

# Development of acoustic computer simulation for performance spaces: A systematic review and meta-analysis

Chao Wang<sup>1</sup>, Jian Kang<sup>1,2</sup> (✉)

1. School of Architecture, Tianjin University, Tianjin 300072, China

2. UCL Institute for Environmental Design and Engineering, The Bartlett, University College London (UCL), Central House, 14 Upper Woburn Place, London WC1H 0NN, UK

## Abstract

This article aims to review the development of acoustic computer simulation for performance spaces. The databases of Web of Science and Scopus were searched for peer-reviewed journal articles published in English between 1960 and 2021, using the keywords for "simulation", "acoustic", "performance space", "measure", and their synonyms. The inclusion criteria were as follows: (1) the searched article should be focused on the field of room acoustics (reviews were excluded); (2) a computer simulation algorithm should be used; (3) it should be clearly stated that the simulated object is a performance space; and (4) acoustic measurements should be used for comparison with the simulation. Finally, twenty studies were included. A standardised data extraction form was used to collect the modelling information, software/algorithm, indicators for comparison, and other information. The results revealed that the most used acoustic indicators were early decay time (EDT), reverberation time ( $T_{30}$ ), strength ( $G$ ), and definition ( $D_{50}$ ). The accuracy of these indicators differed greatly. For non-iterative simulation, the simulation accuracies of most indicators were outside their respective just noticeable differences. Although a larger sample size was required for further validation, simulations of  $T_{30}$ , EDT, and  $D_{50}$  all showed an increase in accuracy with increasing time from 1979 to 2020, except for  $G$ . In terms of frequency, the simulation was generally less accurate at lower frequencies, which occurred at  $T_{30}$ ,  $G$ ,  $D_{50}$  and  $T_{20}$ . However, EDT accuracy did not exhibit significant frequency sensitivity. The prediction accuracy of inter-aural cross-correlation coefficients (IACC) was even higher at low frequencies than it was at high frequencies. The average value of most indicators showed a clear systematic deviation from zero, providing hints for future algorithm improvements. Limitations and the risks of bias in this review were discussed. Finally, various types of benchmark tests were suggested for various comparison goals.

## 1 Introduction

Computer simulation is an effective method to predict the real complex world using mathematical models. It was first used on a large scale during World War II (Winsberg 2010). Then with the improvement of computer technology, it gradually entered households and almost every field in the industry.

Nowadays, computer simulation has been widely used in the prediction of building environments (Hong et al. 2000), including lighting (Jin et al. 2021), heating ventilation and

air-conditioning (HVAC) (Li and O'Neill 2018), airflow (Kong et al. 2015), acoustics etc. Most of them were developed in the 1960s and 1970s (Augenbroe 2002). Room acoustics computer simulation, which predicts sound propagation in buildings, was first applied in the construction of concert halls in 1968 (Krokstad et al. 1968).

Room acoustic computer simulation is widely used in building acoustics. It can predict the acoustic performance of buildings before construction, which is helpful to modify the acoustic design if necessary. It achieves a balance between economics and prediction accuracy and, therefore, is widely

## Keywords

acoustic simulation;  
accuracy;  
systematic review;  
performance space;  
iterative simulation

## Article History

Received: 10 January 2022

Revised: 05 March 2022

Accepted: 26 March 2022

© The Author(s) 2022

E-mail: j.kang@ucl.ac.uk

used. Room acoustic computer simulation is neither as simple as the Sabine formula, which ignores the influence of the shape and position of acoustic materials, nor as time-consuming and expensive as making a scale model (Kuttruff 2009; Rindel 2011).

In the last fifty years, many room acoustic simulation algorithms have been proposed. Some of them are called the geometrical acoustics (GA) model (Vorländer 2013), an algorithm that ignores the wave effect in sound propagation (Savioja and Svensson 2015). It was developed into image-source method (Gibbs and Jones 1972), ray-tracing method (Allred and Newhouse 1958; Schroeder 1970), beam-tracing method (Walsh and Rivard 1982), and surface-based method (Tsingos and Gascuel 1997). Based on the GA model, many commercial software packages have also been developed. The wave-based model (Vercammen 2013), in contrast to the GA model, is good at dealing with some specific problems, such as acoustic prediction in small rooms.

Various hybrid methods (Baines 1983; Lewers 1993) were also proposed to improve calculation efficiency, expand calculation frequency, or improve prediction accuracy. A hybrid method that combined the image and ray-tracing methods was used in many commercial software packages (Naylor 1993; Dalenbäck 1995). Combining the GA algorithm and wave method significantly increased the prediction frequency (Southern et al. 2013).

A series of round robins (Vorländer 1995; Bork 2000, 2005a, 2005b) were held to compare the performance of various simulation software. During each round robin, different phases were set up for further comparison in different situations. The accuracy of computer simulation, which was usually represented by the difference of acoustic indicators between the values of simulation and actual measurement, showed a significant variation in round robins and other researches. For example, In the 14 cases in Round Robin II phase I (Vorländer 1995), the  $T_{30}$  difference at 500 Hz varied from about  $-0.4$  to  $1$  s, whereas in the four cases in recent years (Bustamante et al. 2014; Alfano et al. 2015; Shtrepi et al. 2017), the variation ranged from  $-0.08$  to  $0.05$  s. A similar phenomenon was observed in the comparison of  $D_{50}$ . During the same phase, more than half of the cases reported that the  $D_{50}$  difference at 1000 Hz was located out of the just noticeable difference (JND, 0.05 for  $D_{50}$  based on ISO-3382). But in three recent cases (Bustamante et al. 2014; Shtrepi et al. 2017), all of them were located within the JND. Due to some limitations and deficiencies in previous comparisons, the exploration of the accuracy of the simulation has not stopped. Furthermore, the comparison goals established in these studies are different, resulting in different information being known prior to the simulation. This will also affect the evaluation of the simulation.

The development of room acoustics computer simulation has been reviewed from the perspective of principle in two review articles. Savioja and Svensson (2015) described the main principles and development of techniques based on geometric acoustic principles, especially their ability to model different aspects of sound propagation. Svensson and Kristiansen (2002) provided an overview of computer simulation techniques used in various auralization applications, focusing on the comparison of computational details and their potential impact on accuracy. The uncertainty of the simulation results has also been theoretically analysed by Vorländer (2013). However, all the three review articles described the development of computer simulation based on the difference in the algorithm, but there was still a lack of review focusing on the details about computer simulation used in actual applications, especially based on the simulated results.

Room acoustic simulation was applied to many scenes, such as classrooms (Zannin and Marcon 2007), offices (Etter 2001), historical buildings (Garrido et al. 2012; Alonso et al. 2017) and other spaces (Hodgson 1988; Kang 1996, 1997, 2002; Prawirasasra et al. 2016; Wu et al. 2018; Zhu et al. 2021). Even some simulation algorithms achieved a good prediction accuracy in semi-open spaces (Bo et al. 2018). However, performance space was the earliest application scene of acoustic computer simulation (Krokstad et al. 1968). A systematic review and meta-analysis will help understand the development of computer simulation of room acoustics. How the model was built, which simulation algorithm was used, and what level of accuracy was achieved in the acoustic prediction of the performance space have not been systematically analysed. This analysis will also provide help for the use and improvement of the simulation algorithm.

As a useful statistical tool to combine results of different studies, meta-analysis was considered useful to reduce biases and establish the evidence-based practice in many fields. It is widely used in medicine, astronomy, and other fields (Hedges 1992; Gurevitch et al. 2018). A systematic review and meta-analysis are conducted in this article to review the development of acoustic computer simulation for performance space. It will be analysed from two aspects: (1) research design, including compared indicators and their frequency ranges, simulation setting, and modelling and (2) accuracy, represented by the difference between the simulated and the measured value.

## 2 Method

### 2.1 Search strategy and eligibility criteria

Due to the exploratory nature of current work, there was currently a lack of standard protocols for the systematic

review of computer simulation for performance spaces. Since this was not a standard medical study (Gumaa and Rehan Youssef 2019), the protocol was not registered in PROSPERO or other similar databases. In order to clarify the research object, the performance space in this research was limited to the enclosed spaces related to music performance, such as auditorium, concert hall, opera house, theatre etc.

The following search criteria were used for selecting the articles for review: (1) the searched article should be focused on the field of room acoustics (reviews were excluded); (2) a computer simulation algorithm should be used; (3) it should be clearly stated that the simulated object is a performance space; and (4) acoustic measurements should be used for comparison with the simulation. Only English-language journal articles indexed by Web of Science (WOS) and Scopus were considered. Since the application of room acoustic computer simulation started in the 1960s (Krokstad et al. 1968), the search time was selected from 1960 to 2021. Since the core library of WOS only supports the search after 1985, the actual search period in this library is 1985 to 2021.

This article focused only on the performance space, therefore, the search string chosen was “(TS = ((simulation OR model OR modelling OR modelling OR prediction OR predictive OR computer OR computational ) AND (sound OR acoustic OR acoustics) AND (room OR auditorium OR “concert hall” OR “opera house” OR theatre OR theater OR “performance space”) AND (test OR measure OR valid OR validation) )) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article); Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, ESCI, CCR-EXPANDED, IC; Timespan=All years” in WOS, and “TITLE-ABS-

KEY((simulation OR computer OR computational) AND (sound OR acoustic OR acoustics) AND (room OR auditorium OR {concert hall} OR {opera house} OR theatre OR theater OR {performance space} ) AND (test OR measure OR valid OR validation)) AND DOCTYPE (ar)” in Scopus.

Many studies on auralization (Choi and Fricke 2006; Wang and Vigeant 2008; Pätynen and Lokki 2011) focused on subjective evaluation, which was difficult to compare numerically. So, they were not included in this study. The process of article selection, screening, and exclusion of the systematic review is shown in Figure 1.

The qualification evaluation of this study was independently conducted by two reviewers in a non-blind standardised manner. A few disagreements among reviewers about the inclusion/exclusion of specific items were resolved through consensus.

## 2.2 Data extraction

A standardised data extraction form was used to collect the modelling information (volume, the number of faces which abbreviated as faces,  $F/V$  which is the ratio of the number of faces to the volume), software/algorithm, compared indicators, total sources, total source-receivers, iteration which means whether an iterative simulation was carried out and whether the simulated results were averaged in multiple octave bands. The results were shown in Table 1.

The acoustic indicators were extracted from the simulated results. Since all the acoustic indicator results in the selected articles were expressed in octave bands, all the data in 125 Hz to 4000 Hz octave bands were extracted and analysed. The

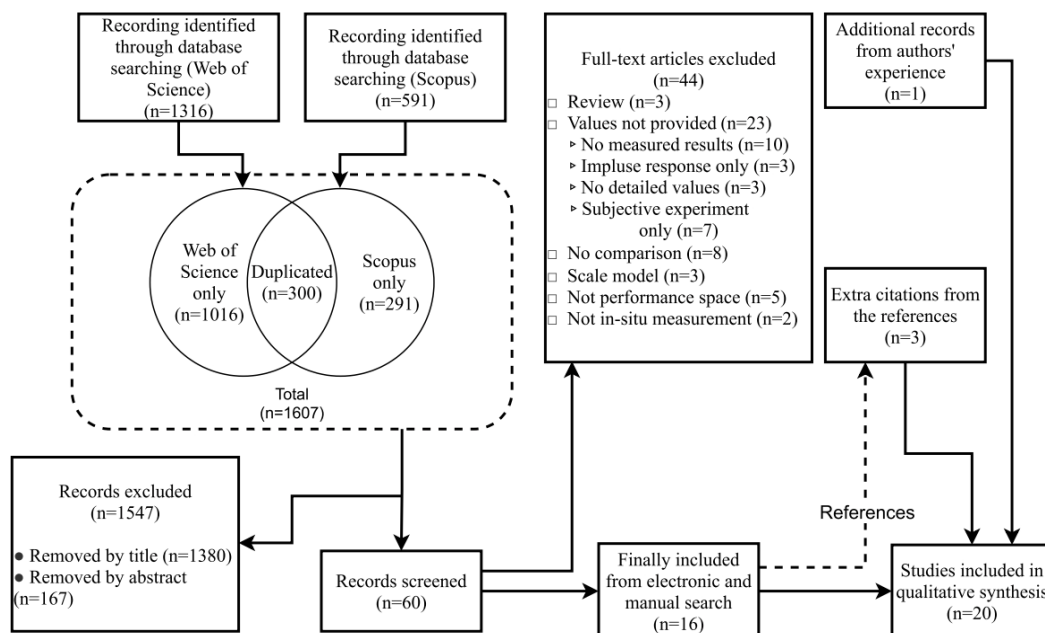


Fig. 1 A flow chart showing article selection, screening, and exclusion in this systematic review

**Table 1** Characteristics of included studies

No.	Study	Volume (m <sup>3</sup> )	Faces	F/V	Volume per seat (m <sup>3</sup> /seat)	Software/algorithm	Compared indicators	Total sources	Total source-receivers	Iteration	Frequency-averaged	
1	D'Orazio et al. 2020	9900	3800	0.3838	7.4	Odeon v12	EDT; C <sub>80</sub>	3	9	Yes	Yes	
2	Ibarra-Zarate and Padilla-Ortiz 2020	1304	\	\	\	EASE 4.3	T <sub>30</sub> ; D <sub>50</sub>	1	3	No	No	
3	D'Orazio et al. 2018	15300	591	0.0386	8.1	Odeon v12	EDT; C <sub>80</sub>	4	240	Yes	Yes	
4	Shtrepi et al. 2017	5270	\	\	15.1	Odeon v13	EDT; T <sub>30</sub> ; C <sub>80</sub> ; D <sub>50</sub>	1	24	Yes	No	
5	Postma and Katz 2016	#1	2500	1300	0.52	\	CATT v9.0	EDT; D <sub>50</sub>	2	54	Yes	Yes
		#2	84000	14700	0.175	\		EDT; D <sub>50</sub>	2	60	Yes	Yes
		#3	22200	2200	0.0991	\		EDT; D <sub>50</sub>	2	48	Yes	Yes
6	Alfano et al. 2015	#1	\	\	\	Odeon	T <sub>30</sub>	1	12	Yes	No	
		#2	\	\	\		T <sub>30</sub>	2	24	Yes	No	
7	Hornikx et al. 2015a	#1	8400	\	\	CATT v9.0; PSTD	T <sub>20</sub> ; G	3	102	No	No	
		#2	10000	\	\	CATT v9.0; PSTD	T <sub>20</sub> ; G	3	111	No	No	
8	Alonso et al. 2014	205524	\	\	\	CATT v9.0	EDT; T <sub>30</sub> ; G; C <sub>80</sub> ; T <sub>5</sub> ; J <sub>LF</sub>	1	9	Yes	No	
9	San Martín et al. 2014	4000	105	0.0263	10.7	Odeon v12	EDT; T <sub>30</sub> ; C <sub>80</sub>	1	375	Yes	No	
10	Bustamante et al. 2014	20571	793	0.0385	11.3	CATT v9.0	EDT; C <sub>80</sub> ; T <sub>5</sub> ; J <sub>LF</sub>	2	28	Yes	No	
11	Koutsouris et al. 2013	11000	\	\	\	CARISM	EDT; T <sub>30</sub> ; G; C <sub>80</sub> ; D <sub>50</sub> ; T <sub>5</sub>	2	6	No	Yes	
12	Picaut and Fortin 2012	924	\	\	\	SPPS	SPL	1	9	No	No	
13	San Martín and Arana 2006	#1	20000	156	0.0078	12.9	Unknown commercial software #1	EDT; T <sub>30</sub> ; SPL; D <sub>50</sub> ; T <sub>5</sub>	1	9	No	No
		#2	20000	156	0.0078	12.9	Unknown commercial software #2	EDT; T <sub>30</sub> ; SPL; D <sub>50</sub> ; T <sub>5</sub>	1	9	No	No
14	Bork 2000	#1-#14	11000	\	\	10	T <sub>30</sub> ; D <sub>50</sub>	2	6	No	No	
15	Gomes et al. 2000	\	\	\	\	Raynoise	EDT; C <sub>80</sub> ; D <sub>50</sub>	1	4	Yes	No	
16	Howarth and Lam 2000	\	\	\	\	Odeon v2.5	EDT; T <sub>30</sub> ; G; C <sub>80</sub> ; J <sub>LF</sub>	1	88	Yes	No	
17	Lam 1996	#1	10500	85	0.0081	\	Odeon v2.5	SPL; C <sub>80</sub> ; J <sub>LF</sub>	1	12	No	No
		#2	10500	85	0.0081	\	SU-model	SPL; C <sub>80</sub> ; J <sub>LF</sub>	1	12	No	No
18	Vorländer 1995	#1-#16	1800	\	\	\	EDT; T <sub>30</sub> ; G; C <sub>80</sub> ; J <sub>LF</sub>	2	6	No	No	
19	Strøm et al. 1985	20000	\	\	13.3	Ray-tracing	RT	1	2	No	No	
20	Marshall 1979	20500	\	\	7.7	Image method	T <sub>15</sub>	1	12	No	No	

values averaged over multiple octave bands were regarded as the result in each octave band.

If a model was simulated by multiple software packages, it was treated as different cases. And if a model was modified and simulated again, which usually occurred in the study of historical buildings, only the initial state was considered. But if the settings of software were modified and simulated again, only the state with higher accuracy was selected and was regarded as a sort of “iterative simulation”.

The calculation time was not included in the analysis because the computing power of the equipment used in each case was usually different, and the calculation time could not represent the computing power of the algorithm.

### 2.3 Quality assessment

The study on acoustic simulation accuracy was different

from the common meta-analysis in the medical field. The variation in room size and shape in each study resulted in a significant difference in the research object. The computer simulation was also very different from medical intervention during operation, making it difficult to evaluate the research quality with the medical paradigm. So, some items in the Evaluation Guidelines for Rating the Quality of an Intervention Study scoring system (MacDermid 2004) were modified to better reflect the accuracy of room acoustic computer simulation. The results are shown in Table A1 in Appendix A, which is available in the Electronic Supplementary Material (ESM) from the online version of this paper.

The article score assessment method (MacDermid 2004) was used in this review. Articles were classified as high quality (HQ, 36–48), medium quality (MQ, 25–35), low quality (LQ, 0–24). 10% of eligible studies ( $n = 2$ ) were high quality,

and the highest quality score was 39/48. 80% ( $n = 16$ ) were medium quality, while 10% were low quality. The lowest quality score was 23/48. It should be noted that these modified evaluation guidelines were only an attempt to evaluate the research quality in this article, and their universality has not been validated by additional research. Furthermore, the evaluation score only represented the evaluation of the article in the dimension of the research question in this review and did not completely reflect the research quality of the article itself. For example, iterative calibration of simulation settings based on test results would result in low scores in this review, but it was a standard step for computer simulation calibration in auralization research for historical buildings.

In ten studies, it was clearly stated in the article that the measured results were known by the operator before the simulation. In seven studies, it was not stated whether it was known in advance. Only three studies clearly stated that the measured results were unknown before the simulation. In 13 studies, the results of simulation and measurement were compared in only one space. In 8 studies, the data of all octave bands from 125 Hz to 4000 Hz were provided. Few statistical analysis was conducted in most studies, and nine studies did not conduct any further statistical analysis after obtaining the difference between the simulated and measured values.

### 3 Results

#### 3.1 Research design of the selected articles

##### 3.1.1 Research type

The simulation results in the twenty selected articles were compared with the measured results, as described in the Method section. However, the purpose of these studies was not always clear. Three types of research were found in the selected articles. *Type 1* was the first to be used (Marshall 1979). It did not affect the modelling process, the selection of acoustic parameters such as sound absorption coefficient, and the settings of the simulation program, which fully respected the user's experience. The simulation operator also did not use the measured results to make any corrections to the simulation process. This type was represented by Twenty cases in nine selected articles in this review, as shown in Table 2.

As the goal of *Type 2* was to compare the performance of the simulation software or algorithm, multiple software packages were used, and some parameters in the simulation process were controlled. The model with the same shape and size was used in the simulation modelling process. Only two articles were classified as *Type 2*. However, considering

**Table 2** Three types of research articles and the total number of cases included

	Description	Article	Case
<i>Type 1</i>	Simulation process was not controlled. All the modelling and parameter settings were determined by the operator without any interference	9	20
<i>Type 2</i>	Some parameters in the simulation process were controlled	2*	21
<i>Type 3</i>	The simulation settings were modified according to the simulation results	10	13

\*As Bork (2000) was included in both *Type 1* and *Type 2*, the total number of articles was 21.

that each of the Round Robin covered 10+ cases and the external conditions of the simulation were well controlled uniformly in the comparison, the two articles were still analysed as a separate type. Twenty-one cases in two selected articles in this review belong to this type. However, the controlled parameters in the modelling process also discarded the experience of simulation operators, such as the processing of fine structures, which had a great impact on the accuracy of the simulation results (Vorländer 2013). In some studies not included in this review, even the choice of sound absorption coefficient and the scattering coefficient had been uniformly regulated, such as in Phase 2 in Round Robin II (Bork 2000) and Round Robin III (Bork 2005a, b). These studies were considered by the author to be too pursuing the comparison of software performance, which were too specific to be included in this review.

In *Type 3*, the simulation operator already knew the values of some measured indicators during the simulation process, so that the simulation settings were modified according to the simulation results. Usually, this kind of correction was carried out many times, which was called "iterative" calibration (Alonso et al. 2014), and finally making most of the simulated indicators within JND. 13 cases in ten selected articles in this review belong to this type. The simulation results of *Type 3* could be expected to be much better than that of the first two types because the calibration of some indicators made the simulation setting more conducive to close to the real situation. Moreover, the correlations among many indicators also provided help for the calibration process.

The number of cases of each type and the year of occurrence are shown in Figure 2. The results show that *Type 1* appeared first, and the related research has continued to this day. *Type 2* consisted of two round robins. This tended to be an organised baseline test, in which the performance of more than ten software/algorithms was compared. However, this kind of round-robin hasn't been carried out for a long time. Some researches gradually shifted to baseline testing of acoustic simulation software (Brinkmann et al. 2021;



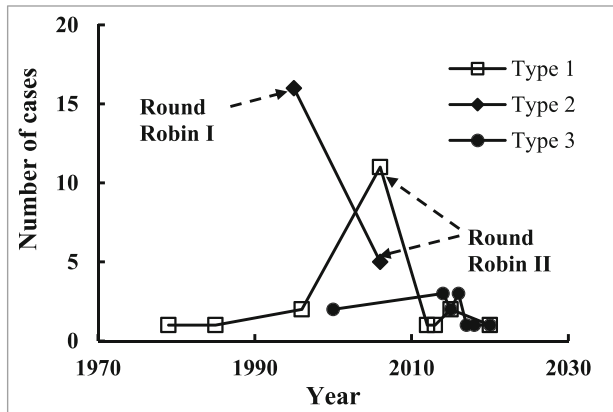


Fig. 2 The number of cases of each type and the year of occurrence

Hornikx et al. 2015b), and some studies focus on the comparison of the perceptual evaluation of the simulation results (Brinkmann et al. 2019). *Type 3* has been in a tepid state of research since about 2000. This type of acoustic simulation was only applicable when some of the test results were known, which obviously limited the scope of its application.

### 3.1.2 Compared indicators and their frequency range

In the selected 20 articles, most of the comparisons between the simulation and the actual measurement were based on the eight acoustic indicators listed in ISO-3382, namely early decay time (EDT), reverberation time ( $T_{30}$ ), strength (G), clarity ( $C_{80}$ ), definition ( $D_{50}$ ), centre time ( $T_s$ ), early lateral energy fraction ( $J_{LF}$ ), and inter-aural cross-correlation coefficients (IACC). These indicators provided a description of the acoustical characteristics of a performance space. However, not all cases compared all the eight indicators on all the six octave bands. The compared indicators and octave bands in all the cases in this review are shown in Figure 3. The results showed that the main indicators selected were EDT,  $T_{30}$ , G, and  $D_{50}$ . These four indicators were selected by more than ten articles. The comparison results will be described in detail in the subsequent analysis in Section 3.2. As a few articles choose  $C_{80}$ ,  $T_s$ ,  $J_{LF}$ , and IACC, the comparison results will be briefly analysed.  $T_{20}$  was selected by several articles and will be discussed briefly in Section 3.2.

The results of simulation and measurement in all cases were recorded in octave bands. The mid-frequency was the most selected octave band, as shown in Figure 3. If frequency-averaging was taken over several octave bands to obtain a single value, many indicators were suggested at mid-frequencies, following the recommendation of ISO-3382. Frequency averaging was conducted in six of the 20 selected articles. Furthermore,  $T_{30}$  and  $D_{50}$ , which were the most selected indicators, were compared in almost all octave bands.

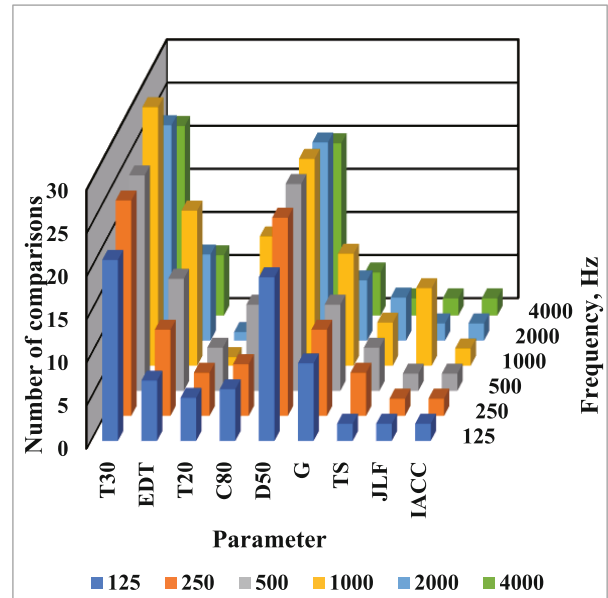


Fig. 3 The number of comparisons of each indicator in the six octave bands in the selected articles

### 3.1.3 Simulation setting

The simulation algorithm of all 54 cases in the review article has been counted and classified, and the results are shown in Table 3. The results showed that although many new simulation algorithms had been developed, the classical GA model remained the primary application in the actual application in the performance space. The simulation algorithm of many software packages was unknown because many comparisons had anonymised the names of software or algorithms, such as Round Robin I and Round Robin II. Another reason for the uncertainty of the simulation algorithm was that some software supported multiple algorithms, but it was not clear which one was used in the article. As can be seen in Table 3, all commercial software used the GA model, and among them, the hybrid model was mainly used. This may be due to its ability to save a lot of computing resources under certain accuracy. The research developments did not show an obvious tendency in algorithm selection because of the anonymity of Round Robin I and Round Robin II.

Because of the difference in simulation algorithms, their settings were difficult to compare, such as the transition order between the image method and the ray-tracing method in the hybrid model. The transition order was set according to the user guide (D'Orazio et al. 2018) or the operator's experience in some cases (Postma and Katz 2016; Shtrepi et al. 2017), and the effect of different values was discussed in some other cases (Lam 1996; Howarth and Lam 2000). Considering the differences in the implementation of algorithms, the reasonable values of the settings varied with different software and cases. In this situation, the

**Table 3** Simulation algorithm used in the cases involved in this review

Algorithm source	Algorithm type		GA model				Wave-based model
	Image	Ray-tracing	Beam-tracing	Surface*	Hybrid	Uncertain	
Commercial software	0	7	4	0	12	7	0
Research development	1	1	0	1	0	20	1

\* Represent surface-based model, such as radiosity model.

operator’s experience had a great influence on the simulation accuracy.

The number of rays was a critical setting in the ray-tracing method. Some software will automatically calculate a recommended value with different degrees of accuracy (Naylor 1993). In the selected cases in the review, the number of rays varied greatly from 5000 (Vorländer 1995) to 2,500,000 (Bustamante et al. 2014), but the accuracy of the simulation results was not reported to be obviously affected. Furthermore, with the rapid development of computer performance, setting a very large number of sound rays required a little calculation time. This value has been rarely discussed in recent cases.

### 3.1.4 The precision level of the modelling

The precision level of modelling could be represented by the ratio of face number to volume ( $F/V$ ). Among the cases in the review, 19 cases gave the volume value, and 9 cases gave the face number of the GA model at the same time. The changes in volume and  $F/V$  with time are shown in Figure 4. The results showed that the volume of the simulation model had not changed significantly in the past 30 years (1985 to 2020), mainly in the range of 1,000 to 100,000. The most common modelled volume is about 10,000 cubic meters. Although  $F/V$  fluctuated greatly with time, it had an obvious growth trend, which indicated that the simulated model was becoming more and more refined. There were two reasons that could not be ignored. One was the development of modelling algorithm. Acoustic software supported the import of more complex models, which made it possible to make it more complex. The second was the improvement

of computer power. Even if the model was complex, the calculation time was also acceptable, so that there was no need to simplify the model in order to realise the calculation.

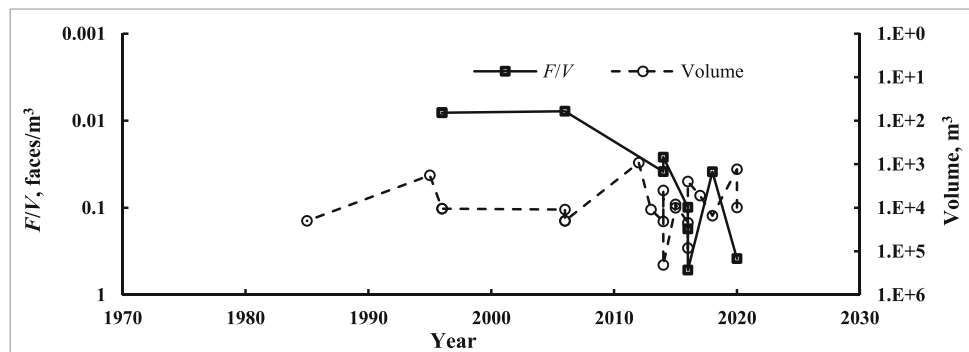
## 3.2 Accuracy

The accuracy was expressed by the difference between the simulated and the measured value (simulated – measured). As mentioned before,  $T_{30}$ , EDT,  $G$ , and  $D_{50}$  were analysed in detail, and  $C_{80}$ ,  $T_s$ ,  $J_{LF}$ , IACC, and  $T_{20}$  were analysed briefly.

### 3.2.1 Reverberation time ( $T_{30}$ )

As one of the most important acoustic indicators, reverberation time is often used in the design of performance space. There were 30 cases in 12 articles that included a comparison of simulated and measured reverberation time ( $T_{30}$ ). The results are shown in Figure 5 and Figure 6. As can be seen in Figure 5, there was a significant difference between the iterative and non-iterative simulated results. However, there was no obvious quantitative difference between the comparison cases of reverberation time in each octave band, and the majority of cases were from Round Robin II. In recent years, the focus has been on the verification of the results of iterative simulations. Only a few cases were still discussing the accuracy of non-iterative simulation results of reverberation time. However, the prediction of  $T_{30}$  still appeared to show a trend of increasing accuracy over time.

A more in-depth analysis was conducted. The differences in all cases were averaged to explore the systematic deviation.



**Fig. 4** The changes in volume and  $F/V$  of the simulated models with time (1985 to 2020)

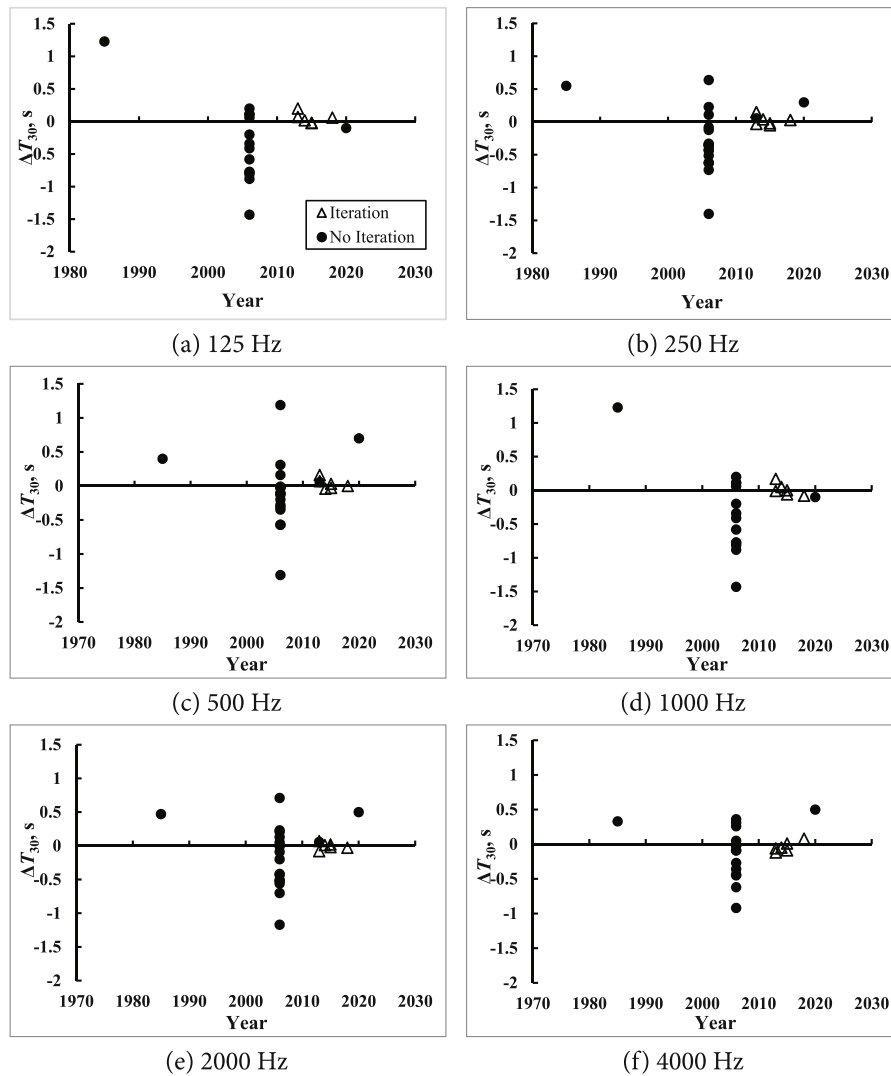


Fig. 5 Prediction difference (simulated – measured) in  $T_{30}$ . Each point represents a case in the review

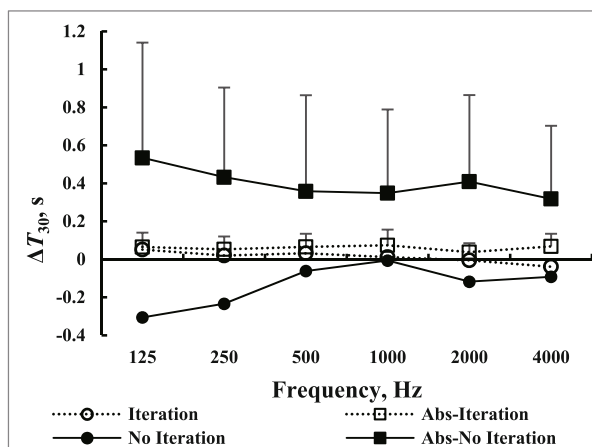


Fig. 6 The average value of  $T_{30}$  prediction differences of all cases. Each point is the average of all the iterative or non-iterative cases. “Iteration” means that the simulated results have been iteratively calibrated. “Abs” means that the differences in all cases were taken as absolute values before averaging

Theoretically, it should be close to zero. The differences in all cases were also taken as absolute values before averaging to evaluate the average level of prediction accuracy. The results are shown in Figure 6. For the “Abs” results, the  $\Delta T_{30}$  without iteration tended to decrease as the frequency increased, indicating that the simulation accuracy increased with the growth of frequency. The error bars in Figure 6 were the standard deviation of the results of all cases. It can be seen that the standard deviation of the simulated difference in the low frequencies was larger than that in the high frequencies, which indicated that the simulation stability grew with the increasing frequency. In most octave bands, there was an order of magnitude difference between the iterative and the non-iterative simulation results.

The average value of  $T_{30}$  prediction differences of all cases was theoretically close to zero, but it can be seen from Figure 6 that the averaged values without iteration at most octave bands were negative, indicating that the simulated



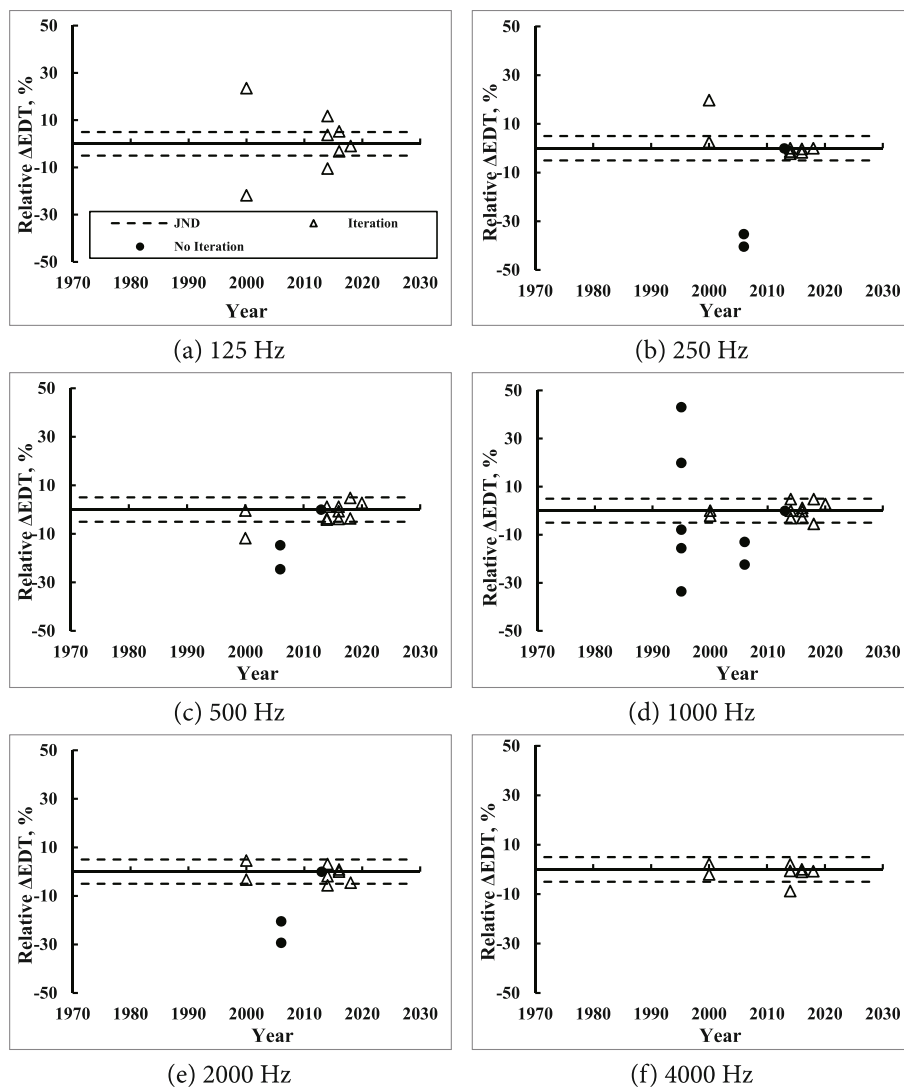
values were generally smaller than the measured ones. This provided clues for the correction of the simulation process in the future.

### 3.2.2 Early decay time (EDT)

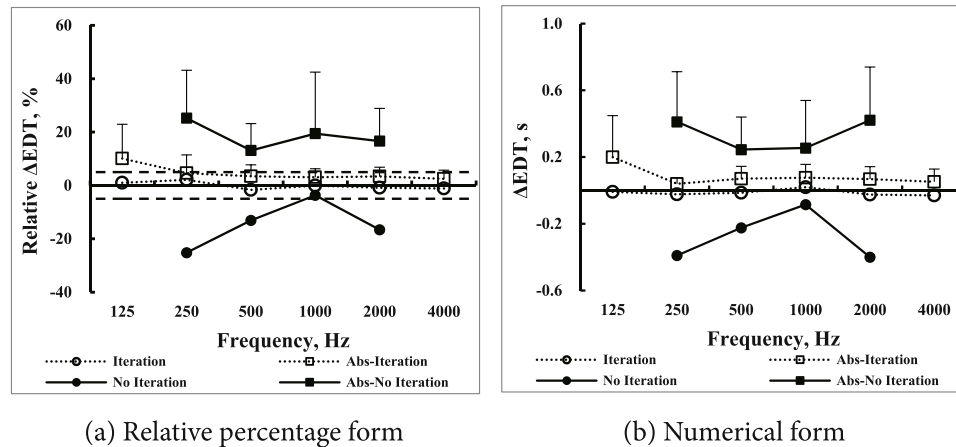
EDT was considered to have a significant relationship with subjective reverberation (Gade 1994). There were 19 cases in 12 articles that included a comparison of simulated and measured EDT. The results are shown in Figure 7 and Figure 8. As the JND of EDT was given as a relative 5% in ISO-3382, all relative  $\Delta$ EDT (relative prediction difference in EDT) values in Figure 7 were presented as percentages. The results showed that the relative  $\Delta$ EDT without iterative simulation was still significantly larger than that after iterative simulation. Most of the relative  $\Delta$ EDT without iterative simulation were outside of JND, and most of the results with iterative simulation were within JND.

There were no significant frequency differences in the relative  $\Delta$ EDT with iterative simulation. The iterative simulation produced more discrete results in the lower octave bands, and some of the data were beyond the JND. This could be because the sound performance at low frequencies was more difficult to predict accurately. Even after the iterative process, it was difficult to achieve all simulated results in all frequency bands within JND. This is most likely because EDT was rarely used as a unique iterative target. The simulation accuracy of EDT has improved significantly over time, whether it is iterative or non-iterative. However, in recent years, the sample size has been relatively small, and more evidence is required for validation.

The averaged simulation difference has been further analysed in both percentage and numerical form. The percentage form is shown in Figure 8(a). The results showed that for the “Abs” results, although the relative  $\Delta$ EDT without



**Fig. 7** Prediction difference (simulated – measured) in EDT. Each point represents a case in the review. The straight dashed line represents the JND of EDT



(a) Relative percentage form

(b) Numerical form

**Fig. 8** The average value of EDT prediction differences of all cases. Each point is the average of all the iterative or non-iterative cases. “Iteration” means that the simulated results have been iteratively calibrated. “Abs” means that the differences in all cases were taken as absolute values before averaging. The straight dashed line represents the JND of EDT

iteration was significantly higher than that with iteration, it was not as large as the  $T_{30}$  difference.

The results in the numerical form are shown in Figure 8(b). Neither the degree of data dispersion nor the values showed a clear trend in frequency. This could be because of the small amount of data in EDT. Because of the short decay time calculated for EDT (0–10 dB decay), it has higher volatility in both measured and simulated results. This could also explain frequency insensitivity. The averaged values without iteration at all octave bands were also negative, indicating that the simulated EDT results were generally smaller than the measured ones.

### 3.2.3 Strength ( $G$ )

$G$  was considered as a suitable normalisation for sound pressure level in an enclosure (Kuttruff 2009). There were 18 cases in 9 articles that included a comparison of simulated and measured  $G$ . The results were shown in Figure 9 and Figure 10. It is worth noting that there was no discernible difference between the  $\Delta G$  with and without iterative simulation, and many of them were within JND. The largest difference, however, appeared at 1000 Hz in the study of *Type 2*, namely the Round Robin II. With the growth of time, the simulation accuracy of  $G$  has not improved significantly except at 1000 Hz.

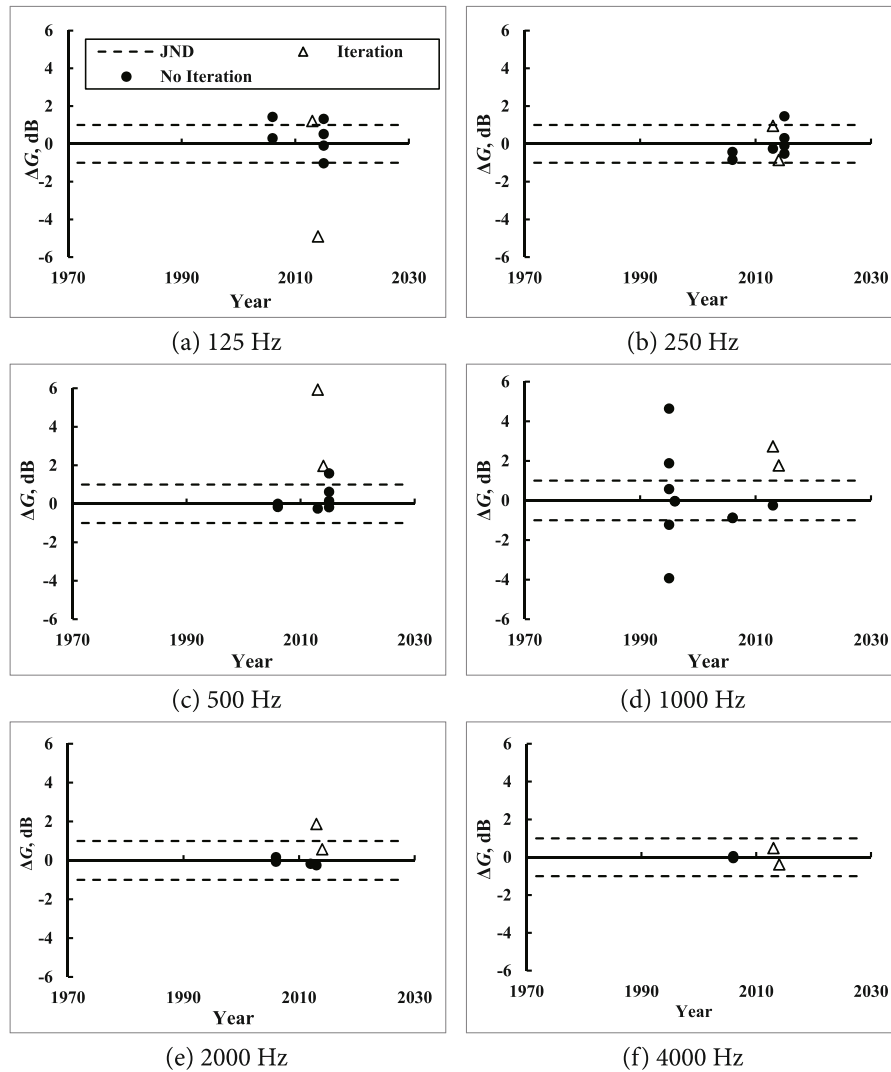
The averaged simulation difference has been further analysed, and the results are shown in Figure 10. The most intriguing phenomenon was that in some octave bands,  $\Delta G$  with iterative simulation was significantly higher than that without iterative, and in some octave bands, values of these two were similar. This contradicts the widely held belief that iteration would bring the simulation results closer to the real situation. One explanation could be that  $G$  was not usually the target of an iteration. The iterations of other

indicators produced overfitting, making the results of  $G$  more discrete and far from the real situation. The  $\Delta G$  with iterative simulation showed a downward trend from low to high frequency, with a trough at 250 Hz, which is very interesting and worthy of further investigation. The error bar shows that the low-frequency data fluctuated greatly, whereas the high-frequency data fluctuated very little.  $G$  demonstrated exceptional accuracy and stability at high frequencies.

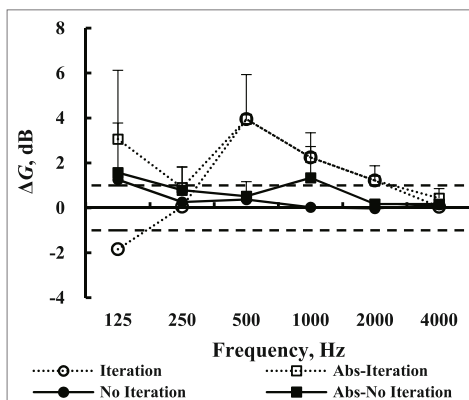
### 3.2.4 Definition ( $D_{50}$ )

There were 24 cases in 8 articles that included the comparison of simulated and measured  $D_{50}$ . The results are shown in Figure 11 and Figure 12. The frequency difference was not obvious, and the point trend was similar in each octave band. Due to the addition of the cases in Round Robin II, the difference between  $\Delta D_{50}$  with and without iterative simulation was very large. Most of the  $D_{50}$  with iterative simulation was within JND, which was also related to the fact that  $D_{50}$  was often used as an iterative target. Two different levels of accuracy appeared in the cases without iterative simulation. One was in the cases of *Type 2*, namely in the Round Robin II, where  $\Delta D_{50}$  was widely distributed within 0.3–0.5. The other cases without iterative simulation showed high accuracy, which was even similar to that with iterative simulation. Benchmarking was needed for further exploration. It can still be seen that as time grew, the simulation accuracy of  $D_{50}$  had improved significantly. But more evidence was needed to confirm.

Further analysis was conducted, and the results are shown in Figure 12. The  $\Delta D_{50}$  without iterative simulation had a downward trend from low to high frequency. The averaged values without iteration at all octave bands were close to zero, which implied a relatively small systematic



**Fig. 9** Prediction difference (simulated – measured) in *G*. Each point represents a case in the review. The straight dashed line represented the JND of *G*



**Fig. 10** The average value of *G* prediction differences of all cases. Each point is the average of all the iterative or non-iterative cases. “Iteration” means that the simulated results have been iteratively calibrated. “Abs” means that the differences in all cases were taken as absolute values before averaging. The straight dashed line represented the JND of *G*

error between the simulated and measured values. The error bars showed similar levels of data fluctuations at each frequency.

### 3.2.5 $C_{80}$ , $T_s$ , $J_{LF}$ , IACC, and $T_{20}$

The remaining five indicators,  $C_{80}$  (clarity),  $T_s$  (centre time),  $J_{LF}$  (early lateral energy fraction), IACC (inter-aural cross-correlation coefficients), and  $T_{20}$ , were compared in fewer cases. So the conclusions in this section need further evidence to support. A brief analysis of these indicators was carried out as below.

There were 15 cases in 10 articles that included the comparison of simulated and measured  $C_{80}$ . The results are shown in Figure 13(a). Most of the cases for  $C_{80}$  used iterative simulation, but most of the differences were still outside of JND. The non-iterative simulations were only conducted in the mid-frequency (500–1000 Hz) and obtained higher

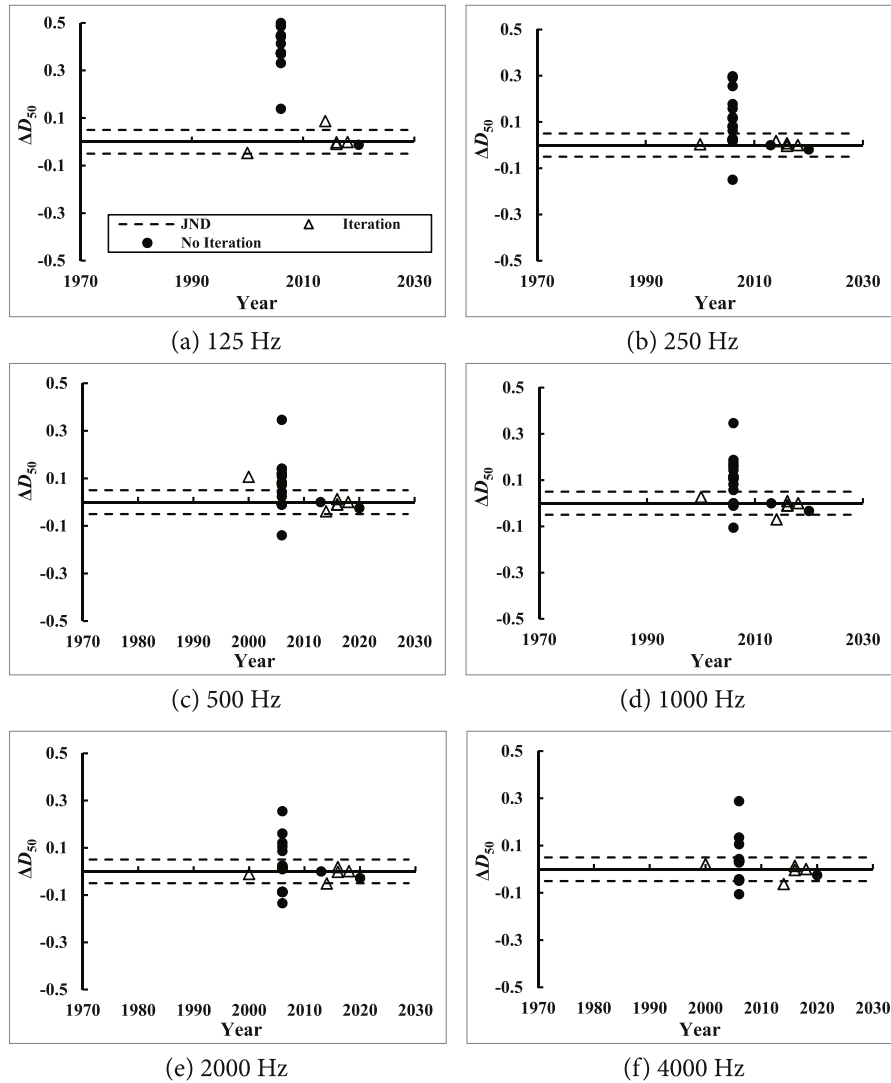


Fig. 11 Prediction difference (simulated – measured) in  $D_{50}$ . Each point represents a case in the review

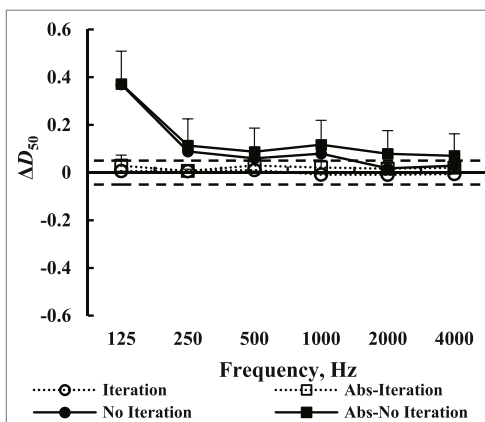
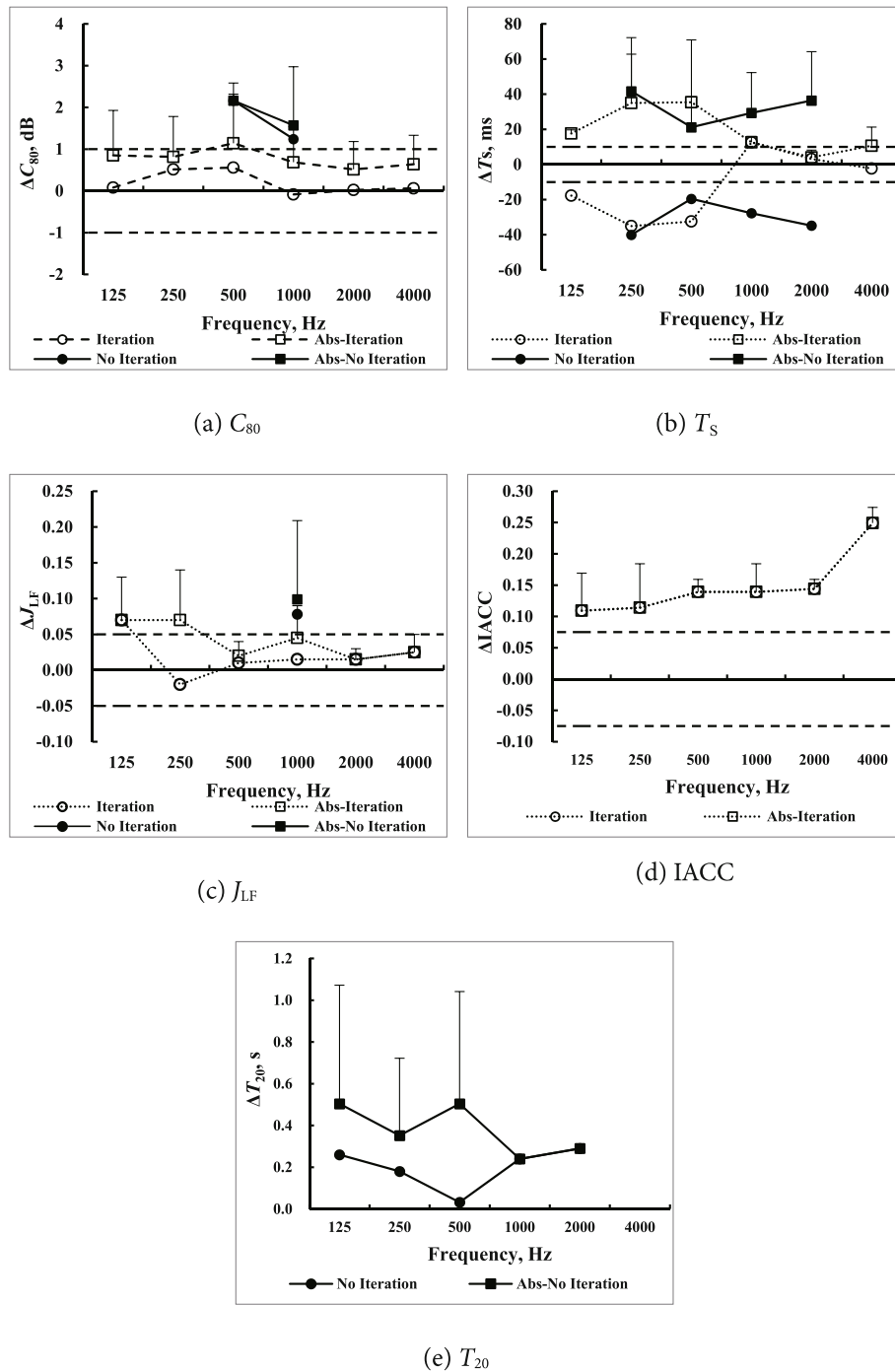


Fig. 12 The average value of  $D_{50}$  prediction differences of all cases. Each point is the average of all the iterative or non-iterative cases. “Iteration” means that the simulated results have been iteratively calibrated. “Abs” means that the differences in all cases were taken as absolute values before averaging. The straight dashed line represented the JND of  $D_{50}$

difference. For the “Abs” results, including iterative and non-iterative simulations,  $\Delta C_{80}$  was positive, indicating that the simulation results tended to be larger than the actual values, which needed more evidence to confirm in further studies.

There were 5 cases in 4 articles that included the comparison of simulated and measured  $T_S$ . The results are shown in Figure 13(b). There was no obvious difference between  $\Delta T_S$  with and without iterative simulation. The accuracy was not high either. Only the results with iterative simulation in high frequencies were within JND. This might be caused by the fact that there were very few cases where  $T_S$  was used as an iterative target.

There were 10 cases in 5 articles that included the comparison of simulated and measured  $J_{LF}$ . The results are shown in Figure 13(c). Most of the cases for  $J_{LF}$  used iterative simulation, and most of the results obtained were within JND. The difference without iterative simulation was



**Fig. 13** The average value of prediction differences of five indicators of all cases. Each point is the average of all the iterative or non-iterative cases. “Iteration” means that the simulated results have been iteratively calibrated. “Abs” means that the differences in all cases were taken as absolute values before averaging. The straight dashed line represented the value of JND

significantly higher than that with iterative simulation and JND.

There were only 2 cases in 2 articles that included the comparison of simulated and measured IACC. The results are shown in Figure 13(d). Only the results with iterative simulation were obtained, and the  $\Delta IACC$  in six octave bands were all outside the JND. Another interesting phenomenon

was that the difference at low frequency was smaller than that at high frequency, which needed more evidence to verify.

There were 5 cases in 2 articles that included the comparison of simulated and measured  $T_{20}$ . The results are shown in Figure 13(e). Only the results without iterative simulation were obtained. The results were mainly concentrated in the middle and low frequencies. Only one case



was found at the high frequency, so the two lines coincided. It can be seen from the error bar that the data of  $\Delta T_{20}$  fluctuated greatly, but the differences were all positive, which might indicate that the predicted values were greater than the actual ones.

## 4 Discussion

The results showed that the main indicators selected were EDT,  $T_{30}$ ,  $G$ , and  $D_{50}$ . It seems difficult to verify the accuracy of all the parameters listed in ISO-3382 in one research. Determining a combination of fewer parameters seems helpful to confirm the current state of the simulation accuracy and provide help for future evaluation of acoustic simulation.

Frequency averaging was conducted in six of the total 20 articles. With the development of recording tools and online sharing of research results, the analysis of more frequency results is possible to the evaluation of the accuracy of the simulation algorithm.

The results indicated that the simulated model was becoming more and more refined. However, overly detailed models have been shown in many studies to be detrimental to accuracy. Therefore, avoiding over-precision modelling may become a new focus, with the rapid growth of computing power in the future.

The accuracy of these indicators varied greatly, and most of them showed obvious systematic deviations in prediction results, which provided clues for future algorithm improvements. The effects of iterative simulation have been compared. If an indicator was chosen as the iterative target, its accuracy could be confident to be within JND. Otherwise, this target may not be achieved. This will be discussed below.

### 4.1 Suggestion of benchmark tests based on different comparison goals

In the cases involved in the review, an important factor affecting prediction accuracy was whether an iterative simulation was conducted. From the results, it could be known that if you didn't know anything about the previously measured results, it was very difficult to make the prediction accuracy of most indicators within 1 JND. Therefore, it is indispensable to determine uniform constraints in the comparison of simulation methods, and different benchmark tests need to be set up to deal with different comparison goals.

(1) *Comparison of software performance.* Classification and comparison based on the algorithm is a very effective method. The differences in the settings of software will affect the consistency of comparison. For comparison

of software performance, the software settings should be unified to avoid the influence of the operator.

(2) *Comparison of accuracy of acoustic simulations for newly built spaces.* In this situation, it is not recommended to limit the role of the operator in the benchmark test because the role of the human is inevitable, which is also in line with reality. Another reason for accepting operator's experience is that it is a useful supplement to simulation software algorithm. Due to the limitations of the algorithms, sometimes some acoustic phenomena cannot be taken into account in the software. However, the operator can make pre-correction during the setting process based on experience to get the result closer to the real situation. For example, in the experience of some acoustic consultants, the sound absorption coefficient measured in the reverberation chamber cannot be directly applied to the simulation software of the GA model but should be multiplied by a certain coefficient in advance to consider the difference between the sound field of the reverberation chamber and the actual situation. This coefficient seems to be related to the volume and shape of the space and is difficult to quantify. In this case, the operator's experience can play a better role. Probably due to updates in simulation algorithms and commercial software versions, the difference between simulated and measured values has decreased significantly in small sample studies in recent years. The need for new Round Robin or benchmark tests has increased significantly.

(3) *Iterative simulation based on the known measured results.* This situation belongs to a niche situation. It is based on the premise that the software used has a high degree of accuracy for all the simulated indicators. The purpose of the iterative simulation is not to compare software packages but more like an iterative calibration for specific cases. However, Iterative modification of the software parameters may cause over-fitting problems, which is far away from reality.

This paper focuses on the overall accuracy development of simulation software. The difference in simulation accuracy can hardly be attributed to the difference in simulation algorithms alone because the volume of the model, the precision level of modelling, and the settings of algorithm details varied greatly in different cases in the involved papers. Simulation accuracy is the result of a combination of operator experience and simulation software algorithms. The sample size of each algorithm is too small to eliminate these differences in this review. Therefore, a more in-depth comparison of accuracy between different algorithms was not conducted. Establishing benchmark tests for similar simulation algorithms is needed for further comparison.

## 4.2 Limitation of the selected articles

Unlike medical research, the sample size in each study was usually very small, and there was no evidence that the randomisation of sample selection had been carried out in all studies, and the selection bias was inevitable. Moreover, only those performance spaces that agreed to be tested would be included in the study, so there was a risk of volunteer bias.

Another difference from medical research was that there was no single-blind or double-blind comparison between computer simulation and actual measurement. Due to the differences between simulation software/algorithms, the absolute standardised operation was difficult to achieve, so there was a risk of performance bias.

Not all the measurements in the cases included in this article were standardised, which will also bring some bias. ISO-3382 provided a standardised testing framework. However, in some studies, it was recorded that ISO-3382 was used in the measurement, but some items were not actually complied with. For example, A minimum of two source positions shall be used according to the standard, but only one source position was used in some cases.

## 4.3 Limitation of this review

It can be seen from the results that the average value of most of the predicted and measured differences deviated from zero, which implied the widespread existence of systematic errors. The addition of more samples will help to determine the presence of systematic errors.

It should also be noted that, due to the differences in the room acoustics of the reviewed cases, the comparison of accuracy level of the software or algorithm was not fully standardised. More benchmark tests are required to give more evidence.

## 5 Conclusion

This article focuses on the development of acoustic computer simulation for performance space. A systematic review and meta-analysis were conducted for this purpose, and the following conclusions were reached.

The studies included in this review could be roughly classified into three types: (1) the simulation process was not controlled; (2) some parameters in the simulation process were controlled; and (3) the simulation settings were modified according to the simulation results. The frequently selected acoustic indicators were EDT,  $T_{30}$ ,  $G$ , and  $D_{50}$  and fewer articles chose  $C_{80}$ ,  $T_s$ ,  $J_{LF}$ ,  $T_{20}$ , and IACC. The mid-frequency was the most selected octave band.

The classical GA model was still the most popular

simulation method. The volume of the simulation model had not changed significantly in the past 30 years, mainly in the range of 1,000 to 100,000. The simulated model was becoming more and more refined with time.

The simulation accuracies of most indicators were outside the respective JNDs for non-iterative simulation, which could represent the comprehensive level of the software and operator, indicating that there was still a high development demand for improving the simulation accuracy. Although a larger sample size was needed for further validation, the available evidence indicated a positive trend in the simulation accuracy. From 1979 to 2020, simulations of  $T_{30}$ , EDT, and  $D_{50}$  all showed an increase with the growth of time in accuracy, except for  $G$ . In terms of frequency, the general trend of accuracy in frequency was that the software simulation was less accurate at lower frequencies, which occurred at  $T_{30}$ ,  $G$ ,  $D_{50}$  and  $T_{20}$ . However, the accuracy of EDT did not show significant frequency sensitivity. The prediction accuracy of IACC was even significantly higher at low frequencies than that at high frequencies, which needed further study. The average value of most indicators showed clear systematic deviation from zero, providing hints for future algorithm improvements.

Finally, the review's limitations and the risk of bias in this review were discussed, and different types of benchmark tests were suggested for different comparison goals.

**Electronic Supplementary Material (ESM):** The Appendix of this paper, detailed quality score for eligible studies, is available in the online version at <https://doi.org/10.1007/s12273-022-0901-4>.

## Acknowledgements

This work was supported by the Second Batch of 2022 MOE of PRC Industry-University Collaborative Education Program (Program No. 202102055010, Kingfar-CES "Human Factors and Ergonomics" Program).

**Open Access:** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory

regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>

## References

- Alfano FRD, Iannace G, Ianniello C, et al. (2015). “Velaria” in ancient Roman theatres: Can they have an acoustic role? *Energy and Buildings*, 95:98–105.
- Allred JC, Newhouse A (1958). Applications of the Monte Carlo method to architectural acoustics. II. *The Journal of the Acoustical Society of America*, 30: 903–904.
- Alonso A, Sendra JJ, Suárez R, et al. (2014). Acoustic evaluation of the cathedral of Seville as a concert hall and proposals for improving the acoustic quality perceived by listeners. *Journal of Building Performance Simulation*, 7: 360–378.
- Alonso A, Suárez R, Sendra JJ (2017). Virtual reconstruction of indoor acoustics in cathedrals: The case of the Cathedral of Granada. *Building Simulation*, 10: 431–446.
- Augenbroe G (2002). Trends in building simulation. *Building and Environment*, 37: 891–902.
- Baines NC (1983). An investigation of the factors which control non-diffuse sound fields in rooms. PhD Thesis, University of Southampton, UK.
- Bo E, Shtrepi L, Pelegrín Garcia D, et al. (2018). The accuracy of predicted acoustical parameters in ancient open-air theatres: A case study in Syracuse. *Applied Sciences*, 8: 1393.
- Bork I (2000). A comparison of room simulation software—the 2nd round robin on room acoustical computer simulation. *Acta Acustica united with Acustica*, 86: 943–956.
- Bork I (2005a). Report on the 3rd round robin on room acoustical computer simulation—Part I: Measurements. *Acta Acustica united with Acustica*, 91: 740–752.
- Bork I (2005b). Report on the 3rd round robin on room acoustical computer simulation—Part II: Calculations. *Acta Acustica united with Acustica*. 91:753–763.
- Brinkmann F, Aspöck L, Ackermann D, et al. (2019). A round robin on room acoustical simulation and auralization. *The Journal of the Acoustical Society of America*, 145: 2746–2760.
- Brinkmann F, Aspöck L, Ackermann D, et al. (2021). A benchmark for room acoustical simulation. Concept and database. *Applied Acoustics*, 176: 107867.
- Bustamante P, Girón S, Zamarreño T (2014). Simulated sound-fields in a multi-configurable auditorium. *Archives of Acoustics*, 39: 365–383.
- Choi YJ, Fricke FR (2006). A comparison of subjective assessments of recorded music and computer simulated auralizations in two auditoria. *Acta Acustica united with Acustica*, 92: 604–611.
- D’Orazio D, de Cesaris S, Morandi F, et al. (2018). The aesthetics of the Bayreuth Festspielhaus explained by means of acoustic measurements and simulations. *Journal of Cultural Heritage*, 34: 151–158.
- D’Orazio D, Fratoni G, Garai M (2020). Enhancing the strength of symphonic orchestra in an opera house. *Applied Acoustics*, 170: 107532.
- Dalenbäck BI (1995). A new model for room acoustic prediction and auralization. PhD Thesis, Chalmers University of Technology, Sweden.
- Etter PC (2001). Recent advances in underwater acoustic modelling and simulation. *Journal of Sound and Vibration*, 240: 351–383.
- Gade A (1994). Acoustic properties of concert halls in the US and in Europe; effects of differences in size and geometry. In: Proceedings of the Wallace Clement Sabine Centennial Symposium, Cambridge, MA, USA.
- Garrido JA, Zamarreño T, Girón S (2012). Virtual models for the prediction of acoustic fields of Manuel de Falla Auditorium in Granada, Spain. *Applied Acoustics*, 73: 921–935.
- Gibbs BM, Jones D (1972). A simple image method for calculating the distribution of sound pressure levels within an enclosure. *Acta Acustica united with Acustica*. 26: 24–32.
- Gomes M, Gerges S, Tenenbaum R (2000). On the accuracy of the assessment of room acoustics parameters using MLS technique and numerical simulation. *Acta Acustica united with Acustica*, 86: 891–895.
- Gumaa M, Rehan Youssef A (2019). Is virtual reality effective in orthopedic rehabilitation? A systematic review and meta-analysis. *Physical Therapy*, 99: 1304–1325.
- Gurevitch J, Koricheva J, Nakagawa S, et al. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555: 175–182.
- Hedges LV (1992). Meta-analysis. *Journal of Educational Statistics*, 17: 279–296.
- Hodgson M (1988). On the prediction of sound fields in large empty rooms. *The Journal of the Acoustical Society of America*, 84: 253–261.
- Hong T, Chou S, Bong T (2000). Building simulation: an overview of developments and information sources. *Building and Environment*, 35: 347–361.
- Hornikx M, Hak C, Wenmaekers R (2015a). Acoustic modelling of sports halls, two case studies. *Journal of Building Performance Simulation*, 8: 26–38.
- Hornikx M, Kaltenbacher M, Marburg S (2015b). A platform for benchmark cases in computational acoustics. *Acta Acustica United With Acustica*, 101: 811–820.
- Howarth M, Lam Y (2000). An assessment of the accuracy of a hybrid room acoustics model with surface diffusion facility. *Applied Acoustics*, 60: 237–251.
- Ibarra-Zarate D, Padilla-Ortiz AL (2020). Acoustic characterization of a room: study case between simulation and a portable method. *Archives of Acoustics*, 45: 153–163.
- Jin Y, Yan D, Zhang X, et al. (2021). A data-driven model predictive control for lighting system based on historical occupancy in an office building: Methodology development. *Building Simulation*, 14: 219–235.
- Kang J (1996). Acoustics in long enclosures with multiple sources. *The Journal of the Acoustical Society of America*, 99: 985–989.
- Kang J (1997). A method for predicting acoustic indices in long enclosures. *Applied Acoustics*, 51: 169–180.
- Kang J (2002). Numerical modelling of the speech intelligibility in dining spaces. *Applied Acoustics*, 63: 1315–1333.
- Kong M, Zhang J, Wang J (2015). Air and air contaminant flows in office cubicles with and without personal ventilation: A CFD modeling and simulation study. *Building Simulation*, 8: 381–392.

- Koutsouris GI, Brunskog J, Jeong CH, et al. (2013). Combination of acoustical radiosity and the image source method. *The Journal of the Acoustical Society of America*, 133: 3963–3974.
- 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: 118–125.
- Kuttruff H (2009). *Room Acoustics*. Boca Raton, FL, USA: CRC Press.
- Lam YW (1996). A comparison of three diffuse reflection modeling methods used in room acoustics computer models. *The Journal of the Acoustical Society of America*, 100: 2181–2192.
- Lewers T (1993). A combined beam tracing and radiative exchange computer model of room acoustics. *Applied Acoustics*, 38: 161–178.
- Li Y, O'Neill Z (2018). A critical review of fault modeling of HVAC systems in buildings. *Building Simulation*, 11: 953–975.
- MacDermid JC (2004). An introduction to evidence-based practice for hand therapists. *Journal of Hand Therapy*, 17: 105–117.
- Marshall AH (1979). Aspects of the acoustical design and properties of Christchurch town hall, new zealand. *Journal of Sound and Vibration*, 62: 181–194.
- Naylor GM (1993). ODEON—Another hybrid room acoustical model. *Applied Acoustics*, 38: 131–143.
- Pätynen J, Lokki T (2011). Evaluation of concert hall auralization with virtual symphony orchestra. *Building Acoustics*, 18: 349–366.
- Picaut J, Fortin N (2012). SPPS, a particle-tracing numerical code for indoor and outdoor sound propagation prediction. In: Proceedings of Acoustics 2012, Nantes, France.
- Postma BNJ, Katz BFG (2016). Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *The Journal of the Acoustical Society of America*, 140: 4326–4337.
- Prawirasasra MS, Sampurna R, Suwandi (2016). The effect of boundary shape to acoustic parameters. *Journal of Physics: Conference Series*, 776: 012068.
- Rindel JH (2011). Room acoustic modelling techniques: A comparison of a scale model and a computer model for a new opera theatre. *Building Acoustics*, 18: 259–280.
- San Martín R, Arana M (2006). Predicted and experimental results of acoustic parameters in the new Symphony Hall in Pamplona, Spain. *Applied Acoustics*, 67: 1–14.
- San Martín R, Arregui A, Machín J, et al. (2014). Comparison of measured and simulated room acoustic parameter values using high resolution grids. *Acoustics Australia*, 42: 27–30.
- Savioja L, Svensson UP (2015). Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America*, 138: 708–730.
- Schroeder MR (1970). Digital simulation of sound transmission in reverberant spaces. *The Journal of the Acoustical Society of America*, 47: 424–431.
- Shtrepi L, Astolfi A, Puglisi G, et al. (2017). Effects of the distance from a diffusive surface on the objective and perceptual evaluation of the sound field in a small simulated variable-acoustics hall. *Applied Sciences*, 7: 224.
- Southern A, Siltanen S, Murphy DT, et al. (2013). Room impulse response synthesis and validation using a hybrid acoustic model. *IEEE Transactions on Audio, Speech, and Language Processing*, 21: 1940–1952.
- Strøm S, Krokstad A, Dahl H, et al. (1985). Acoustical design of the Grieg memorial hall in Bergen. *Applied Acoustics*, 18: 127–142.
- Svensson P, Kristiansen UR (2002). Computational modelling and simulation of acoustic spaces. In: Proceedings of the 22nd International Conference: Virtual, Synthetic, and Entertainment Audio. Audio Engineering Society.
- Tsingos N, Gascuel JD (1997). A general model for the simulation of room acoustics based on hierarchical radiosity. In: Proceedings of SIGGRAPH'97.
- Vercammen ML (2013). Sound concentration caused by curved surfaces. In: Proceedings of Meetings on Acoustics ICA2013, Montreal, Canada.
- Vorländer M (1995). International round robin on room acoustical computer simulations. In: Proceedings of the 15th International Congress on Acoustics, Trondheim, Norway.
- Vorländer M (2013). Computer simulations in room acoustics: Concepts and uncertainties. *The Journal of the Acoustical Society of America*, 133: 1203–1213.
- Walsh JP, Rivard MT (1982). Signal processing aspects of godot: A system for computer-aided room acoustics modeling and simulation. In: Proceedings of Audio Engineering Society Convention.
- Wang LM, Vigeant MC (2008). Evaluations of output from room acoustic computer modeling and auralization due to different sound source directionalities. *Applied Acoustics*, 69: 1281–1293.
- Winsberg E (2010). *Science in the Age of Computer Simulation*. Chicago, USA: University of Chicago Press.
- Wu Y, Kang J, Zheng W (2018). Acoustic environment research of railway station in China. *Energy Procedia*, 153: 353–358.
- Zannin PHT, Marcon CR (2007). Objective and subjective evaluation of the acoustic comfort in classrooms. *Applied Ergonomics*, 38: 675–680.
- Zhu P, Tao W, Lu X, et al. (2021). Optimisation design and verification of the acoustic environment for multimedia classrooms in universities based on simulation. *Building Simulation*, 15: 1419–1436.