Figure 5.67, Table 5.24 and Table 5.25 illustrate the exploratory statistical results for male voice auralizations, applying both numerical approaches. In the Box-and-Whisker plots it is possible to appreciate that every single parameter obtained higher scores in the hybrid approach, for all the positions. Tabular information for this numerical technique indicates that receiver location 1 had the best simulation results for the parameters of localization (4.6), warmth (3.2) and reverberation (2.9), and position 3 presented the highest score for brightness (3.5). In contrast, at the same position, reverberation (2.8) obtained the lowest average response. The tendency results for the GA auralizations are

sorted from the highest to the lowest as follows: localization (4.2) at position 2, brightness (3.1) in receiver location 1, warmth (2.4) at position 2 and reverberation (2.6) in location 1. In this case, reverberation (2.5) has achieved the lowest score at position 2.

As an overall analysis, it is important to note in the hybrid approach regarding the acoustic source and receiver positions assessed, that the bass drum and location 2 were the instrument and place in the room obtaining the most similar simulation results in comparison to the measured reference. In this case, parameters tended to indicate that subjective variables sorted from the highest to the lowest for bass drum are as follows: localization, warmth and reverberation. For the brightness parameter, the highest score tendency in each instrument is approximately the same value; nevertheless, this situation is given in different positions for each acoustic source.

6. Application of the auralization system to evaluate the acoustical conditions of a classroom and cognitive processes

This chapter describes the theoretical background of the methods involved in the two applications researched for the auralization system. The first application consists of the subjective evaluation of intelligibility and listening difficulty, given by the modification of the variables of reverberation time and background noise levels. This section presents the acoustic indicators to assess a classroom, the acoustic treatment design and numerical implementation, the creation of the auralizations and the results and discussion of the subjective tests of Intelligibility and listening difficulty. The second application involves the assessment of the impact on cognitive processes such as attention, memory and executive function, taking into account the same independent acoustic variables. This section considers the methodology to create four groups of auralizations with different conditions of background noise levels and reverberation times, a description of the psychological tests, the pilot study and the results and discussion of the application of the psychological tests evaluating the cognitive processes of attention, memory and executive function by means of auralizations.

6.1 Application of the auralization system to evaluate the acoustical conditions of a classroom

In this section, the techniques used to apply the auralization system to assess the existing and future acoustical conditions of a classroom, are described. First, the acoustic indicators and the correspondent limit values chosen to assess a classroom in acoustic terms are exposed. Afterwards, there is an explanation of the procedures applied to design a theoretical acoustic treatment in order to meet the acoustic limits established previously. Finally, the procedures applied to design a subjective test evaluating intelligibility and listening difficulty for the classroom, taking into account both conditions, are explained.

6.1.1 Acoustic indicators to assess a classroom

Given the negative effects of unfavourable acoustic conditions in learning spaces, which have been reviewed in section 2.3.1, many standards and recommendations for classrooms have been proposed in an important number of countries. As mentioned in the literature, the main acoustic parameter to assess a classroom in terms of teaching-learning practice is given by the speech intelligibility. Taking into account existing evidence suggesting that excellent intelligibility is only possible with a great effort from the listener, a new measure called "listening difficulty" has been proposed, which is defined as the percentage of responses, in an intelligibility test, indicating some level of difficulty. In terms of room acoustics, there are two main parameters related to intelligibility and listening difficult, these are the background noise level and the reverberation time.

Nowadays, there are compulsory limit values recommended for the last two variables in learning spaces. In terms of background noise levels, although there is a range between 30 and 50 dB(A) for maximum interior noise, most of the standards and recommendations set a value of 35 dB(A). The application of NC curves (Noise Criteria) establishing a maximum noise level as a function of frequency, is given occasionally. Regarding reverberation time, most standards and recommendations state a maximum between 0.4 and 0.8 seconds in the octave bands of 500, 1k and 2k Hz, or the arithmetic mean in these bands, which is referred as the mid reverberation time (T_{mid}). However, only Belgium, the American Speaking-Language-Hearing Association (ASHA), the United Kingdom, Australia, New Zealand and the American National standards Institute (ANSI) specify that reverberation time values are given for empty classrooms. In Table 6.1, there is a summary including limit values set in different countries for these two variables in classrooms. In Colombia, the technical standard NTC 4595 of 2006 established acoustic criteria performance of classrooms, defining a maximum background noise level of 40-45 dB(A) and reverberation time between 0.9 and 1.0 seconds.

Table 6.1: Summary of acoustic criteria in terms of background noise levels and reverberation time, given in different countries.

Country (Organization)	Background noise criteria (dB(A))	Reverberation time criteria (s)
Australia	35	0.4-0.6
Belgium	40	0.4
Brazil	40-50	
Canada		0.7
China	40-50	0.9-1.0
Chile	35-40	0.6-0.7
France	33-38	0.4-0.6
Italy	36	
New Zeeland	35	0.4-0.6
Portugal	35	0.6-1.0
United Kingdom	35	0.6-0.8
Sweden	30	
Turkey	45	
USA (ANSI)	35-40	0.6-0.7
USA (ASHA)	30	0.4
USA (ASHRAE)	NC30	
W.H.O	35	0.6

6.1.1.1 Background noise criteria

The noise indicator used to characterize the interior background noise level of the classroom was the equivalent continuous sound pressure level weighted “A”, this acoustic parameter was measured according to the ISO standard 1996:2003 over thirty minutes. The background noise criteria for learning environments is based on the signal-to-noise ratio of +15 dB(A), which is necessary to ensure understanding of a spoken message at average voice level of 50 dB(A) according to appendix B of ANSI standard S12.6 (2010). Even though, most international standards recommend an interior maximum background noise level of 35 dB(A), these are directly related to classrooms in schools, since children are especially sensitive to adverse acoustic conditions. Hence, it is expected that adult students have less difficulty in understanding a spoken message in the same acoustic conditions.

Taking the last into account, in this research the background noise criteria was established according to Sato et al (2012) work, in which a maximum background noise level of 45 dB(A) was defined for a voice level of 60 dB(A) at one meter distance. Nevertheless, it should be noticed that speech level at a receiver position is not necessarily 60 dB(A). According to the above, the reference signal-to-noise ratio value of +15 dB(A) could not be considered if the distance between source and receiver is larger than one meter. Nevertheless,

according to Sato et al (2005) the contribution of early reflections in the sound pressure level at a receiver position in a room, given by a human voice sound source, can significantly increase the effective level of the voice. This is the case of classrooms of small dimensions with a considerable percentage of reflective surfaces, as the classrooms at the San Buenaventura University. It is important to bear in mind that the mentioned background noise does not consider the fluctuating behaviour of pressure with time or tonal characteristics, since it is based on a stable noise pressure signal and a flat frequency response. Nonetheless, the Sato et al study was implemented in Japanese language, results were taken as reference, since studies of listening difficulty for Spanish language have not been documented in the literature.

6.1.1.2 Reverberation time criteria

Reverberation time is an important factor that can affect the quality of speech communication in a room. The excess of reverberation generates a degradation of speech intelligibility, caused by a masking effect and an increase of background noise levels. In this research, the recommendation of the building Bulletin 93 was taken as a reference. This bulletin established a mid-reverberation time of less than 0.8 seconds, estimated as the arithmetic average of the octave bands of 500 Hz, 1 kHz and 2 kHz, for classrooms of no more than fifty people and without any furniture inside the room. Even though this recommendation is given for schools, it is a suitable reference as a start point, given that adult listeners are less sensitive to adverse acoustic conditions. For the lower octave band frequencies, a value less than 1.0 second was established as reverberation time criteria in order to preserve a balance between mid and low frequencies that had not effect on the speech intelligibility. Other reason to establish these time criteria for the lower band frequencies is given by the *just noticeable different* (JND) stated in the ISO standard 3382 for the perceived reverberance, which indicates than a difference of 5% is perceptible for a listener.

6.1.1.3 Speech intelligibility and STI

Speech intelligibility can be defined as the percentage of words or sentences that are correctly understood from a message by a group of listeners. In a room, intelligibility and listening difficulty parameters define the speech transmission quality, which is a function of the signal-to-noise ratio and the architectural acoustics; these two characteristics are related to reverberation time and

background noise. Objectively, one way to assess speech intelligibility is given by the STI.

STI evaluates the effect of a transmission channel on intelligibility, considering the noise, the nonlinear distortion and the signal characteristics in time and frequency (Steeneken & Houtgast, 1980). A speech pressure signal varies in amplitude over time and therefore, has a temporal envelope. The slow fluctuations of the envelope correspond to words or phrase articulation and the rapid variations match the sounds produced by individual phonemes. Therefore, preserving the temporal envelope amplitude is important to obtain excellent intelligibility. The STI determines the degree to which a variation in amplitude over time is affected by a transmission channel using a Modulation Transfer Function (MTF) (International Electrotechnical Commission, 2003).

The IEC 60268-16 standard establishes the necessary and valid methods to calculate STI considering the influence of masking and the gender of the speaker. The MTF required for this calculation can be obtained directly or indirectly. The direct method involves the use of 98 special test signals of ten seconds duration, each one with a modulation frequency for noise in one octave band between 125 Hz and 8 kHz (Steeneken & Houtgast, 1980). Hence, a single measurement using this method requires about 15 minutes. On the other hand, the indirect method involves obtaining the MTF from the impulse response of the transmission channel.

According to ISO standard 9921, STI has a strong direct relationship with intelligibility subjective classification ranges. Although in English language this measure has been extensively studied, little evidence can be found in the literature to verify STI values in comparison to subjective ranges in Spanish language. Rosas & Sommerhoff (2008) gave an example of this relationship in a subjective study applying a list of words with CVC logatoms of Latin American Spanish. The intelligibility classification of CVC logatoms obtained by Rosas & Sommerhoff, can be seen in Table 2.1.

6.1.2 Acoustic treatment design theory

In this section, the acoustic design proposal procedures applied to meet the assessment acoustic indicators of reverberation time and background noise are described. First, the theoretical basics taken into account to estimate the sound

pressure level due to a point source excitation in a room are explained. Second, sound insulation fundamentals are reviewed in order to consider background noise in the acoustic design. The last point considered two main aspects: sound pressure field measurements and reverberation time estimation by means of theoretic Sabine model and numerical GA approach. This section finalizes with the considerations taken into account in the acoustic design and a description of the signal processing applied to include background noise to auralizations.

6.1.2.1 Sound pressure levels estimation in a room

To calculate the sound pressure level generated by a point source at a specific position in a room, the following equation can be used:

$$L_T = 10\log_{10}\left(10^{\frac{L_d}{10}} + 10^{\frac{L_r}{10}}\right) \qquad [6.1],$$

where, L_T is the total sound pressure level at that point, given by the energetic sum of the direct sound pressure level coming from the source L_d and the reverberant sound pressure field contribution L_r. The first term is defined as (Bies & Hansen, 1996):

$$L_d = L_W + 10\log_{10}\left(\frac{Q}{4\pi r^2}\right) \qquad [6.2],$$

where, L_W is the acoustic power level of the source, Q is the directivity factor, which is defined according to equation [3.36], and r is the distance between source and receiver position. The second term in equation [6.1] can be calculated assuming a diffuse sound field in the room, applying the following equation:

$$L_r = L_W + 10\log_{10}\left(\frac{c_0 T_{60}}{13.81V}\right) \qquad [6.3],$$

where, V is the volume of the room.

The distance from the source in which the energy contribution from direct and reverberant field is equal, receives the name of reverberation radius (Kuttruff, 2000) (see equation [6.4]).

$$r_{rev} = 0.0566\sqrt{\frac{V}{T_{60}}} \qquad [6.4].$$

6.1.2.2 Sound insulation basics

When a sound wave is incident upon on a surface, part of the acoustic energy is transmitted to the other side, as can be seen in Figure 3.5. The acoustic power transmission coefficient (τ) is given by the ratio of the incident and the transmitted sound intensities. This coefficient can be expressed on a dB scale and is defined as the Transmission Loss (*TL*), or also known as the Sound Reduction Index (*R*) (Fahy, 2001):

$$TL = 10\log_{10}\left(\frac{1}{\tau}\right) \qquad [6.5].$$

In the case of a composite partition, a total *TL* can be estimated from the following equations:

$$TL = 10\log_{10}\left(\frac{1}{\tau_{tot}}\right) \qquad [6.6],$$

$$\tau_{tot} = \frac{A_1\tau_1 + A_2\tau_2 + A_3\tau_3 + \ldots}{A_{tot}} \qquad [6.7],$$

where, A_1 and τ_1 represents the area and transmission coefficient of the first partition, A_2 and τ_2 from the second, and so on.

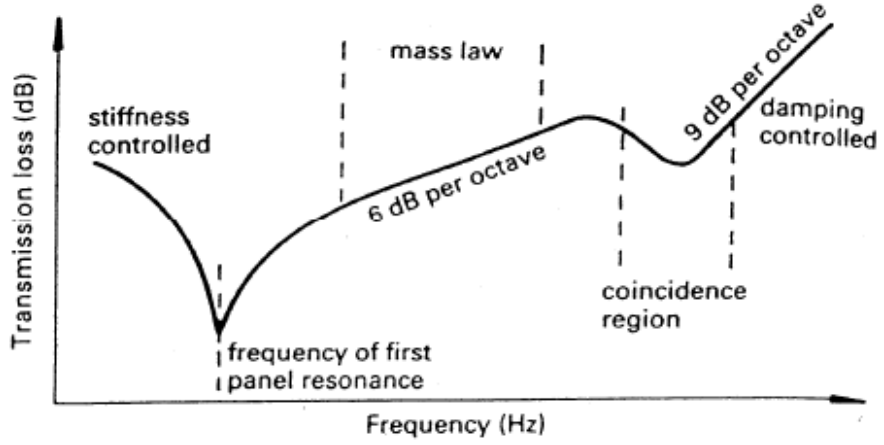


Figure 6.1: Typical Transmission Loss curve. Taken from (Bies & Hansen, 1996).

This coefficient is a function of the first resonant frequency and the critical frequency of the partition. According to Bies & Hansen (1996), three main frequency ranges can be identified when an incident plane wave strikes a surface with a normal angle. The first range is defined by the frequencies below the resonant frequency; the second range is the so called, "*mass law",* given by the frequencies between the first resonant and the critical frequency. The last frequency range is given by the frequencies above the critical frequency, defined as the frequency at which the incident wave coincides with the bending wave of the partition. According to the authors, a typical *TL* curve presents the form appreciated in Figure 6.1.

Below the first resonance frequency, partitions behaviour is controlled by the stiffness of the structure. In this case, the *TL* can be approximated as follows (Fahy, 2001):

$$TL \approx 20\log_{10}\left(\frac{s}{2\omega\rho_0 c_0}\right) \qquad [6.8],$$

where s denotes stiffness per unit area. When the impedance of the partition is large compared to the characteristic impedance of the fluid, such as air, the *TL* can be estimated with the following equation (Fahy, 2001):

$$TL \approx 20\log_{10}\left(\frac{\omega\mu}{2\rho_0 c_0}\right) \qquad [6.9],$$

where, μ is the mass per unit area of the partition. The last is known as the *mass law*, a frequency region in which there is a *TL* increase of 6 dB per octave band every time mass is doubled. For air fluid, expression [5.9] is reduced to (Fahy, 2001):

$$TL = 20\log_{10}(f\mu) - 42 \qquad [6.10].$$

In room acoustics, the incident sound field can be approximated to an ideal "diffuse field", in which plane waves propagating in all directions is assumed. For that reason, it is important to define a *TL* at different angles. In this sense, a *TL* for an incident plane wave at an angle φ can be defined as (Fahy, 2001):

$$TL(\phi) \approx 20\log_{10}\left(\frac{\omega\mu\cos\phi}{2\rho_0 c_0}\right) = TL(0) + 20\log_{10}(\cos\phi) \leq TL(0) \qquad [6.11],$$

where, $TL(0)$ refers to *TL* at normal incidence. Taking into account that a *TL* at any angle φ is less than $TL(0)$, an average TL for field incidence can be approximated as (Fahy, 2001):

$$TL_f = TL(0) - 5 = 20\log_{10}(f\mu) - 47 \qquad [6.12].$$

For frequencies above the critical frequency f_c, the TL_f is defined as (Fahy, 2001):

$$R_d = R(0) + 10\log_{10}\left(\frac{f}{f_c}\right) + 10\log_{10}(\eta) - 2 \qquad [6.13],$$

where, η is a loss factor associated to the stiffness of the partition.

Considering that *TL* varies with respect to frequency, it is convenient to define a single number quantity to characterize the insulation provided by a structure, which facilitates the estimation of internal noise level due to sound transmission through a partition. The sound Transmission Class (STC) is one of the most widely used indices to describe sound insulation in air fluid. It is defined according to the ASTM E413 standard, by adjusting *TL* values measured in 16 third octave bands between 125 and 4 kHz, to a reference STC curve. Another important single quantity to mention is the Weighted Sound Reduction Index R_W, calculated with a similar procedure, this time according to ISO standard 717. As stated in appendix five of BB93 (2004), when the R_W of a façade is known, it is possible to estimate the internal noise level due to sound transmission through the partition, applying the equation established in the standard BS EN 12354:

$$L_I = L_E - R_W + 10\log_{10}\left(\frac{S}{V}\right) + 11 + 10\log_{10} T_{mid} \qquad [6.14],$$

where, L_I is the internal noise level given by sound transmitted through the partition, L_E is the external noise level or outside the room level, S is the surface area of the element, V is the volume of the room and T_{mid} is the mid reverberation time.

6.1.3 Acoustic design

After finishing with the acoustic diagnostic of the classroom, the respective calculations aiming to reduce reverberation time and background noise levels were carried out, in order to come closer to the acoustic criteria established in section 6.1.1 of background noise levels less than 45 dB(A) and T_{mid} less than 0.8 seconds. Regarding T_{60} criteria, it is important to note that different octave band values were assigned, having as a goal to obtain 1 second T_{60} for the octave bands of 125 and 250 Hz, and 0.8 seconds for the other octave bands. Afterwards, using Sabine equations [3.30] and [3.31] and T_{60} measurements results, the average absorption coefficient and surfaces areas necessary to achieve the desired reverberation times were determined.

Regarding background noise, equation [6.14] was used to estimate the variation of background noise level given by the changes of reverberation time. Taking into account these values and the background noise levels measured, a new internal background noise level was estimated for each octave band. At this point it is important to state that, although expression [6.14] is given for cases in which external noise is coming from partitions next-door the façade of a building, it provides a good approximation in cases where room partitions have neighbouring corridors.

Regarding the reverberation time control, the first step after estimating the necessary absorption areas with equations [3.34] and [3.35], was to select from libraries materials having appropriate acoustic absorption coefficients to add in the room. The next step consisted in locating the materials chosen in the room, taking into account the room geometry and the corresponding absorption areas.

6.1.4 Acoustic treatment design and numerical implementation

The next step consisted in calculating the area of acoustic materials to be added, according to its corresponding octave band absorption coefficients and the geometry of the classroom. The materials selected were fiberglass of 4 inches thick protected by a decorative veil and a membrane resonator composed of a 4 mm plywood sheet, with a 7.5 cm cavity and 25 mm of mineral wool on the partition. In order to place the acoustic material on room walls, three hypothetic panels (A, B and C) were designed. Panels A and B corresponded to the fiberglass supported in a 5 cm width frame, with the same thickness of the absorbent

material and the following dimensions for panel A, 1 m x 2.16 m and 0.7 m x 2.16 m for panel B. Panel C was given by the membrane resonator with dimensions of 1m x 2.16m. The projected location for the panels can be seen in Figure 6.2. Table 6.2 describes the surfaces presented in the classroom after the acoustic treatment has been considered, taking into account the materials and their corresponding areas and absorption coefficients. The analytical Reverberation times, which are the result of the implementation of the acoustic treatment, are shown in Figure 6.4.

In order to estimate the variation of interior background noise levels due to the acoustic treatment designed, the procedure explained in section 6.1.2 was applied. Equation 6.14 was used with the purpose of calculating the new noise levels given by the modification of the T_{60} in octave bands. The resulting background noise level was 43.6 dB(A), which meets the limit determined by local directive of 45 dB(A), as described in section 6.1.1. For that reason, no additional calculation regarding the sound reduction index provided by the façade was implemented. The differences between noise levels measured and estimated after the hypothetic acoustic treatment can be seen in Figure 6.3.

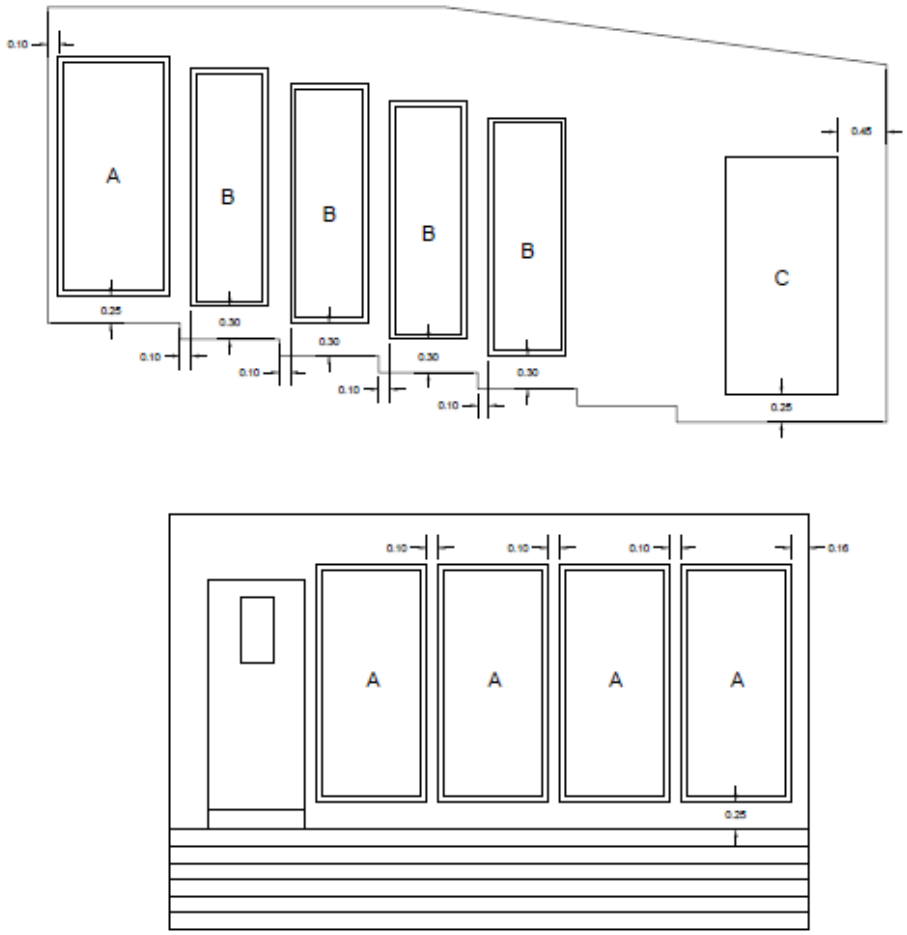


Figure 6.2: Projected location of absorbent acoustic panels. Above, placement configuration for left and right walls. Below, panels position on back wall.

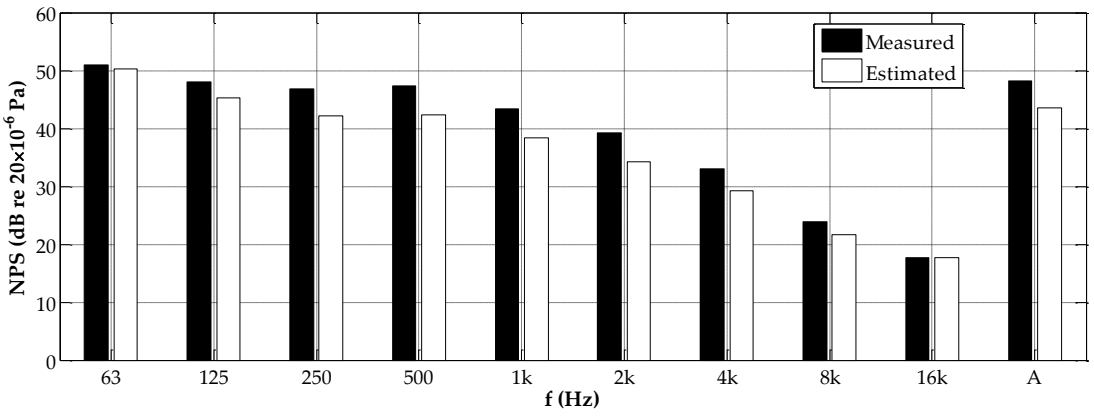


Figure 6.3: Background noise levels measured and estimated after the application of the hypothetical acoustic treatment.

Table 6.2: Materials, areas and absorption coefficients used on each surface considering the hypothetical acoustic treatment, for Sabine and GA models.

Surface	Material	Area (m²)	Reference	Octave Band Centre Frequency (Hz)					
				125	250	500	1000	2000	4000
				Absorption Coefficients					
Floor	Tile	48.99	Cox	0.01	0.01	0.02	0.02	0.02	0.02
Doors	Wood	4.06	Petersen	0.14	0.10	0.06	0.08	0.10	0.10
Windows	Glass	0.84	Karlen	0.35	0.25	0.18	0.12	0.07	0.04
Board	Acrylic	2.91	Kutruff	0.02	0.02	0.03	0.03	0.04	0.05
Panel	Foam	0.50	Bies & Hansen	0.08	0.22	0.55	0.70	0.85	0.75
Air conditioning	Plastic	1.38	Kutruff	0.02	0.02	0.03	0.03	0.04	0.05
Video projector	Plastic	0.30	Kutruff	0.02	0.02	0.03	0.03	0.04	0.05
Lights	Metal	5.04	Kutruff	0.02	0.02	0.03	0.03	0.04	0.05
Right wall	Plaster	12.04	Karlen	0.12	0.10	0.08	0.06	0.06	0.06
Left wall	Painted concrete	14.06	Petersen	0.01	0.01	0.01	0.02	0.02	0.02
Back wall	Plaster	5.90	Karlen	0.12	0.10	0.08	0.06	0.06	0.06
Front wall	Painted concrete	12.75	Petersen	0.01	0.01	0.01	0.02	0.02	0.02
Ceiling	Plaster	39.63	Karlen	0.12	0.10	0.08	0.06	0.06	0.06
Panels A and B	Fiberglass (4 inches)	21.01	Kinsler	0.45	0.90	0.95	1.00	0.95	0.85
Panel C	Membrane resonator	4.32	Petersen	0.58	0.28	0.08	0.04	0.05	0.01

6.1.4.1 GA simulations of the classroom including the acoustic treatment

The creation of the GA model with the projected acoustic treatment considered the theoretical methods explained in Chapter 3 and the same absorption coefficients described in Table 6.2. Scattering coefficients were assigned according to the dimensions of the elements presented in the model leaving a default coefficient of 0.1 in all the frequency bands for large surfaces, as recommended by the software user manual (CATT, 2007). The GA model including the acoustic treatment can be seen in Figure 6.5 and the T_{60} obtained by means of GA simulations in Figure 6.4.

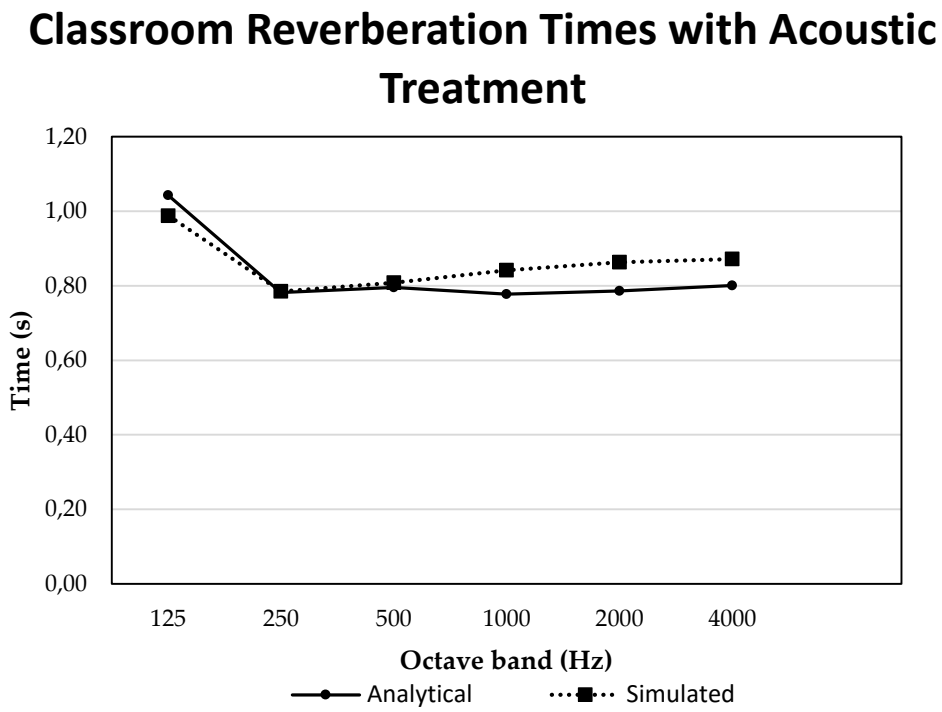


Figure 6.4: Classroom Reverberation times estimated by means of Sabine model and GA numerical method, after considering the designed acoustic treatment.

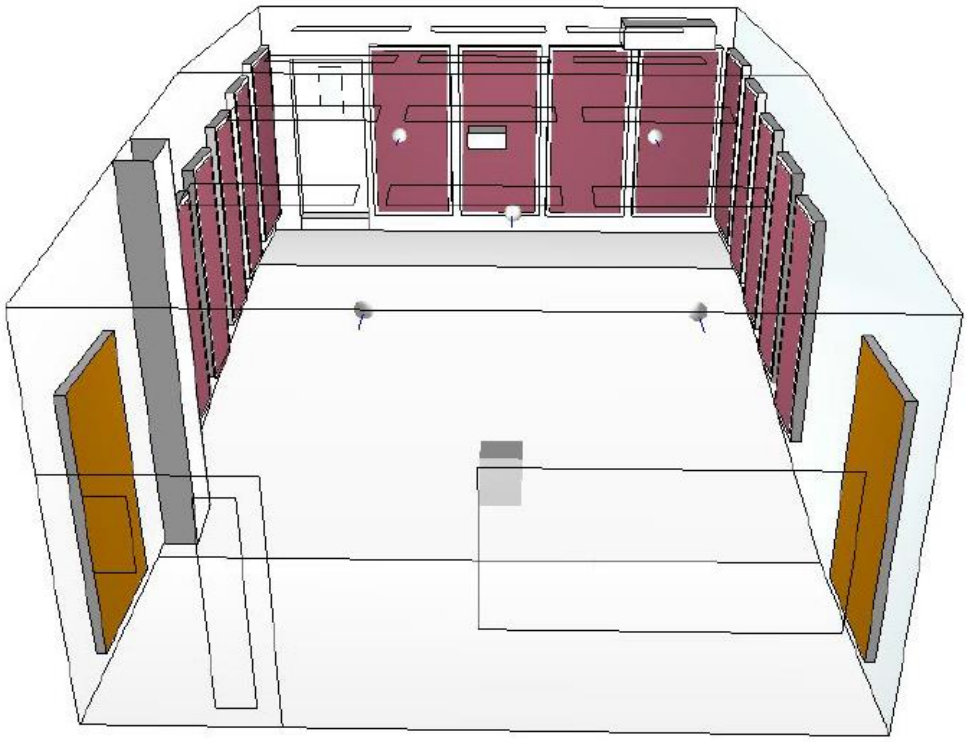


Figure 6.5: GA model of the classroom including the hypothetic acoustic treatment, simulated in CATT-Acoustic software.

6.1.4.2 Procedure to add background noise to auralizations

To consider the ambient noise in the auralizations of the classroom, a practical approach had to be proposed to add background noise, taking into account that the same signal-to-noise ratio for each source-receiver combination had to be achieved for both acoustic conditions: existing and calculated. For a particular source-receiver combination, equation [6.1] was used to calculate the direct and reverberant sound pressure levels in the classroom. In order to do that, the first step consisted of estimating the acoustic power level of the source considering a typical spectrum and directivity factor of a male voice speaking at normal loudness. With this value, it was possible to estimate direct and reverberant sound pressure levels of expressions [6.2] and [6.3] at any position in the room. Afterwards, the signal to noise ratio was estimated using the spatially averaged background noise measurement results (see chapter 4).

The background noise was recorded on the center position of the classroom (position 5), using a binaural recording head of 01dB – Cortex MK2B during 23 minutes. Then, for the purpose of having a continuous background noise level, the sound level variations per second of the signal (A Weighted) were found using the MATLAB's function '*diff*'. Knowing every moment for which the signal has a variation greater than 4 dB, and deleting each part of the signal using a digital audio station software (DAW), Reaper® v.5. Figure 6.6 shows the sound level variation of the recorded signal, which exceeds a difference of 4dB, Figure 6.7 shows the resulting signal without variations greater than 4 dB, satisfying main goal of having a continuous temporal behavior for the background noise.

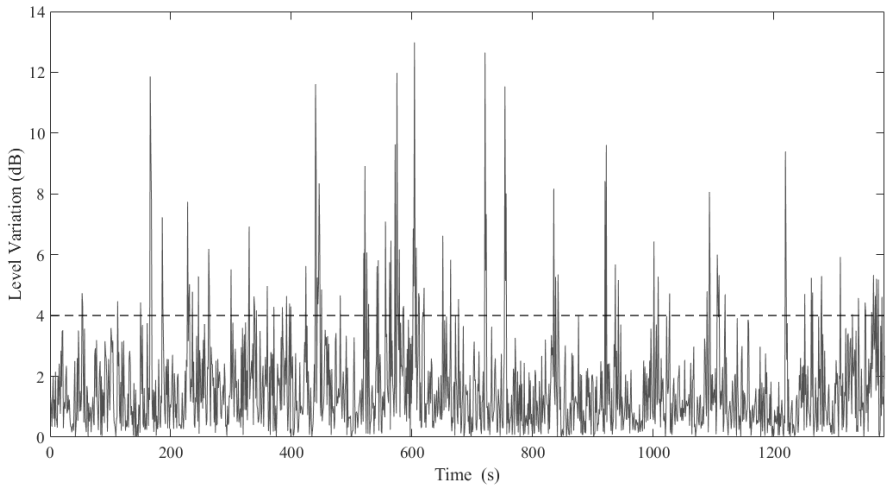


Figure 6.6: Sound level variations of the recorded signal.

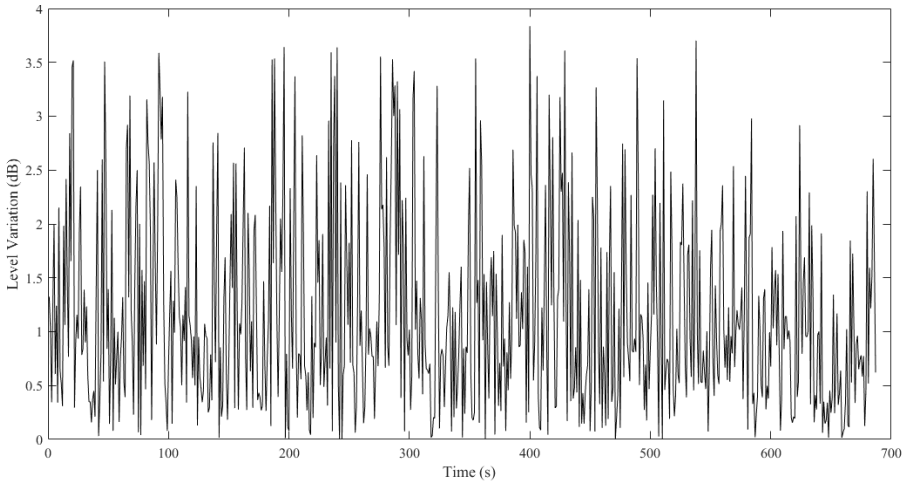


Figure 6.7: Sound level variation of the edited signal.

The next step was the inclusion of the noise in the auralization. For this procedure was used the OPSODIS Marantz (Model No: ES7001/U1B – Serial No: 20000803001018) as sound reproduction system. The standard ANSI S12.60 (Acoustical Society of America, 2002) recommends a background noise level of 35 $dB_{LAeq,\ 1h}$ for learning spaces, in this order the ideal condition must be the one which satisfy the international standards requirements. The worst condition would be the one which simulate the acoustics behaviour of the classroom with the acoustic treatment, in that order the sound pressure level for the background noise was 43.6 dBA.

The binaural background noise signal was reproduced using Reaper® V5 on a laptop, from where the amplitude of the signal was controlled. The equivalent continuous sound pressure level was measured at 1.8 m in front of reproduction system, using class 1 sound level meter CESVA SC310 for 11 minutes. The amplitude of the signal was changed until the L_{Aeq} satisfy the sound pressure level of desired conditions. Figure 6.8 shows the measurement set up for checking noise levels in the reproduction system.

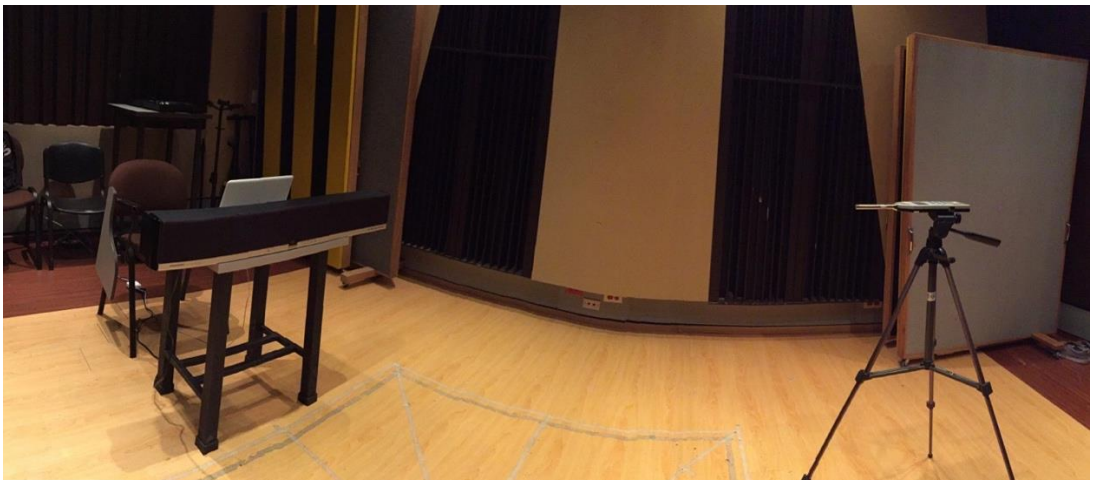


Figure 6.8: Measurement set up to include noise in the auralizations.

Consequently, the auralizations were reproduced according to signal-to-noise ratio calculated for the central position. The amplitude level of the binaural noise condition was set as a reference point. The auralizations were reproduced using Reaper® V5. Without modifying the amplitude of the continuous noise conditions, the auralization were added and reproduced simultaneously with each condition. The auralizations signal amplitude were manipulated with Reaper® in order to achieve the signal-to-noise ratio criteria for both conditions. Figure 6.9 shows the amplitude level of each signal on Reaper®.

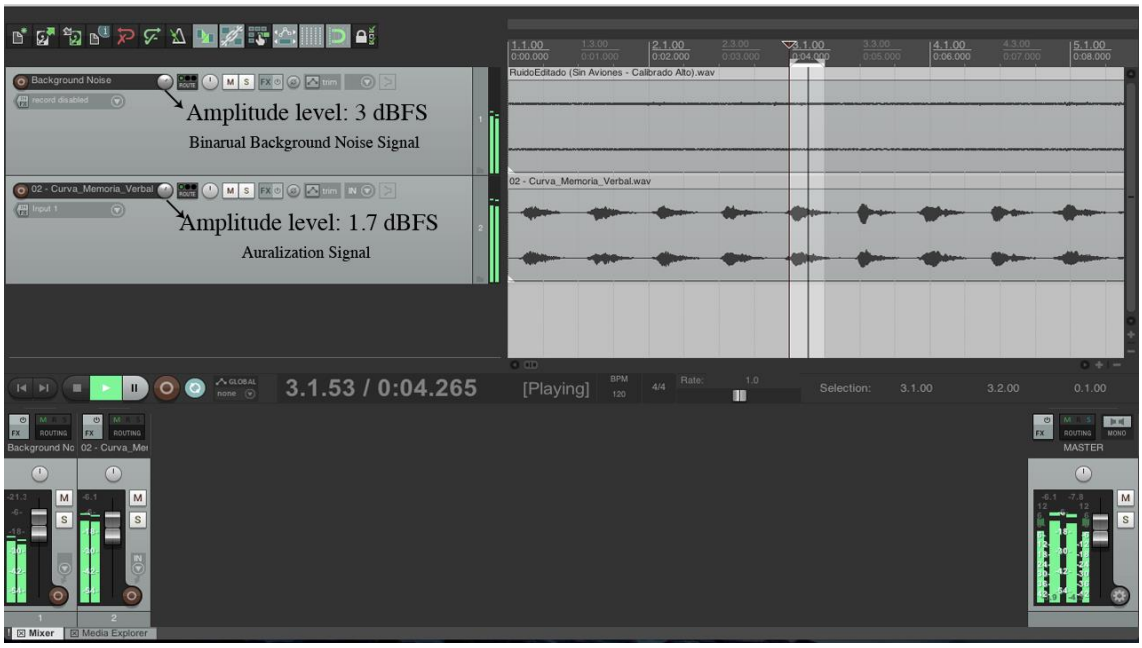


Figure 6.9: Amplitude of binaural background noise and auralization signals in Reaper®.

6.1.5 Intelligibility and listening difficulty subjective test

To assess subjectively the present conditions of a classroom and the impact of implementing an acoustic treatment, a comparative exploratory study was conducted using auralizations and a sample of 40 people. The classroom acoustics was taken as an independent variable in two different situations (present and with acoustic treatment), and as dependent variables, intelligibility and listening difficulty were assessed. The first situation corresponded to the current acoustic conditions, characterized by BIR measurements, at five different receiver positions distributed inside the classroom. The second condition considered the same receiver positions in a numerical GA simulation, but this time an acoustic treatment has been included in the classroom. In order to evaluate the influence of background noise over the dependent variables, the study was carried out again; nonetheless, this time background noise was added to the auralizations.

In the intelligibility test, each participant was assigned one of the five receiver positions in the classroom, thus eight people evaluated each source-receiver combination. The auralizations corresponding to both conditions were reproduced in the recording studio, by means of binaural reproduction system OPSODIS. In the test form (see Appendix D: "Intelligibility and listening difficulty subjective test"), the participant wrote the logatom they were able to understand. Intelligibility was assessed according to the percentage of correctly written words. The five logatom lists used are shown in Appendix E: "Lists of Logatoms". At the same time, participants were asked to rate the listening difficulty of each word, according to the following scale:

Table 6.3: Listening difficulty scale. Adapted from Sato´s (2005).

0	No difficulty
1	Little difficulty
2	Moderate difficulty
3	Much difficulty

The listening difficulty was assessed as the percentage of responses different to "0".

6.1.6 Auralizations of the classroom

Four groups of auralizations were created in order to evaluate the influence of the acoustical conditions of background noise and Reverberation Times over the variables of intelligibility and listening difficulty. The first group of auralizations were considered as the reference ones, created by means of BIR measurements in order to have a characterization of the classroom with the existing acoustical conditions. The second group consisted of the auralizations of the classroom considering the acoustic treatment (see section 6.1.4). The last two groups considered the same auralizations created in the first two groups with background noise included for both conditions. The general procedure to create the virtual sound environments was explained in Section 3. The main aspects taken into account to create the classroom auralizations considered the generation of the logatoms and the addition of background noise.

In the sound generation stage, a male voice reading six lists of 40 phonetically balanced logatoms (Rosas & Sommerhoff, 2008) was used to create the sound signals. The lists were recorded at the Recording studio of San Buenaventura University, as explained in section Chapter 3. In the transmission stage, the BIR of the classroom with existing conditions were obtained by means of acoustic measurements (see Chapter 4). The BIR were measured with a 01dB dummy head having one source location and five different receiver positions (see Figure 6.10). To create the simulated BIR the HRTFs from MIT database were used. Two sets of numerical BIR were created, the first with the room current conditions and the second with the acoustic treatment. The first group were used to compare the intelligibility of the measured BIR against the GA BIR in a critical listening exercise. The BIR considering the acoustic treatment were created for the same source-receiver combinations, according to the procedures explained in section Chapter 3 and 6.1.4. The auralizations including background noise were generated following the steps described in section 6.1.4.2. Table 6.4 shows a summary of the four acoustic conditions created to evaluate subjectively both intelligibility and listening difficulty, specifying the T_{mid} and background noise level of the classroom and the source level and SNR at each receiver position.

6.1.7 Intelligibility and listening difficulty tests results

This section presents the results obtained by the application of the subjective tests assessing intelligibility and listening difficulty, all based on auralizations. According to the test design, both parameters (INT and LDFF from now on), were evaluated with existing acoustical conditions (denoted PRE) and considering an acoustic treatment (symbolized by POS). The test was applied twice, including background noise (denoted NOI) in the second one. Figure 6.11 and Figure 6.12 illustrate the spatially averaged estimates of INT and LDFF subjective assessments of the classroom. The corresponding exploratory statistics for both dependent variables can be seen in Table 6.5 and Table 6.7. Table 6.6 and Table 6.8 describe the correlation between each pair of data assessed for INT and LDFF, respectively. Finally, Figure 6.13 illustrates the proportion of students assessing more than 50% of LDFF for all the situations considered.

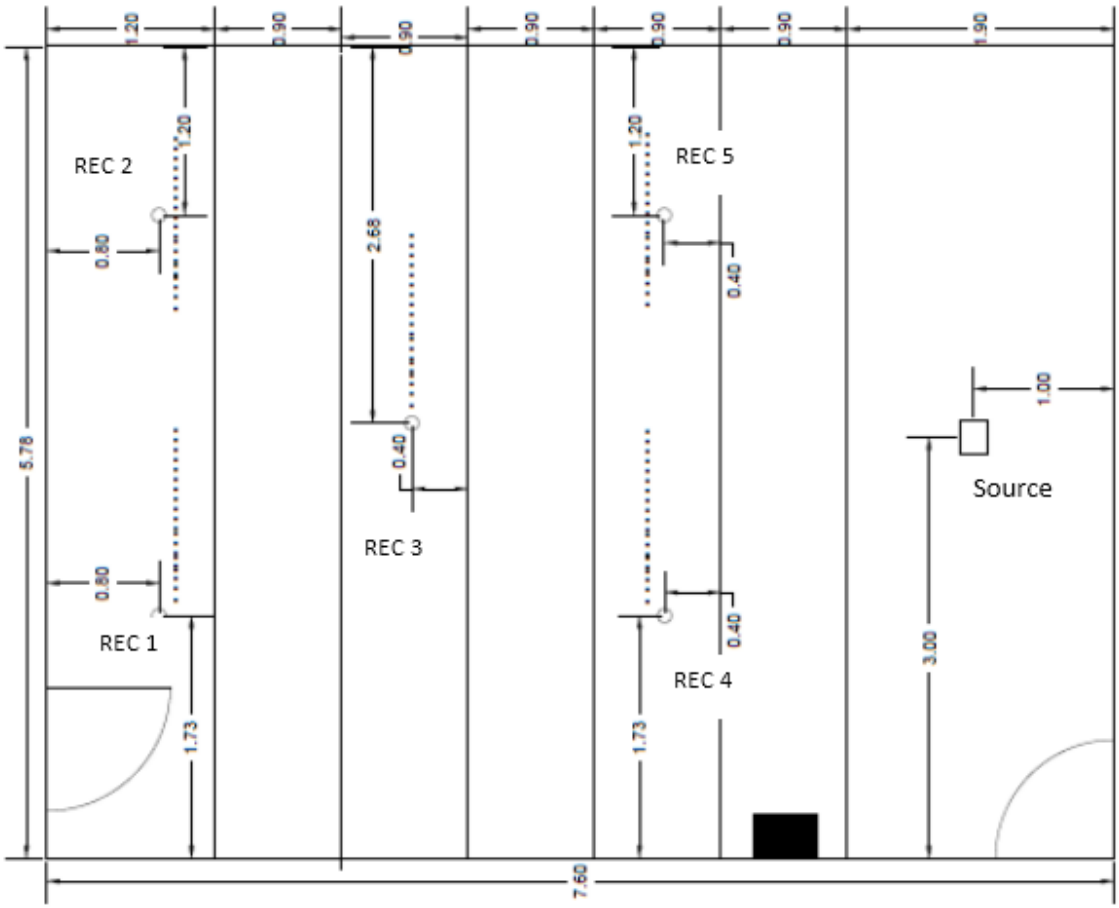


Figure 6.10: Top view of the classroom indicating source and binaural receiver positions, both measured and simulated. The relative height to the floor for source and receivers was 1.5 m.

Table 6.4: Summary of the four acoustic conditions created at each receiver position studied in the classroom, including T_{mid}, source level at receiver position, background noise level and SNR. The existing acoustical conditions and taking into account the hypothetical acoustic treatment denoted as PRE and POS, respectively. The conditions including background noise symbolized as PRE_NOI and POS_NOI.

Condition		PRE	POS	PRE_NOI	POS_NOI
T_{mid} (s)		2.6	0.8	2.6	0.8
Background noise level (dBA)		NA		48.4	43.6
Source level at receiver position (dBA)	1	58.4	53.3	58.4	53.3
	2	58.4	53.3	58.4	53.3
	3	59.2	55.4	59.2	55.4
	4	59.4	55.7	59.4	55.7
	5	58.8	54.3	58.8	54.3
SNR at receiver position (dBA)	1	NA		10	9.7
	2			10	9.7
	3			10.8	11.8
	4			11	12.1
	5			10.4	10.7

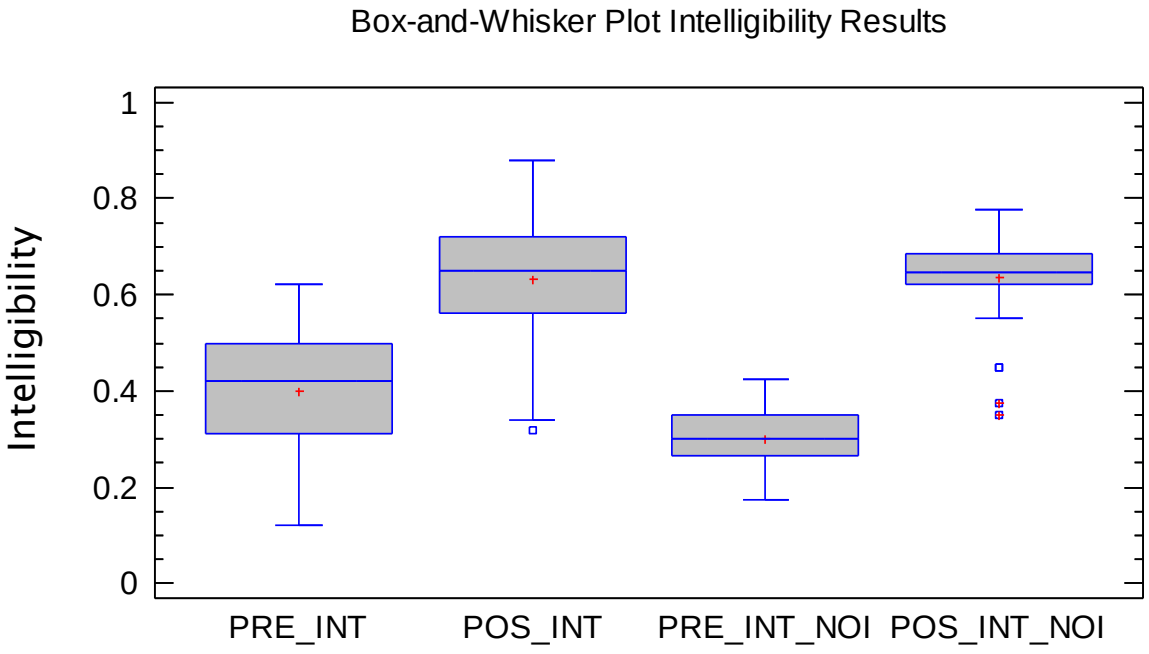


Figure 6.11: Spatially averaged estimates of Intelligibility subjective assessments of the classroom. The results for existing acoustical conditions and taking into account the hypothetical acoustic treatment denoted as PRE_INT and POS_INT, respectively. The results including background noise for both conditions symbolized as PRE_INT_NOI and POS_INT_NOI.

Table 6.5: Summary of exploratory spatially averaged statistics obtained for the Intelligibility subjective assessments of the classroom.

Statistic Indicator	PRE_INT	POS_INT	PRE_INT_NOI	POS_INT_NOI
Count	40	40	40	40
Average	0.4	0.63	0.3	0.64
Median	0.42	0.65	0.3	0.65
Mode	0.44	0.72	0.35	0.63
Standard deviation	0.12	0.13	0.06	0.09
Coeff. of variation	29.21%	21.04%	21.44%	14.32%
Minimum	0.12	0.32	0.18	0.35
Maximum	0.62	0.88	0.43	0.78
Lower quartile	0.31	0.56	0.27	0.62
Upper quartile	0.5	0.72	0.35	0.69
Skewness	-0.58	-0.49	-0.16	-1.62
Kurtosis	-0.13	-0.02	-0.46	3.11

Table 6.6: Pearson correlation coefficient between each pair of Intelligibility tests results, the number of pairs of data values used to compute each coefficient and the *p-value* testing the statistical significance of the estimated correlations.

	PRE_INT	***POS_INT***	***PRE_INT_NOI***	***POS_INT_NOI***
PRE_INT		0.7438	0.1045	-0.2828
		(40)	(40)	(40)
		0.0000	0.5209	0.0770
POS_INT	0.7438		-0.247	-0.5182
	(40)		(40)	(40)
	0.0000		0.1245	0.0006
PRE_INT_NOI	0.1045	-0.247		0.5218
	(40)	(40)		(40)
	0.5209	0.1245		0.0006
POS_INT_NOI	-0.2828	-0.5182	0.5218	
	(40)	(40)	(40)	
	0.0770	0.0006	0.0006	

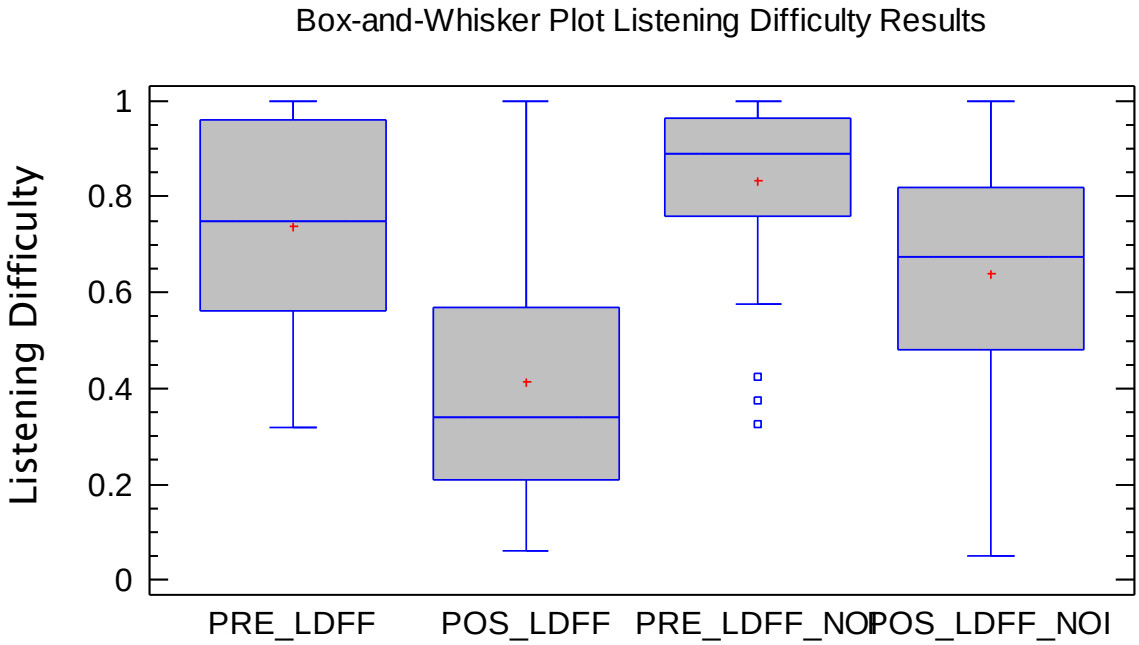


Figure 6.12: Spatially averaged estimates of Listening Difficulty subjective assessment of the classroom. The results for existing acoustical conditions and taking into account the hypothetic acoustic treatment denoted as PRE_LDFF and POS_ LDFF, respectively. The results including background noise for both conditions symbolized as PRE_ LDFF _NOI and POS_ LDFF _NOI.

Table 6.7: Summary of exploratory spatially averaged statistics obtained for the Listening Difficulty subjective assessments of the classroom.

	PRE_LDFF	POS_LDFF	PRE_LDFF_NOI	POS_LDFF_NOI
Count	40	40	40	40
Average	0.74	0.41	0.83	0.64
Median	0.75	0.34	0.89	0.68
Mode	1.00	0.24	1.00	
Standard deviation	0.21	0.28	0.17	0.23
Coeff. of variation	28.92%	67.70%	20.79%	35.62%
Minimum	0.32	0.06	0.33	0.05
Maximum	1.00	1.00	1.00	1.00
Lower quartile	0.56	0.21	0.76	0.48
Upper quartile	0.96	0.57	0.96	0.82
Skewness	-0.30	0.83	-1.42	-0.40
Kurtosis	-1.21	-0.38	1.84	-0.12

Table 6.8: Pearson correlation coefficient between each pair of Listening Difficulty tests results, the number of pairs of data values used to compute each coefficient and the P-value testing the statistical significance of the estimated correlations.

	PRE_LDFF	POS_LDFF	PRE_LDFF_NOI	POS_LDFF_NOI
PRE_LDFF		0.6918	0.1746	0.1163
		(40)	(40)	(40)
		0.0000	0.2813	0.4749
POS_LDFF	0.6918		0.1497	0.0963
	(40)		(40)	(40)
	0.0000		0.3566	0.5543
PRE_LDFF_NOI	0.1746	0.1497		0.8069
	(40)	(40)		(40)
	0.2813	0.3566		0.0000
POS_LDFF_NOI	0.1163	0.0963	0.8069	
	(40)	(40)	(40)	
	0.4749	0.5543	0.0000	

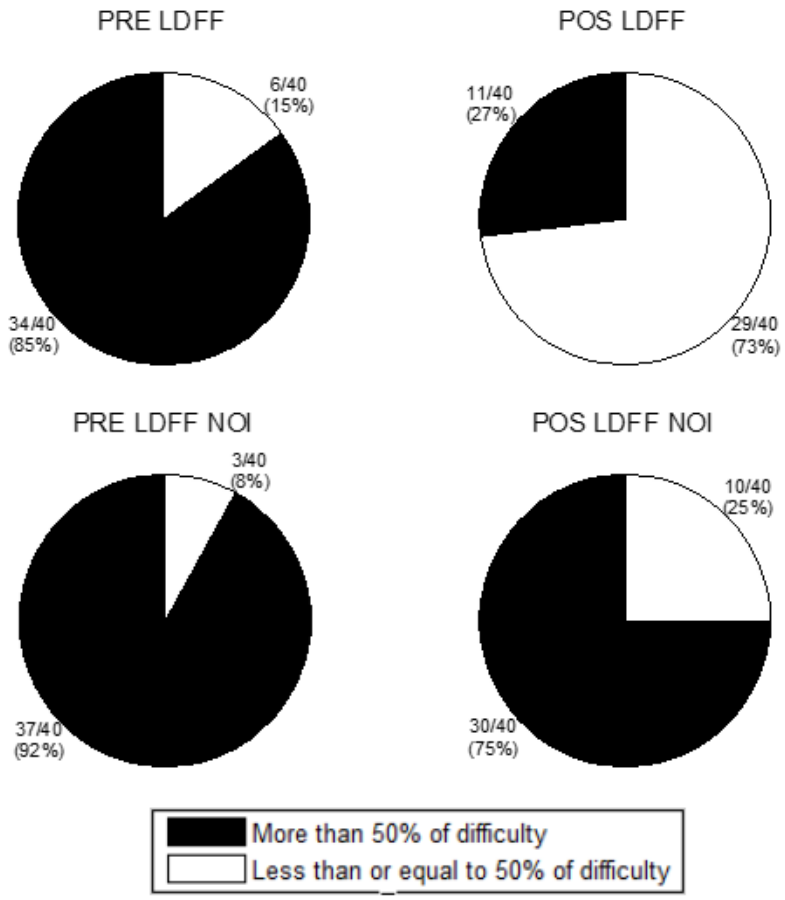


Figure 6.13: Listening difficulty test results, showing the proportion of students assessing more than 50% of difficulty for all conditions.

6.1.8 Discussion of intelligibility and listening difficulty tests results

Figure 6.11 and Table 6.5 clearly illustrate a significant improvement of intelligibility when the subjective assessment includes the designed acoustic treatment. According to the rating scale of ISO 9921, when the scenario without background noise is considered, the spatially averaged intelligibility with existing conditions is assessed as poor (40%); in contrast with the results obtained when the acoustic treatment is considered, which is evaluated as fair (63%). When the background noise is included in the auralizations, similar results are perceived between current and hypothetic acoustical conditions, scoring once again as poor (30%) and fair (64%) respectively; although a more significant difference is given by the numerical implementation of the acoustic treatment, which improves in this case by 10% more the intelligibility assessment.

Table 6.6 shows an analysis of correlation coefficients between each pair of variables in order to quantify the strength of their linear relationship. The underlined number is the *p-value*, which below 0.05 indicates, with a confidence level of 95.0%, a statistically significant non-zero correlation. Considering this, it is possible to distinguish an acceptable positive correlation between PRE_INT and POS_INT variables, having a Pearson correlation value of 0.7438. A linear regression analysis gives a coefficient of determination of 0.5532, in order to explain the variability of the intelligibility in the classroom given by the change of the reverberation times, taking into account the existing conditions and the virtual implementation of an acoustic treatment.

In terms of listening difficulty, Figure 6.12 and Table 6.7 show the positive effect of the implementation of a virtual acoustic treatment for both situations. Without background noise, the Listening Difficulty is reduced from a spatially averaged of 74% to 41%, which gives an improvement of 33%. In the second scenario, the addition of background noise increases the Listening Difficulty to 83% when existing acoustical conditions are considered. In this case, a spatially averaged of 64% is obtained with the acoustic treatment, having a less significant decrease of 19%. It is possible to see by looking at the statistical indicators variation, how the dispersion increases when the hypothetical treatment is considered no matter the presence of background noise. The last ideas suggest that both acoustic dependent variables affect the listening difficulty; although, when the proportion of student´s rating in all conditions is more than 50% of difficulty as

illustrated (see Figure 6.13), it is possible to distinguish that the presence of background noise along with the reverberation have a significant influence on this dependent variable. In this aspect, it is important to note that all the receiver positions analysed are in the reverberant field of the room, which means that the signal-to-noise ratio is dependent of background noise levels and the corresponding reverberant field contribution. Considering this, the signal-to-noise ratio estimates for both scenarios provide similar results of about ±1dB at each source-receiver combination (see Table 6.4), which indicates that background noise presents a similar behaviour with existing and hypothetical acoustic conditions.

The analysis of correlation coefficients between each pair of Listening Difficulty test results can be seen in Table 6.8. In this case, a positive significant statistical correlation is distinguished between PRE_LDFF_NOI and POS_LDFF_NOI variables, with a Pearson correlation value of 0.8069. In this case, a linear regression analysis gives a coefficient of determination of 0.651, which is a statistical measure indicating how well the variability of the listening difficulty in the classroom, might be explained by the change of the reverberation times in the presence of background noise, taking into account existing conditions and the virtual implementation of an acoustic treatment.

6.2 Application of the auralization system to evaluate cognitive processes

This section explains the psychological tests applied to evaluate the cognitive processes and the statistical analysis implemented in order to measure the impact of the acoustic independent variables over the dependent ones. The first part describes the psychological tests used in the subjective tests and the criteria applied to select them. The statistical analysis theory is focused on finding significant differences between acoustic conditions for the components of the psychological tests, evaluating the cognitive processes of study, in order to determine the significance of the variances on those components.

6.2.1 Selection of psychological tests to evaluate cognitive processes

The psychological tests used were selected taking into account the criteria of the duration of the test, the measurement of dependent variables and their

components, and the validity of the tests in a given population. With regard to duration, so that the study was cross-sectional, short-term tests were selected in the application that would ensure the variables' measurement at one time, and that its duration did not produce adverse effects such as onset of fatigue, wear out or boredom, which would have affected the measurement of dependent variables. In the second criteria, tests that made possible the evaluation and measurement of components in attention, memory and executive function were selected, such as response time, long and short term recall, progressive and regressive repetition and comprehension, among others. In the final selection principle, it was made sure that the tests used have measurable relevance (scales) in the Colombian population, to ensure the validity of these in the target population, and to guarantee that the tests used really evaluate the desired variables. The scales are measures used to compare the individual assessed with a regulatory group, in this case the population to which the person belongs. Considering the above criteria, the selected psychological tests are described below.

6.2.1.1 Trail Making Test (TMT)

Partington & Leiter (1949) originally developed the Trail Making Test (TMT) after a while, this test was part of the Halsted-Reitan Neuropsychological Battery test (Reitan & Wolfson, 1985). The TMT was designed to assess sustained attention, visual search speed, mental flexibility and motor function. The test consisted of two parts called A and B, in each one, the participant must locate elements and follow sequences. For the application of the test, the assessor explained to the participant the task to be performed, which was to connect a sequence of numbers in the sheet of paper provided as quickly as possible, following two main directions: the participant could only use straight lines without lifting the pencil from the paper and could not cross out or surpass the circles at any time. Part A consisted of circles that were numbered from 1 to 25 and part B, contained numbers from 1 to 8 and letters from A to G (see Appendix F: Trail Making Test (TMT)). The participant's task was to draw straight lines from the number to the letter, in a sequential way until the end of the exercise (1-A, 2-B, etc.). Each part was marked separately and considered the time in seconds that the participant spent to complete them, omission errors, commission errors and correct answers, where one point was awarded for each right answer (Periáñeza, et al., 2007).

6.2.1.2 Continuous Performance Test (CPT)

The Conners´ Continuous Performance Test (CPT) is a standardized test developed by Conners (1995) to analyse the assessed sustained attention in tasks requiring continuous work. It is designed to assess sustained attention, visual tracking and activation of quick answers. The cancellation test of "A" consists of two parts, a hearing and a visual part, both composed of a matrix of letters with 20 columns and 8 rows, with 160 letters of which 16 are the letter "A" (see Appendix G: Continuous Performance Test (CPT)). The main task in the first part of the test is for the participant to knock once on the table every time they hear the letter A. In the second part, the participant must cross out this letter as quickly as possible. For this research exercise, a Spanish language version validated with the Colombian population is to be used (Ardila, et al., 1994). The marking considers the time in seconds, omission errors, commission errors and correct answers.

6.2.1.3 Wechsler Memory Scale - III

The Wechsler Memory Scale (WMS-III) (Breslow, et al., 1980) is a test to obtain a quick and practical memory assessment, which appears as a result of ten years of research. The WSM - III is one of the tests with the largest standardized and representative database to evaluate memory, including a representative stratified sample of the general population, where 1,250 people ages between 16 and 89 were included (Ardila & Ostrosky, 2012). In 1997, the third version was published, which brings significant changes in relation to the two previous versions. WMS - III includes subtests and composite marks, which are aimed at measuring the functions of memory and attention, using visual and auditory stimuli. Although this scale is composed of nine main indices, for this research exercise, five of them are not used since they do not evaluate the selected variables of interest for this study (Personal Information and Guidance - Visual Memory - Visual Associative Memory – Visual Reproduction I and II - Volume of Visual Memory). The indices to be taken into account are:

- Mind Control: The participant must count numbers from 20 to 1 in descending order, say the alphabet and starting with 1 add 3 to the proceeding number until reaching 40; it evaluates automatic language.

- Logical Memory: It evaluates the ability to immediately recall two stories heard by the participant.
- Digit Span: It evaluates the ability of the participant to immediately recall a list of numbers in the order and reverse order they were heard.
- Paired-Associate: Evaluates the ability of immediately recalling associated pairs of words; the participant has three opportunities to recall them correctly.

For this research, the WMS - III adapted to Spanish (Breslow, et al., 1980) is to be used (see Appendix H: Wechsler Memory Scale – III Test). The Wechsler - III, gives two main scores of memory, an index of immediate memory and a general index of memory. Each subtest has a separate score which enables the use of these for individual analysis.

6.2.1.4 Verbal memory curve (VMC)

It is a commonly used memory test, as described by Lezak (1995), Spreen & Strauss (1998) and Ardila & Rosselli (1992). In this test, participants are read a list of 10 words, which they must recall and repeat back in the order heard. The ultimate goal is to repeat all 10 words in the correct order and the participant has ten opportunities to do this. In addition to this, the participants are given two different time frames of three and twenty minutes within which to recall the words. The initial amount (first recall), maximum amount (the largest number of words that the subject manages to recall), number of trials, the shape of the curve and the delay recall (three and twenty minutes) are scored.

6.2.2 Statistical analysis theory background

This section explains the statistical theory implemented in order to analyse the cognitive test results. The first part includes a nonparametric test such as the Kruskal-Wallis, which is used to determine if there are statistically significant differences between groups of independent variables given by the acoustic conditions, on dependent cognitive variables of attention, memory and executive function. The second part contains an Analysis of variance (ANOVA) in order to compare the means between groups and determine the differences among them.

6.2.2.1 Nonparametric statistics

The nonparametric statistical tests are not based on the assumption that the population sample data belongs to any particular parametric distribution. It is assumed that data in nonparametric procedures is freely statistically distributed, which make them a suitable choice for data distribution not near normal. These techniques do not require the data to be quantitative, it may be categorical, one of the main advantages of nonparametric statistics. The measurement in categorical scales and the sorting by data ranges give some characteristics of these procedures.

6.2.2.1.1 Kruskal-Wallis test

The Kruskal-Wallis test is a rank-based nonparametric statistic technique to assess for significant differences on a dependent variable, between three or more groups of an independent variable. This test is based on the comparison of medians or mean ranks, depending on shape distribution in each group, to test the null hypothesis saying there are no significant differences between the medians or means of different groups. It is based on the calculation of the statistical *H*, which is defined by the following equation (Montgomery & Runger, 2003):

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^{\,2}}{n_i} - 3(N+1) \qquad [6.15],$$

where *R* is the rank of group *i*, *k* is the number of groups, *N* is the data total number, n_i is the overall data number of group *i*. If the data is linked, that is, two or more groups with the same rank, a correction factor is applied whereby the H_c statistic is used instead of H (see equation [6.16],]) (Wayne, 2009).

$$H_c = \frac{H}{C} \qquad [6.16],$$

where, *C* is the correction factor given by the following equation (Wayne, 2009):

$$C = 1 - \frac{\sum_{i=1}^{M} (t_i^{\,3} - t_i)}{N^3 - N} \qquad [6.17],$$

where, t_i is the number of linked ranks in each group and *M* is the number of ranks.

The distribution of statistic H_c approximates a chi-square distribution with *k-1* degrees of freedom where n_i should be greater than five. If the calculated value of H_c is greater than the critical chi-square value, then it is possible to reject the null hypothesis and say that the sample comes from a different population. In this case, the *p-value* or asymptotic significance should be smaller than 0.05 in order to have enough evidence to say that there are significant differences between the medians or means ranks of different groups.

6.2.2.2 Analysis of variance (ANOVA)

The ANOVA is a statistical method for studying sampled-data variance relationships. The test can be used to quantify the degree to which two or more sample means differ in an experiment. The ANOVA assumes that distribution of each group is normally distributed and there is approximately equal variance on the scores for each group. In this case, the statistic *F-ratio* tests the null hypothesis saying mean values for all the samples is the same. The statistical significance of the *F-ratio* is much easier to judge by its *p-value*. If the *p-value* is less than 0.05, the null hypothesis of equal means is rejected at the significance level of 5%. This does not imply that each means is significantly different from each other. It simply suggests that all means are not equal. To determine which sample means are significantly different from what others, it is possible to perform a *multiple range test*. This statistical procedure indicates homogeneous pair groups, which means are significantly different. A graphical way to compare multiple group means is given by the analysis of means (ANOM). This is similar to a standard control chart, where each mean is plotted along a centre line and upper and lower limits decision. The centre line is located at the all data mean. The graphic tests the null hypothesis saying all group means are equal to the overall mean. Any mean that falls outside the decision limits indicates that the corresponding mean differs significantly from the overall one.

In order to assess by means of auralizations the influence of a virtual acoustic treatment and the impact of acoustic variables on cognitive processes, a "Classroom" situated in the San Buenaventura University was selected as the second case study. The first application presents the methodology of assessment and results of current acoustical conditions, followed by a description of the virtual classroom state taking into account the hypothetic acoustic treatment, and a discussion of subjective tests results of intelligibility and listening difficulty comparing existent and simulated conditions. The second auralization system implementation describes the methodology used to apply the psychological tests, the main outcomes of the pilot study experiment and the results and discussion of the final subjective test application evaluating the impact of acoustic variables on cognitive processes of attention, memory and executive function.

6.3 Application of the auralization system to assess the impact of acoustic variables on cognitive processes

This section describes the methodology, results and discussion of the application of the auralization system to assess the impact of acoustic variables on cognitive processes. First, the methodology of the experiment in order to assess the dependent cognitive variables is explained. Afterwards, the procedure and outcomes of the pilot study is presented. Finally, the implementation and results of the subjective tests are described, followed by a discussion of assessing the impact of acoustic variables on cognitive processes of attention, memory and executive function.

6.3.1 Experiment methodology

The approach of the experiment was empirical analytical and the explicative level is transversal. An experimental 2x2 factorial design with four independent groups was applied. The independent acoustic variables manipulated were the reverberation time and the background noise levels, which were studied in two categories: Long - short and high - low, respectively. The details of the acoustical conditions created are described in section 6.3.1.1. The dependent variables measured were the performance on cognitive tasks of attention, memory and executive function, all measured at one time. A description of the components

evaluating the dependent cognitive processes of study is presented in section 6.3.1.2. In order to guarantee internal validity in the experiment due to the influence of external variables, the following aspects were considered:

- The groups were balanced by assigning equal number of participants by sex and by using the same cognitive assessment tests.
- The participants in each group were assigned using a stratified random method, in order to equilibrate the individual differences between them.
- In order to control the influence of environmental variables on the organismic variables in the experiment results, the following was kept constant in the four groups: the assessment schedule, using the same recording studio, lighting of the place, the level of education of the participants, sex and age.
- The dependent variables of attention, memory and executive function were measured by using validated tests in the target population.
- The research was conducted in the recording studio A of the University of San Buenaventura, which has favourable acoustical conditions to carry out subjective tests (see Figure 6.14). This room allowed presenting the auralizations and the subjective tests, since it has a much lower background noise level and shorter reverberation time than the classroom investigated (see Chapter 4).
- The same equipment and set up were used to reproduce the auralizations for all groups. In addition, a short acoustic measurement was carried out with a sound level meter every day the tests were applied, in order to check the SNR.

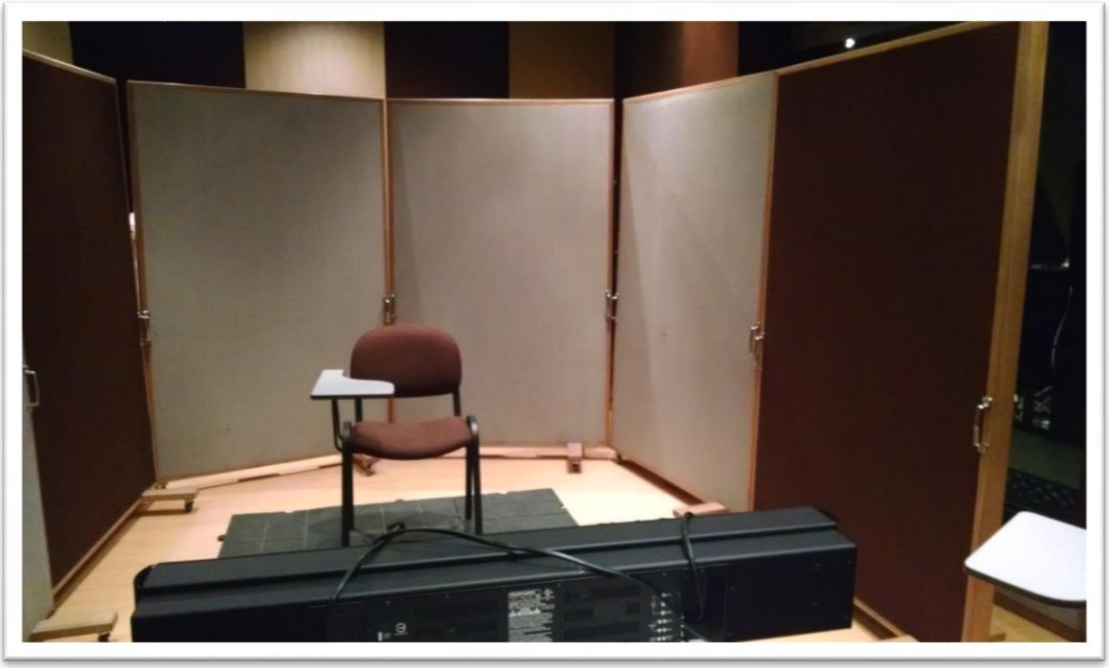


Figure 6.14: Picture of the place where the subjective tests were carried out.

6.3.1.1 Independent acoustic variables

The independent acoustic variables manipulated were the reverberation time and the background noise level. In both cases, with the purpose of establishing two categories denoting good and no proper acoustical conditions, the acoustic criteria described in section 6.1.1 was taken into account. For the first variable, the existing and simulated conditions of the classroom were taken as the long and short T_{mid} of study (see section 6.1). For the background noise level, the upper limit of Colombian regulation was taken as the high level reference or no proper condition, although higher limits for interior noise are found in other countries. The pilot study evidenced that this value was adequate given the difficulty of the questionnaires. The low level reference was set in 30 dB(A) as it is the lowest limit value according to acoustic criteria discussed in section 6.1.1, having a significant difference of 15 dB(A) between high and low level references which ensures the influence of this variable on the dependent ones. This way, the categories of T_{mid} and background noise level for the experiment were established as follows: Long (2.6 s) - short (0.8) and high (45 dBA) - low (30 dBA), respectively. These categories were taken into account to create the auralizations recreating four different conditions for the classroom.

In this experiment, the auralizations were created using only one source – receiver combination, with the latter placed on position 3 (see Figure 6.10), just

in front of the source. The simulation of the four acoustic conditions was done by the combination of the values of T_{mid} and background noise level mentioned above, in the following way:

- Condition A given by short T_{mid} (0.8 s) and low noise level (30 dBA).
- Condition B with long T_{mid} (2.6 s) and low noise level (30 dBA).
- Condition C having short T_{mid} (0.8 s) and high noise level (45 dBA).
- Condition D given by long T_{mid} (2.6 s) and high noise level (45 dBA).

6.3.1.2 Dependent cognitive variables

The cognitive variables of attention, memory and executive function are assessed by means of the psychological tests described in section 6.2.1. It is important to note that each test contains a number of components, which are evaluating one or two dependent variables. For instance, the TMT (see section 6.2.1.1) contains parts A and B, each one having five (5) different components according to the characteristics of the test and the marking process. In this case, all the components are assessing the cognitive processes of attention and executive function. The CPT (see section 6.2.1.2) consists of two parts, CPT auditory and CPT visual, each one with its corresponding number of components, which in this circumstance are evaluating only the variable of attention. The following test, called the WMS-III (see section 6.2.1.3), is the one including the maximum number of components with seventeen (17). In this case, all of them are assessing memory and ten (10) of them attention. The VMC test (see section 6.2.1.4) has seven (7) components that are assessing memory, with one of them also evaluating attention. Table 6.9 presents a summary of the psychological tests, their corresponding components and the dependent variables assessed by each one.

6.3.2 Pilot study

A pilot study was conducted in order to give strength to the proposed methodology, in order to analyse with greater internal validity, the impact of noise and reverberation time in classrooms on cognitive performance. The sample characteristics, procedure and results of the study are described below. Based on this report the methodological support for the completion of the final tests was developed.

6.3.2.1 The sample

For the pilot study, 24 university students, 6 for each group were evaluated. These were selected by probability sampling, considering the following inclusion criterion evaluated from a screening questionnaire (see Appendix I: Screening questionnaire):

- Adult university students.
- No previous diagnoses of mental disorders.
- No use of psychoactive substances.
- No use of medications that could affect cognitive performance.

6.3.2.2 Description of the procedure

To conduct the pilot study, 24 university students were assessed in order to identify the clarity of the instructions that were given, relevance of the tests used, logic and suitability of the procedure performed during the experiment and finally, to identify the behaviour, correlation and influence between variables. First, a protocol for the experiment was drafted, which was based on the guidelines established by the authors of the questionnaires chosen (see section 6.3.1). The protocol included the instructions to be given to the students being evaluated, in which the aim was to establish fixed indications, so that all participants would receive the same instructions and in this way to avoid biases/confusion. Once the protocol was established, 12 people were evaluated to identify the extent to which instructions were clear, verify effectiveness of the tests and identify if it was necessary to modify, remove or add other elements.

Table 6.9: Summary of the components included in each psychological test and the corresponding cognitive process assessed.

Psychological test	Component	Cognitive Process		
		Attention	Memory	Executive Function
TMT	TMT A Commissions	X		X
	TMT A Errors	X		X
	TMT A Omissions	X		X
	TMT A Time	X		X
	TMT A Total	X		X
	TMT B Commissions	X		X
	TMT B Errors	X		X
	TMT B Omissions	X		X
	TMT B Time	X		X
	TMT B Total	X		X
CPT	CPT(A) Successes	X		
	CPT(A) Commissions	X		
	CPT(A) Omissions	X		
	CPT(V) Time in Seconds	X		
	CPT(V) Successes	X		
	CPT(V) Omissions	X		
	CPT(V) Commissions	X		
WMS	WMS Progressive digits		X	
	WMS ABC errors	X	X	
	WMS ABC Points	X	X	
	WMS ABC Time	X	X	
	WMS Achievements Associated Pairs		X	
	WMS Associated Pairs Total		X	
	WMS Associated Pairs Trials		X	
	WMS Counting errors 1-40	X	X	
	WMS Counting Points 1-40	X	X	
	WMS Counting Time 1-40	X	X	
	WMS Points Stories		X	
	WMS Regressive digits		X	
	WMS Regressive errors 20-1	X	X	
	WMS Regressive Points 20-1	X	X	
	WMS Regressive Time 20-1	X	X	
	WMS Total digits		X	
	WMS Total Mind Control	X	X	
VMC	VMC Deferred 20 Min		X	
	VMC Deferred 3 Min		X	
	VMC Trials		X	
	VMC Organizational Index		X	
	VMC Curve Type		X	
	VMC Initial Amount	X	X	
	VMC Maximum Amount		X	

During the evaluation, it was found that although the tests were assessing what was expected, the cognitive variables that were supposed to be measured could be interfered by other factors, which should be clearly defined and/or controlled before starting the experiment. The intention was to separate the factors related to the exclusion criteria and avoid making wrong conclusions. To this extent the need of extending the initial questionnaire was identified. The purpose was to recognise, through it, the current status of persons regarding existence of psychopathology, presence of emotional issues, physical and/or environmental factors that interfere with the performance and prolonged exposure to noisy environments, among others.

Furthermore, it was found that besides the initial questionnaire, the order of presentation of the tests should be modified because the participants were anxious when doing the first questionnaire of Verbal Memory Curve (see section Appendix H: Wechsler Memory Scale – III Test). This could interfere with the performance and development of the following tests. Therefore, it was considered appropriate to define the questionnaire Visual and Auditory Continuous Performance (see section Appendix G: Continuous Performance Test (CPT)) as the first test to be presented, to enable the adaptation of the subject to the space. Then it was decided to present the verbal memory curve, leaving the other tests in the order originally proposed.

The pilot study also identified difficulties relating to the understanding of certain words by the participants, because there were level differences between them, causing intelligibility of almost zero for some keywords. It was found that the problem came from the original recording or reverberation free and not from auralizations as such, therefore, it was considered necessary to modify the "dry" audio files using the recording software ProTools®, by conducting automation that would enable a normalization of levels. The above was done so that when the respective auralizations were created, the loss of intelligibility would be proportional for all the words. Finally, to ensure that the modifications made would contribute to the evaluation of variables, 10 volunteers were invited and they were presented with only the modified words to confirm the intelligibility.

Furthermore, changes in the time history greater than 4 dB(A) were observed in the recordings of background noise, which affected and interfered with the performance of specific tests. This caused a bias that could not be controlled

during the course of the test, therefore, the change of levels and the noise sources that had tonal content not common in a classroom were eliminated, in order to generate a temporary continuous noise during the experiment. Finally, during the second part of the pilot study, the remaining 12 people were evaluated and it was noted that the above changes had positive effects in the experiment.

6.3.2.3 Pilot study preliminary results

In order to analyse the results of the cognitive tests, the nonparametric Kruskal-Wallis technique was implemented using the software "*Statistical Package for the Social Science (SPSS®)"*. The application of this statistic intended to determine if there were statistically significant differences between the groups of independent variables for any of the components assessing the dependent cognitive variables. Considering this, the asymptotic significance was checked for all the components, looking for values smaller than 0.05, in order to have evidence of the impact of acoustic variables on the cognitive processes studied. In this sense, for the variable of attention, the components that presented an asymptotic significance smaller than 0.05 were the *TMT B Commissions*, the *CPT(A) Successes* and the *CPT(A) Omissions* (see Figure 6.15). For the cognitive process of memory, the component that met that condition was the *VMC Curve Type* (see Figure 6.6.16). In the executive function case, the component with an asymptotic significance smaller than 0.05 was the *TMT B Commissions* (see Figure 6.6.17). It is important to note that no further analysis was necessary at this point given that to quantify the impact of the acoustic conditions on the dependent variables; a bigger sample was required. For this reason, a more exhaustive statistical analysis was implemented in the next section discussing the results of the final cognitive tests.

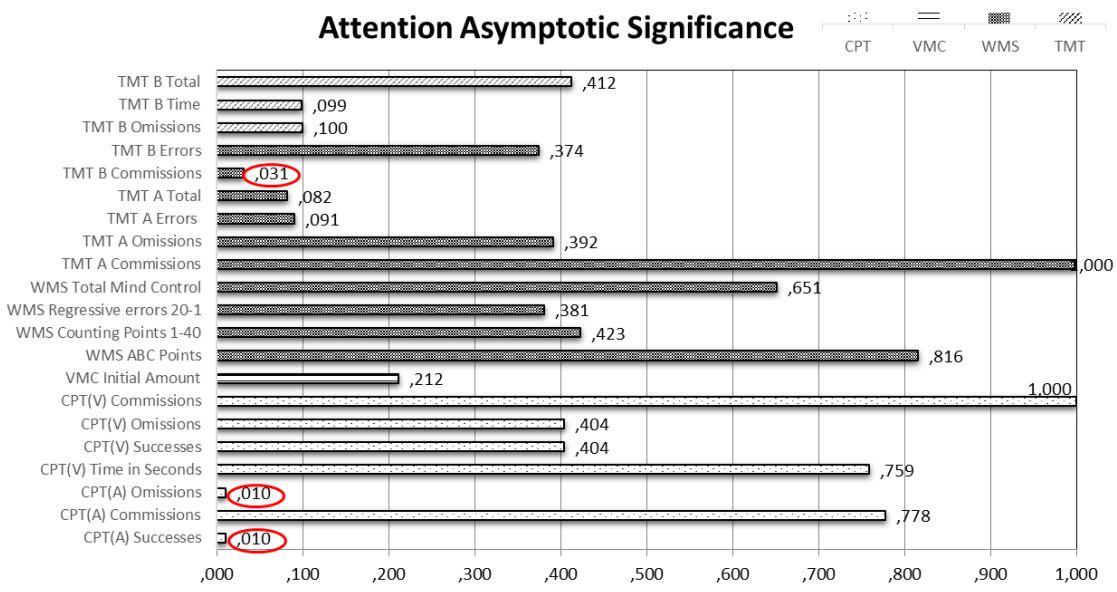


Figure 6.15: Asymptotic significances obtained for each component assessed in the cognitive variable of attention in the pilot study.

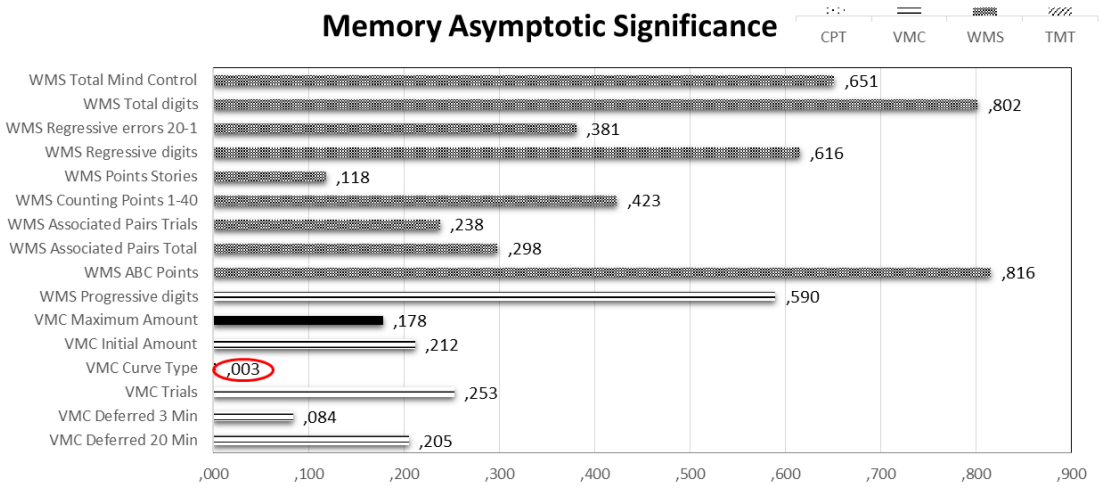


Figure 6.6.16: Asymptotic significances obtained for each component assessed in the cognitive variable of memory in the pilot study.

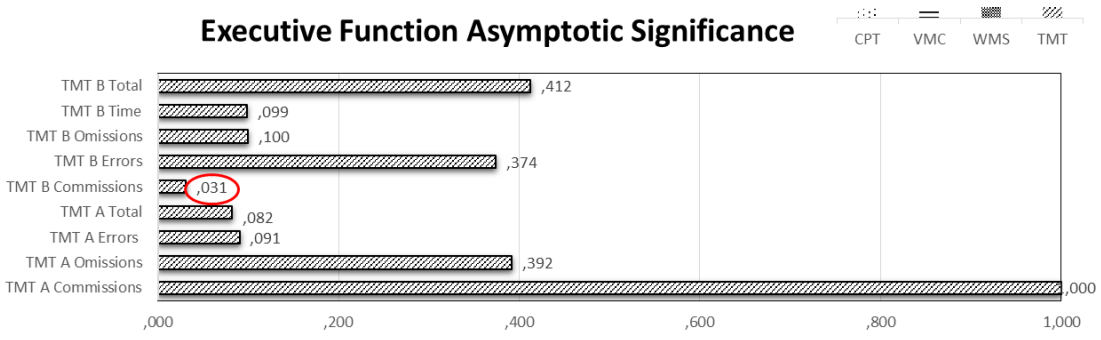


Figure 6.6.17: Asymptotic significances obtained for each component assessed in the cognitive variable of executive function in the pilot study.

6.3.3 Final cognitive tests results

The final experiment took account the aspects named in the pilot study as necessary to be modified, so that the test protocol was completed and a larger sample was evaluated to corroborate the results obtained in the pilot study. The methodology used in the experiment was similar to the one used in the pilot study, having similar criteria for the sample selection, psychological tests implemented and analysis of results. The main change was given by the size of the sample, which for this experiment included 60 university students, 15 for each group. Figure 6.18, Figure 6.19 and Figure 6.20 show the asymptotic significance results obtained for each component for the cognitive variables of attention, memory and executive function, respectively. The values highlighted with red circles denote the components with asymptotic significance values smaller than 0.05.

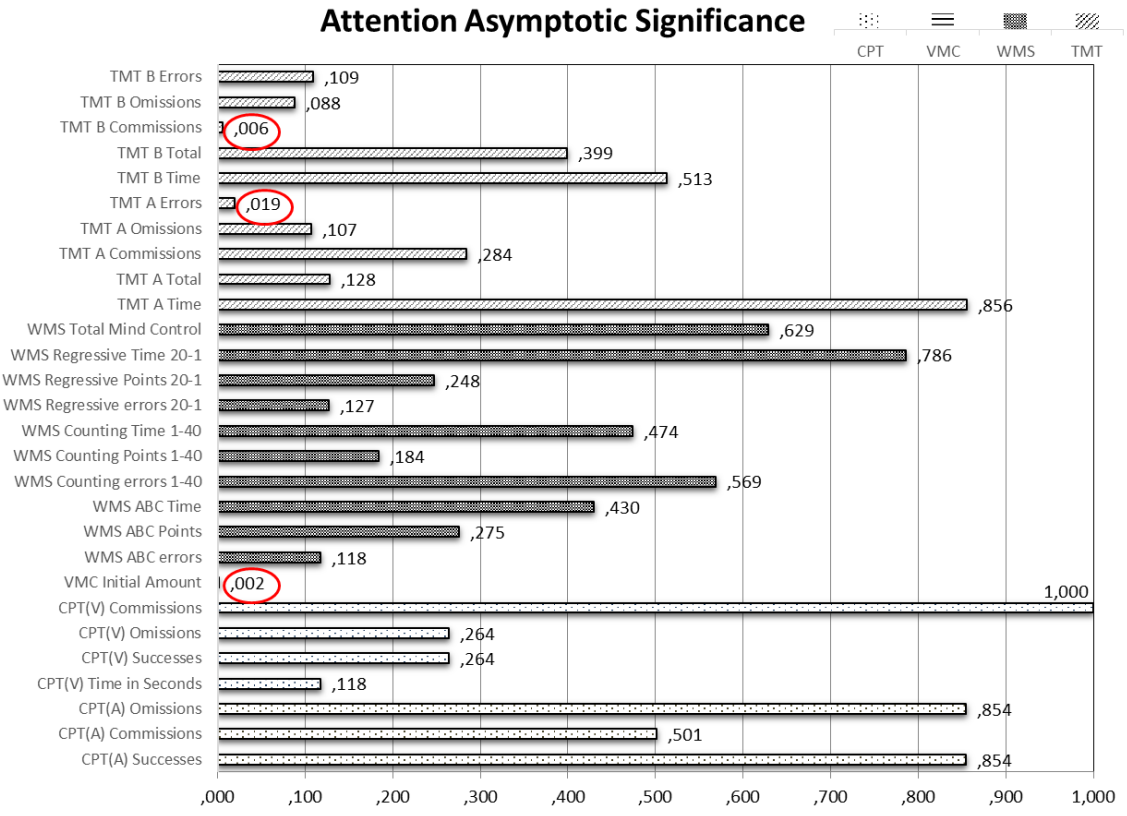


Figure 6.18: Asymptotic significances obtained for each component assessed in the cognitive variable of attention.

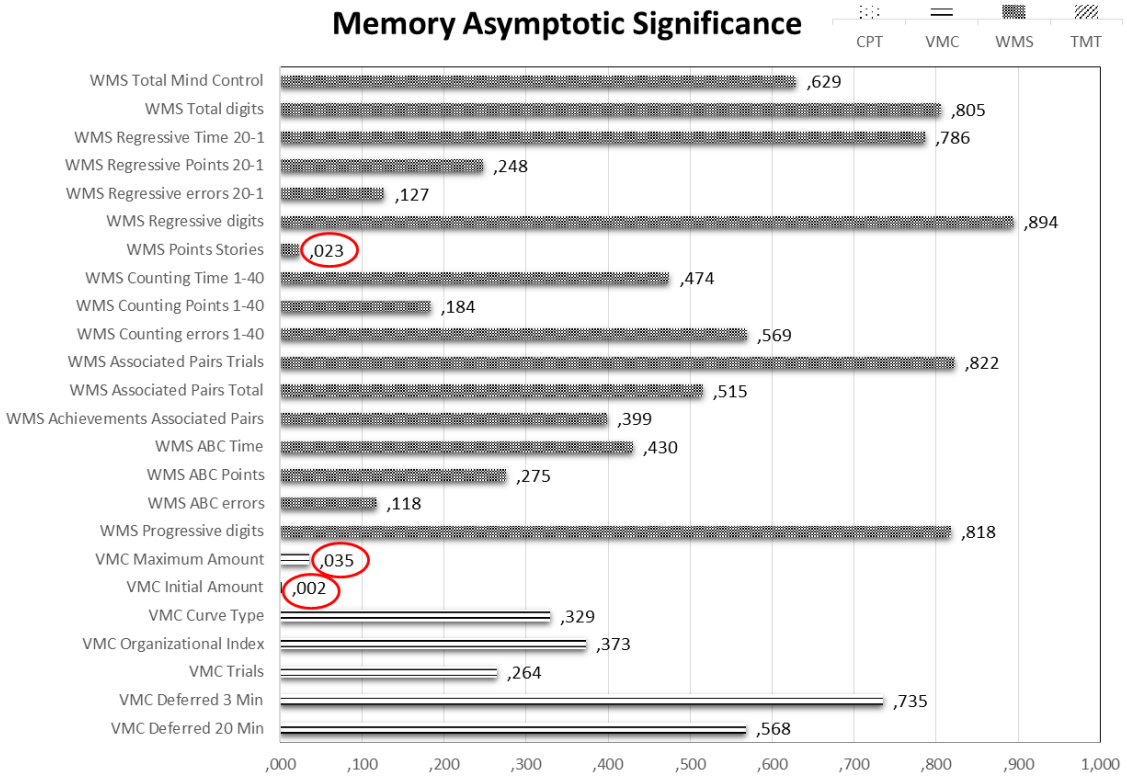


Figure 6.19: Asymptotic significances obtained for each component assessed in the cognitive variable of memory.

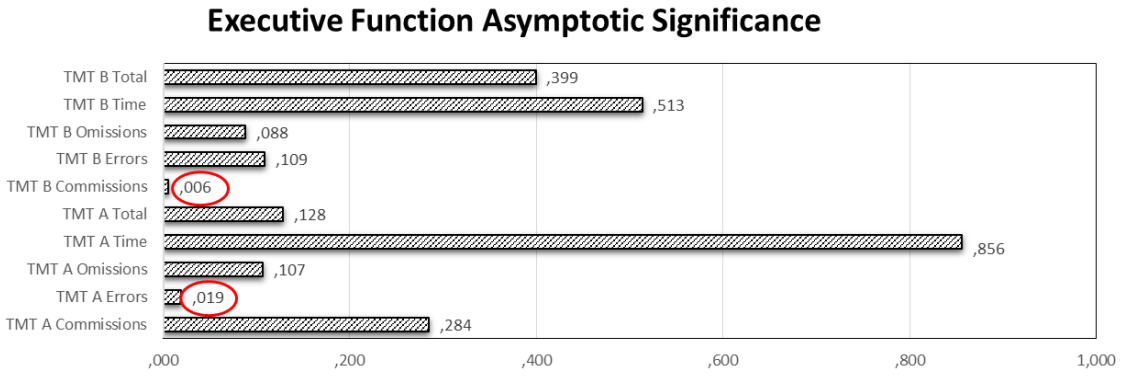


Figure 6.20: Asymptotic significances obtained for each component assessed in the cognitive variable of executive function.

6.3.3.1 Mean ranks results

This section presents the mean ranks obtained for the psychological components in which the asymptotic significance evidenced an impact of the independent variables on the cognitive processes of study (see Figure 6.18, Figure 6.19 and Figure 6.20), according to each acoustic condition of assessment. The following figures show the mean ranks results of the following components: VMC Initial Amount (see Figure 6.21), VMC Maximum Amount (see Figure 6.22), TMT A Errors (see Figure 6.23), TMT B Commissions (see Figure 6.24) and WMS Points Stories (see Figure 6.25).

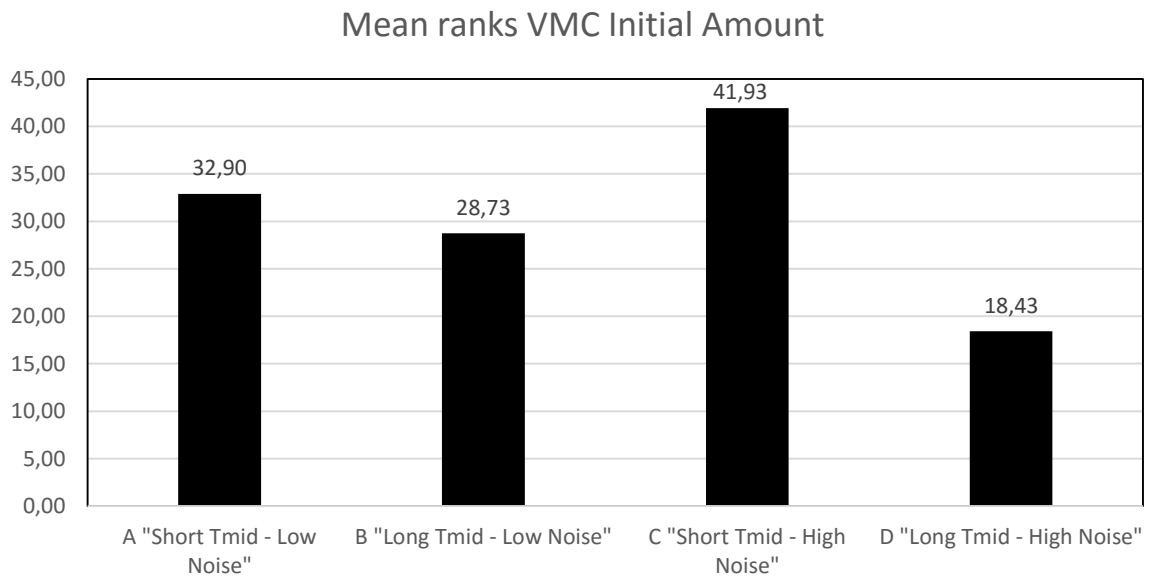


Figure 6.21: Mean ranks obtained in the VMC Initial Amount component for each acoustic condition.

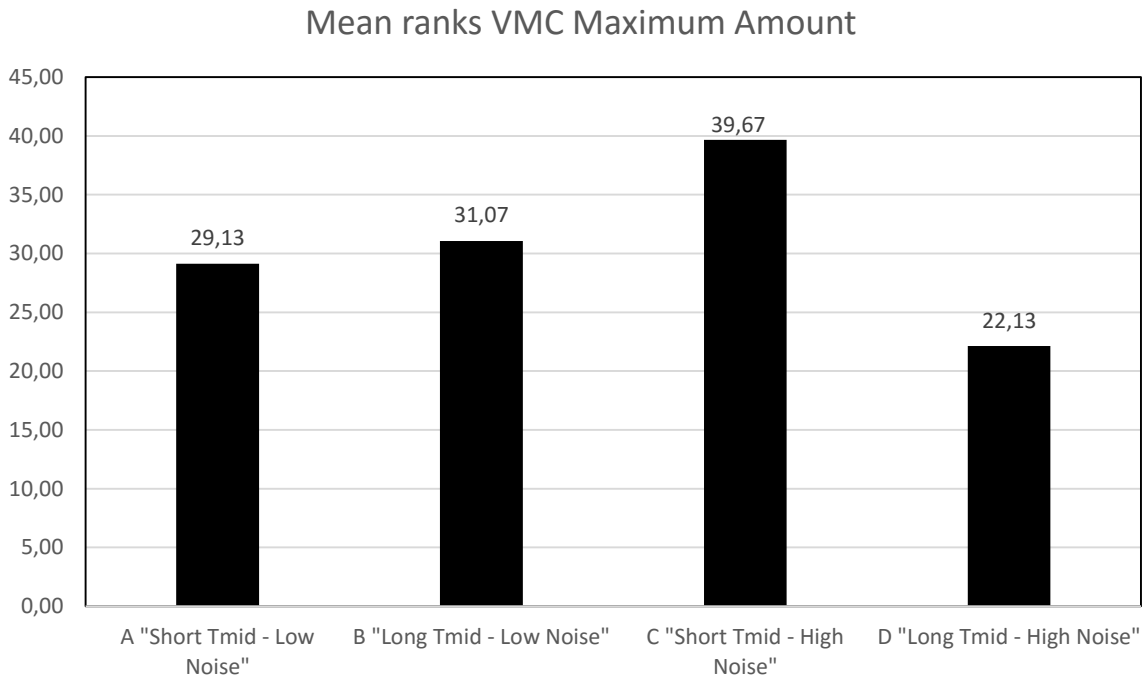


Figure 6.22: Mean ranks obtained in the VMC Maximum Amount component for each acoustic condition.

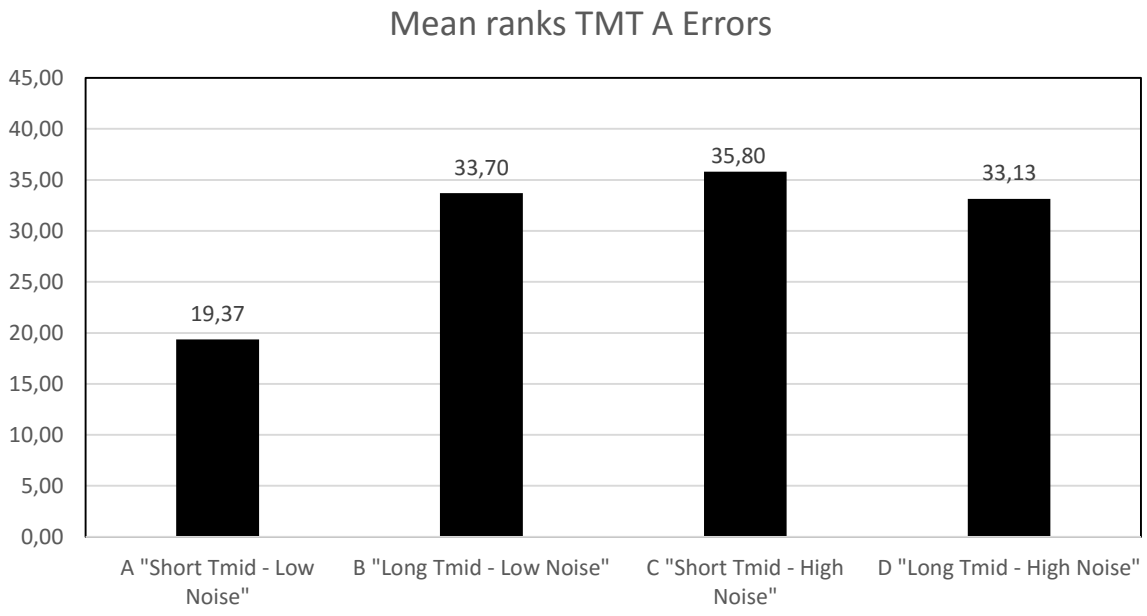


Figure 6.23: Mean ranks obtained in the TMT A Errors component for each acoustic condition.

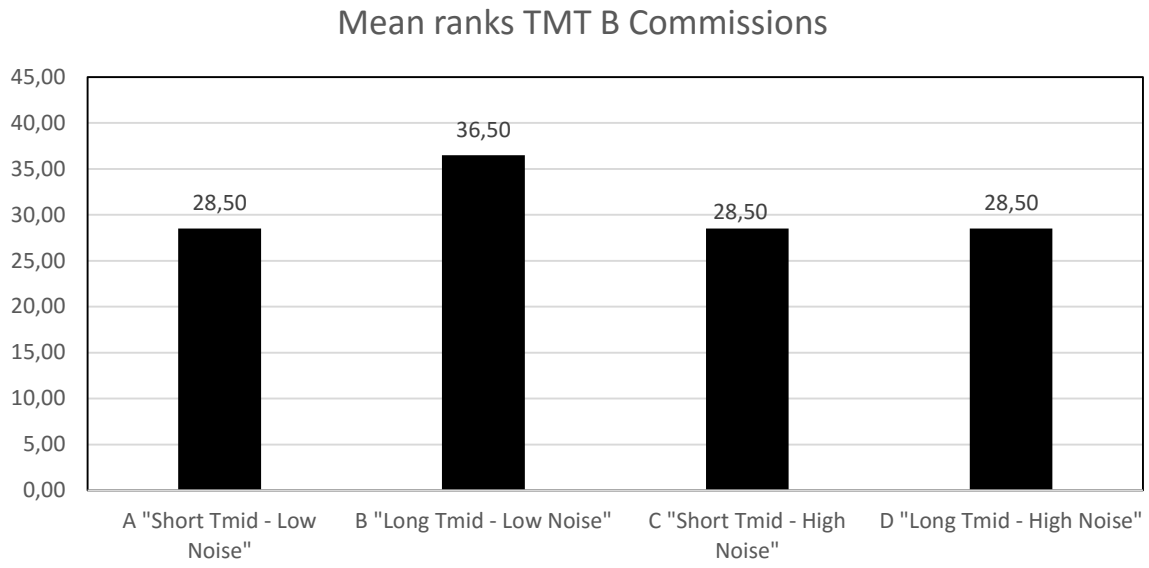


Figure 6.24: Mean ranks obtained in the TMT B Commissions component for each acoustic condition.

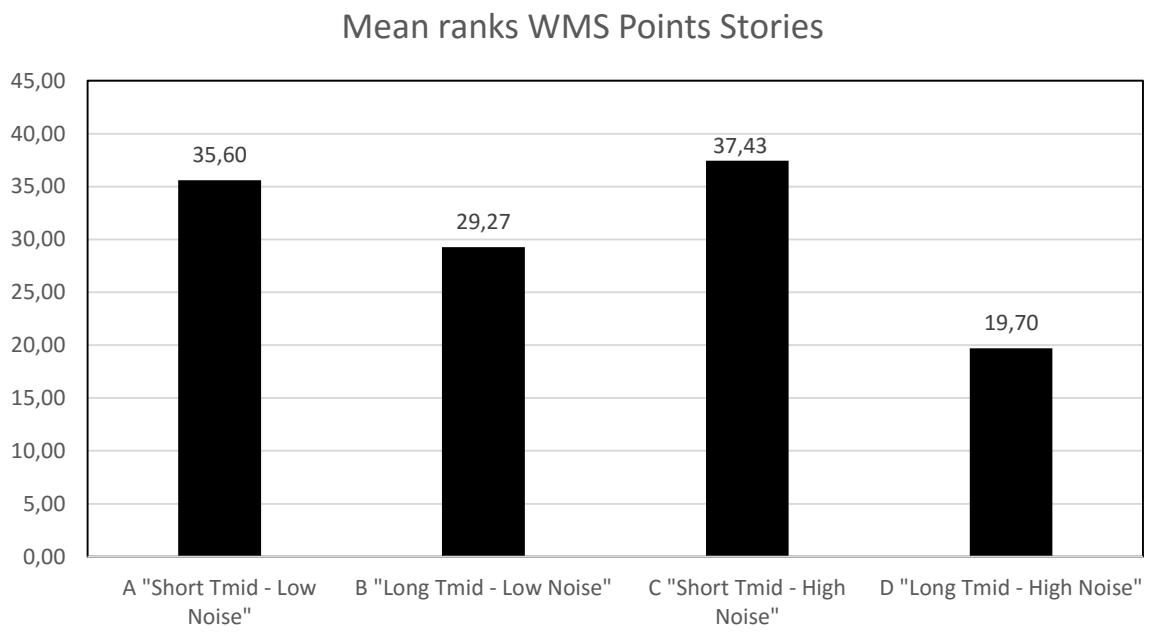


Figure 6.25: Mean ranks obtained in the WMS Points Stories component for each acoustic condition.

6.3.3.2 ANOVA of mean ranks by acoustic condition

This procedure performs an ANOVA for the mean ranks results obtained in the last section, in order to compare their mean values for the four acoustic conditions of study. The *F-test* in the ANOVA table is testing whether there are any significant differences amongst the means (see Table 6.10). The *F-ratio*, which in this case equals 4.65, is a ratio of the between-group estimate to the

within-group estimate. Since the *P-value* of the *F-test* is less than 0.05, there is a statistically significant difference between the mean ranks from one acoustic condition to another at the 95% confidence level. The Multiple Range Tests are indicating which means are significantly different from which others (see Table 6.11). The bottom half of the table shows the estimated difference between each pair of means. An asterisk has been placed next to three pairs, indicating that these pairs show statistically significant differences at the 95% confidence level. At the top of the table, three homogenous groups are identified using columns of X's. Within each column, the levels containing X's form a group of means within which there are no statistically significant differences. The method currently being used to discriminate among the means is Fisher's least significant difference (LSD) procedure. With this method, there is a 5% risk of calling each pair of means significantly different when the actual difference equals zero.

Table 6.10: ANOVA table for Mean Ranks by acoustic conditions.

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-value*
Between groups	396.368	3	132.123	4.65	0.016
Within groups	454.514	16	28.4071		
Total (Corr.)	850.882	19			

Table 6.11: Multiple Range Tests for Mean Ranks by acoustic conditions using the 95% Least Significant Difference (LSD) method.

Level	Count	Mean	Homogeneous Groups
D	5	24.38	X
A	5	29.1	XX
B	5	31.8533	XX
C	5	36.6667	X

Contrast	Sig.	Difference	+/- Limits
A - B		-2.75333	7.14597
A - C	*	-7.56667	7.14597
A - D		4.72	7.14597
B - C		-4.81333	7.14597
B - D	*	7.47333	7.14597
C - D	*	12.2867	7.14597

* denotes a statistically significant difference.

In order to see the practical significance of the ANOVA results, a scatter plot is shown in Figure 7.6.26. The Analysis Of Means (ANOM) plot tests the null hypothesis stating that all the acoustic conditions means are equal to the overall mean of all observations (see Figure 7.6.27). In this plot, each mean is connected to a centre line (CL) and upper (UDL) and lower (LDL) decision limits are defined. Any mean falling outside the decision bounds indicates that the corresponding acoustic condition differs significantly from the overall mean. It is easy to see from the ANOM plot that the acoustic condition C has a general mean rank higher than average, while the acoustic condition D has a mean considerably lower.

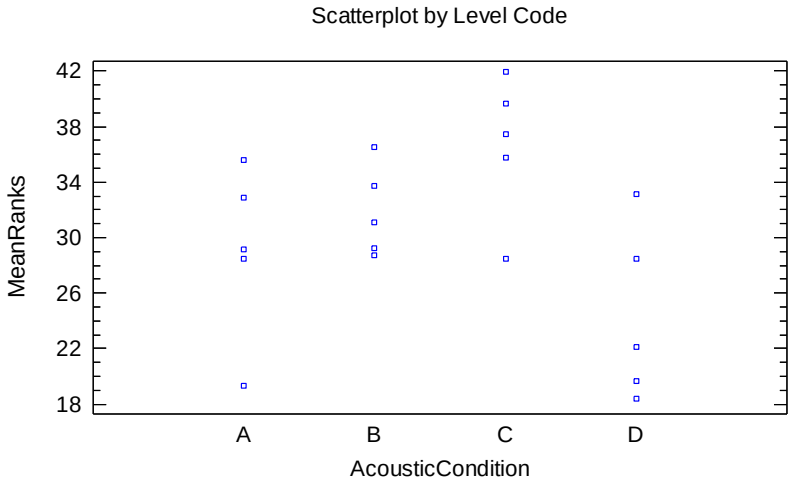


Figure 7.6.26: Scatter plot of mean ranks by acoustic condition.

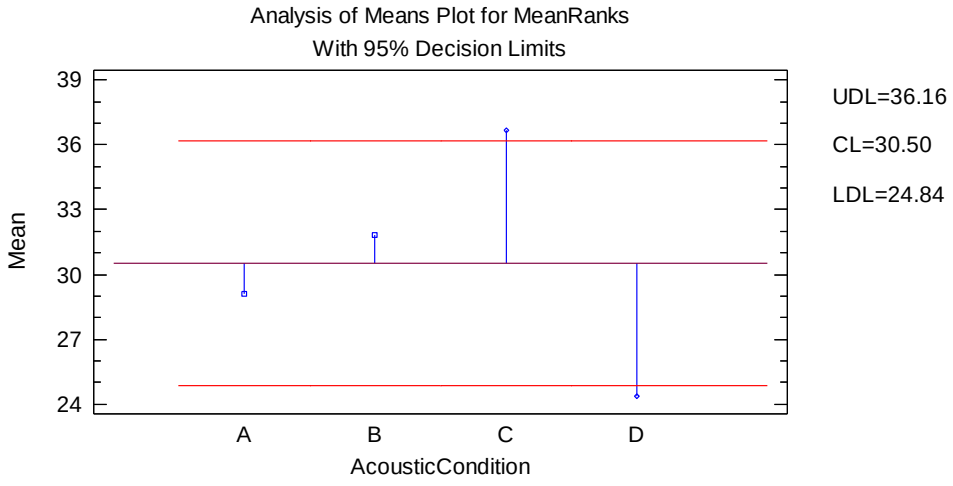


Figure 7.6.27: Analysis Of Means (ANOM) plot for mean ranks by acoustic condition, with 95% decision limits.

6.3.4 Discussion of cognitive tests results

According to the asymptotic significance results of section 6.3.3, the acoustic variables of reverberation time and background noise present an impact on the cognitive processes of attention, memory and executive function. For the dependent variable of memory, three components evidenced significance differences between the independent groups; these were the WMS Points Stories, the VMC Maximum Amount and the VMC Initial Amount (see Figure 6.19). The last one is a component that evaluates also attention, a variable that showed significant differences in two other components, the TMT A Errors and the TMT B Commission errors (see Figure 6.18). The last two components also evaluated the variable of executive function (see Figure 6.20).

The mean ranks of the component WMS Points Stories (see Figure 6.25) show that in terms of memory, the worst acoustic condition was D (Long T_{mid} - High Noise) and the best scores were obtained in the condition C (Short T_{mid} - High Noise). Comparing the mean ranks of conditions A (Short T_{mid} - Low Noise) and C, it is possible to say that there were not significant differences given by the change of background noise levels. On the other hand, conditions B (Long T_{mid} - Low Noise) and D presented the worst scores, which is indicating that reverberation time had a more negative influence on memory, which worsened when it was combined with a high noise level.

The mean ranks of the component VMC Maximum Amount in Figure 6.22, presented similar results to the last component with the worst and the best acoustic conditions given by D and C, respectively. This corroborates that the memory was not affected by high levels of background noise as far as the reverberation was short and the cognitive tasks had short duration. However, in this case condition A had comparable results to condition B, which infers once again that background noise does not have a significant influence on memory. It is important to bear in mind that the cognitive tasks evaluated were of short duration, so that the effects of mental fatigue have not been taken into account at this point.

The results of the component VMC Initial Amount (see Figure 6.21) assessed two dependent variables, memory and attention. This component evaluated the number of words a student managed to recall the first time he/she listened to a list of words. This case presented similar results in comparison to the first

component analysed, with the best scores in condition C and the worst in condition D, nonetheless, having a more significant difference between conditions A and C. The mean ranks of this component denoted that reverberation time not only had more impact on memory, but also attention. It is important to note the results obtained for the last three components in condition C, this situation could be indicating that the student increased the level of attention in the presence of high noise levels. The ANOVA of mean ranks by acoustic condition described in section 6.3.3.2 has corroborated the results of last three components, in which the comparison between acoustic condition groups stated that condition C obtained a general mean rank significantly higher than average, while condition D had a general mean rank considerably lower.

The next component was able to assess two cognitive processes, attention and executive function. The mean ranks of TMT A Errors (see Figure 6.23) presented a relative high number of errors for conditions B, C and D, in comparison with condition A. This could be interpreted as any unfavourable acoustic condition, whether reverberation time or background noise, has a significant impact on these cognitive variables. It is important to note that this test did not have a spoken message to be analysed, so that the intelligibility had no influence in this case. The importance of these results on the executive function lies on the negative impact of the acoustic variables on the capacity related to solving problem and decision-making.

The results of the component TMT B Commissions (see Figure 6.24) assessed the same cognitive variables of attention and executive function. In this case, the worst scores were obtained in condition B and similar results were found in the other three conditions. It is important to note that the scores found in the last two components are relatively high in comparison with results obtained without the influence of acoustic variables.

7. Conclusions

In this research, a hybrid numerical approach combining a FEM based wave equation method with the well-established GA was applied to estimate the sound propagation in all the audibly frequency range, with the purpose of assessing the advantages and drawbacks of the hybrid model to create auralizations of non-built spaces. The FEM was applied to estimate the sound propagation in the low frequency range and the GA methods to calculate the mid and high frequencies. The results of the numerical simulations obtained by the application of both methods (FEM and GA) were combined by using digital filters and a crossover frequency. In order to have a hybrid model for non-constructed rooms, the use of material parameter databases was considered with the purpose of defining the boundary conditions in the numerical modelling. In this sense, the GA absorption coefficient databases provides reasonable input data to construct the GA models. Nevertheless, FE boundary conditions data is limited and more difficult to access than GA information. According to Aretz (2009), to specify the impedance in FE simulations for extended reaction materials, an approach consisting in defining a real valued frequency dependent impedance, corresponding to the absorption coefficient can be considered. To assess the application of this approach, numerical models of two different rooms were implemented in order to evaluate the sound propagation estimation. In order to have a reference to compare the numerical results, sound field measurements were carried out in the rooms investigated.

The idea of analysing two rooms with different conditions pretended to establish the scope of application of the hybrid approach considering the definition of FE impedance boundary conditions, as a frequency dependent real valued related to GA absorption coefficient databases. The rooms investigated were a meeting a room and a classroom. The differences between the rooms were given by the size, the geometry, the materials and the presence of furniture within the enclosure. For instance, the meeting room is a space of 67 m³ that had a table and wood furniture inside at the moment of realization of the acoustic measurements. On the other hand, the classroom is an enclosure that doubles the size of the meeting a room and had no furniture inside at the moment of applying the acoustic measurements. In terms of material, both rooms mostly have the so-called sound hard boundaries or materials considered as extended

reaction materials. The only surface that can be considered as a local reaction material is the carpet on the floor of the meeting room. Nevertheless, the approach of defining in the FE model an acoustic impedance boundary condition related to the GA absorption coefficient was also considered in this case.

The FE simulation results of the meeting room evidenced an improvement in the sound wave propagation estimation in the low frequency range, as it can be appreciated Chapter 5. This situation was expected given the size of the room and the fact that GA technique provides a suitable sound propagation estimation if the wavelength is small compared to the dimensions of the room. Moreover, this room highlighted the drawbacks of GA methods in a number of ways. The first one was given by the shape of the room, in which the eigenmodes played an important role in the frequency response in the room, according to the positions analysed (Chapter 4). This aspect represented an advantage for the FEM, so it was able to simulate the phenomena of diffraction and interference. Another disadvantage of GA was related with the acoustic near field modelling. These numerical methods were based on far field assumptions that cannot be met in small rooms with furniture in it. This situation was accentuated by positioning the source on the table and the receivers around it, the last with the intention of estimating the sound wave propagation in the real positions where the persons normally are located. This caused an overestimation of the acoustic energy reaching the receivers; hence, an overestimate of acoustic indicators might be obtained in GA when source and/or receivers are located near to obstacles.

In the case of the classroom, the FE simulation results no evidenced an improvement in the sound wave propagation estimation in the low frequency range, as it can be appreciated in section 5.4. This was an unexpected situation. Although the size of the room doubles the size of the meeting room, a Schroeder frequency of approximately 280 Hz meant that an improvement was expected in the octave bands of 125 Hz and 250 Hz, which was not the case according to the results exposed in section 5.4.3.1. A number of reasons could explain this situation, nevertheless, the generation of a diffuse field at low frequencies could be one of them. The last could be given by the shape of the classroom, which has no parallel partitions, and also because there was not absorbent material or furniture inside at the moment of the acoustic measurements. Nevertheless, this only indicates that GA simulations had the potential of generating more accurate

results at low frequencies, not necessarily that FE simulations obtained les accurate results. Another reason could be the approach of defining the impedance boundary condition in the FE method as a real valued related to the absorption coefficient. In this sense, it is important to consider the imaginary part of the acoustic impedance, which is related with the mechanical properties of the boundaries. However, this consideration not allow the use of GA absorption coefficients databases and to obtain this information infers the application of acoustic measurements in the place, which is not useful if the idea is to have an approach to create auralizations of non-built enclosures.

In general terms, the use of a hybrid numerical approach improved the acoustic indicator estimates in the meeting room for the low frequency range and the GA method showed better results in the classroom. According to the indicator results comparison presented in section 5.4.3.1.4 and the errors estimates obtained by means of both numerical approaches (hybrid and GA) compared to the measured reference, the EDT, T_{20}, C_{80}, and D_{50} indicators estimates was more accurate for the hybrid approach in the meeting room and more precise for GA in the classroom. The estimation of the IACC presented results more similar with respect to binaural impulse responses measurements for the hybrid approach in both rooms, provided a cube simulating a head has been included in the FE model. From this it can be concluded that for the low frequency range, a basic model of a cube with average size of human head is enough to obtain the binaural cues needed for binaural simulation and reproduction.

The implementation of the hybrid numerical approach to create auralizations of the meeting room improved the subjective perception of localization, warmth and reverberance in comparison with GA methods. The discussion of subjective test results of sections 5.5.3 indicated that for the three sources of study (saxhorn, bass drum and male voice), all the subjective parameters evaluated obtained higher scores when the auralizations simulated by means of the hybrid approach were assessed. It was observed that localization was judged as the most similar with respect to measured reference auralizations, which indicates the importance of the transient responses for this psychoacoustic parameter. Moreover, the bass drum was the source with the best localization assessment that points out how the hybrid approach improved the transient response given by the low frequency range.

In terms of the application of auralizations to evaluate acoustical conditions of a classroom, this research evidenced a simple approach to incorporate binaural background noise to auralizations applying binaural technology for reproduction. In this sense, it was possible to include and control the sound level reproduction of binaural recorded background noise. It is important to bear in mind that this approach not allow to include synthesized binaural noise.

According to the discussion of intelligibility and listening difficulty tests results of section 6.1.8, the intelligibility was more affected by reverberation time than background noise level. Taking into consideration the rating scale of ISO 9921, the intelligibility was assessed as poor with existing acoustic conditions and fair, when the acoustic treatment was considered, no matter whether background noise was included or not in the auralizations. On the other hand, the listening difficulty assessment results suggested the influence of both acoustic variables, although, a more significant impact was given by the presence of background noise, since a high percentage average was obtained even with the hypothetical acoustic condition having a short reverberation time.

In the analysis of correlation coefficients between each pair of variables in the subjective assessment of intelligibility and listening difficulty (see section 6.1.8), it was demonstrated that an acceptable positive correlation between existing conditions and the virtual implementation of an acoustic treatment was evidenced. In the first case, the variability of the intelligibility in the classroom given by the change of the reverberation times could be explained by a linear regression, when background noise is not included in the auralizations. In the second case, the variability of the listening difficulty in the presence of background noise might be explained by a linear regression, having the reverberation time as an independent variable. These ideas indicate that there is a potential in order to study the variability of intelligibility and listening difficulty, given by the modification of acoustic variables, by means of statistical models based on subjective assessment results of virtual sound environments.

Another aspect regarding the application of auralizations to evaluate acoustical conditions of classroom is given by possibility of having as independent variables background noise levels and reverberation times to study the impact of these variables on cognitive processes of attention, memory and executive function. The discussion of the cognitive test results in section 6.3.4

corroborated the negative impact of the acoustic variables of reverberation time and background noise on the cognitive processes of memory, attention and executive function. For the former two, it was possible to verify that there were not significant differences given by the change of background noise levels when a short reverberation time was included. On the other hand, the reverberation time evidenced a more negative influence on these variables, which worsened when it was combined to a high noise level. Another aspect to note was the apparent increase of attention when the student was exposed to an acoustic condition given by a high noise level and short reverberation time. This situation was concurrent in three different components evaluating memory and attention, meaning the unnecessarily need of maintaining high attention levels when there are adverse acoustic conditions, which eventually would provoke fatigue at long periods. For the last cognitive process of executive function, it was found that any unfavourable acoustic condition, whether long reverberation time or high background noise level has a significant impact on this variable, which affects the capacity of solving problem and decision-making.

7.1 Future work

In terms of numerical simulation, this research evidenced the need of investigating different approaches to define impedance boundary conditions for non-constructed enclosures in FE simulations. In this regard is important to consider the use of information of GA databases as an initial input.

Although a novel procedure was implemented to include binaural background noise recordings in auralizations applying binaural technology, it is not clear how to include synthesized binaural background noise without affecting the binaural hearing experience of the auralizations reproduced by binaural technology. The application of this procedure could be applied to create auralizations including environmental noise.

In terms of auralization implementation to evaluate the impact of background noise and reverberation time on cognitive processes, it is relevant to study the influence of background noise including tonal components and transients.

Taking advantage of having the acoustic variables under control, another subject of study is the objective measurement of cognitive processes when are

influenced by the independent variables of background noise and reverberation time.

References

Acoustical Society of America; American National Standard., 2010. *Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools.*

Acoustical Society of America, 2002. *Acoustical Performance Criteria, Design Requirements, And Guidelines For Schools,* Melville.

Acoustical Society of America, 2010. *ANSI/ASA S12.60-2010/Part 1, Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, Part 1: Permanent Schools,* Melville, NY: American National Standards Institute.

Adrian James Acoustics Limited, 2004. *Workshop on room acoustics measurements - analysis of comparative measurements,* Technical Report 9506/2: Commissioned by National Physical Laboratory for the National Measurement System Acoustic Metrology Programme.

Ali, S. A., 2013. Study Effects of School Noise on Learning Achievement and Annoyance in Assiut City, Egypt. *Applied Acoustics,* 74(4), pp. 602-606.

Allen, J. B. & Berkley, D. A., 1979. Image Method for Efficiently Simulating Small-Room Acoustics. *Journal of Acoutical Society of America,* April, 65(4), pp. 943-950.

Ardila, , A. & Rosselli, , M., 1992. *Neuropsicologia Clinica [Clinical neuropsychology].* Medellin (Colombia): Prensa Creativa.

Ardila, A. & Ostrosky, F., 2012. *Guía para el Diagnóstico Neuropsicológico,* México D.F.: Universidad Nacional Autónoma de México.

Ardila, A. & Rosselli, M., 1992. *Neuropsicologia Clinica [Clinical neuropsychology].* 1st ed. Medellin (Colombia): Prensa Creativa.

Ardila, A., Rosselli, M. & Puente, A., 1994. *Neuropsychological Evaluation of the Spanish Speaker.* 1st ed. New York: Plenum Press.

Aretz, M., 2009. Specification of Realistic Boundary Conditions for the FE Simulation of Low Frequency Sound Fields in Recording Studios. *Acta Acustica united with Acustica,* August, 95(5), p. 874 – 882.

Aretz, M., 2012. *Combined Wave and Ray Based Room Acoustic Simulations of Small Rooms.* Aachen: RWTH Aachen University.

Aretz, M., Nöthen, R., Vorländer, M. & Schröder, D., 2009. *Combined Broadband Impulse Responses Using FEM and Hybrid Ray-Based Methods.* Espoo, Finland, Jun 15th-17th, in Proc. EAA Auralization Symp.

Aretz, M. & Vorländer, M., 2010. Efficient Modelling of Absorbing Boundaries in Room Acoustic FE Simulations. *Acta Acustica united with Acustica,* 96(6), p. 1042 – 1050.

Astley, R., 2010. *Discrete Models for Acoustics - Lecture Notes.* Southampton: University of Southampton.

Astolfi, A., Corrado, V. & Griginis, A., 2007. Comparison between measured and calculated parameters for the acoustical characterization of small classrooms. *Applied Acoustics,* Volume 69, p. 966–976.

Bai, M. R. & Chang, S., 1996. Active Noise Control of Enclosed Harmonic Fields by Using BEM-Based Optimization Techniques. *Applied Acourlics,* 48(1), pp. 1-32.

Bai, M. R. & Lee, C.-C., 2006. Objective and Subjective Analysis of Effects of Listening Angle on Crosstalk Cancellation in Spatial Sound Reproduction. *J. Acoust. Soc. Am.,* 120(4), pp. 1976-1989.

Barry, G., 2010. *Finite Difference Time Domain for 3D Sound Reproduction.* Southampton: University of Southampton.

Basner, M. et al., 2014. Auditory and Non-Auditory Effects of Noise on Health. *The Lancet,* 383(9925), pp. 1325-1332.

Beranek, L. L. & Ver, I. L., 1992. *Noise and Vibration Control Engineering Principles and Applications.* Illustrated ed. New York: John Wiley & Sons, Inc..

Berkhout, A., 1988. A holographic approach to acoustic control. *Journal of the Audio Engineering Society,* p. 977–995.

Berkhout, A., de Vries, D. & Vogel, P., 1993. Acoustic Control by Wave Field Synthesis. *J. Acoust. Soc. Am.,* 93(5), pp. 2764-2778.

Berkhout, A. J., 1988. A Holographic Approach to Acoustic Control. *Journal of the Audio Engineering Society,* 36(12), p. 977–995.

Bies, D. A. & Hansen, C. H., 1996. *Engineering Noise Control: Theory and Practice.* 2nd ed. London: Spon Press.

Bies, D. A. & Hansen, C. H., 2009. *Engineering Noise Control: Theory and Practice.* Fourth ed. London and New York: Spon Press.

Bistafa, S. & Bradley, J., 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), p. 861–875.

Borish, J., 1984. Extension of the Image Model to Arbitrary Polyhedra. *J. Acoust. Soc. Am.,* 75(6), pp. 1827-1836.

Botteldooren, D., 1995. Finite-Difference Time-Domain Simulation of Low-Frequency Room Acoustic Problems. *J. Acoust. Soc. Am.,* 98(6), pp. 3302-3308.

Bradley, J. S., 1986. Speech Intelligibility Studies In Classrooms. *Journal of the Audio Engineering Society,* 80(3), p. 846–854.

Bradley, J. S., 1998. Relationships among Measures of Speech Intelligibility in Rooms. *Journal of the Audio Engineering Society,* 46(5), pp. 396-405.

Brancati, A. & Aliabadi, M., 2012. Boundary Element Simulations for Local Active Noise Control Using an Extended Volume. *Engineering Analysis Boundary Elements,* 36(2), p. 190–202.

Breslow, R., Kocsis, J. & Belkin, B., 1980. Memory deficits in depression: Evidence utilizing the Wechsler Memory Scale. *Perceptual and Motor Skills,* Volume 51, pp. 541-542.

Buen, A., 2008. How Dry Do the Recordings for Auralization Need to Be?. *Proceedings of the Institute of Acoustics,* 30(3), pp. 107-114.

Campos, G. R. & Howard, D. M., 2005. On the Computational Efficiency of Different Waveguide Mesh Topologies for Room Acoustic Simulation. *IEEE Transactions on Speech and Audio Processing,* 13(5), pp. 1063-1072.

Cantor Cutiva, L. C., Vogel, I. & Burdorf, A., 2013. Voice Disorders in Teachers and Their Associations with Work-Related Factors: A Systematic Review. *Journal of Communication Disorders,* 46(2), p. 143–155.

Cardoso, M. et al., 2012. Vocal Effort and Voice Handicap Among Teachers. *Journal of Voice,* 26(6), pp. 820.e15-820.e18.

CATT, 2007. *CATT-Acoustic v8.0 User's Manual.* 1st ed. Gothenburg: CATT.

CENELEC European Committee for Electrotechnical Standardization, 2011. *Sound system equipment Part 16: Objective rating of speech intelligibility by speech transmission index,* s.l.: BSi.

Clark, C. & Stansfeld, S. A., 2007. The Effect of Transportation Noise on Health and Cognitive Development: A Review of Recent Evidence. *International Journal of Comparative Psychology,* 20(2-3), pp. 145-158.

Conners, C. K., 1995. *Conners' Continuous Performance Test User's Manual,* Toronto, Canada: Multi-Health Systems.

Corteel, E., 2007. Synthesis of Directional Sources Using Wave Field Synthesis, Possibilities and Limitations. *EURASIP Journal on Advances in Signal Processing,* 2007(1).

D'Antonio, P., 2001. Desktop Auralization. In: *Master Handbook of Acoustics.* Fourth Edition ed. Santa Barbara: McGraw-Hill, pp. 566-570.

Dalenbäck, B.-I., 1995. *A New Model for Room Acoustic Prediction and Auralization.* Gothenburg: Chalmers University of Technology.

Dalenbäck, B.-I., 1995. The Importance of Diffuse Reflection in Computerized Room Acoustic Prediction and Auralization. *Proceedings of the Institute of Acoustics,* Volume 17, p. 27-34.

Dalenbäck, B.-I., 1996. Room Acoustic Prediction Based On a Unified Treatment of Diffuse and Specular Reflection. *J. Acoust. Soc. Am.,* 100(2), pp. 899-909.

Department of Education and Skills UK, 2004. *Acoustic Design of Schools. Building Bulletin 93,* London: Architects and Building Branch.

Dockrell, J. E. & Shield, B. M., 2006. Acoustical Barriers in Classrooms: The Impact of Noise on Performance in the Classroom. *British Educational Research Journal,* 32(3), pp. 509-525.

Escolano, J., Jacobsen, F. & Lopez, J. J., 2008. An Efficient Realization of Frequency Dependent Boundary Conditions in an Acoustic Finite-Difference Time-Domain Model. *Journal of Sound and Vibration,* 316(1-5), p. 234–247.

Escolano, J. & Lopez, J. J., 2007. Directive Sources in Acoustic Discrete-Time Domain Simulations Based on Directivity Diagrams. *J. Acoust. Soc. Am.,* June.121(6).

Escolano, J., Pueo, B. & Bleda, S., 2004. *Finite-Difference Time-Domain Acoustic Analysis of Fibrous Sound-Absorbing Materials.* Berlin, Alemania, Presented at the 116th AES Convention.

EUROPEAN COMMITTEE FOR STANDARDIZATION, 2010. *BS EN ISO 3744:2010.* Brussels: s.n.

Fahnline, J., 2009. Numerical Difficulties with Boundary Element Solutions of Interior Acoustic Problems. *Journal of Sound and Vibration,* 319(3-5), p. 1083–1096.

Fahy, F., 2001. *Foundations of Engineering Acoustics.* London: Elsevier.

Farina, A. & Ugolotti, E., 1999. *Subjective Comparison between Stereo Dipole and 3d Ambisonic Surround Systems for Automotive Applications.* Rovaniemi, AES 16th International conference.

Fastl, H. & Zwicker, E., 2007. *Psycho-Acoustics - Facts and Models.* 3rd ed. New York: Springer.

Frank, M., Zotter, F. & Sontacchi, A., 2015. *Producing 3D Audio in Ambisonics.* Hollywood, CA, Presented at the 57th AES international conference.

Gauthier, P.-A. & Berry, A., 2008. Adaptive Wave Field Synthesis for Active Sound Field Reproduction: Experimental Results. *J. Acoust. Soc. Am.,* 123(4).

Gerzon, M., 1976. *Multidirectional Sound Reproduction Systems.* UK, Patent No. 3 997 725.

Giner, J., Militello, C. & Garcia, A., 1999. Ascertaining Confidence within the Ray-Tracing Method. *J. Acoust. Soc. Am.,* 106(2), pp. 816-822.

Giner, J., Militello, C. & Garcia, A., 2001. The Monte Carlo Method to Determine the Error in Calculation of Objective Acoustic Parameters within the Ray-Tracing Technique. *J. Acoust. Soc. Am.,* 110(6), pp. 3081-3085.

Granier, E., Kleiner, M., Dalenbäck, B. I. & Svensson, P., 1995. *Experimental Auralization of Car Audio Installations.* Paris, France, Feb 25-28, in Proc. of the 130th AES Convention.

Haapakangas, A. et al., 2014. Effects of Unattended Speech on Performance and Subjective Distraction: The Role of Acoustic Design in Open-Plan Offices. *Applied Acoustics,* Volume 86, pp. 1-16.

Hacıhabiboglu, H. & Murtagh, F., 2008. Perceptual Simplification For Model-Based Binaural Room Auralisation. *Applied Acoustics,* 69(8), p. 715–727.

Hamilton, B. & Webb, C. J., 2013. *Room Acoustics Modelling Using GPU-Accelerated Finite Difference and Finite Volume Methods on a Face-Centered Cubic Grid.* Maynooth, Ireland, in Proc. of the 16th International Conference on Digital Audio Effects (DAFx-13).

Hammond, J. & White, P., 2008. Signals and Systems. In: D. Havelock, S. Kuwano & M. Vorländer, eds. *Handbook of signal processing in Acoustics.* New York: Springer, p. 13.

Harari, I. & Hughes, T. J. R., 1992. A Cost Comparison of Boundary Element and Finite Element Methods for Problems of Time-Harmonic Acoustics. *Computer Methods in Applied Mechanics and Engineering,* May, 97(1), pp. 77-102.

Hartmann, W. M. & Wittenberg, A., 1996. On the Externalization of Sound Images. *J. Acoust. Soc. Am.,* 99(6), pp. 3678-3688.

Heckl, M., 1992. Numerical Methods. In: *Modern Methods in Analytical Acoustics Lecture Notes.* London: Springer-Verlag, pp. 283-284.

Heller, A. J., Lee, R. & Benjamin, E. M., 2008. *Is my Decorder Ambisonic?.* San Francisco, CA, Presented at the 125th AES Convention.

Hodgson, M., 1991. Evidence of Diffuse Surface Reflections in Rooms. *J. Acoust. Soc. Am.,* 89(2), pp. 765-771.

Hodgson, M., 2002. Rating, Ranking, and Understanding Acoustical Quality in University Classrooms. *The Journal of the Acoustical Society of America,* 112(9), pp. 568-575.

Hygge , S. & Knez, I., 2001. Effects of Noise, Heat and Indoor Lighting on Cognitive Performance and Self-Reported Affect. *Journal of Environmental Psychology,* 21(3), pp. 291-299.

International Electrotechnical Commission, 2003. *Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index,* Geneva: BSI.

International Electrotechnical Commission, 2011. *Electroacoustics - Simulators of Human Head and Ear - Part 7: Head and Torso Simulator for the Measurement of Hearing Aids,* Geneva: BSI.

International Electrotechnical Commission, 2011. *Sound System Equipment – Part 16: Objective Rating of Speech Intelligibility by Speech Transmission Index,* Geneva: BSI.

International Organization for Standardization, 2003. *BS EN ISO 9921:2003, Ergonomics - Assessment of speech,* Geneva: BSI.

International Organization for Standardization, 2008. *EN ISO 3382-2:2008, Acoustics - Measurement of Room Acoustic Parameters - Reverberation Time in Ordinary Rooms,* Geneva: BSI.

International Organization for Standardization, 2009. *EN ISO 3382-1:2009 Acoustics - Measurement of Room Acoustic Parameters - Part 1: Performance Spaces,* Geneva: BSI.

International Organization for Standardization, 2010. *BS EN ISO 3744:2010 , Acoustics - Determination of Sound Power Levels and Sound Energy Levels of Noise Sources Using Sound Pressure,* Geneva: BSI.

ITU Radiocommunication Assembly, 1990. *Recommendation ITU-R BS.562-3. Subjective Assessment of Sound Quality,* Geneva: Radiocommunication.

Jeong, H. & Lam, Y. W., 2010. *FDTD Modelling of Frequency Dependent Boundary Conditions for Room Acoustics.* Sydney, 20th International Congress on Acoustics.

Jeong, H. & Lam, Y. W., 2012. Source Implementation to Eliminate Low-Frequency Artifacts in Finite Difference Time Domain Room Acoustic Simulation. *J. Acoust. Soc. Am.,* 131(1), pp. 258-268.

Jorgensen, M., Ickler, C. B. & Jacob, K. D., 1991. Judging the Speech Intelligibility of Large Rooms via Computarized Audible Simulations. *Journal of the Audio Engineering Society,* pp. 3126 (C-8).

Kennedy, S. M. et al., 2006. Subjective Assessment of Listening Environments in University Classrooms: Perceptions of Students. *The Journal of the Acoustical Society of America,* 119(1), p. 299–309.

Kim, S.-M. & Wonjae, C., 2005. On the Externalization of Virtual Sound Images in Headphone Reproduction: A Wiener Filter Approach. *J. Acoust. Soc. Am.,* 117(6), pp. 3657-3665.

Kinsler, L. E., Frey, A. R., Coppens, A. B. & Sanders, J. V., 2000. *Fundamentals of Acoustics.* 4th ed. New York: John Wiley & Sons.

Kirkeby, O. & Nelson, P. A., 1998. Local Sound Field Reproduction Using Two Closely Spaced Loudspeakers. *J. Acoust. Soc. Am.,* 104(4), pp. 1973-1981.

Kirkeby, O., Nelson, P. A., Rubak, P. & Farna, A., 199. *Design of Cross-Talk Cancellation Networks by using Fast Deconvolution.* Munich, Germany, Presented at the 106th AES Convention.

Kjellberg, A., 2004. Effects of Reverberation Time on the Cognitive Load in Speech Communication: Theoretical Considerations. *Noise & Health,* 7(25), pp. 11-21.

Klatte, M., Hellbrück, J., Seidel, J. & Leistner, P., 2010. Effects of Classroom Acoustics on Performance and Well-Being in Elementary School Children: A Field Study. *Journal of Environment and Behavior,* 42(5), pp. 659-692.

Klatte, M., Lachmann, T. & Meis, M., 2010. Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise & Health,* Volume 12, pp. 270-282.

Kleiner, M., 1981. Speech Intelligibility in Real and Simulated Sound Fields. *Acta Acustica united with Acustica,* 47(2), pp. 55-71.

Kleiner, M., Dalenback, B.-I. & Svensson, P., 1993. Auralization An Overview. *J. Audio Eng. Soc.,* Volume 41, pp. 861-874.

Kopuz, S. & Lalor, N., 1995. Analysis of Interior Acoustic Fields Using the Finite Element Method and the Boundary Element Method. *Applied Acoustics,* 45(3), pp. 193-210.

Krebber, W., Gierlich, H.-W. & Genuit, K., 2000. Auditory Virtual Environments: Basics and Applications. *Signal Processing,* 80(11), pp. 2307-2322.

Kristiansen, J. et al., 2013. Effects of Classroom Acoustics and Self-Reported Noise Exposure on Teachers' Well-Being. *Environment & Behavior,* 45(2), pp. 283-300.

Krokstad, A. & Strøm, S., 1979. Acoustical Design of hte Multi-Purpose "Hjertnes" Hall in Sandefjord. *Applied Acoustics,* 12(1), pp. 45-63.

Krokstad, A., Strom, S. & Sørsdal, S., 1968. Calculating the Acoustical Room Response by the Use of a Ray Tracing Technique. *Journal of Sound and Vibration,* 8(1), pp. 118-125.

Krokstad, A., Strom, S. & Sørsdal, S., 1983. Fifteen Years' Experience with Computerized Ray Tracing. *Applied Acoustics,* 16(4), p. 291 312.

Kulowski, A., 1982. Error Investigation for the Ray Tracing Technique. *Applied Acoustics,* 15(4), p. 263 274.

Kulowski, A., 1985. Algorithmic Representation of Ray Tracing Technique. *Applied Acoustics,* 18(6), pp. 449-469.

Kumar, S., 2009. *Acoustic Design of Classrooms.* Munich, AES 126th Convention, pp. 41-42.

Kuttruff, H., 2000. *Room Acoustics.* Fourth Edition ed. London and New York: Spon Press.

Kuttruff, H., 2007. *Acoustics: An Introduction.* London and New York: Taylor & Francis.

Kuttruff, H., 2009. *Room Acoustics.* 5th ed. London and New York: Spon Press.

Kwan Ryu, J. & Yong Jeon, J., 2008. Subjective and Objective Evaluations of a Scattered Sound Field in a Scale Model Opera House. *J. Acoust. Soc. Am.,* 124(3), pp. 1538-1549.

Kyriakakis, C., 1998. Fundamental and Technological Limitations of Immersive Audio Systems. *Proceedings of the IEEE,* 86(5), pp. 941-951.

Lacouture Parodi, Y., 2010. *A Systematic Study of Binaural Reproduction Systems Through Loudspeakers: A Multiple Stereo-Dipole Approach.* Aalborg: Section of Acoustics, Aalborg University.

Laine, S., Siltanen, S., Lokki, T. & Savioja, L., 2009. Accelerated Beam Tracing Algorithm. *Applied Acoustics,* 70(1), pp. 172-181.

Lee, H. & Lee, B.-H., 1988. An Efficient Algorithm for the Image Model Technique. *Applied Acoustics,* 24(2), pp. 87-115.

Lehmann, E. A. & Johansson, A. M., 2008. Prediction of Energy Decay in Room Impulse Responses Simulated With an Image-Source Model. *J. Acoust. Soc. Am.,* 124(1), pp. 269-277.

Lehmann, E. A., Johansson, A. M. & Nordholm, S., 2007. Reverberation-Time Prediction Method for Room Impulse Responses Simulated with the Image-Source Model. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,* pp. 159-162.

Lehnert, H. & Blauert, J., 1992. Principles of Binaural Room Simulation. *Applied Acoustics,* 36(3-4), pp. 259-291.

Lentz, T., 2007. *Binaural Technology for Virtual Reality.* Berlin: Aachen University.

Lentz, T. & Behler, G. K., 2004. Dynamic Crosstalk Cancellation for Binaural Synthesis in Virtual Reality Environments. *J. Audio Eng. Soc,* 54(4), pp. 283-194.

Lentz, T., Schroder, D., Vorlander, M. & Assenmacher, I., 2007. Virtual Reality Systemwith Integrated Sound Field Simulation and Reproduction. *EURASIP Journal on Advances in Signal Processing,* Issue 70540, p. 19 S.

Lewers, T., 1993. Combined Beam Tracing and Radiant Exchange Computer Model of Room Acoustics. *Applied Acoustics,* 38(2-4), pp. 161-178.

Lezak, M. D., 1995. *Neuropsychological Assessment.* 3rd ed. New York: Oxford University Press.

Liitola, T., 2006. *Headphone Sound Externalization.* Tampere: Helsinki University of Technology.

Lindau, A. et al., 2014. A Spatial Audio Quality Inventory (SAQI). *ACTA Acustica united with Acustica,* 100(5), pp. 984-994.

Ljung , R. & Kjellberg, A., 2009. Long Reverberation Time Decreases Recall of Spoken Information. *Journal of Building Acoustics,* 16(4), pp. 301-312.

Ljung, R., Sörqvist, P., Kjellberg, A. & Green, A., 2009. Poor Listening Conditions Impair Memory for Intelligible Lectures: Implications for Acoustic Classroom Standards. *Building Acoustics,* 16(3), pp. 257-265.

Long, M., 2014. *Architectural Acoustics.* 2nd edition ed. Burlington(MA): Elsevier Academic Press.

López, J. J., Carnicero, D., Ferrando, N. & Escolano, J., 2013. Parallelization of the finite-difference time-domain method for room acoustics modelling based on CUDA. *Mathematical and Computer Modelling,* 57(7-8), pp. 1822-1831.

Mahesh, B., Feistel, S. & Ahnert, W., 2005. *First Approach to Combine Particle Model Algorithms with Modal Analysis using FEM.* Barcelona, Spain, May 28-31, in Proc. of the AES 118th Convention.

Mallardo, V., Aliabadi, M., Brancati, A. & Marant, V., 2012. An Accelerated BEM for Simulation of Noise Control in the Aircraft Cabin. *Aerospace Science and Technology,* 23(1), pp. 418-428.

Marburg, S. & Schneider, S., 2003. Performance of Iterative Solvers for Acoustic Problems. Part I. Solvers and Effect of Diagonal Preconditioning. *Engineering Analysis with Boundary Elements,* 27(7), p. 727–750.

Mechel, F., 2000. The Corner Source Model. *Acta Acustica united with Acustica,* 86(6), pp. 957 - 969.

Mechel, F., 2002. Improved Mirror Source Method in Room Acoustics. *Journal of Sound and Vibration,* 256(5), pp. 873-940.

Mechel, F. P., 2009. *Formulas of Acoustics.* Second ed. Berlin: Springer.

Middlebrooks, J. C. & Green, D. M., 1990. Directional Dependence of Interaural Envelope Delays. *Journal of hte Acoustical Society of America,* 87(5), pp. 2149-2162.

Møller, H., 1992. Fundamentals of Binaural Technology. *Applied Acoustics,* 36(3-4), pp. 171-218.

Montgomery, D. & Runger, D., 2003. *Applied Statistics and Probability for Engineers.* 3rd ed. New York: John Wiley & Sons.

Murphy, D., Kelloniemi, A., Mullen, J. & Shelley, S., 2007. Acoustic Modeling Using the Digital Waveguide Mesh. *IEEE signal processing magazine,* March.pp. 55-66.

Murphy, D. T. & Howard, D., 1999. *Digital Waveguide Modelling of Room Acoustics- Comparing Mesh Topologies.* Milan, in Proc. of the 25th EUROMICRO conference, pp. 82-89.

Murphy, D. T., Southern, A. & Savioja, L., 2014. Source Excitation Strategies for Obtaining Impulse Responses in Finite Difference Time Domain Room Acoustics Simulation. *Applied Acoustics,* Aug., Volume 82, pp. 6-14.

Mydlarz, C. et al., 2013. Comparison of Environmental and Acoustic Factors in Occupied School Classrooms for 11-16 Year Old Students. *Building and Environment,* Volume 60, pp. 265-271.

Nachbar, C., Nistelberger, G. & Zotter, F., 2010. *Listening to the Direct Sound of Musical Instruments in Freely Adjustable Surrounding Directions.* Paris, France, in Proc. of the 2nd International Symposium on Ambisonics Symposium and Spherical Acoustics.

Nagy, A. et al., 2006. Prediction of Interior Noise in Buildings Generated by Underground Rail Traffic. *Journal of Sound and Vibration,* 293(3-5), p. 680–690.

Naylor, G., 1993. ODEON Another Hybrid Room Acoustical Model. *Applied Acoustics,* 38(2-4), pp. 131-143.

Nelson, P. A., 1998. An Introduction to acoustics. In: *Fundametals of Noise and Vibration.* London and New York: Spon Press, p. 50.

Nelson, P. A., 1998. An Introduction to acoustics. In: *Fundametals of Noise and Vibration.* London and New York: Spon Press, pp. 35-50.

Nelson, P., Kirkeby, O. & Takeuchi, T., 1997. Sound Fields for the Production of Virtual Acoustic Images. *Journal of Sound and Vibration,* 204(2), pp. 386-396.

Nelson, P. & Rose, J., 2005. Errors in Two-Point Sound Reproduction. *J. Acoust. Soc. Am.,* 118(1), pp. 193-204.

Otondo, F. & Rindel, J. H., 2004. The Influence of the Directivity of Musical Instruments in a Room. *Acta Acustica united with Acustica,* 90(6), p. 1178 – 1184.

Partington, J. E. & Leiter, R. G., 1949. Partington's Pathways Test. *Psychological Service Center Journal,* Volume 1, pp. 11-20.

Pelzer, S., Masiero, B. & Vorländer, M., 2014. *3d Reproduction of Room Auralizations by Combining Intensity Panning, Crosstalk Cancellation and Ambisonics.* Berlin, Germany, in Proc. of the EAA Joint Symposium on Auralization and Ambisonics.

Peng, J., 2005. Feasibility of Subjective Speech Intelligibility Assessment Based on Auralization. *Applied Acoustics,* 66(5), pp. 591-601.

Periáñeza, J. A. et al., 2007. Trail Making Test in Traumatic Brain Injury, Schizophrenia, and Normal Ageing: Sample Comparisons and Normative Data. *Archives of Clinical Neuropsychology,* 22(4), p. 433–447.

Petyt, M. & Jones, C., 2004. Numerical Methods in Acoustics. In: F. Fahy & J. Walker, eds. *Advanced applications in acoustics, noise and vibration.* London: Taylor & Francis, pp. 54, 55, 62.

Polack, J.-D. & Figueiredo, L., 2012. *Room Acoustic Auralization with Ambisonics.* Nantes, France, in Proc. of the Acoustics 2012.

Pollow, M., Behler, G. & Masiero, B., 2009. *Measuring Directivities of Natural Sound Sources with a Spherical Microphone Array.* Graz, Ambisonics Symposium.

Proakis, J. G. & Manolakis, D. G., 2007. *Digital Signal Processing.* Fourth ed. New Jersey: Prentice Hall.

Pulkki, V. & Merimaa, J., 2005. *Spatial Impulse Response Rendering - A Tool for Reproducing Room Acoustics for Multi-Channel Listening.* Finland: University of Helsinki.

Pulkki, V. & Merimaa, J., 2005. *Spatial Impulse Response Rendering: Listening Tests and Applications to Continuous Sound.* Barcelona, Spain, Presented at the 118th AES Convention.

Randall, R. B., 2008. Spectral Analysis and Correlation. In: D. Havelock, S. Kuwano & M. Vorländer, eds. *Handbook of signal processing in acoustics.* New York: Springer, p. 35.

Reitan, R. M. & Wolfson, D., 1985. *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation.* 2nd ed. Tucson, Ariz: Neuropsychology Press.

Rosas, C. & Sommerhoff, J., 2008. Acoustic Intelligibility in Spanish: a Proposal for its Measurement. *Estudios Filológicos,* Volume 43, pp. 179-190.

Rozenn, N. et al., 2014. *A roadmap for assessing the quality of experience of 3d audio binaural rendering.* Berlin, Proc. of the EAA Joint Symposium on Auralization and Ambisonics.

Sato, H., Bradley, J. & Morimoto, M., 2005. Using Listening Difficulty Ratings of Conditions for Speech Communication in Rooms. *The Journal of the Acoustical Society of America,* 117(3), p. 1157–1167.

Sato, H., Morimoto, M. & Ota, R., 2007. Acceptable Range Of Speech Level for both Young and Aged Listeners in Reverberant and Quiet Sound Fields. *The Journal of the Acoustical Society of America,* 122(3), p. 1616–1623.

Sato, H., Morimoto, M. & Ota, R., 2011. Acceptable Range of Speech Level in Noisy Sound Fields for Young Adults and Elderly Persons. *The Journal of the Acoustical Society of America,* 130(3), p. 1411–1419.

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(6), p. 4596–4605.

Savioja, L. & Svensson, U. P., 2015. Overview of Geometrical Room Acoustic Modeling Techniques. *J. Acoust. Soc. Am,* Aug., 138(2), pp. 708-730.

Savioja, L. & Välimäki, V., 2003. Interpolated Rectangular 3D Digital Waveguide Mesh Algorithms with Frequency Warping. *IEEE transactions on speech and audio processing,* November, 11(6), pp. 783-790.

Schneider, J. B., Wagner, C. L. & Broschat, S. L., 1998. Implementation of Transparent Sources Embedded in Acoustic Finite-Difference Time-Domain Grids. *J. Acoust. Soc. Am,* January, 103(1), pp. 136-142.

Sheaffer, J. D., Fazenda, B. M., Angus, J. A. S. & Murphy, D. T., 2011. *A Simple Multiband Approach for Solving Frequency Dependent Problems in Numerical Time Domain Methods.* Aalborg, Denmark, in Proc. of EAA Forum Acusticum.

Sheaffer, J. & Fazenda, B., 2010. *FDTD/K-DWM Simulation Of 3d Room Acoustics On General Purpose Graphics Hardware Using Compute Unified Device Architecture (CUDA).* s.l., s.n., p. Vol. 32. Pt. 5..

Sheaffer, J. & Fazenda, B., 2010. *FDTD/K-DWM Simulation Of 3d Room Acoustics On General Purpose Graphics Hardware Using Compute Unified Device Architecture (CUDA).* Salford, Proceedings of the Institute of Acoustics, p. Vol. 32. Pt. 5..

Sheaffer, J., van Walstijn, M. & Fazenda, B., 2014. Physical and Numerical Constraints in Source Modeling for Finite Difference Simulation of Room Acoustics. *J. Acoust. Soc. Am,* January, 135(1), pp. 251-261.

Shield, B. M. & Dockrell, J. E., 2003. The Effects of Noise on Children at School: A Review. *Building Acoustics,* 10(2), pp. 97-106.

Siltanen, S., Lokki, T., Kiminki, S. & Savioja, L., 2007. The Room Acoustic Rendering Equation. *J. Acoust. Soc. Am.,* September, 122(3), pp. 1624-1635.

Siltanen, S., Lokki, T., Tervo, S. & Savioja, L., 2012. Modeling Incoherent Reflections from Rough Room Surfaces with Image Sources. *J. Acoust. Soc. Am.,* 131(6), pp. 4606-4614.

Smith, H. M., 2004. *Geometric Acoustic Modeling of the LDS Conference Center.* Provo(UT): Brigham Young University.

Sommerhoff, J. & Rosas, C., 2007. The evaluation of the intelligibility of speech in Spanish. *Estudios filológicos,* Volume 42, pp. 215-225.

Sommerhoff, J. & Rosas, C., 2008. Inteligibilidad acústica en español: una propuesta para su medición. *Estudios filológicos,* Volume 43, p. 179–190.

Sommerhoff, J. & Rosas, C., 2011. Study of the Correlation between STI and Subjective Intelligibility Test. *Estudios filológicos,* Volume 47, pp. 133-147.

Southern, A. & Murphy, D., 2009. *Low Complexity Directional Sound Sources for Finite Difference Time Domain Room Acoustic Models.* May 7-10 Munich, Germany, Presented at the 126th AES Convention.

Southern, A., Murphy, D. T. & Savioja, L., 2012. Spatial Encoding of Finite Difference Time Domain Acoustic Models for Auralization. *IEEE Transactions on audio, speech, and language processing,* November, 20(9), pp. 2420-2432.

Southern, A., Siltanen, S., Murphy, D. T. & Savioja, L., 2013. Room Impulse Response Synthesis and Validation Using a Hybrid Acoustic Model. *IEEE transactions on audio, speech, and language processing,* 21(9), pp. 1940-1952.

Spa, C., Garriga, A. & Escolano, J., 2010. Impedance Boundary Conditions for Pseudo-Spectral Time-Domain Methods in Room Acoustics. *Applied Acoustics,* 71(5), p. 402–410.

Spors, S., Rudolf, R. & Jens, A., 2008. *The Theory of Wave Field Synthesis Revisited.* Amsterdam, The Netherlands, Presented at the 124th AES Convention.

Spreen, O. & Strauss, E., 1998. *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* 2nd ed. New York: Oxford University Press.

Stanef, D., Hansen, C. & Morgans, R., 2004. Active Control Analysis of Mining Vehicle Cabin Noise Using Finite Element Modelling. *Journal of Sound and Vibration,* 277(1-2), p. 277–297.

Steeneken, H. J. M. & Houtgast, H., 1980. A Physical Method for Measuring Speech-Transmission Quality. *The Journal of the Acoustical Society of America,* 67(1), p. 318–326.

Stefanakis, N., Jacobsen, F. & Sarris, J., 2010. Effort Variation Regularization in Sound Field Reproduction. *J. Acoust. Soc. Am.,* 128(2), pp. 740-750.

Stephenson, U. M., 1996. Quantized Pyramidal Beam Tracing - a New Algorithm for Room Acoustics and Noise Immission Prognosis. *Acta Acustica united with Acustica,* 82(3), pp. 517 - 525.

Strøm, S., Dahl, H., Krokstad, A. & Eknes, E., 1985. Acoustical Design of the Grieg Memorial Hall in Bergen. *Applied Acoustics,* 18(2), pp. 127-142.

Strøm, S., Krokstad, A., Sørsdal, S. & Stensby, S., 1986. Design of Room Acoustics and a MCR Reverberation System for Bjergsted Concert Hall in Stavanger. *Applied Acoustics,* 19(6), pp. 465-475.

Sutherland, L. C. & Lubman, D., 2001. *The Impact of Classroom Acoustics on Scholastic Achievement.* Rome, ICA, p. 71.

Suzuki, H., Omoto, A. & Fujiwara, K., 2007. Treatment of Boundary Conditions by Finite Difference Time Domain Method. *Acoust. Sci. & Tech.,* 28(1), pp. 16-26.

Takeuchi, T., 2010. *3D Sound Reproduction with "OPSODIS" and its Commercial Applications.* Southampton: OPSODIS Limited.

Takeuchi, T. & Nelson, P., 2002. Optimal Source Distribution for Binaural Synthesis over Loudspeakers. *J. Acoust. Soc. Am.,* 112(6), pp. 2786-2797.

Takeuchi, T. & Nelson, P., 2008. Extension of the Optimal Source Distribution for Binaural Sound Reproduction. *Acta Acustica united with Acustica,* 94(6), p. 981 – 987.

Takeuchi, T. & Nelson, P. A., 2001. Robustness to head misalignment of virtual sound imaging systems. *J. Acoust. Soc. Am.,* pp. 958-971.

Takeuchi, T. & Nelson, P. A., 2002. Optimal source distribution for binaural synthesis over loudspeakers. *J. Acoust. Soc. Am.,* pp. 2786-2797.

Terai, T. & Kawai, Y., 1991. BEM Applications in Architectural Acoustics. In: *Boundary Elements Methods in Acoustics.* London and New York: Elsevier, pp. 193-195.

Thompson, D. & Nelson, P., 2015. Fundamentals of Acoustics. In: F. Fahy & D. Thompson, eds. *Fundamentals of Sound and Vibration.* Second ed. London and New York: Taylor & Francis Group, pp. 38, 69.

Toole, F., 1970. In-Head Localization of Acoustic Images. *The Journal of the Acoustic Society of America,* 48(4), pp. 943-949.

Trimmel , M., Atzlsdorfer, J., Tupy, N. & Trimmel, K., 2012. Effects of Low Intensity Noise from Aircraft or from Neighborhood on Cognitive Learning and Electrophysiological Stress Responses. *International Journal of Hygiene and Environmental Health,* 215(6), p. 547– 554.

Trombetta Zannin, P. H. & Reich Marcon, C., 2007. Objective and Subjective Evaluation of the Acoustic Comfort in Classrooms. *Applied Ergonomics,* 38(5), pp. 675-680.

Trombetta Zannin, P. H., Zwirtes, Z. & Petri, D., 2009. Evaluation of the Acoustic Performance of Classrooms in Public Schools. *Applied Acoustics,* 70(4), pp. 626-635.

Vigeant, M. C., Wang, L. M. & Rindel, J. H., 2011. Objective and Subjective Evaluations of the Multi-Channel Auralization Technique as Applied to Solo Instruments. *Applied Acoustics,* 72(6), pp. 311-323.

Vigran, T. E., 2008. *Building Acoustics.* London: Taylor & Francis.

Vorländer, M., 1989. Simulation of the Transient and Steady-State Sound Propagation in Rooms Using a New Combined Ray-Tracing/Image-Source Algorithm. *J. Acoust. Soc. Am.,* 86(1), pp. 172-178.

Vorländer, M., 2008. *Auralization.* Berlin: Springer.

Vorländer, M., 2013. Computer Simulations in Room Acoustics: Concepts and Uncertainties. *J. Acoust. Soc. Am,* March, 133(3), pp. 1203-1213.

Vorländer, M. & Mommertz, E., 2000. Definition and Measurement of Random-Incidence Scattering Coeficients. *Applied Acoustics,* 60(2), pp. 187-199.

Wayman, J. L. & Vanyo, J. P., 1977. Three-Dimensional Computer Simulation of Reverberation in an Enclosure. *J. Acoust. Soc. Am.,* Volume 62, pp. 213-215.

Wayne, D., 2009. *Biostatistics: A foundation for Analysis in the Health Sciences.* 9th ed. New York: John Wiley & Sons.

Wightman, F. L. & Kistler, D. J., 1992. The Dominant Role of Low-Frequency Interaural Time Differences in Sound Localization. *Journal of the Acoustical Society of America,* 91(3), pp. 1648-1661.

Wu, H., Liu, Y. & Jiang, W., 2012. A Fast Multipole Boundary Element Method for 3D Multi-Domain Acoustic Scattering Problems Based On the Burton–Miller Formulation. *Engineering Analysis with Boundary Elements,* 36(5), p. 779–788.

Xiangyang, Z., Ke'an, C. & Jincai, S., 2003. On The Accuracy of the Ray-Tracing Algorithms Based On Various Sound Receiver Models. *Applied Acoustics,* 64(4), p. 433–441.

Yang, W. & Hodgson, M., 2006. Auralization Study of Optimum Reverberation Times for Speech Intelligibility for Normal and Hearing-Impaired Listeners in Classrooms with Diffuse Sound Fields. *The Journal of the Acoustical Society of America,* 120(2), p. 801–807.

Yang, W. & Hodgson, M., 2006. Validation of the Auralization Technique: Comparative Speech Intelligibility Tests in Real and Virtual Classrooms. *Acta Acustica united with Acustica,* 93(6), p. 991–999.

Yee, K. S., 1966. Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media. *IEEE Transactions on Antennas and Propagation,* 14(5), pp. 302-307.

Yesenia , L. P. & Per , R., 2011. Sweet Spot Size in Virtual Sound Reproduction: A Temporal Analysis. In: Y. Suzuki, et al. eds. *Principles and Applications of Spatial Hearing.* Singapore: World Scientific Publishing Co Pte Ltd, pp. 292-302.

Appendix B: "Subjective test to assess virtual sound environments"

Date:

Identification – Tick or write an answer

- Gender: Female [ ] male [ ]

Age: 16-20 21-25 26-30 31-35 36-40

Introduction

This study is researching different aspects involved in the creation of virtual sound environments, also called: auralizations. Auralization is the process of audibly rendering the sound field created by a source in a simulated space, in order to reproduce the binaural listening experience at a given position. The aim of this project is to develop an auralization system able to provide realistic sound environments, taking into account three essential stages: Generation, transmission and reproduction. The importance of this system lies on its capacity of predicting the consequences of modifying parameters such as room shape, material selection, or source placement on the acoustic response at receiver position and hence, on acoustical variables at specific points inside of a non-constructed room.

This test aims to rate the accuracy of a numerical approach used in this research in order to create auralizations. The evaluation is conducted by comparing the simulated auralizations with reference auralizations obtained by means of acoustic measurements. In this assessment, three different sound sources (male voice, bass drum and saxhorn) are used in the auralizations. The audio samples have been recorded in a recording studio of the San Buenaventura University, after that, these signals were digitally processed with the acoustic responses of the room of study. A number of source-receiver combinations are presented, with the intention of evaluating the influence of distance and position within the room. In order to assess the subjective quality of the auralizations, a comparison between numerical approaches is based on the assessment of four specific parameters, which are explained in the following section.

Note: All information will be kept in the strictest confidence and individuals will not be identifiable in any output from this work.

Please read carefully the following instructions and if you have any problem with the questionnaire, please do not hesitate to ask the person running the experiment.

How does the test work?

In the test, auralizations are presented to you by pairs with the purpose of realising an AB comparison of four parameters. The reproduction is made through a 3D system called OPSODIS, the sound bar in front of you. A tablet is provided in order to play the audio samples, each one does not last more than 30 seconds and you are free to play the samples as many times as you consider necessary. For each pair of samples, you have to assess the similarity of sample B with respect to sample A, using the following scale:

RATING	ASSESSMENT
Not different	5,0
Slightly not different	4,0
Slightly different	3,0
Rather different	2,0
Completely different	1,0

The meaning of the parameters to compare are as follows:

- Localization: attribute associated to a subjective perception of the direction indicating the origin of sound and the relative position of the source.
- Sense of space: Similar to reverberation, this parameter refers to a subjective permanence of reflected sound within the enclosure. In other words, it indicates a subjective size impression of the room in acoustic terms.
- Warmth: attribute denoting a subjective perception of loudness at low frequencies of the corresponding source.
- Brightness: parameter indicating a subjective perception of loudness at high frequencies of the corresponding source.

Previous listening

In order to become familiar with the devices, the test environment, the attributes of study and the assessment scale, please listen to the tracks marked as "previous listening" (top of the interface) according to the following table:

Number	Prev. Listening
1	Saxhorn "dry"
2	Male voice "dry"
3	Bass drum "dry"
4	Bass drum pos1
5	Bass drum pos2
6	Bass drum pos3
7	Localization L
8	Localization R

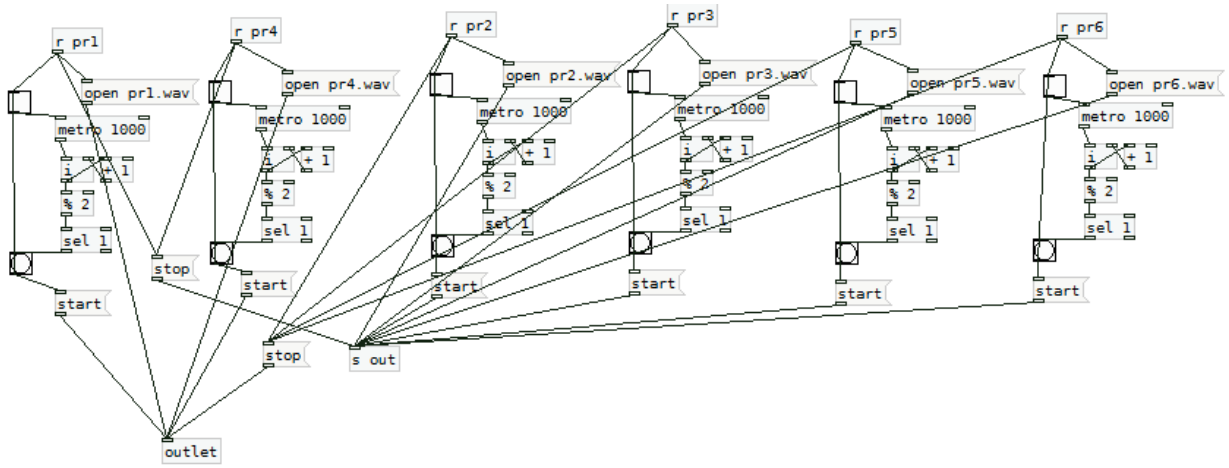
Listening Test

Now, you are going to assess 18 pair samples (A and B). Please listen to the first pair samples A and B and evaluate the similarity between them, in terms of localization, sense of space, warmth and brightness, writing your ratings in the following table. Once you have assessed the pair, press the "Next" button to continue with the following one. Please remember you can play each sound as many times as you want.

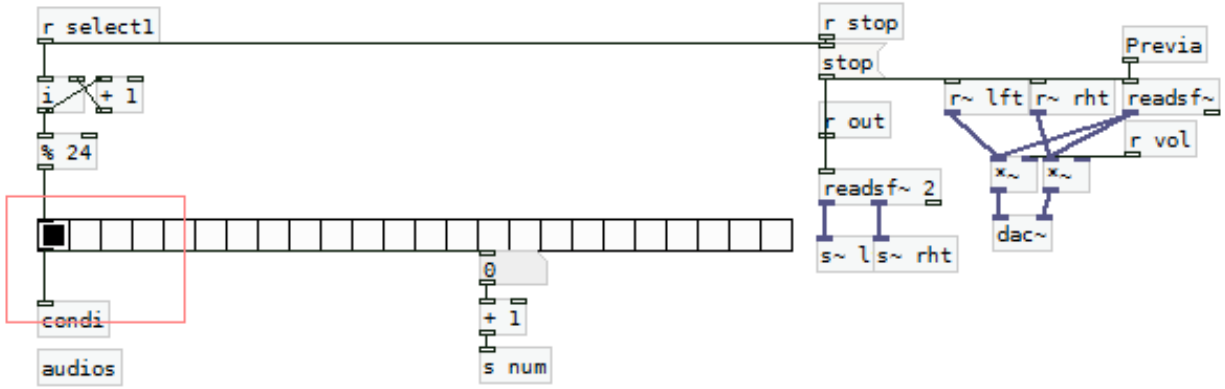
Assessment Form

Pair	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Localization																		
Sense of space																		
Warmth																		
Brightness																		

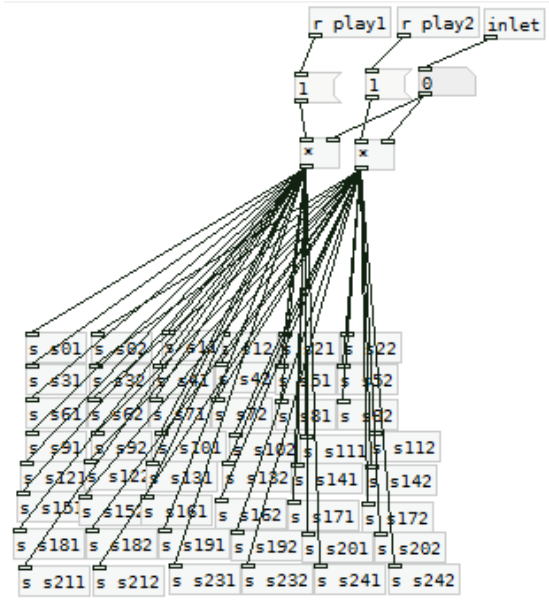
Appendix C: “Block diagrams of the subjective test in Pure Data”



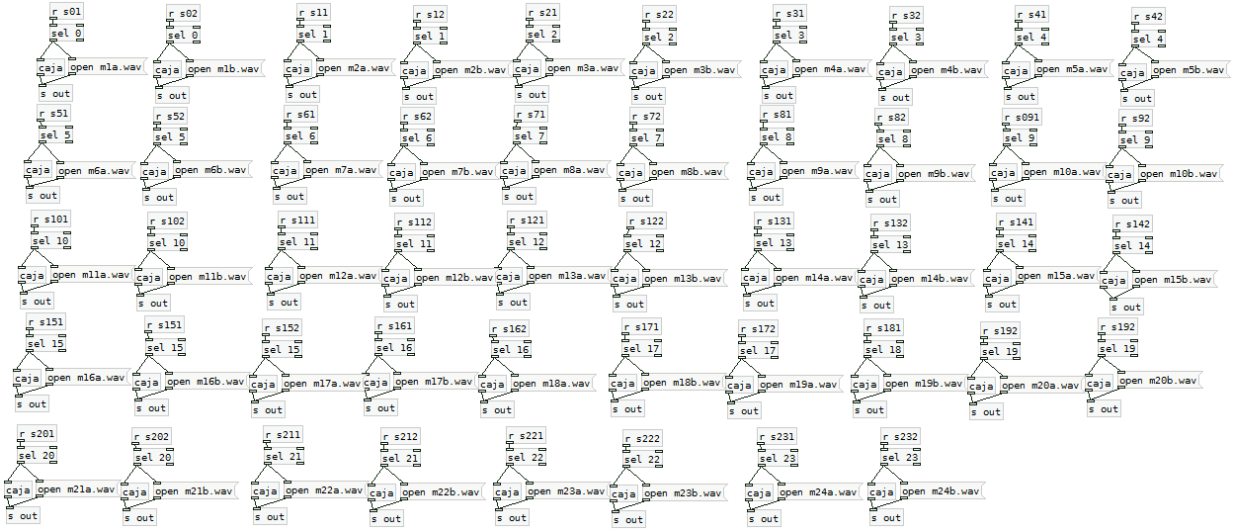
Selector



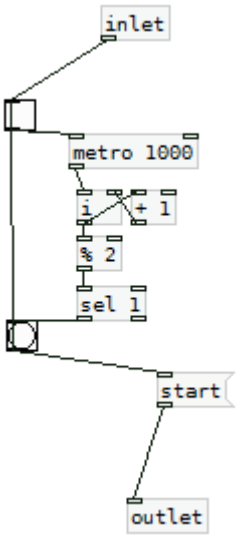
Reproduction conditions



Audio files organization



Visual interface



Appendix D: "Intelligibility and listening difficulty subjective test"

Date:

Identification – Tick or write an answer

- Gender: Female [ ] male [ ]
- Age: 16-20 21-25 26-30 31-35 36-40

We will now present you two lists of 50 words consisting of a single syllable, which lack of sense in Spanish language. Each word is included in a sentence with the following structure: "The word (number) is …". Please write in the attached tables the word you listened and evaluate for each one, the listening difficulty you perceived on a scale from 0 to 3, where:

0 No difficulty

1 Little difficulty

2 Moderate difficulty

3 Much difficulty

LIST 1

Word	Difficulty (0 to 3)	Word	Difficulty (0 to 3)
1.		26.	
2.		27.	
3.		28.	
4.		29.	
5.		30.	
6.		31.	
7.		32.	
8.		33.	
9.		34.	
10.		35.	
11.		36.	
12.		37.	
13.		38.	
14.		39.	
15.		40.	
16.		41.	
17.		42.	
18.		43.	
19.		44.	
20.		45.	
21.		46.	
22.		47.	
23.		48.	
24.		49.	
25.		50.	

LIST 2

Word	Difficulty (0 to 3)	Word	Difficulty (0 to 3)
1.		26.	
2.		27.	
3.		28.	
4.		29.	
5.		30.	
6.		31.	
7.		32.	
8.		33.	
9.		34.	
10.		35.	
11.		36.	
12.		37.	
13.		38.	
14.		39.	
15.		40.	
16.		41.	
17.		42.	
18.		43.	
19.		44.	
20.		45.	
21.		46.	
22.		47.	
23.		48.	
24.		49.	
25.		50.	

Appendix E: "Lists of Logatoms"

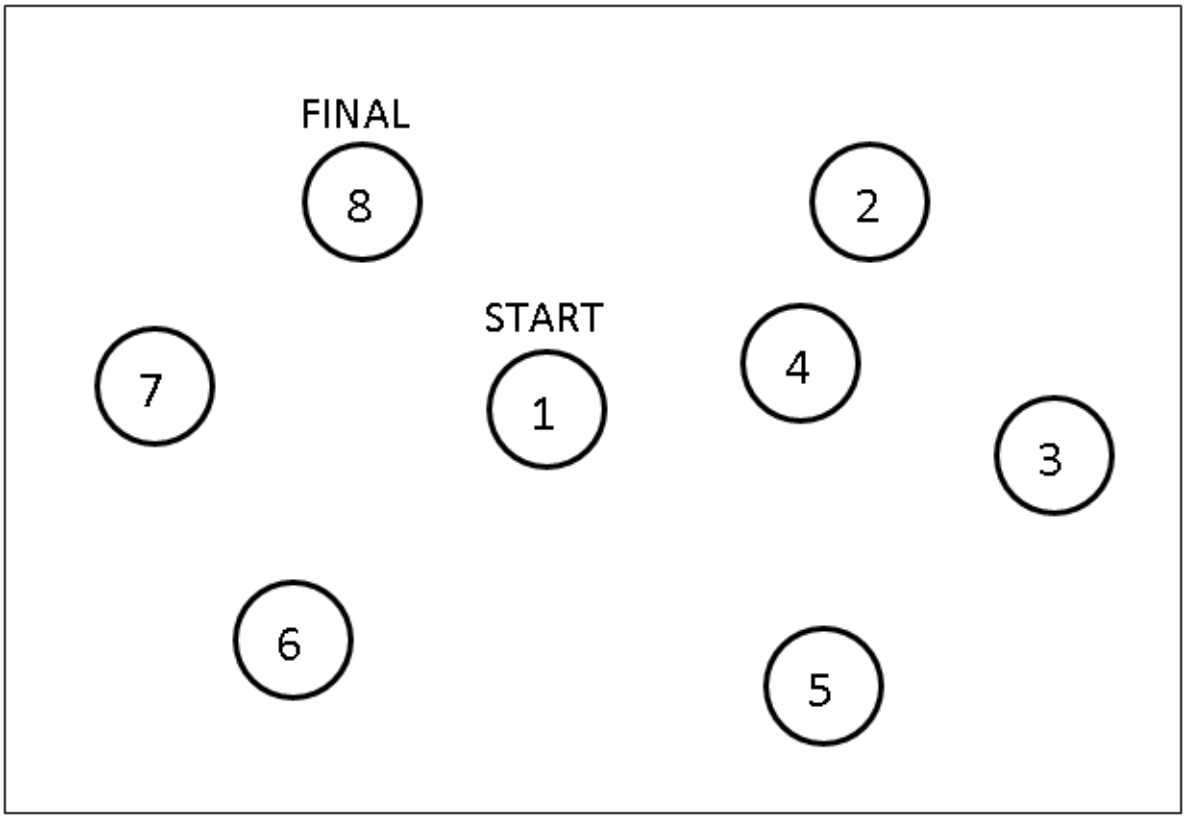
	List 1	List 2	List 3	List 4	List 5
1	JAR	FUK	REL	LUF	GUIN
2	RES	KAR	FUCH	GUIM	CHEJ
3	TICH	LUD	GOL	COCH	RUN
4	JUD	SEP	FOS	GOS	ÑAT
5	MEX	LIT	GUET	CHECH	REX
6	SEN	NEX	PAS	SEK	RACH
7	DOL	SEB	JEJ	MIR	LUS
8	GUEN	FON	SIF	KUS	DOCH
9	DOG	FUM	ÑAM	LON	PAK
10	PEM	RECH	YEN	JUF	LIB
11	FIT	LAN	KON	ÑUR	BUG
12	YOJ	TUT	LUL	RAS	TEK
13	SER	YAJ	GUIK	TIJ	BOJ
14	TAJ	ROT	MIT	CHAL	JIP
15	FOR	ÑOP	KUF	JUK	PAR
16	LAK	MAK	SEM	MAP	JOS
17	MAM	FOL	JUS	GUCH	NAP
18	ÑEL	RIX	MAT	NAN	MAX
19	GUF	NACH	NAK	PACH	MEJ
20	POT	FOG	KAX	TEM	NUL
21	ÑIJ	TIK	CHEF	YUT	YEM
22	PIX	NAM	BUX	MUL	JUCH
23	DUK	GOR	GUEB	SOP	BUP
24	SUT	JON	ÑEX	CHUF	PIT
25	SIK	KAP	RIJ	BIN	YECH
26	JAL	JAX	SIX	KUJ	MON
27	TIM	REM	YEP	NAT	BIS
28	MOP	GUES	CHAK	LAT	TEL

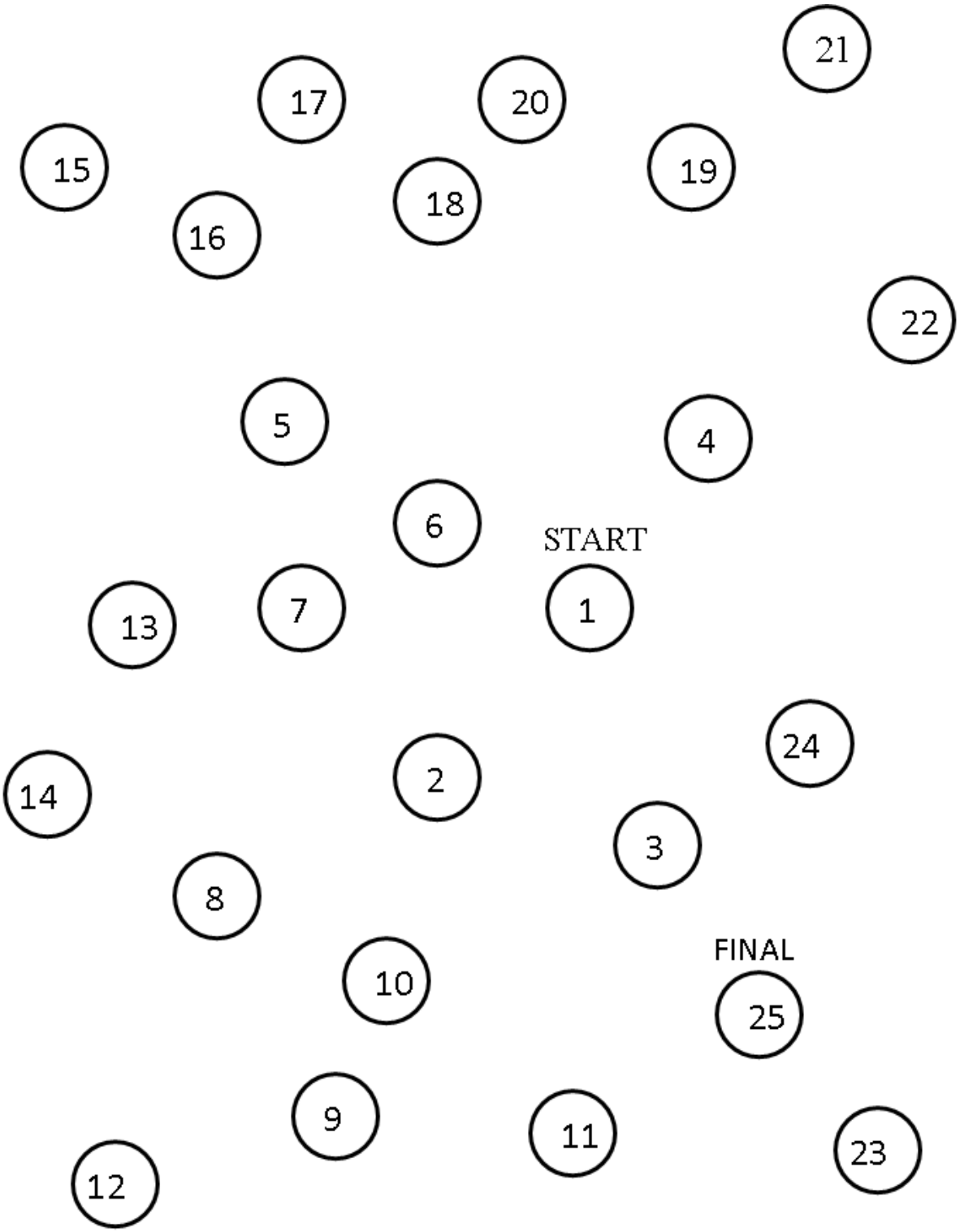
29	MACH	KUL	JOR	JIB	YIX
30	DUM	NUN	ÑUN	LEJ	DOK
31	BOS	YICH	YER	PUN	RAM
32	PEB	YOX	TES	PAM	NAL
33	ÑIF	CHACH	LAP	ÑUM	KUK
34	KAT	GUICH	GOG	SEL	YIF
35	REP	SES	JUJ	DIP	LUM
36	MUM	PIJ	MAN	NIT	MUF
37	ÑEK	PEL	NUR	NAR	PUR
38	PECH	TIF	ÑACH	TUR	FUX
39	NAS	ÑAS	YIK	TIX	DAF
40	BUCH	REK	SECH	DUX	SAS
41	DAP	PIF	BOCH	ÑAK	KOL
42	GAX	PEK	GUK	GUEX	CHAM
43	CHIX	YIM	POP	BOK	LUJ
44	JUL	DOS	LAX	MAL	ÑAN
45	KAN	TOT	BOM	PEX	CHUCH
46	TOX	TER	SOT	JUM	FOM
47	CHAF	MUR	GUS	TECH	FIR
48	DUCH	KAL	LAL	PUG	LOR
49	LUR	NOL	MAR	JOL	YOT
50	FUS	BUJ	GUM	YES	SUS

Appendix F: Trail Making Test (TMT)

PART A

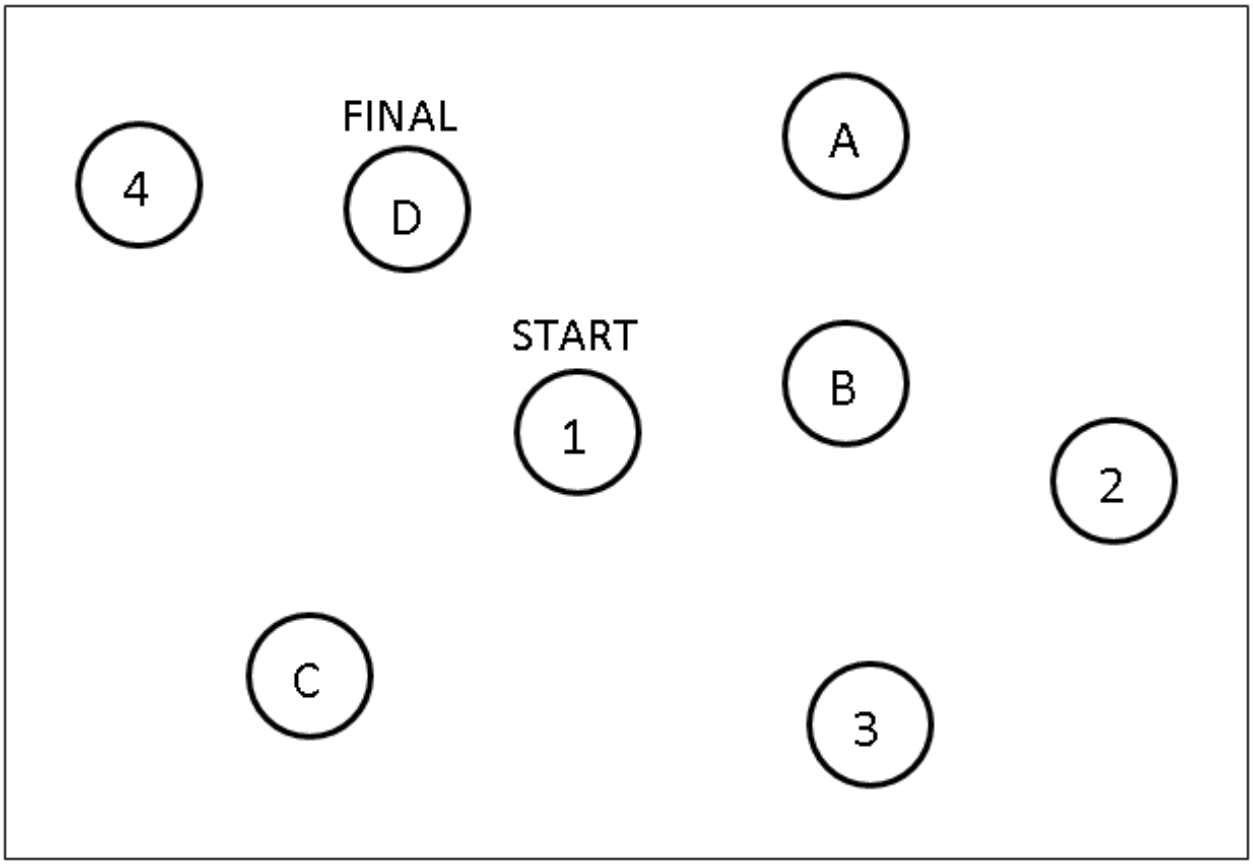
EXAMPLE

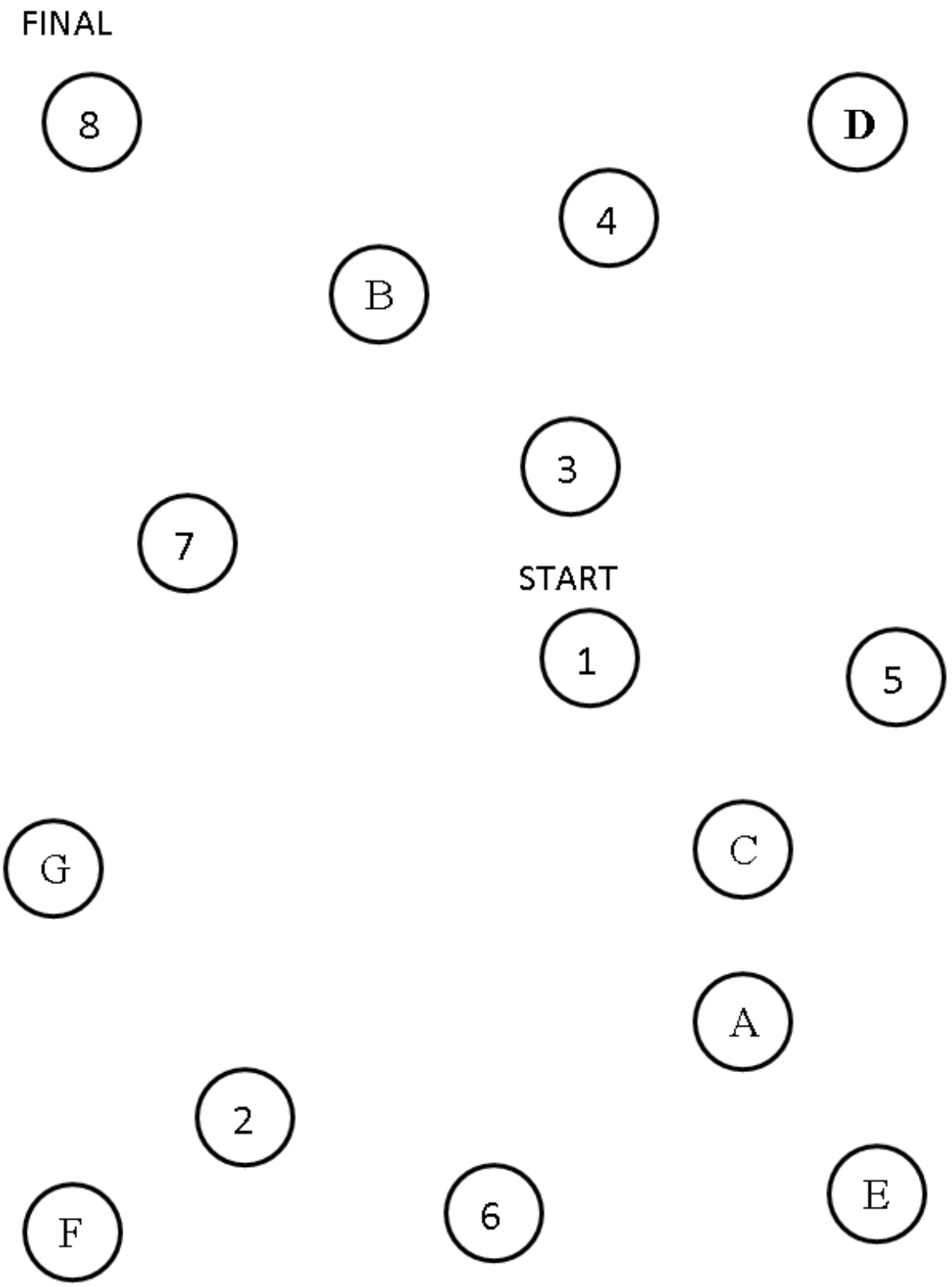



21
17
20
15
19
18
16
22
5
4
6
START
7
1
13
24
14
2
3
8
FINAL
10
25
9
11
23
12

PART B

EXAMPLE




FINAL
8
D
4
B
3
7
START
1
5
G
C
A
2
E
6
F

Appendix G: Continuous Performance Test (CPT)

Evaluator Form

NAME			
CONDITION EVALUATED		TIME	

HEARING

Listen carefully and indicate when the letter A is mentioned:

B	S	**A**	P	G	Q	T	V	E	X	**A**	C	B	Y	W	P	K	N	**A**	F
O	T	M	C	L	N	D	U	V	C	H	M	G	T	R	**A**	B	D	V	X
Z	L	S	Y	W	**A**	N	N	T	E	G	**A**	K	O	**A**	V	S	J	C	E
W	D	Q	Z	B	H	R	Z	D	U	S	Y	**A**	L	I	Z	**A**	B	D	P
A	N	C	U	F	G	R	**A**	F	J	Q	H	R	F	M	G	W	F	T	C
Q	W	N	P	L	C	I	T	V	K	U	E	Z	L	C	H	S	H	I	O
V	**A**	X	R	B	J	C	**A**	W	E	S	C	U	F	I	**A**	R	Z	**A**	I
G	O	U	**A**	N	G	U	Z	H	W	D	T	Q	C	J	N	V	W	K	E

Correct answers:___ Commission errors:___ Omission errors:_____

Verbal memory curve

We will now present you with a list of words via audio recording, please pay close attention, as you will be asked to repeat them. You will need to try to recall them in the same order as they were presented.

		I	II	III	IV	V	VI	VII	VIII	IX	X	3´	20´
1	**Dice**												
2	**Moon**												
3	**Glass**												
4	**House**												
5	**Lima**												
6	**Focus**												
7	**Knot**												
8	**Whistle**												
9	**Rose**												
10	**Candle**												
	Total												

TIME 3'		
TIME 20'		

- Initial volume.____
- Maximum volume.____
- Type of curve.____
- Organizational index.____
- Deferred evocation. 3´____ 20´____

- Pathological phenomenon._____

VISUAL

Cross out all the *As* that you see in the following table.

B	S	A	P	G	Q	T	V	E	X	A	C	B	Y	W	P	K	N	A	F
O	T	M	C	L	N	D	U	V	C	H	M	G	T	R	A	B	D	V	X
Z	L	S	Y	W	A	N	N	T	E	G	A	K	O	A	V	S	J	C	E
W	D	Q	Z	B	H	R	Z	D	U	S	Y	A	L	I	Z	A	B	D	P
A	N	C	U	F	G	R	A	F	J	Q	H	R	F	M	G	W	F	T	C
Q	W	N	P	L	C	I	T	V	K	U	E	Z	L	C	H	S	H	I	O
V	A	X	R	B	J	C	A	W	E	S	C	U	F	I	A	R	Z	A	I
G	O	U	A	N	G	U	Z	H	W	D	T	Q	C	J	N	V	W	K	E

Appendix H: Wechsler Memory Scale – III Test

Mind control

I will now ask you to please count numbers from 20 to 1 in descending order. For example 20, 19, etc. Please do it as quickly as possible.

1. (30") 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
   Time.____ Errors.____ Points.____

In this exercise, you will need to say the letters of the alphabet from A to Z. I would like to check how quick you can tell me the alphabet: A, B, C, D, E, F, G..... and so on up to the letter Z.

2. (30") A B C D E F G H I J K L M N Ñ O P Q R S T U V W X Y Z
3. Time.____ Errors.____ Points.____

For this exercise, I will ask you to mention the numbers starting with 1 add 3 to the proceeding number until reaching 40, as fast as you can. Example: 1, 4..... Please start now

4. (45") 1 4 7 10 13 16 19 22 25 28 31 34 37 40
5. Time.____ Errors.____ Points.____

Logical Memory

I will now present you with a story via audio recording; please pay close attention because at the end, you will be asked to tell the story as similar as possible.

A. A Ma'am/ Maria Moreno/ 55 years old/ who worked/ cleaning floors/ in an office building/ in the address/ Easter Avenue,/ one day/ when leaving work/ at 6/ in the evening/ in the Caracas Street/ she was mugged/ by two men/ and a woman/ and robbed/ 50.000 pesos./ She went to the policy/ reported the incident,/ the police men were touched/ and gave her 10.000 pesos.

B. The Colombian/ ship/ Gloria/ crashed/ against a rock/ near/ Cartagena/ on Monday night./ Despite the storm/ and darkness,/ the 60 passengers,/ including 18 women,/ were all rescued,/ even when the lifeboats/ were moving from side to side/ like corks/ in the ocean./ The following day,/ they were moved/ to the port/ by a Venezuelan ship.

A. Number of ideas=____ B. Number of ideas=____
Total score: (A + B)/2= ____

Digits

I will now ask you to listen carefully the numbers that are going to be presented via audio recording as you will need to repeat them in the same order they were presented.

I will now ask you to listen carefully the numbers that are going to be presented via audio recording as you will need to repeat them from the last one to the first one.

In Progression	Score	**Countdown**	Score
6-4-3-9	4	2-8-3	3
7-2-8-5	4	7-1-6	3
4-5-1-6-3	5	8-6-3-2	4
8-4-1-5-6	5	2-6-1-7	4
2-4-1-7-5-8	6	6-3-5-9-1	5
8-3-6-2-7-1	6	3-8-1-6-2	5
2-6-1-7-3-9-3	7	9-5-3-1-6-4	6
3-9-6-4-8-5-2	7	1-9-6-2-7-8	6
6-1-7-3-2-8-6-9	8	6-5-1-4-8-2-7	7
4-1-5-7-2-9-6-3	8	2-6-1-8-3-4-5	7
Total points	____	Total points	____

Paired Associate

We will now tell you a list of pairs of words, then I will tell you the first word and you will need to tell me the word with which it is associated

First Presentation	**Second Presentation**	**Third presentation**
Metal-Iron	Rose-flower	Baby-Cry
Baby-Cry	Obey-Centimetre	Obey-Centimetre
Accident-Darkness	North-South	North-South
North-South	Cabbage-Pencil	College-Market
College-Market	Up-Down	Rose-Flower
Rose-Flower	Fruit-Apple	Cabbage-Pencil
Up-Down	College-Market	Up-Down
Obey-Centimetre	Metal-Iron	Fruit-Apple
Fruit-Apple	Accident-Darkness	Accident-Darkness
Cabbage-Pencil	Baby-Cry	Metal-Iron

First Test	Easy	Difficult	Second Test	Easy	Difficult	Third Test	Easy	Difficult
North	____		Cabbage		____	Obey		____
Fruit	____		Baby	____		Fruit	____	
Obey		____	Metal	____		Baby	____	
Rose	____		College		____	Metal	____	
Baby	____		Up	____		Accident		____
Up	____		Rose	____		College		____
Cabbage		____	Obey		____	Rose	____	
Metal	____		Fruit	____		North	____	
College		____	Accident		____	Cabbage		____
Accident		____	North	____		Up	____	
Total	____	____	**Total**	____	____	**Total**	____	____

Score: Easy/2 + Difficult: _____

Appendix I: Screening questionnaire

UNIVERSIDAD DE SAN BUENAVENTURA MEDELLÍN	INITIAL QUESTIONNAIRE	DATE VERSION: November 15, 2014

Date: Time:

Full name:						**ID**							
Programme of study at USB:								**Semester**					
Age		**Sex**	**M**	**F**	**Telephone number**			**Mobile phone number**					
Previous studies	Technician		Technologist		Professional		Specialist		Master		PhD		
Occupation	**Study**	___	**Work**	___	**Specified position:**								

Carefully read and answer each of the following questions honestly, remember that your identity will be protected rigorously.

Have you ever been exposed to intense noise for a month or more	**NO**	**YES**	**Specify:** Place _____ Type of noise _____ Since when _____ How long for_____
Do you use headphones?			**Specify:** Frequency: Always/ almost always / Sometimes/ Almost never Level: High/ Medium/ Low
Do you suffer from hearing loss?			**Specify:**
Do you have any difficulty / disability?			**Specify:**
Do you have any current medical conditions?			**Specify:**
Do you have or have you had any psychological conditions?			**Specify:**
Are you taking any medication?			**Specify:** What? _____ Since when? _____
Has anything happened in the last month that could affect your concentration or memory?			**Specify:**
In the last 12 hours, have you consumed any medication or			**Specify:**

anything else that could affect or improve your concentration?			
Do you consider your memory is: **Deficient __** **Acceptable __** **Good __** **Excellent __**			**Specify:**
Do you consider your concentration is: **Deficient __** **Acceptable __** **Good __** **Excellent __**			**Specify:**
Do you consider that your cognitive performance is better in: **Morning __** **Afternoon __** **Evening __**			**Specify:**