Audio Engineering Society

Convention Express Paper 221

Presented at the AES 156th Convention
2024 June 15–17, Madrid, Spain

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Calibrating geometric room acoustic models using a gradient descent algorithm

Samuel D. Bellows[1] and Brian F. G. Katz[1]

[1] *Sorbonne Université, CNRS, Institute Jean le Rond d'Alembert, UMR 7190, 4 Place Jussieu, 75005 Paris, France*

Correspondence should be addressed to Samuel D. Bellows (samuel.bellows11@gmail.com)

ABSTRACT

One aspect of obtaining realistic results in room acoustic simulations for virtual reality environments or room acoustic design is careful material property adjustments to achieve agreement between the simulation output and measured results. Nonetheless, calibrating geometric room acoustic simulations presents several challenges in selecting absorption and scattering coefficient values. This work aims to improve model calibration efforts by developing a gradient-descent algorithm that minimizes differences between simulated and measured results. The algorithm follows by taking the derivative of room acoustic metrics such as reverberation time or clarity with respect to the absorption coefficient values. The resultant gradient thus derives from a single deterministic ray-tracing result represented by a set of ray paths and a set of absorption coefficients values for each surface. This formulation consequently enables efficient computational realization in calibrating absorption coefficients. Incorporating scattering coefficient adjustments would require further repeated simulations, currently not optimized. A reverberation-time optimized initialization and inequality constraints help to maintain realistic absorption coefficient values. Application of the calibration algorithm to an abbey church demonstrates the robustness and efficiency of the method.

1 Introduction

Room acoustic design, auralizations, and virtual reconstructions of heritage spaces commonly employ numerical methods based on geometric acoustics (GA) to predict a space's response [1, 2]. As with any other numerical simulation, proper material-property assignment is essential to achieving realistic results [3]. Calibrating a GA model requires carefully adjusting surface scattering and absorption characteristics to achieve agreement between simulated and measured results. Nonetheless, the number of available numerical algorithms to perform model calibration remain limited.

The aim of GA simulation methods is to estimate a room's energetic-decay response based on the laws of specular and diffuse reflection applied to rays or parti-

cles emitted from a source and detected by a receiver. Algorithms to determine the source-to-receiver path include image-source methods (ISM), ray-tracing techniques, or cone or beam-tracing approaches (see [4] for a review). Typically, random-incidence absorption coefficients determine the amount of energy lost at each reflecting surface, whereas random-incidence scattering coefficients determine the ray's trajectory after a reflection. The absorption coefficients have a strong influence on the output metrics. However, uncertainties in their values have long been known to be a challenge in room acoustic modeling. How to rigorously adjust these values to achieve agreement between simulation output and measurement results has remained even more elusive, although several techniques exist. The approaches broadly fall into guess-and-check manual adjustments or complicated non-linear optimization approaches.

One of the most commonly used techniques is to adjust values "manually in an iterative process" [5]. Martellotta et al. [6] considered "case by case" modification of materials to achieve less than 1 JND between the predicted and measured reverberation time (RT), with adjustments beginning "from surfaces with the most uncertain behavior and covering the largest areas." Postma and Katz [3] manually adjusted absorption coefficients of the material with the largest surface area until obtaining 1 JND of T_{20} and the EDT. They then used scattering coefficient adjustments to calibrate C_{50}, and lastly applied a fine-tuning step taking into account observations of local variations due to source-receiver locations. Their approach required about six iterations when applied to a lecture hall and church.

Rather than manual adjustments to absorption coefficients, Christensen et al. [5] applied a genetic algorithm to optimize the absorption coefficients of a GA model. The algorithm had some success at calibrating to within 1 JND several of the target acoustic parameters for several of the target frequency bands. However, without careful initialization and parameter constraints, the calibrated values "did not lead to realistic materials in most cases" [5]. In addition, the large computational expense led the authors to conclude that "we do not conisder the calibration utility ... fast enough to be a useful tool" [5]. A later work by Pilch [7] compared the effectiveness of three different non-linear optimization algorithms in calibrating GA models, with results comparable to those of Christensen et al. [5].

The proposed optimization algorithms of Christensen et al. [5] and Pilch [7] highlight the utility of using numerical algorithms over manual adjustments. While manual adjustment has the benefit that an experienced practitioner can ensure physically meaningful calibrated absorption coefficient values, it is tedious and prone to human error and bias. On the other hand, while complicated non-linear optimization algorithms provide a programmatic approach, their precise underlying behavior may be unclear, necessitating careful initialization and parameter ranges to ensure reliable results. As a result, there remains a need for calibration approaches which are simple, computationally efficient, and intuitive to control while remaining numerically robust to avoid human errors and bias.

This work develops a gradient-descent based optimization algorithm to improve GA model calibration. The approach derives from taking the gradient of room-acoustic metrics with respect to the absorption coefficient values. The method consequently allows controllable iterative adjustments to absorption coefficient values. However, because the gradient applies to a fixed set of source-receiver paths, it cannot optimize scattering coefficient values. In addition, a maximum-likelihood calibration applied to the Sabine equation provides an initialization procedure which incorporates statistical uncertainty in material properties. Calibration of the abbey church Saint-Germain-des-Prés validates the proposed technique.

2 Methods

2.1 Geometric Acoustics and the Echogram

Models applying GA methods estimate the temporal energetic response of a source within a room through the use of acoustic rays. Each ray carries some proportion of the total emitted sound power, while the GA method determines the ray's path. Although the exact ray trajectories may slightly vary between GA methods and implementations of diffuse scattering, all pure GA algorithms result in a set of rays, each with their own amplitude and ray path.

For a room defined by N surfaces, each with absorption coefficient α_n, the energetic response (echogram) may be expressed as

$$E(t) = \sum_{m=1}^{M} W_m \left(\prod_{n=1}^{N} (1-\alpha_n)^{p_{nm}} \right) \delta(t-\tau_m) \qquad (1)$$

where M is the number of rays and τ_m is the impact time of the mth ray with the receiver position. The integer p_{nm} represents the number of times the mth ray reflected from the nth surface. The ray paths, and thus the variable p_{nm}, implicitly depend on the room geometry, the scattering coefficients, and the scattering method employed. The sum of p_{nm} over the index n represents the order of the ray, e.g., $\sum_n p_{nm} = 0$ indicates the direct sound while a ray with $\sum_n p_{nm} = 1$ resulted from a first-order reflection. The initial ray energetic amplitude

$$W_m = \frac{W}{M} Q(\theta_m, \phi_m) \tag{2}$$

where W is the sound power of the source and $Q(\theta_m, \phi_m)$ is its directivity factor function, represents the fraction of the total sound power each ray carries. For cone or beam tracing algorithms or for image source methods, an additional $1/r_m^2$ factor must appear to compensate for spherical spreading.

The echogram allows calculation of many room acoustic parameters. For example, to compute various reverberation time metrics, one needs the energy decay curve

$$\begin{aligned} \mathrm{EDC}(t) &= \int_t^\infty E(t')dt' \\ &= \int_t^\infty \sum_{m=1}^{M} A_m \delta(t' - \tau_m)dt' \\ &= \sum_{\substack{m \\ \tau_m \geq t}}^{M} A_m, \end{aligned} \tag{3}$$

where

$$A_m = W_m \left(\prod_{n=1}^{N} (1 - \alpha_n)^{p_{nm}} \right). \tag{4}$$

On a normalized logarithmic scale, the EDC becomes

$$L(t) = 10\log_{10} \left(\sum_{\substack{m \\ \tau_m \geq t}}^{M} A_m \right) - 10\log_{10} \left(\sum_{m=1}^{M} A_m \right). \tag{5}$$

Let t_1 be the first time that $L(t)$ drops below -5 dB and t_2 be the first time that $L(t)$ drops below -25 dB. Then the RT metric T_{20} becomes

$$\begin{aligned} T_{20} &= \frac{60(t_2 - t_1)}{L(t_1) - L(t_2)} \\ &= \frac{6\ln(10)(t_2 - t_1)}{\ln(\mathrm{EDC}(t_1)) - \ln(\mathrm{EDC}(t_2))}. \end{aligned} \tag{6}$$

Other RT metrics follow by simply replacing the corresponding levels for t_1 and t_2. For example, for T_{30}, the corresponding levels are -5 dB and -35 dB, while for EDT, the corresponding levels are 0 dB and -10 dB.

Clarity metrics likewise derive from the echogram. For example, let t_{50} be the time 50ms after the first ray arrives, i.e., $t_{50} = \tau_1 + 50$ms. Then

$$\begin{aligned} C_{50} &= 10\log_{10} \left(\frac{\int_{\tau_1}^{t_{50}} E(t)dt}{\int_{t_{50}}^{\infty} E(t)dt} \right) \\ &= 10\log_{10} \left(\sum_{m=1}^{\tau_m \leq t_{50}} A_m \right) - 10\log_{10} \left(\sum_{\substack{m \\ \tau_m > t_{50}}}^{M} A_m \right) \end{aligned} \tag{7}$$

For the case of C_{80}, the value $t_{80} = \tau_1 + 80$ms replaces t_{50} in Eq. (7).

2.2 Gradient-Descent Optimization

Model calibration requires adjusting the material properties expressed by the absorption and scattering coefficients so that simulated GA output matches measured results. Accomplishing model calibration through numerical optimization requires defining a suitable objective function, determining constraints for the free variables, and a numerical technique to perform the optimization.

A reasonable objective function applied in previous works is to minimize the squared error between a measured and simulated room acoustic metric. Ranges of published values can guide upper and lower-bound constraints for the absorption coefficients [8], whereas realistic ranges of scattering coefficient values may be estimated from the geometric variations of the reflected surface [3]. Choosing a suitable numerical technique to perform the optimization is less straightforward.

If the objective function was convex, then any optimization algorithm would likely be suitable. However, both the Sabine and Eyring equations demonstrate that many different combinations of absorption coefficients applied to different surfaces can lead to the same RT. In fact, these RT equations represent an under-determined system of equations, where many free variables determine a single output value. Although GA models can add a few more metrics such as clarity, there remains the problem that only a few output metrics inform material assignment of numerous unknowns. Consequently, one anticipates that objective functions based on a few

derived room acoustic metrics will be non-convex with many local minima. That Christensen et al. [5] found that the genetic algorithm produced unrealistic absorption coefficient combinations while still minimizing their objective function further corroborates this claim.

In the case of an objective function with many local minima, one reasonable approach is to choose the local minima closest to the initial data. The use of manual calibration motivates this choice because practitioners often try to make small adjustments to maintain reasonable material properties. Developing an algorithm which can make these minor adjustments in a programmatic rather than guess-and-check fashion would greatly facilitate GA model calibration. The natural choice of algorithm to accomplish this choice is a gradient-descent approach.

In a gradient-descent algorithm, one begins with an initial set of absorption coefficients $\boldsymbol{\alpha}^{(0)} = [\alpha_1^{(0}, \alpha_2^{(0)}, \cdots, \alpha_N^{(0)}]^T$ and makes iterative updates to these values using the local gradient of the objective function as

$$\boldsymbol{\alpha}^{(k+1)} = \boldsymbol{\alpha}^{(k)} - \mu \nabla J(\boldsymbol{\alpha}^{(k)}), \tag{8}$$

where J is the objective function and μ is the step size or learning rate.

Gradient-descent methods, while powerful and computationally efficient, require the gradient of various metrics with respect to the absorption coefficients. Because the metrics all make use of the echogram, it is convenient to derive its gradient here. The derivative of the ray's energetic amplitude A_m with respect to the nth absorption coefficient is

$$\frac{\partial}{\partial \alpha_n} A_m = -\frac{p_{nm} W_m}{(1-\alpha_n)} \prod_{n'=1}^{N} (1-\alpha_{n'})^{p_{n'm}} = B_{nm} \tag{9}$$

where

$$B_{nm} = -\frac{p_{nm}}{(1-\alpha_n)} A_m. \tag{10}$$

Consequently,

$$\frac{\partial}{\partial \alpha_n} E(t) = \sum_{m=1}^{M} B_{nm} \delta(t - \tau_m). \tag{11}$$

Because the ray path p_{nm} remains fixed over the optimization, the gradient does not alter the temporal values of the echogram, only the ray amplitudes. Importantly, because B_{nm} relates to the already known ray amplitude A_m by two multiplies, these coefficients may be computed with relative efficiency. In fact, to compute the entire gradient for a room with N surfaces (or materials) requires $2NM$ multiplications.

2.2.1 Gradient of Reverberation Time Metrics

Reverberation-time metrics follow from the energy decay curve. Its derivative with respect to α_n, referred to for convenience as DEDC_n, is

$$\mathrm{DEDC}_n(t) = \frac{\partial}{\partial \alpha_n} \mathrm{EDC}(t) = \sum_{\substack{m \\ \tau_m \geq t}}^{M} B_{nm} \tag{12}$$

which is simply a summation over the same temporal indices as in the $\mathrm{EDC}(t)$ but with new ray amplitudes given by B_{nm}.

For the purpose of optimization, it is convenient to optimize with respect to reverberation rate, i.e.,

$$R_{20} = 1/T_{20} = \frac{\ln(\mathrm{EDC}(t_1)) - \ln(\mathrm{EDC}(t_2))}{6\ln(10)(t_2 - t_1)}. \tag{13}$$

so that a $R_{20} = 1\ \mathrm{s}^{-1}$ means that a room decays 60 dB per second. This choice places the EDC in the numerator and yields a simpler expression for the gradient.

Before proceeding, note that

$$\frac{\partial}{\partial \alpha_n} \ln[\mathrm{EDC}(t)] = \frac{\mathrm{DEDC}_n(t)}{\mathrm{EDC}(t)}, \tag{14}$$

so that, under the approximation that t_1 and t_2 may be treated as constants, the gradient of R_{20} with respect to the nth absorption coefficient becomes

$$\frac{\partial R_{20}}{\partial \alpha_n} = \frac{1}{6\ln(10)(t_2 - t_1)} \left(\frac{\mathrm{DEDC}_n(t_1)}{\mathrm{EDC}(t_1)} - \frac{\mathrm{DEDC}_n(t_2)}{\mathrm{EDC}(t_2)} \right). \tag{15}$$

For other reverberation time metrics, such as EDT or T_{30}, one simply needs to change the time indices t_1 and t_2 to correspond to those associated with the desired metric.

2.2.2 Gradient of Clarity Metrics

The gradient of clarity follows as

$$\frac{\partial C_{50}}{\partial \alpha_n} = \frac{10}{\ln(10)} \frac{\partial}{\partial \alpha_n} \left(\ln \left(\sum_{m=1}^{\tau_m \le t_{50}} A_m \right) - \ln \left(\sum_{\tau_m > t_{50}}^{M} A_m \right) \right) \tag{16}$$

$$= \frac{10}{\ln(10)} \left(\frac{\sum_{m=1}^{\tau_m \le t_{50}} B_{nm}}{\sum_{m=1}^{\tau_m \le t_{50}} A_m} - \frac{\sum_{\tau_m > t_{50}}^{M} B_{nm}}{\sum_{\tau_m > t_{50}}^{M} A_m} \right).$$

For computing derivatives of C_{80}, one need only replace t_{50} with t_{80}.

2.2.3 Optimization Algorithms

With closed-form solutions for the derivatives, it is now possible to develop gradient-descent algorithms to optimize absorption coefficient values in GA model calibrations. First, define a squared-error objective function between the measured R_{20} and simulated $\tilde{R}_{20}$ as

$$J_{R_{20}}(\boldsymbol{\alpha}) = (\tilde{R}_{20}(\boldsymbol{\alpha}) - R_{20})^2 \tag{17}$$

where the dependence of $\tilde{R}_{20}$ on the absorption coefficients is made explicit. Taking the derivative with respect to the absorption coefficients gives

$$\frac{\partial}{\partial \alpha_n} J_{R_{20}} = 2(\tilde{R}_{20} - R_{20}) \frac{\partial}{\partial \alpha_n} \tilde{R}_{20}, \tag{18}$$

where Eq. (15) gives the value of $\frac{\partial}{\partial \alpha_n} \tilde{R}_{20}$. After a single iteration, the nth absorption coefficient becomes

$$\alpha_n^{(k+1)} = \alpha_n^{(k)} - 2\mu(\tilde{R}_{20}^{(k)} - R_{20}) \frac{\partial}{\partial \alpha_n} \tilde{R}_{20}^{(k)}. \tag{19}$$

After computing a new set of absorption coefficients, the new $\tilde{R}_{20}^{(k+1)}$ follows by recomputing the echogram using the same p_{nm} but with the updated absorption coefficient values.

Likewise for clarity, defining a squared-error objective function as

$$J(\boldsymbol{\alpha}) = (\tilde{C}_{50}(\boldsymbol{\alpha}) - C_{50})^2 \tag{20}$$

gives an update of

$$\boldsymbol{\alpha}^{(k+1)} = \boldsymbol{\alpha}^{(k)} - 2\mu(\tilde{C}_{50}^{(k)} - C_{50}) \nabla_{\boldsymbol{\alpha}} C_{50}^{(k)}(\boldsymbol{\alpha}). \tag{21}$$

2.2.4 Combining Metrics and Source-Receiver Pairs

The various room acoustic metrics may be added into a single objective function. As discussed by Pilch [7], each metric should first be normalized by a value of 1 JND to ensure equal weighting. Of course, one may also choose to weight certain metrics over others, e.g., prioritizing RT metrics over clarity. In addition, multiple source-receiver pairs may be added to optimize over many positions. Because the gradient is a linear operator, averaging across source-receiver pairs amounts to averaging each of their individual gradients in a global update step. For more details on normalization procedures, see [7].

2.3 Initialization Through Reverberation-Time-Equation Calibration

Because of anticipated non-convex objective functions with many local minima, choosing a good initial estimate is essential to obtaining the most reliable results [5]. Although the best initial values would follow from careful measurements, using tabulated values available in handbooks is often the only feasible option. However, these published values may significantly vary from actual ones. Consequently, it would be beneficial to apply an initial calibration step to ensure a reasonable initial value $\boldsymbol{\alpha}^{(0)}$ to feed into the gradient-descent algorithm. This work proposes the use of a maximum-likelihood constrained optimization applied to the Sabine equation to perform this task.

The Sabine equation expresses the RT as

$$T_{60} = \frac{0.164V}{\mathbf{S}^T \boldsymbol{\alpha}} \tag{22}$$

where $\mathbf{S}$ is a vector containing the surface areas of each wall. While GA models allow more complex physics compared to the simplifications made to derive this result, the Sabine estimate is often reasonable, especially for simply shaped rooms [9]. Rearranging terms shows that all possible combinations of $\boldsymbol{\alpha}$ which yield the same value of T_{60} lie on the hyperplane defined by

$$\mathbf{S}^T \boldsymbol{\alpha} = b, \tag{23}$$

where

$$b = \frac{0.164V}{T_{60}} \tag{24}$$

is a constant.

The aim of the RT pre-calibration is to ensure that the gradient-descent algorithm's initial values $\boldsymbol{\alpha}^{(0)}$ satisfy the Sabine equation as a preliminary validity check. Constructing a probabilistic model for each absorption coefficient allows one to choose the absorption-coefficient-value combination which is mostly likely to occur. This choice is necessary as there are an infinite number of combinations which produce the same target RT.

Because absorption is limited to $0 \leq \alpha \leq 1$, the truncated normal distribution (see [10]), serves as a convenient parameterized model of the probability. Using the parameters μ and σ, the mean and standard deviation of the untruncated normal distribution, one may express the relative certainty of different absorption coefficient values. In this way, one speaks of the likelihood or probability of the absorption coefficient falling between certain ranges. For example, a probability density function (PDF) defined by $\mu = 0.5$ and $\sigma = 0.01$ would represent a material whose absorption coefficient value falls very near $\alpha = 0.5$, whereas values of $\mu = 0.1$ and $\sigma = 0.5$ could be used to model a material with low absorption but whose precise value is known with little certainty. Tabulated datasets can assist in choosing appropriate choices of μ and σ [8].

With a probabilistic model of each individual absorption coefficient, one may describe the likelihood of any combination of $\boldsymbol{\alpha}$ through a likelihood function

$$L_N(\boldsymbol{\alpha}) = \prod_{n=1}^{N} f_n(\alpha_n; \mu_n, \sigma_n), \tag{25}$$

where f_n are the individual probability density functions of each absorption coefficient. Calibration requires that the absorption coefficients produce the measured T_{60} according to the Sabine equation. Consequently, one seeks to find the maximum likelihood of this function on the hyperplane defined by $\mathbf{S}^T\boldsymbol{\alpha}$ = b. In practice, it is easier to maximize the log-likelihood

$$l_N(\boldsymbol{\alpha}) = \log[L_N(\boldsymbol{\alpha})] = \sum_{n=1}^{N} -\frac{1}{2}\left(\frac{\alpha_n - \mu_n}{\sigma_n}\right)^2 + C_n \tag{26}$$

where C_n is a scale-factor constant. The optimization problem may be expressed as

$$\begin{aligned} &\text{maximize} && l_N(\boldsymbol{\alpha}) \\ &\text{subject to} && \mathbf{S}^T\boldsymbol{\alpha} = b \\ & && 0 \leq \alpha_n \leq 1. \end{aligned} \tag{27}$$

Loosely speaking, among the set of absorption coefficients which yield the desired T_{60}, the maximum-likelihood estimate (MLE) seeks to find the combination $\boldsymbol{\alpha}^{(0)}$ which is most likely to occur given the probabilistic model.

Momentarily setting aside the inequality constraints, the solution follows from Lagrange multipliers. The Lagrangian is

$$\mathscr{L}(\boldsymbol{\alpha}, \lambda) = \left(\sum_{n=1}^{N} -\frac{1}{2}\left(\frac{\alpha_n - \mu_n}{\sigma_n}\right)^2 + C_n\right) + \lambda(\mathbf{S}^T\boldsymbol{\alpha} - b). \tag{28}$$

Differentiating with respect to $\boldsymbol{\alpha}$ and setting the result equal to zero leads to

$$\boldsymbol{\alpha}^{(0)} = \boldsymbol{\mu} - \left(\frac{\tilde{b} - b}{\mathbf{S}^T\mathbf{K}\mathbf{S}}\right) S_n \sigma_n^2, \tag{29}$$

where

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix} \tag{30}$$

is the covariance matrix and $\tilde{b} = \mathbf{S}^T\boldsymbol{\mu}$. Importantly, this form of calibration modifies absorption coefficients with the highest uncertainty and that are attached to the largest surface areas. This closed-form solution assumes that the inequality constraints are inactive. In the case a constraint is active, one must apply a numerical method to find the minimum. However, the simple expression of the objective function and constrains allows straightforward implementation into commonly available optimizers.

3 Results

The GA model under consideration is that of the abbey church Saint-Germain-des-Prés (Fig. 1). Postma and Katz [3] contains more details on the GA model creation and room acoustic measurements for this building. Although that work considered manual adjustments of absorption and scattering coefficients to calibrate averaged metrics across source-receiver pairs, the present work applies the proposed method as part of an automated calibration procedure.

First, the initial absorption and scattering coefficient values were the same initial values appearing in [3], Table 4. While [3] adjusted scattering coefficient values

during their calibration procedure, the present work can only modify absorption coefficients so that the scattering coefficients remained fixed. Next, an in-house ray-tracing program evaluated the church's energetic response for two source and six receiver locations as shown in Fig. 1. The green dotted curves in Fig. 1(b)-(e) indicates the averaged T_{20}, EDT, C_{50}, and C_{80} across the twelve source-receiver pairs using the initial values contained in [3] with no effort at calibration. Similar to [3], the initial values underestimate RT metrics at low frequencies.

Second, the pre-calibration step applied the RT-initialization using the MLE constrained optimization. A three-tiered ranking determined the material uncertainties σ. Materials with high uncertainty values, such as the alter, pulpit, organ, or pews, used a value of $\sigma = 0.1$. Those with lower uncertainty, such as the stone floor, used a value of $\sigma = 0.01$, while all other materials used $\sigma = 0.05$. The MLE step then calibrated the initial absorption data against the averaged T_{20} across the twelve source-receiver pairs. Applying these adjusted absorption coefficient values to the ray-tracing results produced the metrics appearing as blue dash-dot curves in Fig. 1(b)-(e). This simple step, following from a closed-form solution, already yielded significant improvement; all metrics fell within 1.5 JND and both clarity metrics fell within 1 JND.

Lastly, the gradient-descent calibration procedure simultaneously optimized across the two sources, six receivers, and four room acoustic metrics (T_{20}, EDT, C_{50}, and C_{80}) for each octave band. Using a metric weighting of [2, 0.1, 1, 1], respectively, favored adjusting metrics which had the worst agreement after the MLE step. The optimization required roughly three minutes (CPU: Intel i7, RAM: 32 GB). Results after twenty iterations appear as the dashed red curves in in Fig. 1(b)-(e). For all cases, the averaged value fell within 1 JND of the averaged measured value across the twelve source-receiver pairs.

4 Discussion

The two primary challenges of calibrating a GA model are first, ensuring agreement between measured and simulated data and second, ensuring reasonable calibrated absorption coefficient values. To achieve the latter, the proposed method employs two key approaches: an initialization based on maximum-likelihood estimation using statistical models of material uncertainties and a gradient-descent algorithm which provides minor iterative adjustments to the absorption coefficient values.

Although the method also employed inequality constraints as in [5, 7], during the calibration procedure these constraints remained inactive as the adjustments were very small. For example, the maximum change in absorption coefficient values between the initial material properties and the MLE initialization was only five percent absorption. The gradient-descent approach modified the MLE coefficients further by at most another four percent absorption. The final calibrated values all fell within four percent absorption of the initial values used in [3]. Consequently, one strength of the proposed algorithm is its ability to tune the model using only minor adjustments. This essentially shifts the focus of model calibration onto obtaining the most realistic initial absorption coefficient values and properly quantifying material uncertainty, rather than hoping that algorithms producing widely fluctuating values will yield a meaningful end result.

5 Summary

This work derived a gradient-descent optimization approach to calibrating geometric acoustic models. The technique followed by differentiating room metrics with respect to the absorption coefficients. An initialization procedure based on maximum-likelihood estimation employed a statistical model of material uncertainty to improve calibration results. Application of the model to the abbey church Saint-Germain-des-Prés demonstrated its ability to tune absorption coefficient parameters in order to acheive 1 JND agreement between measurements and simulations. Future work includes validating the method in different rooms, developing analytic methods to optimize scattering coefficients, and improving statistical modeling of material-property uncertainty.

6 Acknowledgements

This work was carried out in part in the context of the SONICOM project (https://www.sonicom.eu) that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 101017743.

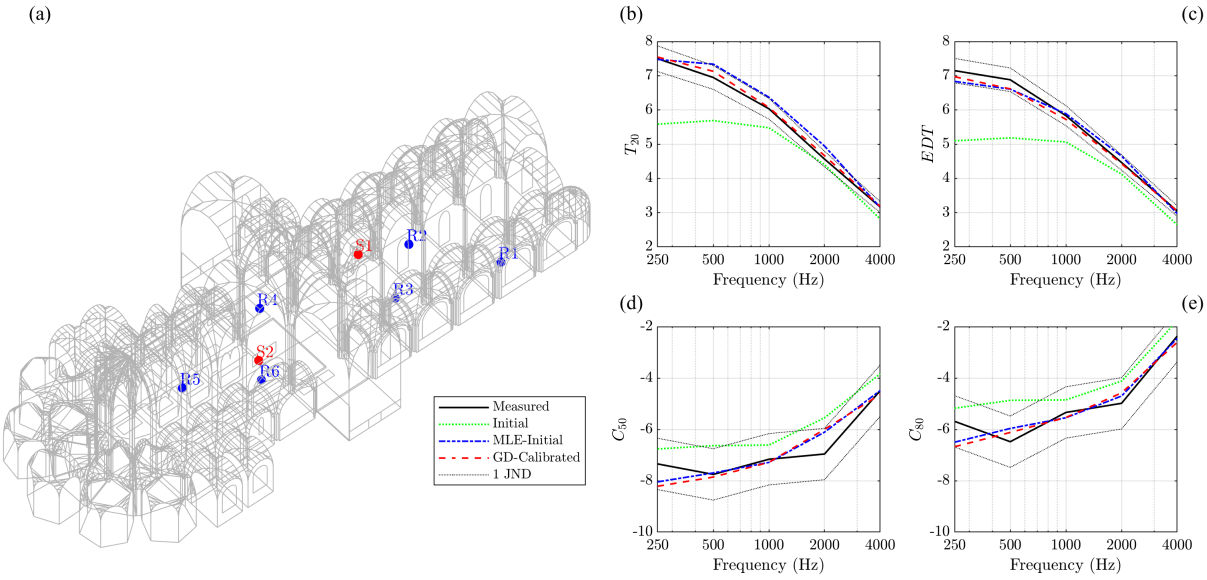


Fig. 1: (a) GA model of Saint-Germain-des-Prés with the source-receiver locations used in the calibration. Averaged values of room metrics including (b) T_{20}, (c) EDT, (d) C_{50}, and (e) C_{80}, during the calibration procedure.

References

[1] Kleiner, M., Dalenbäck, B.-I., and Svensson, P., "Auralization-An Overview," *J. Audio Eng. Soc*, 41(11), pp. 861–875, 1993.

[2] Vorländer, M., *Auralizations: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Springer Berlin, Heidelberg, 2008.

[3] Postma, B. N. J. and Katz, B. F. G., "Creation and calibration method of acoustical models for historic virtual reality auralizations," *Virtual Reality*, 19(3–4), pp. 161–180, 2015, doi:10.1007/s10055-015-0275-3.

[4] Savioja, L. and Svensson, U. P., "Overview of geometrical room acoustic modeling techniques," *J. Acoust. Soc. Am.*, 138(2), pp. 708–730, 2015, doi:10.1121/1.4926438.

[5] Christensen, C. L., Koutsouris, G., and Rindel, J. H., "Estimating absorption of materials to match room model against existing room using a genetic algorithm," in *Proceedings of Forum Acusticum*, Krakow, Poland, 2014.

[6] Martellotta, F., Crociata, S. D., and D'Alba, M., "On site validation of sound absorption measurements of occupied pews," *Applied Acoustics*, 72(12), pp. 923–933, 2011, doi:10.1016/j.apacoust.2011.06.001.

[7] Pilch, A., "Optimization-based method for the calibration of geometrical acoustic models," *Applied Acoustics*, 170, pp. 1–14, 2020, doi:10.1016/j.apacoust.2020.107495.

[8] Deetz, N. and Boren, B., "Improving auto-calibration of GA-based simulations through a statistical absorption database," in *Proceedings of Forum Acusticum*, Torino, Itay, 2023, doi:10.61782/fa.2023.0328.

[9] Prawda, K., Schlecht, S. J., and Välimäki, V., "Calibrating the Sabine and Eyring formulas," *J. Acoust. Soc. Am.*, 152(2), pp. 1158–1169, 2022, doi:10.1121/10.0013575.

[10] Johnson, N. L., Kotz, S., and Balakrishnan, N., *Continuous Univariate Distributions*, volume 1, John Wiley & Sons, 2 edition, 1994.