ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013
Engineering Acoustics
Session 2pEAb: Controlling Sound Quality

2pEAb8. Predicting speech transmissibility using ray tracing and statistical energy analysis
Sascha Merz*, Vincent Cotoni and Phil Shorter

*Corresponding author's address: ESI Group, San Diego, CA 92130, sascha.merz@esi-group.com

Statistical Energy Analysis (SEA) is used across many different industries to predict interior noise, for example the background noise in vehicle interiors due to steady-state exterior sources. SEA describes the reverberant response, however some applications exist where the direct field from interior sources is also important. This includes prediction of audio system sound quality in automobiles, or speech intelligibility for public address systems in trains, coaches, and aircrafts. Computing these indices requires both the background noise level and the impulse response at the listener location due to a source representing a loudspeaker. Geometrical methods such as ray tracing are often used for describing the early impulse response where only few reflections of the direct field are involved; however they are not well suited for describing the late impulse response due to the increasing number of reflections. Instead a statistical description similar to transient SEA is typically used for the late time reflections. An SEA model of background noise therefore contains most of the information needed for predicting speech transmissibility: the background noise prediction, a simplified geometry of the acoustic space, and accurate wall impedance models. This paper demonstrates prediction of speech transmissibility by using ray tracing with typical SEA models.

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

Statistical Energy Analysis (SEA) is the industry standard for predicting car, truck, train and aircraft interior noise due to broadband exterior sources at mid to high frequencies. Typical exterior noise sources include turbulent boundary layer noise, tire noise, and machinery noise. Traditional SEA models predict the spatial average of the reverberant sound field which is a valid approximation for most exterior noise sources. This noise is typically perceived by the passengers as background noise and represents an important factor in predicting the speech transmissibility for PA and entertainment audio systems.

Speech transmissibility defines how well a transient modulated signal is transmitted from source to receiver. Indices for speech transmissibility that are frequently used are the STI (Speech Transmission Index) and the RASTI (RApid Speech Transmission Index). Speech transmissibility from a given source to a receiver location depends on the relative strength of the direct field from the source, the reverberant field from the source, and the background noise. All information related to the direct and reverberant fields from the source is contained in the transfer function or impulse response between source and receiver. The impulse response can be obtained via testing or using predictions from numerical models. Such models are typically based on ray or beam tracing for the direct field and early reflections, and on statistical models for the late reflections which arrive after the so called transition time.

In this paper a process is presented, where the impulse response between source-receiver pairs is obtained by running a ray tracing algorithm based on the information contained in an existing SEA model of a cavity. SEA models are particularly well suited because the ray tracing method typically requires a coarse geometric description of the cavity, accurate models of sound package impedance on the cavity walls, and a statistical model of sound decay in the room; all of which are typically contained in an SEA model. In addition, the background noise required for computing speech transmissibility is obtained from the exact same SEA model. The process therefore allows predicting STI, RASTI, and other room acoustics indices using a single model. Detailed sound package properties such as the incidence angle dependence of impedance can be taken into account, which is particularly useful for applications with multi-layered sound package. The method is validated against the Boundary Element Method (BEM) and published results of a round robin for room acoustics software.

PROCESS

The overall simulation process is summarized in Fig. 1. The process requires a baseline SEA model with an SEA cavity and noise control treatments (NCTs). SEA is initially used to determine the background noise due to exterior sources. The background noise is assumed to be spatially uniform throughout the cavity and is used later in the process to compute the STI and RASTI. The transfer functions between a number of interior sources, for example loudspeakers, and a number of receiver locations are then predicted as follows:

![Diagram](image)

FIGURE 1. Speech transmissibility prediction process.

In the first step (1), a number of source locations are added to the baseline model. A frequency independent ray tracing procedure is then performed up to a certain reflection order which depends on the transition time. Adding information about receiver locations, the ray tracing results are used to find the source receiver specular paths based
on a hybrid ray tracing and image source method (2). Based on the determined paths and with the information about the wall impedance, narrowband FRFs are computed for the direct field and early reflections with coherent summation of the ray contributions at a receiver location. The portion of diffused sound energy is subtracted from the ray amplitude. Inverse FFTs of the FRFs are then performed to obtain the early time impulse response before the transition time, without diffused sound.

The results from the initial ray tracing procedure (1) are then employed to determine the diffused energy (4) and the remaining sound energy in the cavity at the transition time (5). Diffused sound energy is redirected into time bins that correspond to specific time windows before the transition time. The remaining sound energy at the transition time is stored in a separate final bin. Transient SEA is then used for each bin to compute the wideband energy response of the cavity based on the initial energy in the bin. For each bin, a contribution to the total cavity impulse response in the time domain is synthesized based on the transient SEA parameters (10).

In steps (3-5), two different models of the reflection coefficient of sound package can be used. The first model is called “locally reacting” and is used in most traditional ray tracing codes: with that model, the angle of wave incidence is not taken into account for computing the wall impedance. Instead, the normal impedance is used to determine the reflection coefficient. The second model called “extended” (6, 7) accounts for the angle of incidence in the impedance calculation. Figure 2 shows the dependence of the impedance on the angle of incidence for a typical 2-layer sound-package.

![Figure 2: Variation of surface impedance of a 2-layer sound package with incidence angle](image)

a) Locally reacting impedance. b) Extended impedance.

**FIGURE 2.** Variation of surface impedance of a 2-layer sound package with incidence angle

From the results of steps (3-5), early specular response, early diffuse response, and late response are combined to yield the full impulse response for each source-receiver pair. Note that the diffused and late responses only depend on the source location and not the receiver location. For each source-receiver combination, the acoustic indices such as T60 as well as the speech transmissibility indices STI and RASTI can then be computed based on the impulse response and background noise (6).

**VALIDATION**

In the following, three validation cases are presented. The first validation case considers a monopole source near a trimmed plate, where the receivers ‘see’ the image of the source through different reflection angles with respect to the plate. The SEA / ray tracing results for the SPL are compared with results obtained from a BEM model for a locally reactive trim. The SEA / ray tracing model is then used to demonstrate the effect of using locally reactive impedance vs. the extended impedance formulation. The second validation case represents a monopole in a passenger compartment section, where sound package is applied to the seats and the carpet. SEA / ray tracing results for the RASTI are compared against results obtained from BEM. The third validation case compares SEA / ray tracing results for the reverberation time T30, the definition D50 and the center time TS against results obtained during a round robin on room simulation software.
Monopole Near a Trimmed Wall

A monopole source is placed near the center of a rigid plate with a side length of 2 m at a distance of 0.25 m. A 2-layer treatment is applied to the plate. The layer against the plate consists of 3 cm light glass wool, and the top layer is a limp septum with 0.05 kg/m² area mass. Air under atmospheric conditions is considered. 4 microphones are placed at 0.5 m distance from the wall. If edge diffraction is neglected, then the microphones are hit by a direct ray from the source and by a ray reflected once on the plate. The microphones are placed such that the reflection angle of the image source varies between 0° and 58°. A frequency domain from 400 to 2,000 Hz using 1/12th octave bands was investigated. The microphone SPL responses to a unit source is predicted by the SEA / ray tracing model and by a Fast Multipole (FMM) BEM model with about 4,000 degrees of freedom.

Figure 3 shows the results obtained using SEA / ray tracing and the results obtained from BEM for the locally reacting impedance formulation. Good agreement for all microphones can be observed. Figure 4 compares SEA / ray tracing results for the locally reacting impedance formulation with results from the extended impedance formulation. The responses for microphone 1 for which the ray reflects with 0° are identical. For increasing angle of incidence, the differences become larger. The results for microphone 4 at an angle of 58° show the largest difference of more than 3 dB.

![Figure 3](image1.png)

**FIGURE 3.** SEA / ray tracing results vs. BEM results for locally reacting impedance formulation.

![Figure 4](image2.png)

**FIGURE 4.** Local vs. extended impedance, where microphone 1 is at 0°, microphone 2 is at 19°, microphone 3 is at 39°, and microphone 4 is at 58°.
It was shown by Hodgson that the impact of the impedance model is particularly important for multilayer treatments, specifically when one layer is air. This kind of treatment is often encountered in transportation industry designs.

**Partially Trimmed Passenger Compartment**

The second validation broadly represents a PA system speaker in a passenger compartment. The room is irregularly shaped with a volume of 14 m$^3$. An omnidirectional point source is located near a top corner of the room. Multi-layered sound-package is applied to the carpet and the seats, where the seat absorption dominates the carpet absorption. RASTI results are compared between an SEA / ray tracing model and a Fast Multipole BEM model with about 70,000 degrees of freedom. Only the 500 Hz $1/3$rd octave band was taken into account. Narrowband BEM results were computed in the 500 Hz band at a frequency resolution of 1 Hz. This required several hours on a 6-core machine with 3.2 GHz clock frequency. SEA / ray tracing results were computed using about 20,000 rays with 6 reflections corresponding to a transition time of 20 ms. No diffusion was included in order to match the BEM model assumption. The calculation took few seconds on the same machine.

The SEA / ray tracing and the FMM models are shown in Fig. 5 a) and b), respectively. Figure 5 b) indicates the source as a purple object in the top left corner and the microphones as blue sensor objects. The RASTI results are presented in Fig. 6. Both the SEA / ray tracing results and the FMM results indicate the same trend: the sensors 1 and 3 which are closest to the source predict the highest values for RASTI as the direct field dominates. With increasing distance to the source, the RASTI drops. Due to the high absorption, the RASTI is at around 0.9 for most sensors which corresponds to excellent speech transmission. The discrepancies for sensors 4 and 5 are attributed to diffraction effects which are not considered in the SEA / ray tracing model. These effects are however artificially enhanced by the sharp edges used in the BEM model.

![Fig. 5. Models of a trimmed cavity.](image)

a) SEA / ray tracing model.  

b) FMM model.
Recording Studio

The third validation case is taken from a room acoustics software round robin, which was organized by the Physikalisch-Technische Bundesanstalt, Germany, in 2005. The case considers an irregularly shaped room with 10% wall absorption and scattering. A frequency domain from 125 Hz to 4 kHz in octave bands is investigated. The acoustic fluid is air at 20°C with a humidity of 50% at 1000 hPa atmospheric pressure. There are two monopole sources and three receiver locations. Refer to reference 8 for more details on the problem configuration. The predicted quantities are the decay time $T_{30}$, the definition $D_{50}$ and the center time $T_{S}$. The SEA / ray tracing model of the case is shown in Fig. 7.

The $T_{30}$ for source 1 and receiver 1 are shown in Fig. 8. The graph indicates significant variability in the round robin results. The frequency dependence of the air absorption has been taken into account in the SEA / ray tracing model, but there are two round robin results that do not consider the frequency dependence of the air absorption. Hence the $T_{30}$ is constant over frequency due to the constant absorption and diffusion coefficients. The SEA / ray tracing results are in good agreement with most curves that also consider the frequency dependence of air absorption. Note that $T_{30}$ is primarily determined by the late impulse response. The results for $D_{50}$ are shown in Fig. 9 for the 1 kHz band. Again, there is some variability in the round robin results, but the SEA / ray tracing results agree well with the general trend. Note that for $D_{50}$, correct prediction of the direct field and early reflections is crucial, as the $D_{50}$ is defined as the ratio of sound energy within the first 50 ms of the impulse response with respect to the total energy of the impulse response. For this example, the transition time is 60 ms. Results for the center time $T_{S}$ at the 1 kHz band are compared in Fig. 10. The center time is defined as the time of the center of gravity of the squared impulse response, and it requires good prediction of both the early and late parts of the impulse response. The SEA / ray tracing results follow the general trend which is about 125 ms for all source-receiver paths.
FIGURE 7. SEA / ray tracing model of irregularly shaped room.

FIGURE 8. Comparison of decay time $T_{30}$. The thick blue curve indicates the results from SEA / ray tracing and the remaining curves indicate results from the round robin.
FIGURE 9. Comparison of definition D50 for all source-receiver paths at 1 kHz. The thick blue curve indicates the results from SEA / ray tracing and the remaining curves indicate results from the round robin.

FIGURE 10. Comparison of center time TS for all source-receiver paths at 1 kHz. The thick blue curve indicates the results from SEA / ray tracing and the remaining curves indicate results from the round robin.

CONCLUSIONS

A new process was presented to predict speech transmissibility indices such as STI and RASTI by including ray tracing in an SEA model. The impulse response is computed for a number of source-receiver pairs in an SEA cavity or SEA semi-infinite fluid, where ray tracing is used for predicting the early reflections and SEA is used for predicting diffusion, the late response as well as the background noise due to exterior noise sources. If the SEA cavity represents a passenger compartment of an aircraft, train or coach, then the sources are typically speakers of a PA system and the receivers are the passenger’s ears. The angle of incidence can be taken into account for computing the reflection coefficient of multi-layered sound-package which is often encountered in passenger compartments. A number of validation examples have been presented, where good agreement between the results obtained from the new method, BEM results, and published round robin results was observed.

ACKNOWLEDGMENTS

The authors would like to acknowledge Benoit Minard for helpful discussions.
REFERENCES