CALIBRATED AURALIZATION SIMULATION OF THE ABBEY OF SAINT-GERMAIN-DES-PRÉS FOR HISTORICAL STUDY

BNJ Postma Audio Acoustics Group, LIMSI-CNRS, Orsay, France
A Tallon Art Department, Vassar College, Poughkeepsie, USA
BFG Katz Audio Acoustics Group, LIMSI-CNRS, Orsay, France

1 INTRODUCTION

Over recent decades, auralizations have become more prevalent in historic research and archaeological acoustics. With these techniques it is possible to explore the acoustic conditions of buildings which have been significantly modified over time, providing that the original geometry and the acoustic characteristics of their surfaces are known.[1] In this manner, historians are provided with the opportunity to explore lost acoustic environments of important buildings.

Calibration of auralizations is necessary if one wishes to build a scientific tool rather than a simple audio novelty. In this context, a study was carried out on the Parisian Saint-Germain-des-Prés. The abbey church was begun in the 11th century, with major modifications undertaken in the 12th and again 17th centuries which resulted in changes in the acoustic conditions.[2]

A geometrical acoustic (GA) model of the church was created and calibrated, as discussed in the following section. Sec. 3, describes the validation of the calibration by means of an auralization listening test. The acoustic environment of the church as it stood before the 17th century modifications was compared to that of the current Saint-Germain-des-Prés. The calibrated GA model was modified to reflect the church's configuration during this period in Sec. 4. When simulation results of the current and pre-modern configurations were compared, it was observed that the abbey church of Saint-Germain-des-Prés used to have perceptually shorter reverberation (T20 and EDT) and higher clarity (C50 and C80), especially in the principally occupied areas.

2 CALIBRATION CONCEPT

As with any scientific simulation, it is necessary to calibrate GA models. An overview of the calibration procedure is presented in what follows. Subsequently, acoustic measurements which served as references for the calibration are discussed. Finally, the creation and calibration of the GA model are considered. The calibration was performed according to the previously reported 7 step procedure.[3]

1. RIR measurements are carried out in the studied venue. The results of these measurements are used as a reference for the calibration.
2. The geometrical model is created and remains unchanged during calibration.
3. Preliminary acoustical properties are assigned to all surfaces, resulting in a GA model.
4. Since stochastic implementations of Lambert scattering in GA software leads to run-to-run variations, the GA model's repeatability is quantified. These variations are then taken into account when simulation and measurement are compared.
5. The sensitivity of the GA model to adjustments of scattering coefficients is quantified.
6. Acoustical surface properties are modified, taking into account the determined sensitivities, in order to arrive at global mean differences between measurement and simulated results for reverberation and clarity parameters of less than 1 just noticeable difference (JND).
7. Acoustic properties of local key surfaces are adjusted to minimize the standard deviation (SD) of the differences for reverberance and clarity parameters.

This procedure relies on several assumptions:

- Selected acoustic parameters for reverberation and clarity are sufficient metrics.
- Calibration according to objective parameters within 1 JND results in a valid simulation.
- Calibration of a GA model for a sufficient ensemble of discrete points (source and receiver positions) provides sufficient confidence in the quality of the simulated RIR at other positions.

Additionally, since previous studies reported variations in RIR analysis algorithm implementations[4], it is important to use a single analysis tool for estimating the acoustic parameters when comparing results from measured and simulated RIRs. In order to avoid issues regarding automatic noise detection algorithms when applied to the ideal background noise-free simulated RIRs, a low level white Gaussian noise (65 dB) was added to the simulated RIRs prior to parameter analysis.

This calibration procedure shows similarities to a recent study on another church.[5] However that study did not address calibration for auralization, but was concerned with calibration of the GA model for reported parameters by the simulation software only, in order to predict the effect of changes in acoustic conditions. The selected metrics were global mean EDT and C80 with 1 JND of measured values. While these parameters agreed well, it is worth noting that T30 values differed by over 20% for some positions. No local calibration was carried out to reduce variance.

2.1 Associated measurement

Acoustic measurements were carried out in order to serve as a reference for the calibration with the following details. Additional details can be found in Postma and Katz[3].

Signal- The Exponential Swept Sine method was employed. The sweep frequency went from 20 Hz to 20 kHz, duration of 10 s.

Sound source- The audio output was sent to an amplifier (Servo 120a, SAMSOM) and sequentially to two miniature dodecahedral sound sources (model 3D-032, Dr-Three).

Microphones- Four omnidirectional microphones (model 4006, DPA) and an artificial head (Neuman KU 80 equipped with model 4060, DPA) were used.

Measurement positions- Fig. 1 shows the measurement plan. Two source positions were chosen representing typical usage (pulpit:S1 and altar:S2). 32 receiver positions covered normal attendance (1-2, 4-7, 9-12, 14-17, 19-20), the high altar (21-22, 24-25), and the sanctuary (26-27, 29-30). For the dummy head, center-line positions (3, 8, 13, 18, 23, 28) were measured.

Measured parameters- RIR's were analyzed using LIMSI's in-house MatLab impulse response analysis (IRA) toolkit. For the purpose of this study, six ISO Standard[6] parameters were calculated: T20, EDT, C50, C80, IACC early (e), and IACC late (l).

2.2 GA model

CATT-Acoustic (v.9.0.c, TUCT v1.1a) was employed to create the GA model and perform simulations.[7] The geometry of the Saint-Germain-des-Prés was determined from a 3D laser scan point cloud as well as architectural plans and sections (see Fig. 2). The surface materials in the abbey church were determined by means of visual inspection. Absorption coefficients were adopted

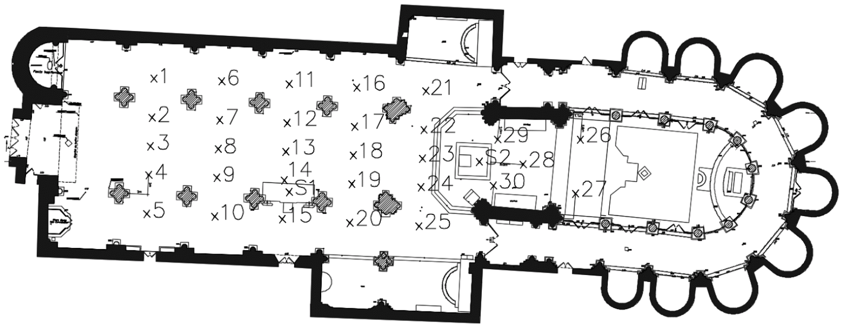

Fig.1: Source and receiver measurement plans in the Saint-Germain-des-Prés.

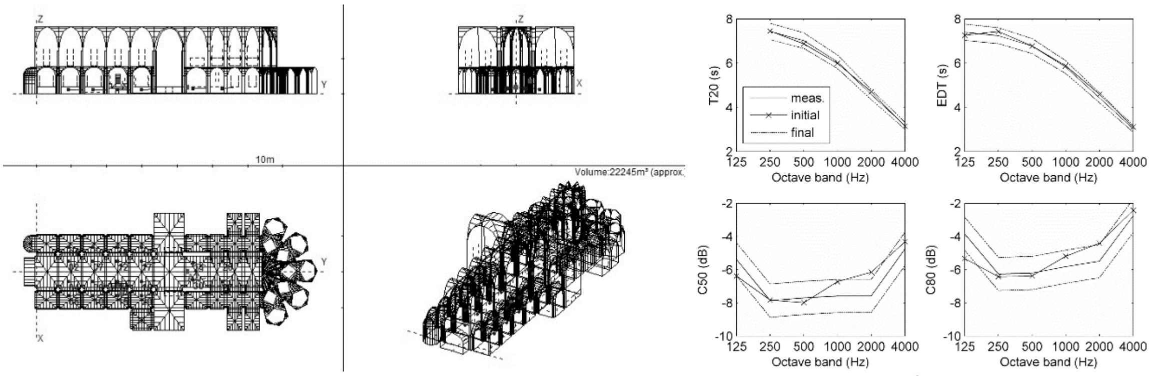

Fig.2: (Left:) GA model of the Saint-Germain-des-Prés. (Volume ~22200 m^{3}, 4518 surfaces). (Right:) Comparison between measured and simulated mean T20, EDT, C50, and C80 (±1 JND).

from publicly available databases. Scattering coefficients of surfaces were generally modeled using the option *estimate* which provides a simple estimation of these coefficient based on a given characteristic depth representative of the surface's roughness. The binaural GA simulation incorporated the previously measured HRTF of the dummy head used during the measurement.

Ten repetitions were run of the initial configuration. Analysis of the SD calculated per position for each acoustic parameter quantified the run-to-run variation. The sensitivity of the GA model to the scattering coefficient was studied by running simulations of the initial GA model followed by simulations with all scattering coefficients set to 0%, then to 99%, with absorption coefficients unchanged. Then, absorption coefficients of the materials with the largest surface areas were adjusted, since small variations lead to a considerable effect. After the mean reverberation parameters (T20 and EDT) were adjusted to within 1 JND of the measured values, scattering coefficients were adjusted to achieve mean parameters (C50 and C80) within 1 JND. Fig. 2 compares the mean measured T20, EDT, C50, and C80 to those of the calibrated GA model. Finally, acoustical properties of local key surfaces were adjusted to minimize the SD of the differences between measured and simulated results for the reverberance and clarity parameters.

3 SUBJECTIVE LISTENING TESTS

To evaluate the assumptions on the validity of the calibration procedure, a listening test was carried out comparing measured and simulated RIR auralizations. It should be noted that prior to commencing listening tests, some additional processing is required concerning the measured RIR.

3.1 Preparation of the measured RIR

The frequency response characteristics of the measurement system were compensated for by creating an equalization filter. The measurement chain (one microphone only) was installed in an anechoic room (IRCAM, Paris) and the RIR of the omnidirectional speaker was measured at 5º increments in the horizontal plane. The resulting RIRs were time-windowed to 512 pt, to remove any reflection artifacts, from which the FFT was calculated and the mean over all directions of the magnitude determined. A filter was generated to match the inverse of this response using the recursive filter design *yulewalk* method. Non-linear frequency weighting followed a bark scale approximation, constraining the filter's level of detail to follow human hearing sensitivity. The resulting filter was applied to all measured RIRs prior to any spectral analysis.

Subsequently, differences in SNR between frequency bands were compensated for. The RIR was decomposed into 1/3rd-octave band components (spanning 100 – 16000 Hz). The noise floor was detected by determining the SNR for each 1/3rd-octave band. The signals were then windowed at the point 5 dB above the noise floor, eliminating the trailing noise. The decay rate (reverberation time) was then calculated over the entire window, and this decay rate was used to synthesize the continuation of a noise-free reverberant tail. Since a reliable decay estimate reasonably requires at

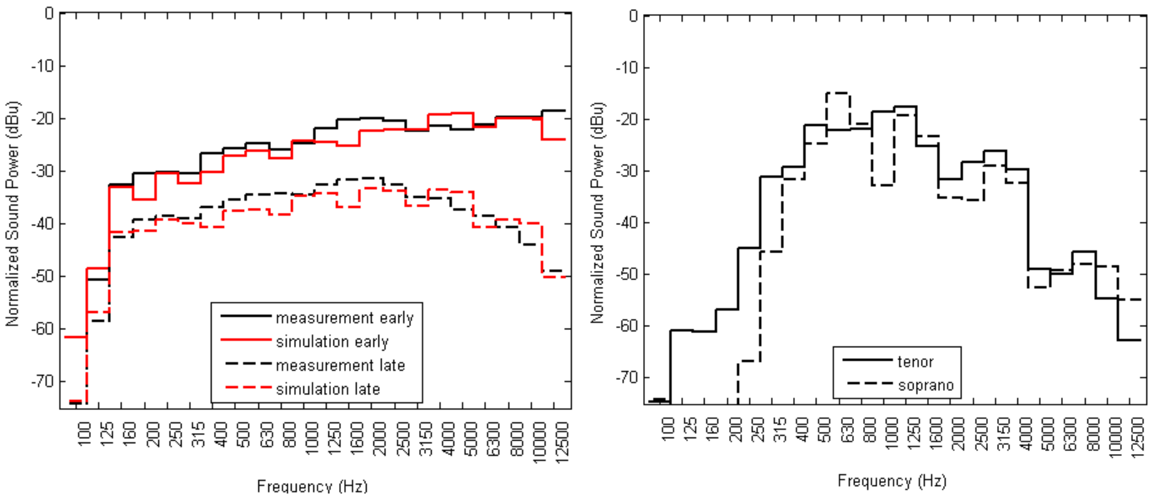

Fig 3: 1/3-octave RMS power spectrum for (Left) measured and simulated RIR for the early (0-200 ms) and late (200-3000 ms) parts of the RIR: S2R02. (Right) anechoic stimuli.

least 15 dB SNR, 1/3rd-octave bands with SNR < 20 dB were discarded (muted in the final RIR). This typically resulted in omitting the lowest two 1/3rd-octave bands: 100 and 125 Hz (these were therefore also omitted from the simulated RIRs). An equal power cross-fade between the measured and synthesized responses was applied over the last 10 dB decay of the measured and the first 10 dB decay of the synthesized response to provide a smooth transition, limiting audible artifacts. Fig. 3 depicts the spectral magnitudes of the simulated and measured RIRs on measurement position of source 2 combined with receiver 02 (S2R02) after this procedure.

3.2 Stimuli

The resulting measured and simulated RIRs were convolved with two anechoic audio extracts appropriate to the acoustic function of the room: female soprano singing *Abendempfindung*, by W.A. Mozart; male tenor performing *A Chloris,* by R. Hahn. Fig. 3 depicts the spectral composition of the chosen extracts. As the two lowest 1/3rd-octave bands were omitted from the RIRs, the extracts were chosen to have minimal energy in these bands. RMS of the measured and simulated convolutions was used for normalization.

3.3 Test protocol

The test was set up as an AB comparison. Stimuli were compared for both omnidirectional and binaural receiver configurations for two source and receiver positions (Omni: S1R02, S1R12, S2R02, S2R12; Binaural: S1R03, S1R13, S2R03, S2R13) resulting in 16 tested pairs. Four configurations were repeated to monitor the repeatability of responses, resulting in 20 pairs. Additionally, participants were given three training pairs to ensure they understood the task. Results for these three training pairs were not analyzed.

Participants were asked to rate the similarity of samples according to *Reverberance*, *Clarity, Distance to the source*, *Coloration,* and *Plausibility*. For binaural receiver pairs, participants were asked to additionally rate the similarity of *Apparent Source Width (ASW)* and *Listener Envelopment (LEV)*. It should be noted that the binaural head orientation in all configurations was towards S2. Participants responded using a continuous graphic 100 pt scale, ranging from ‘A is much more ...’ to ‘B is much more ...’ corresponding respectively to values of –50 and +50, with a center 0 response indicating no perceived difference. Presentation order and AB correspondence to simulation and measurement were randomized. Participants were able to listen to the compared pairs as many times as desired. Auralizations were presented via headphones (Sennheiser model HD 600) at an RMS level of 75 dBA. The experiment took place in an isolation booth, ambient noise level <30 dBA. The 12 participants (mean age: 39.6 SD: 16.7) all reported normal hearing.

3.4 Test results

Initial attention is given to the repeatability of responses, determined from the absolute difference between the 4 repeated configurations (see Fig. 4a). The mean difference between repetitions over attributes was 9.7 pt. Individual attribute repeatability mean values were used as tolerance ranges to estimate whether a subjective acoustic attribute differed between measurement and simulation. Overall results were then compared (see Fig. 4b). Considering repetition tolerance, measured binaural auralizations were judged slightly brighter and the mean slightly *ASW* wider than simulated RIR auralizations. Results for the remaining attributes were near 0, within repetition tolerances. Comparing omnidirectional and binaural configurations, a one-way ANOVA indicated significant differences for *Clarity*, *Coloration,* and *Plausibility* (*Reverberance*: F=0.02, p= 0.88; *Clarity*: F=8.80, $p<10^{-2}$; *Distance*: F: 0.03, p=0.87; *Coloration*: F=52.00, $p<10^{-2}$; *Plausibility*: F=7.42, p=0.01).

In an attempt to understand the foundations of the subjective responses, Table 1 presents objective parameter averages for measured and simulated RIRs at the auralized positions for EDT, C80, IACC(e), and IACC(l), which relate to *Reverberance*, *Clarity*, *ASW*, and *LEV* respectively.[6] Perceptual results agree with these objective parameters.

Subsequently, the results were analyzed per position. Figs. 4c and 4d show that the degree of variance is smaller for the omnidirectional receiver condition. The omnidirectional auralizations are centered on 0. The exception to this observation is S1R12 for attributes *Coloration*, *Clarity,* and *Distance* which could be explained by the difference in C80 (see Table 1). The binaural condition responses exhibit more variation while the majority of the attributes were within the repetition tolerances (~10 pts). Specifically, the measured RIR at S1R13 was perceived slightly more unclear, further away, brighter, and more enveloping. The measured RIR at S2R03 was found to be brighter and to have a wider *ASW*, and the measured S2R13 was judged to have a wider *ASW* than the simulated RIR counterparts. It should be noted that the exact positions between omnidirectional and binaural receiver conditions differed slightly (approx. 2 m). Taking this into consideration, significant differences in perceptual similarity judgements between receiver types were found for *Clarity* at S1R12/13 (F=5.91, p=0.02) and S2R02/03 (F=9.74, $p<10^{-2}$) as well as for *Coloration* at S1R02/03 (F=13.53, $p<10^{-2}$), S1R12/13 (F=27.18, $p<10^{-2}$), and S2R02/03 (F=22.55, $p<10^{-2}$).

In general, binaural auralizations for *Clarity* agreed slightly better than omnidirectional auralizations. This is due to the outlier position S1R12. Source and receiver for this position are near the highly ornate pulpit and several columns. It is possible that the scattering properties for these elements were erroneous, leading to excessive early reflections and subsequently a perceivable higher *Clarity*. *Coloration* similarity results agreed better for omnidirectional auralizations than binaural auralizations. A possible explanation is slight misalignments of the dummy head during the measurement relative to S2. Concerning *Plausibility*, both omnidirectional and binaural auralizations were judged equally plausible for measured and simulated responses, within the repetition

Table 1: Measured and simulated single number frequency average objective parameters (EDT, C80, IACC (e), IACC (l)) per position (Differences in **bold** are higher than 1 JND).

Position	EDT (s) – JND: 0.31 (500-1000Hz)			C80 (dB) – JND: 1.0 (500-1000Hz)			IACC(e) – JND: 0.075 (500-4000Hz)			IACC(l) – JND: 0.075 (500-4000Hz)		
	Meas	Sim.	Diff.	Meas	Sim.	Diff.	Meas	Sim.	Diff.	Meas	Sim.	Diff.
S01R02	5.76	5.56	+0.20	-6.9	-5.9	-1.0	-			-		
S01R12	5.27	5.12	+0.15	-1.2	1.7	**-2.9**	-			-		
S02R02	6.96	7.02	-0.06	-8.9	-8.6	-0.3	-			-		
S02R12	6.76	6.73	+0.03	-8.1	-7.8	-0.3	-			-		
S01R03	5.65	5.53	+0.11	-5.4	-6.5	**+1.1**	0.351	0.296	+0.055	0.089	0.304	**-0.215**
S01R13	5.04	4.55	**+0.49**	-0.0	1.4	**-1.4**	0.387	0.514	**-0.127**	0.093	0.308	**-0.216**
S02R03	7.11	7.20	-0.09	-9.5	-8.1	**-1.4**	0.463	0.730	**-0.267**	0.104	0.359	**-0.255**
S02R13	6.98	6.77	+0.21	-7.8	-6.1	**-1.7**	0.766	0.863	**-0.096**	0.093	0.381	**-0.288**

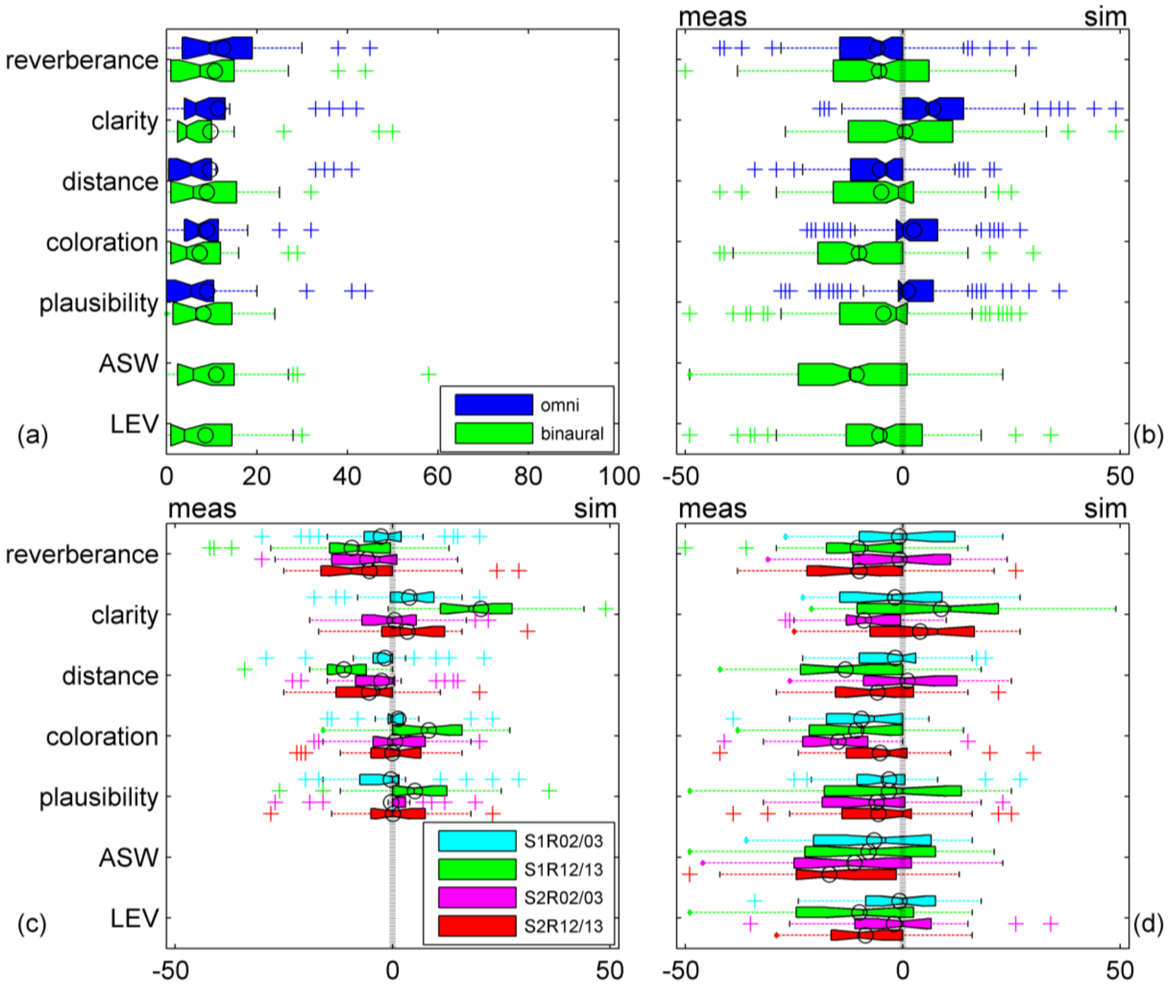

Fig 4: Perceptual results for all subjects on subjective similarity of measured and simulated RIR auralizations. (a) Absolute differences for repeated pairs. (b) Omnidirectional vs. binaural receivers. Measured on the left, Simulated on the right. Results by position for (c) omnidirectional and (d) binaural receivers. Box limits represent 25% and 75% quartiles, (+) outliers, (O) median, and (|) mean values.

tolerance. Finally, *ASW* was judged slightly wider for the measured binaural auralizations than for their simulated counterparts.

4 PRE-MODERN AND CURRENT SAINT-GERMAIN-DES-PRES

With the GA model validated, modifications were made to represent its pre-modern configuration. A 17th century plan, the only direct evidence of the state of the building during the Enlightenment, was used as a reference for geometrical modifications. Beside these geometrical changes, the use and position of liturgical adornments are different relative to today's minimal usage.

Fig. 5 depicts the pre-modern plan of the abbey church of Saint-Germain-des-Prés. The architecture changed in the 17th century, with the easternmost bay of the south nave aisle converted into a chapel dedicated to Saint Maur. Furthermore, the sanctuary, the high altar, and the center of the fifth bay of the nave were fully enclosed by screens, defining the principal liturgical zones of the pre-modern building. A final screen, positioned to separate the sanctuary from the choir, was probably acoustically and visually transparent.

Due to lack of information about adornments in Saint-Germain-des-Prés during the 16th and 17th century, the Cathedral of Notre Dame de Paris, for which the documentary evidence is relatively

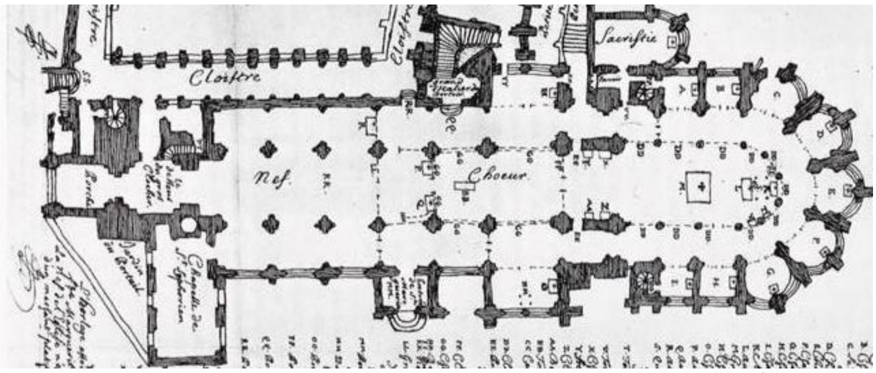

Fig.5: 17th century plan of the Saint-Germain-des-Prés.[8]

rich, was selected as a reference in order to make an ‘educated guess' as to which materials were used and where they were positioned.[9] In Notre Dame, screens also enclosed the sanctuary and high altar, presumably extending upward halfway the shaft of the sanctuary's columns and covered with draperies. Notre Dame was furnished for festivities especially in the area of the sanctuary by tapestries and the floor immediately in front of the altar was covered with rugs. Finally, it was assumed that the painted plaster was already present since this was typical for churches of that age. Details of the GA model material definitions can be found in Postma and Katz[3].

The principal performers and auditors at Saint-Germain-des-Prés were monks, situated in the principal areas of the high altar and sanctuary. Therefore, current and historical listening conditions are compared when a source is positioned in these areas and the receivers positioned inside and outside these principal areas. Fig. 6 shows that the pre-modern configuration had a perceptually shorter EDT and higher C50, an aspect which is more prominent in the principal areas.

Shorter reverberation times and higher clarity are associated with higher speech intelligibility. It is probable, therefore, that the performance of the liturgy in the choir of Saint-Germain-des-Prés changed in synchrony with the acoustics. For example, it would have been possible, with this increased clarity, to have performed the standard chant repertory at a higher tempo. The increased clarity of the performance space may have also encouraged the composition of new musical forms inspired by, and adapted to, the changed acoustical environment. Further research on the implications of these observations is the focus of ongoing studies.

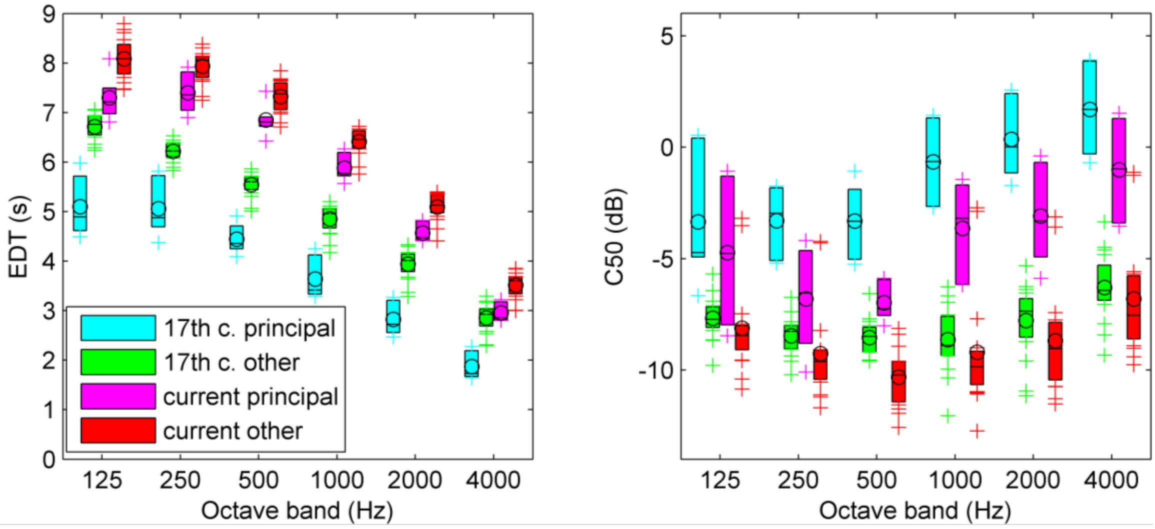

Fig. 6: Summary of EDT and C50 results for the 17th century and the current configuration GA model considering receivers inside the principal areas (S2-R17,19,22,24,26,27) and those at the other positions (S2-R1-2,4-7,9-12,14-16,20-21,25) (number of rays: auto (average = 127400, SD=0), length IR: auto (average=7850 ms, SD=74 ms)). (See legend of Fig. 4 for notations.)

5 CONCLUSION

A GA model of the abbey church Saint-Germain-des-Prés was calibrated according to a standardized method. The calibration was performed with the assumption that when reverberation and clarity parameters are within 1 JND of the measured value a valid model is created. To evaluate the validity of this assumption a listening test was carried out. Finally, the model was adjusted to represent its architecture during festivities in the Enlightenment era.

The listening test showed that using the methodical calibration procedure led to a perceptually valid GA model of the Saint-Germain-des-Prés, in addition to the objective measure validity based on reverberation and clarity metrics. Some trends on perceptual attribute differences between measured and simulated auralizations were found which slightly exceeded participant repeatability tolerances, specifically *Coloration* and *ASW* for binaural auralizations. Additional studies are necessary to understand the reason for these differences. As *Coloration* was judged sufficiently similar for omnidirectional auralizations, HRTF interpolation and processing employed to convert the measured data to CATT format should be investigation. Studies are currently underway to examine the suitability of the calibration method for other venue types, such as theatres.

Exploration of the estimated acoustical conditions of the pre-modern state of the abbey church of Saint-Germain-des-Prés as compared to the calibrated model, through analysis of the omnidirectional simulated RIRs indicate that it is reasonable to assume that the church had a shorter reverberation time and higher clarity in the principal liturgical areas, especially during the most important celebrations when greater quantities of sound absorptive materials were deployed. The space today is thus acoustically unlike that of the Middle Ages, and performance practices of monks who used the space daily were adapted to an acoustical space that favoured greater clarity.

Acknowledgments - This work was funded in part by the ECHO project (ANR-13-CULT-0004, echo-projet.limsi.fr).

6 REFERENCES

1. B.F.G. Katz and E. Wetherill, 'Fogg art museum lecture room, a calibrated recreation of the birthplace of room acoustics.' Proc. Forum Acusticum, 2191–2196, Budapest (2005).
2. A. Tallon, "L'espace acoustique de l'abbatiale de Saint-Germain-des-Prés" [The acoustic space of the abbey of Saint-Germain-des-Prés], in *Saint-Germain des Prés. Mille ans d'une abbaye à Paris*, ed. Matthieu Guyot, Paris: Académie des Inscriptions et Belles-Lettres, (forthcoming).
3. B.N.J. Postma and B.F.G. Katz, "Creation and calibration method of virtual acoustic models for historic auralizations," *Virtual Reality*, no. SI: Spatial Sound, pp. 1–20, 2015, doi:10.1007/s10055-015-0275-3.
4. B.F.G. Katz, 'International round robin on room acoustical response analysis software 2004.' J. Acoust. Soc. Am. 4:158-164, doi:10.1121/1.1758239 (2004).
5. G. Iannace, U. Berardi, and C. Ianniello, 'Study of a Historical Church Based on Acoustic Measurements and Computer Simulation.' Proc. 22nd ICSV, 1-8, Florence (2015).
6. ISO, Geneva, Switzerland: International Organization for Standardization: ISO 3382-1:2009(E). Measurement of the Reverberation Time of Rooms with reference to other Acoustical Parameters (2009).
7. B. Dalenbäck: CATT-A v9: User's Manual CATT-Acoustic v9. CATT, Gothenburg (Sweden): (2011).
8. Bibliothèque national de France: Plans manuscrits de l'église de Saint-Germain-des-Prés en 1644, "l'année de son rétablissement" [Plan documents of the abbey of Saint-Germain-des-Prés in 1644 "year of reinstatement"]. In: Annales de l'abbaye de Saint-Germain-des-Prés (555-1743), et pièces annexes jusqu'en 1753. MS. FR. 18816, FOL. 72, vol 265, Fondation de L'Abbey (1898).
9. C. Wright: Music and Ceremony at the Notre Dame of Paris. Cambridge University Press (2008).