Acoust Aust
DOI 10.1007/s40857-016-0055-6

TECHNICAL NOTE

Calculating Reverberation Time from Impulse Responses: A Comparison of Software Implementations

Densil Cabrera[1] · Jianyang Xun[1] · Martin Guski[2]

Received: 1 February 2016 / Accepted: 16 April 2016

Abstract In room acoustics measurement, calculating reverberation time from room impulse responses is often done, aided by software. This paper compares the performance of nine software implementations for calculating octave band reverberation time, including two written by the authors. Synthetic impulse responses are used to test decays without and with a steady noise floor, and an impulse response from a real measurement is also used for comparison. Results indicate no significant reverberation time calculation problems for noise-free exponential decays, and for exponential decays leading to a steady noise floor. Frequency selectivity is identified as an area for potential improvement in filter-bank design, and a highly selective octave band filter-bank is shown to be effective without introducing errors. Testing with a real measured impulse response, which had been used in a 2004 study comparing reverberation time analysis implementations, showed greater agreement between software than was found previously. This might reflect an improvement in software performance in the years between the two studies. However, it also might reflect the smaller scale of the present study. Nevertheless, the results can contribute to confidence in current software implementations.

Keywords Room acoustics · Reverberation time · Software

1 Introduction

Reverberation time is the most used and best-known parameter in room acoustics and has a long history of development in terms of theory, measurement and application. It is established in key standards, such as the ISO3382 series, and almost all projects concerned with room acoustics use reverberation time in some way. Measurement of reverberation time can be done in various ways, and is commonly done by recording a room impulse response, often indirectly (e.g. from a swept sinusoid signal). Deriving reverberation time from a room impulse response relies on software, and while some researchers implement their own analysis software, it is quite common for researchers and professionals in architectural acoustics to use software available for purchase or distributed freely from external sources. Hence, with a range of software available, measured values of reverberation time might be affected by software choice.

The majority of published work comparing reverberation time calculation from an impulse response has focussed on the efficacy of algorithms, often proposing refinements or novel approaches to the calculation method (e.g. [1–3]). Related to this, research has examined the errors associated with certain types of input, most often examining how to deal with a steady-state noise floor, and proposing required impulse response criteria for results within acceptable error margins (e.g. [2,4–6]). Another less common type of comparison study is between software, without particular knowledge of the algorithm, and there is some value in this considering that many acousticians rely on closed source software. The present study is primarily a software comparison, but it also considers some algorithm-related matters as they arise from the comparison.

✉ Densil Cabrera
densil.cabrera@sydney.edu.au

1 Faculty of Architecture, Design and Planning, The University of Sydney, Sydney, NSW 2006, Australia

2 Institute of Technical Acoustics, RWTH Aachen University, Aachen, Germany

Published online: 04 May 2016

Lundeby et al. [7] and Bradley [8] conducted studies in the 1990s comparing systems for reverberation time measurement and analysis. In Lundeby's study the whole measurement and analysis system (including physical equipment) was included, whereas Bradley used an electronic artificial reverberator to remove physical variability. However, both studies were comparisons of measurement and analysis systems, rather than isolating just the analysis part. Subsequent studies by Ikegami and Uchida [9] and Katz [10] focussed only on analysis by using an impulse response as the test input for software. Katz's impulse response (a recording of a balloon pop in a large room) is of particular interest because of the large scale of the study and because the impulse response has a significant noise floor. The impulse response was sent to researchers and practitioners to analyse, using software of their choice. Results showed a wide range of reverberation time values in the lowest octave band (125 Hz), for which signal-to-noise ratio was poorest. Katz's study is also of particular interest because some of the software that was used for the tests is still in use today, albeit in updated versions. Since Katz's study, some smaller scale software comparison studies have been published, for example, by Tronchin et al. [11] (four software implementations), Ţopa et al. [12] (three implementations), Machín et al. [13] (three implementations, two external and one by the authors) and Mansilla et al. [14] (four implementations, three external and one by the authors). The conclusions of these studies vary: Tronchin et al., using impulse responses measured in an opera theatre, found average T_{30} (reverberation time) deviations of up to 8.3 % between software, with a maximum deviation of almost 25 %; Ţopa et al. tested three rooms (a school hall, a church and a university auditorium) and found that the software did not affect the results; the main aim of the study by Machín et al. was validate software developed by the authors through a set of measurements in one auditorium, and they found only minor deviations in reverberation time between the tested software; and Mansilla et al. tested 59 diverse impulse responses, finding average octave band deviations of about 0.3 s in T_{30} between the external software tested. Concurrent with the present study, Álvarez-Morales et al. [15] conducted a detailed software comparison based on a large set of in-situ measurements in a university auditorium, using four software implementations, with extensive statistical analysis. They found that the implementations mostly agreed when used as recommended by the software documentation. Hence the extent to which software agrees or disagrees appears to depend on the type of testing done, and the present paper focuses on testing specific potential vulnerabilities of software, along with a reassessment using Katz's impulse response with present day software.

1.1 Issues in the Calculation of Reverberation Time

In general, the procedure for calculating reverberation time from a single measured room impulse response will involve the following steps:

1. Truncation of the impulse response at the start, although this should have little effect on reverberation time calculation and is mainly important for energy ratio parameters such as clarity index (ISO3382-1 suggests starting where the energy first rises significantly above the background noise, but is more than 20 dB below the peak);
2. Applying a filter-bank (e.g. octave band);
3. Truncation of the impulse response at the end, usually at a different point in each band;
4. Squaring the impulse response;
5. Optionally compensating for the decay energy lost from end-truncation and/or subtraction of the excess energy due to a steady-state noise floor;
6. Integrating the impulse response in reverse time, and converting to decibels; and
7. Finding reverberation time parameters by linear regression over the appropriate evaluation ranges (-5 to -25 dB for T_{20}; -5 to -35 dB for T_{30}).

Steps 2 and 3 and the optional step 5 have some scope for variations in approach, which could affect the resulting values.

Filter-banks for reverberation time measurement can influence the result for two reasons. Firstly, the time-response of the filter-bank can have its own decay, which could artificially increase the measured reverberation time [16]. This is only likely to be noticeable when reverberation time is very short, and especially in the lower octave or 1/3-octave bands. However, this can be avoided by using reverse-time filtering [3,17]. Secondly, the selectivity of the filter-bank is only loosely defined by standards [18–20], and cases where there is significant variation in reverberation as a function of frequency will be better represented by the results of a highly selective filter-bank. Venturi et al. [21] suggested the use of highly selective filters (144 dB/octave skirts, equivalent to 24th order) for octave band reverberation time measurement, finding that high-order filtering always improved the accuracy reverberation time results compared to sixth-order filters (both of which comply with IEC Class 0 response limits from the standard at the time [19]). However, they speculated that lower selectivity might be needed for very short reverberation times and for short evaluation ranges such as early decay time and T_{10}.

Much has been written about end-truncation, because it has a large effect on the results in practical measurements

[4,6,7,22]. If end-truncation is omitted, then the background noise will accumulate in the reverse integration, potentially yielding artificially high reverberation time values. Following ISO3382-1, the end-truncation point should be at the intersection between a horizontal line representing the background noise and a sloping line representing the level decay (prior to reverse integration) [18], but the method for finding this could vary between implementations, and it is likely that variation in results between software for impulse responses that include noise will be partly due to the way in which end-truncation is done.

Compensation for the energy lost from end-truncation is an optional step in ISO3382-1, which helps straighten the reverse-integrated decay. Without doing this, there may be an underestimation of reverberation time. The specifics of how this is done are open to interpretation, but ISO3382-1 recommends a value derived from the decay rate in the final 10 dB. A variation of this, by Lundeby et al. [7], provides a 5–10 dB safety margin in finding this late decay rate, to reduce the influence of the noise floor on the late decay rate calculation.

Noise power subtraction is another technique that helps straighten the reverse-integrated decay [2]. Although this approach is not compliant with ISO3382, noise compensation is implemented (although not necessarily described) in commonly used software. Guski and Vorländer [6] recommended that this, together with truncation energy compensation, be included in future modification of ISO3382. Where available, it was included in the software tests reported here.

There are many non-standard alternative approaches to calculate reverberation time from an impulse response. For example, Xiang [23] proposed non-linear regression accounting for the reverse-integrated noise floor. Morgan [4] proposed using an evaluation range that depended upon the noise floor (finishing 5 dB above the noise, in conjunction with end-truncation at the elbow) instead of using fixed evaluation ranges. However, mostly reverberation calculation software aims to follow a standard approach (often with the non-standard option of noise power subtraction), and it is common to see references to ISO3382 in such software.

2 Software and Test Concept

The test concept of the present study is that the software settings are fixed (the same settings used for every impulse responses) so that human decision making is removed. Hence, some results may be inferior if there are cases for which an expert might have optimised the analysis by changing the settings, by editing the impulse response, by choosing the most appropriate value to report or by not reporting a value due to lack of validity. Where available, noise compensation is applied (including in cases where the impulse response is essentially noise-free), and where reverse-time filtering is chosen if it is a user setting. For the sake of succinctness, only T_{30} values are reported (i.e. reverberation time calculated over the −5 to −35 dB evaluation range in the reverse-integrated decay), as these are more vulnerable than T_{20} to background noise, truncation and filter response. Indeed, Hak et al. [5] show that T_{30} is the most sensitive of the ISO 3382-1 parameters to signal-to-noise ratio.

Seven software implementations were tested, together with two of the authors' own implementations. The seven external software instances in this study include prominent reverberation time calculation software that tends to be used in professional and research work: in alphabetic order, the software is ARTA, Aurora, Dirac, EASERA, Odeon, Smaart and WinMLS. Software that required a superseded operating system or particular hardware was excluded. All of this software has been included in previous software comparisons by: Katz (Aurora, Dirac, Smaart and WinMLS) [10], Tronchin et al. and Ţopa et al. (Aurora, Dirac and WinMLS) [11,12], Machín (Odeon, WinMLS) [13], Mansilla et al. (Aurora, Dirac, EASERA) [14] and Álvarez-Morales (ARTA, Dirac, EASERA, WinMLS) [15]. One of the concerns with software selection was that the software be recognised as a serious analysis tool, and inclusion in previous software comparisons done by acoustics researchers provides one indication of this. Two other software implementations, which were not in these previous comparison studies, were tested in preparing the present paper, but were excluded from this study due to the combination of deficiencies in performance (yielding outlying values in the tests) and a lack of external evidence that the software was recognised as a serious analysis tool (such as multiple citations in room acoustics literature or authorship associated with acoustics research). Although Smaart tends to be used more for professional audio system tuning than for room acoustics, it was included in the study because of its inclusion in Katz's. Every software version tested was current in December 2015. Like some previous software comparison studies in room acoustics, including Katz's, the external software is de-identified, and is referred to arbitrarily using the letters A to G.

The two implementations that were developed by the authors and their colleagues are the ITA Toolbox (developed in the Institute of Technical Acoustics, RWTH University Aachen) [24] and AARAE (developed at the University of Sydney) [25]. The code for these was written independently. Both of these projects are freely available as open-source Matlab-hosted projects, aiming to support teaching and research in room acoustics and related areas. These projects have a wide range of possible settings for reverberation time calculation, but the results reported in this paper are from the respective authors' preferred settings. In both cases, this included automatic end-truncation, compensating for decay energy lost from truncation and background noise power subtraction. However, the implementations differ in details, such as the filter-bank design, and the methods to find the truncation point and to do the abovementioned adjustments. For the ITA Toolbox, the settings are exactly those described by Guski and Vorländer [6] as 'Method E'. The most distinctive feature of the chosen AARAE settings was that they applied the suggestion by Venturi et al. [21] for highly selective octave band filtering (implemented as a reverse-time, or maximum phase, filter-bank in the frequency domain, with 144 dB/octave skirts).

3 Tests and Results

3.1 Noise-Free Artificial Impulse Responses

A simple and trivial test of reverberation time software performance is to calculate the reverberation time of an exponentially decaying waveform. To avoid random fluctuations in the envelope, we use exponentially decaying sinusoidal waves. Equation 1 expresses the waveform, y, as a function of time, t (starting at 0 s and finishing at a duration sufficiently long for reverberation time calculation). The lowest tone frequency in the present study is $f_0 = 125$ Hz, and seven octaves spanning 125 Hz to 8 kHz are used ($k = 1$–7). In Eq. 1, T_k is the octave band reverberation time, and so the error in the reverberation time calculation software can be found by comparison to this.

$$y(t) = \sum_{k=1}^{7} \sin\left(2^k \pi f_0 t\right) e^{-3\ln(10)t/T_k}. \tag{1}$$

The synthetic impulse responses tested for this paper were generated in Matlab using a sampling rate of 48 kHz, and saved as 16-bit wav files, which were then analysed by the reverberation time software. Using a reverberation time of 1 s in all seven bands and a waveform duration of 2 s, all software implementations tested yield negligible errors. Similarly, negligible errors are found with longer and shorter uniform reverberation times. For a very short reverberation time of 0.1 s, results mostly show negligible errors, although software implementation A only returned results in the two lowest octave bands. Implementation C returned an increased result in the 125 Hz band (of 0.112 s), which is the expected effect of a filter's own reverberation contaminating the result (e.g. for a minimum phase filter). Considering that 0.1 s is a shorter reverberation time than we would normally encounter in non-anechoic rooms, the conclusion is that no substantial issues were revealed by testing noise-free artificial impulse responses with uniform octave band reverberation times in the 125–8000 Hz range.

3.2 Exponential Decay with Steady Noise Floor

Apart from the relatively trivial task of calculating reverberation time from a noise-free impulse response, probably the most telling simple test is to introduce a steady-state noise floor. This means that automatic end-truncation must be done by the analysis software, as well as associated adjustments. An exponentially decaying envelope with a perfectly steady noise floor can be generated by adding a tonal noise term to Eq. 1, as expressed by Eq. 2, where *PNR* is the peak-to-noise ratio expressed in decibels. Note that the tonal noise is in quadrature with the tonal decay, so that they add similarly to incoherent noises in terms of the resulting envelope, but without random fluctuations. The noise term is multiplied by the square root of 2 because PNR is a ratio of maximum to root-mean-square amplitude.

$$y(t) = \sum_{k=1}^{7} \sin\left(2^k \pi f_0 t\right) e^{-3\ln(10)t/T_k} + \sqrt{2} \times 10^{PNR/20} \cos\left(2^k \pi f_0 t\right). \tag{2}$$

Results from testing this are more interesting than without noise, and they indicate that the tested software is generally well behaved in dealing with steady noise. Figure 1 shows results for T_{30} with a PNR of 30, 35, 40 and 45 dB. In interpreting this, it should be borne in mind that a signal-to-noise ratio of 45 dB or greater is required for standard calculation of T_{30}, so the lower PNR values do not meet this requirement. Admittedly there is some subtlety in signal-to-noise ratio metrics for impulse responses as explained by Hak et al. [5], but in this paper we use PNR because of its simplic-

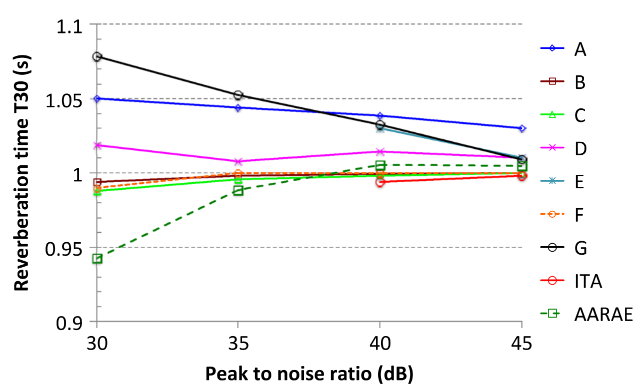

Fig. 1 Reverberation time (T_{30}) calculated from synthetic impulse responses consisting of octave-spaced exponentially decaying sine tones mixed with an octave-spaced cosine tone noise floor, for peak-to-noise ratios from 30 to 45 dB. The synthesised reverberation time in all bands for all impulse responses was exactly 1 s and the waveform duration was 2 s. Calculated reverberation time values shown are averaged over the seven octave bands 125–8000 Hz

ity as a parameter for synthesis and analysis. In two cases (implementation E and the ITA Toolbox), the software did not output a result for PNR of 30 and 35 dB. In other cases (including AARAE), although a result is returned at a PNR of 30 dB, its validity is counter-indicated by other output parameters such as measured PNR, correlation coefficient of the linear regression, and/or the ratio of T_{20} to T_{30}. Nevertheless, some implementations (B, C and F) are remarkably robust in this test, yielding errors of only about 1 % in the worst case PNR.

A 5 % error margin is generally considered acceptable in reverberation time measurement, and almost all of the results in Fig. 1 are within this margin. At the required PNR of 45 dB, all but one result (implementation A) have an error of 1 % or less. The results do not reveal any significant problems with any implementations, and provide confidence that end-truncation and associated adjustments by the software are generally functioning well, at least for simple exponential decay with added steady-state noise.

3.3 Frequency Selectivity

To examine the effect of octave band filter-bank selectivity, decaying octave-spaced sinusoids without added noise floor were generated, with reverberation time varying markedly between adjacent octave bands. The two examples presented here have reverberation time alternating between 0.1 and 1 s over the seven octave bands from 125 to 8000 Hz, or between 0.5 and 1 s over the seven bands. The first of these is an extreme case, with a 10:1 ratio that is highly unlikely to be encountered in octave band measurements of any real room. The second has a large but not implausible ratio of 2:1 between adjacent bands. The results in Fig. 2 show that in many cases the implementations return a value much higher than the short reverberation time (0.1 or 0.5 s) in the 125, 500, 2000 and 8000 Hz bands, due to leakage from the adjacent band or bands. On the other hand, the long reverberation time (1 s) in the 250, 1000 and 2000 Hz bands is accurately analysed in all cases. The filter-bank for the AARAE implementation was designed to be highly selective (following the suggestion of Venturi et al. [21]), and returns results with negligible error in both examples. The ITA Toolbox implementation was not configured for extreme selectivity for this test (and used filters equivalent to 10th order), but nevertheless is one of the most selective. Some implementations, especially A, B and G, yield results indicating low selectivity. Cabrera et al. [21] previously compared software for speech transmission index measurement, and observed that the modulation transfer functions of A and B were consistent with that expected for sixth-order filters (A and B in the present study are labelled F and C, respectively, in the previous study), which comply with IEC Class 1 octave band filter criteria [20]. Considering the case with the 2:1

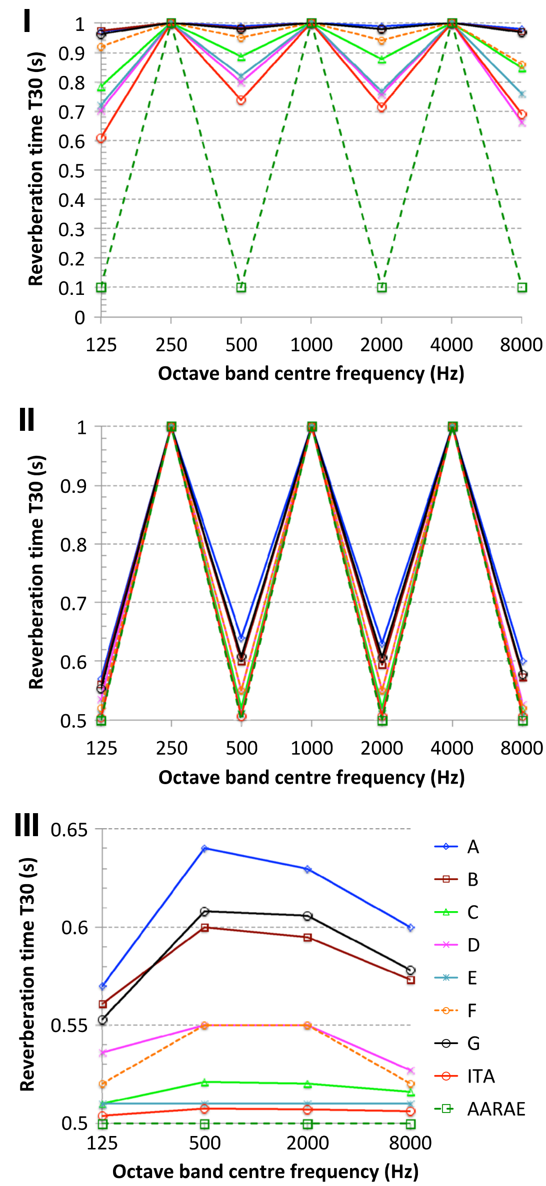

Fig. 2 Reverberation time (T_{30}) calculated from a synthetic impulse response consisting of octave-spaced exponentially decaying sine tones with a synthesised reverberation time alternating between 0.1 and 1 s over the 125–8000 Hz bands (*top chart*, labelled I) or 0.5 s and 1 s over the 125–8000 Hz bands (*middle* and *bottom chart*, labelled II and III). For visual clarity, the *bottom chart* shows the same data as the *middle chart*, but omits the bands that have a reverberation time of 1 s

ratio of reverberation time between adjacent bands, three of the nine implementations return errors of less than 5 % in all bands. If the 24th-order filter-bank is considered to be unnecessarily selective, a 12th-order filter-bank (tested using AARAE) yields a maximum octave band T_{30} error of 0.3 % (or $T = 0.5017$ s in the 0.5 s bands) in the 2:1 ratio case (not shown in the figure).

The waveforms tested here have unrealistically high contrast between bands, and so the effect of frequency selectivity will be less important in most real cases. However, the interesting point here is that very high selectivity is achievable without causing problems with the measurement of short reverberation times. To examine this further, a synthetic tonal impulse response with octave band reverberation times of 0.01 s was tested with the AARAE 24th-order filter-bank implementation. The measured result for T_{30} is 0.01 s in all seven bands (using 'double' precision format for the input waveform, instead of the 16-bit signed integer wav file format that was used for the other tests reported here). This confirms the proposition by Venturi et al. [16] that a highly selective filter-bank provides robust analysis, and it extends this by showing that the filter-bank does not introduce significant time-related errors for the shortest conceivable reverberation time, at least for T_{30} calculation from an exponential decay.

3.3.1 Further Implications of Filter-Bank Design: Other Parameters

In the present work we focus on reverberation time, but a filter-bank ideal for that might not be ideal for another parameter. For example, for speech transmission index (STI) calculation, zero or linear phase filters are ideal because phase distortion within each octave band must be minimised to avoid changing the octave band envelopes [26]. Furthermore, excessive filter order artificially reduces the modulation transfer function [27]. Therefore, Cabrera et al. [26] selected a linear phase 12th-order octave band filter-bank for STI analysis. On the other hand, with regard to the clarity and definition energy ratio parameters in ISO3382-1, the filter phase response and order do not fundamentally affect results (apart from controlling spectral leakage) because the relevant time periods are extracted prior to filtering, so a high-order reverse-time filter can be used. However to avoid error, the filters' reverse-time decay should be captured, which may require zero-padding the relevant time periods extracted from the waveform prior to filtering. The theoretical clarity index of a simple exponential decay is known from Eq. 3, where t_E is the limiting time between early and late (0.05 s for C_{50}, 0.08 s for C_{80}), and T is reverberation time.

$$C_{1000t_E} = 10\log_{10}\left(\frac{1 - \mathrm{e}^{-t_E 6\ln 10/T}}{\mathrm{e}^{-t_E 6\ln 10/T}}\right). \tag{3}$$

This was tested using AARAE, with a reverse-time octave band filter-bank spanning 125–8000 Hz. Exponentially decaying octave-spaced tones over a range of reverberation times (0.1–2 s) yielded negligible errors (<0.2 dB) in the seven octave bands for C_{50} and C_{80}, and increasing the filter order from 6 to 24 did not increase the error (in fact it slightly reduced it).

Conceptually, centre time may be affected by reverse-time filtering, especially at high orders in the low frequency bands, because it results in a shift of energy backwards in time. Hence zero-phase filtering is conceptually ideal for centre time, but not ideal for reverberation time. The theoretical centre time, T_s, of a simple exponential decay is proportional to its reverberation time (Eq. 4).

$$T_s = \frac{T}{6\ln(10)}. \tag{4}$$

Tests using AARAE with ideal exponential decay of $T = 1$ s (for which $T_s = 72.38$ ms) using a time-reverse filter of order 6 to 24 yielded octave band centre time values with errors no more than 0.1 ms, which can be neglected. A shorter reverberation time is more vulnerable to filter effects, and a decay of $T = 0.1$ s (for which $T_s = 7.238$ ms) was also tested with filter orders 6 to 24. In that case, the errors were higher in absolute terms and much higher in relative terms, with a maximum error of 0.27 ms. This error can still probably be regarded as acceptable, especially since T_s is unlikely to provide any insights to situations with very short reverberation times. These maximum errors occurred in the 125 Hz band, and errors were an order of magnitude smaller in high frequency bands. Errors did not increase as the filter order increased. Considering this, not using a zero-phase filter-bank for octave band centre time appears to be unproblematic from a practical standpoint, at least for the bands centred on 125 Hz and higher.

Although the aim of this study is to examine T_{30} results of the surveyed instances of software, the considerations in this section of the paper naturally raise the question of software performance for clarity index and centre time. Figure 3 shows the C_{50} and T_s errors from a noise-free synthetic impulse response with $T = 1$ s in all seven bands. For the most part, the errors in C_{50} are negligible, although one software implementation (F) deviates by more than 1 dB in the 125 Hz band. On the other hand, T_s errors in excess of 5 % are common in the lower two octave bands, which could be due to forward temporal smearing by the octave band filters.

3.4 Realistic Measured Impulse Response

Real measured impulse responses are much more complicated than the synthetic waveforms tested here. While they may raise multiple diverse challenges for analysis, they do not have known reverberation time values *a priori*, and so there is not an exactly correct result for error analysis. We use the same impulse response as Katz [10] in the present paper, more than 1 decade after Katz's study. It should be noted that Katz allowed expert analysts to contribute to the analysis—for example, they might manually choose the start and end times to prepare for the analysis. By contrast, the present study excluded pre-manipulation by a human expert, and relied entirely on the automatic processes in the software

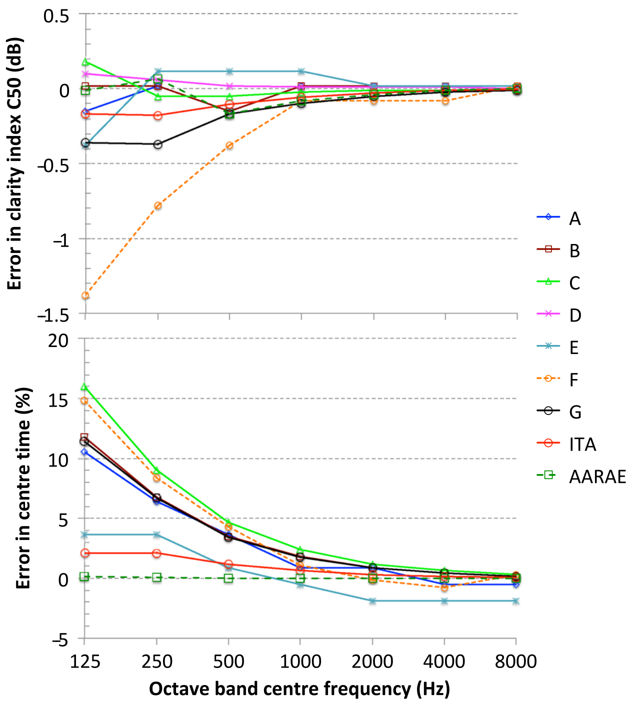

Fig. 3 Octave band error in clarity index (C_{50}, *upper chart*) and centre time (T_s, *lower chart*) for a synthetic impulse response consisting of octave-spaced exponentially decaying sinusoids from 125 Hz to 8 kHz with a uniform reverberation time of exactly 1 s. The theoretical C_{50} value is −0.02 dB, and the theoretical T_s value is 72.38 ms. The AARAE values were calculated using a 24th order reverse-time octave band filterbank

for the results. Visualising the impulse response in octave bands provides some insight into the challenges faced by the software. Figure 4 shows the squared impulse response (expressed in decibels) in each of the octave bands. In all bands, the noise floor is essentially steady over the long term, and the decay appears to be linear. Greater short-term fluctuations are seen in the lower bands, both in the decay and in the noise floor. The 500–4000 Hz bands have a distinct peak from the direct sound, and have a visually smoother decay than the lower two bands.

Figure 5 shows the T_{30} results for the software implementations, together with deviations from the median in each octave band. The ITA Toolbox and AARAE results are nearly identical, and implementations E and F also return near-identical results. Like the results reported by Katz, the largest variability between implementations is in the lowest octave band. Implementations A and C have the most distinctive results. In the 125 Hz band, implementation C returns a value substantially lower than the others, but is in general agreement in the remaining five bands. In the 500–2000 Hz bands, the results are in close agreement, except that implementation A returns outlying values at 500 and 1000 Hz. Implementations B, C and F, which had very similar results with minimal errors for the synthetic test of PNR, have a wide spread of results in the 125 Hz band for the real impulse response, suggesting that there is more behind the divergence of results than just the PNR in each band. Unlike the ideal decay lead-

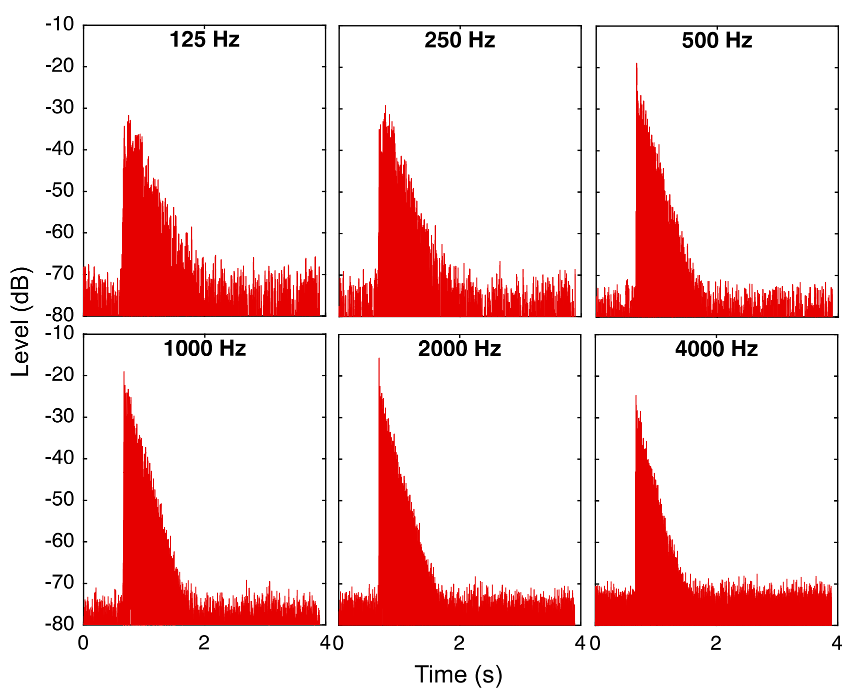

Fig. 4 Octave band instantaneous level of the impulse response published by Katz [10]. Values, in decibels, are with reference to full-scale amplitude of the wav file format

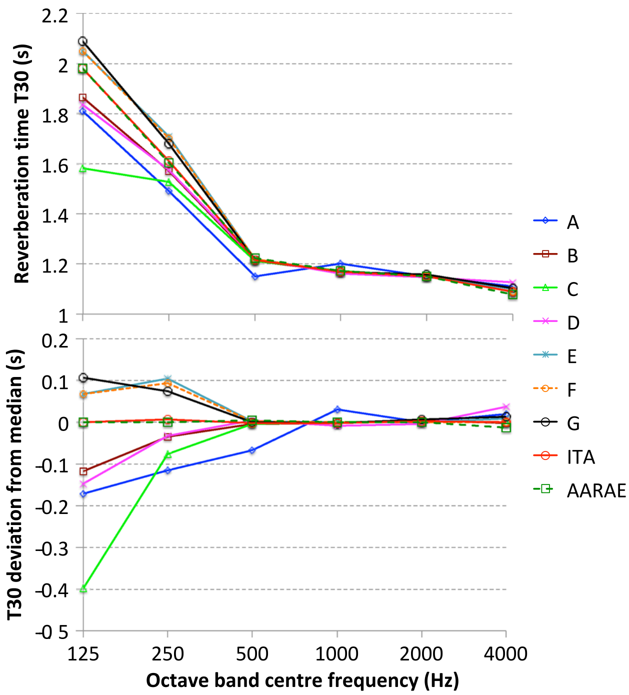

Fig. 5 The *upper chart* shows octave band reverberation time (T_{30}) calculated from the impulse response published by Katz [10]. The *lower chart* shows the deviation from median (i.e. the calculated value minus the median value) in each octave band

ing to steady noise that was used for the synthetic test of PNR, the real impulse response has fluctuations in both the decay and the noise floor. Hence the differences exhibited here by B, C and F may be due to the sensitivities of their particular algorithm used to identify the truncation point, and perhaps the way in which the late decay rate or the noise floor values are calculated. The lower frequency selectivity of B's filterbank might also contribute to its divergence from C and F, which is supported by the similarity of its results with those of G (which has similarly low selectivity) in the lower two bands.

Apart from the 4 kHz band, the ITA Toolbox and AARAE implementations return results that are essentially at the median. Although AARAE returns an outlier at 4 kHz, its deviation in that band returns only 0.01 s from the median.

Table 1 provides a comparison between the results of the present study and those of Katz. The median values are similar, but the standard deviations and interquartile ranges in the present study are much smaller. The interquartile range excludes outliers, such as implementations C and A in the lower bands. This reduced spread of results in the present study could be from a combination of factors, including improvements in software implementations, the selection of software, the human contribution to analysis in Katz's study (which was not in the present study) and statistical issues relating to sample size. The end-truncation method in the current standard (ISO3382-1-2009) provides less flexibility than in the version at the time of Katz's study (ISO3382-1997) [28], which may have contributed to greater agreement between implementations.

To put the statistical dispersion of results in context, the last row of Table 1 shows the greatest absolute value deviation from the median in the present study, expressed as a percentage. In other words, this represents the furthest outlier (either above or below the median), and provides a simple way of comparing with the 5 % acceptable deviation rule of thumb. This is not to say that the median is correct—it is merely used as a reference for comparison. The PNR in the 125 Hz band is almost 45 dB, and is greater than this in the other bands (see Fig. 4), and so the signal-to-noise ratio is essentially that required for T_{30} calculation. Of the nine software implementations, five had deviations from the median greater than 5 % in the 125 Hz band, three in the 250 Hz band, and one in the 500 Hz band.

In comparing dispersion of T_{30} with the PNR in each octave band, PNR appears to be a plausible predictor of variation between the implementations' results. Figure 6 shows this comparison for the present study and for Katz's study, with dispersion expressed in relative terms, as interquartile ratio (interquartile range divided by median). This concurs

Table 1 Octave band median reverberation time (T_{30}) and indicators of statistical dispersion calculated from the impulse response published by Katz [10], with results from this study and the values reported in Katz's study

	125 (Hz)	250 (Hz)	500 (Hz)	1 (kHz)	2 (kHz)	4 (kHz)
Median	1.98 s	1.61 s	1.22 s	1.17 s	1.15 s	1.09 s
Median from Katz	1.92 s	1.62 s	1.22 s	1.18 s	1.16 s	1.12 s
STD	0.16 s	0.08 s	0.02 s	0.01 s	0.00 s	0.02 s
STD from Katz	0.35 s	0.17 s	0.08 s	0.07 s	0.09 s	–
IQR	0.22 s	0.11 s	0.01 s	0.00 s	0.00 s	0.01 s
IQR from Katz	0.49 s	0.16 s	0.03 s	0.02 s	0.02 s	–
Greatest deviation from median	20.1 %	7.6 %	5.5 %	2.6 %	0.6 %	3.4 %

Standard deviations (STD) and interquartile ranges (IQR) from both studies are shown, along with the greatest absolute value deviation from median (expressed as a percentage) in results of the present study. Katz did not report STD or IQR at 4 kHz, and 8 kHz was not included at all in his study. In the present study not all software returned a result at 8 kHz, and so it is omitted

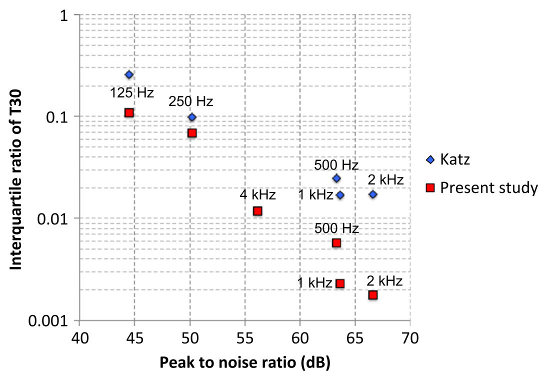

Fig. 6 Interquartile ratio of octave band reverberation time (T_{30}) calculated from the impulse response published by Katz [10], shown in relation to PNR. Interquartile ratio is the interquartile range divided by the median, and so is a measure of dispersion relative to the median. Results are shown for the present study and for Katz's study. PNR values were measured using a 12th-order octave band zero-phase filter-bank, with the period from 2.5 s to the end of the waveform taken as the noise floor

with the pattern of results for the synthetic impulse responses tested in the present study, which also exhibited increased dispersion at lower PNR (except that the synthetic impulse response's dispersion of results is smaller, e.g. IQR of 0.01 s for PNR of 45 dB). However, it must be borne in mind that for Katz's impulse response, PNR is correlated with octave band centre frequency, which can plausibly contribute to standard deviation of T_{30} in itself, due to the increased irregularity of the reverse-integrated decay at lower frequencies (this irregularity was avoided in the synthetic impulse responses by using decaying tones instead of decaying noise). Apart from the interquartile ratio result for the 4 kHz band (which is not available from Katz's paper), there is nothing to disentangle the influence of PNR from the influence of frequency in this real impulse response.

4 Conclusions

The main conclusion of these tests is to confirm that the tested software mostly provides good reverberation time results. This is confirmed for noise-free impulse responses, including very short reverberation times—the results for noise-free impulse responses were essentially in perfect agreement with the synthesised decay. Good performance is also confirmed for when a steady noise floor is added, even at peak-to-noise ratio (PNR) values significantly lower than required for standard analysis. With noise present, results deviated noticeably from the ideal, but were all well within the generally accepted 5 % error margin at the required 45 dB signal-to-noise ratio for standard T_{30} analysis, and mostly within 1 %. A further positive indicator is that the results for the real recorded impulse response show strong agreement in most bands, and substantially reduced variability compared to a software comparison conducted more than one decade prior to this one. Nevertheless, the range of results still exceeded ±5 % from the median in the low frequency bands (including bands for which the signal-to-noise ratio requirement for T_{30} was met), and this indicates that improvements in the software would be useful. Now that indirect methods for measuring impulse responses are well developed and widely implemented, it is usually possible to achieve high PNR values in room acoustics measurements (with appropriate equipment and expertise), certainly higher than those in a balloon pop recording. Considering that PNR is clearly one of the sources of error and deviation between software, current indirect methods should result in reduced variability in reverberation time calculated from well-made impulse responses.

There is evidently some room for improvement, arguably in all of the software implementations tested, including those of the authors. Mostly these improvements would yield minor increases in robustness, or reductions of errors that are already well within acceptable margins. This paper has focussed especially on filter selectivity, and results support the suggestion by Venturi et al. [21] that highly selective octave band filters improve results when reverberation time varies between adjacent bands. Implemented as reverse-time filters, a highly selective octave band filter-bank did not have any unwanted side effects in the tests performed. That is not to say that a 24th-order octave band filter-bank is necessary—results indicate that a significantly lower order filter-bank will return negligible errors for a 2:1 reverberation time ratio between adjacent bands (with energy at the centre of each band). Filter artefacts seem to be evident in the centre time values returned by some of the software implementations tested, including the ones with relatively low frequency selectivity. This suggests that there is room for improvement in filter-bank design in some of the external software implementations tested, not only with regard to selectivity, but also with regard to time-response—and that improving one will not necessarily damage the other.

This study is complemented by other room acoustics measurement software comparison studies that use more realistic measured impulse response sets—for example full sets of measurements within an auditorium, allowing spatial statistics to be examined. Both types of studies are limited, but together they provide guidance on software vulnerabilities and the degree of uncertainty.

References

1. Schroeder, M.: New method of measuring reverberation time. J. Acoust. Soc. Am. **37**, 409–412 (1965)

2. Chu, W.: Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. J. Acoust. Soc. Am. **63**, 1444–1450 (1978)
3. Jacobsen, F., Rindel, J.: Time reversed decay measurements. J. Sound Vib. **117**, 87–190 (1987)
4. Morgan, D.R.: A parametric error analysis of the backward integration method for reverberation time estimation. J. Acoust. Soc. Am. **101**, 2686–2693 (1997)
5. Hak, C., Wenmaekers, R., van Luxemburg, L.: Measuring room impulse responses: impact of the decay range on derived room acoustic parameters. Acta Acust. United Acust. **98**, 907–915 (2012)
6. Guski, M., Vorländer, M.: Comparison of noise compensation methods for room acoustic impulse response evaluations. Acta Acust. United Acust. **100**, 320–327 (2014)
7. Lundeby, A., Vigran, T., Bietz, H., Vorländer, M.: Uncertainties of measurements in room acoustics. Acustica **81**, 344–355 (1995)
8. Bradley, J.S.: An international comparison of room acoustic measurement systems. Institute for Research in Construction, IRC Internal Report IRC-IR-714 (1996)
9. Ikegame, M., Uchida, K.: Comparison of various types of acoustic indexes calculated from the same impulse response (in Japanese). In: Proceedings of the Symposium on Impulse Response Measurements and Evaluation of Acoustic Indexes, pp. 7–13, Architectural Institute of Japan (2001)
10. Katz, B.F.G.: International round robin on room acoustical impulse response analysis software 2004. Acoust. Res. Lett. Onl. **5**, 158–164 (2004)
11. Tronchin, L., Tarabusi, V. and Bortilitti, A.: On the uncertainties of calculations of acoustical parameters in room acoustics measurements. In: Proceeding of 6th AMTA/WSEAS Conference, Buenos Aires (2005)
12. Ţopa, M.D., Toma, N., Kirei, B.S., Homana, I., Neag, M., De Mey, G.: Comparison of different experimental methods for the assessment of the room's acoustics. Acoust. Phys. **57**(2), 199–207 (2011)
13. Machín, J., Arregui, A., San Martín, R., Arana, M.: Areto 1.0: Software para la caracterización acústica de recintos (Areto 1.0: Software for the characterization of room acoustics). In: Techniacustica 2013: 44° Congresso Español de Acústica, Valladolid (2013)
14. Mansilla, J., Sato, S.I., Rodiño, L., Arias, A., Masson, F., De Palma, I.C.: Comparison of room acoustic parameters calculated by different types of software. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings. Institute of Noise Control Engineering vol. **250**(3), pp. 3333–3343 (2015)
15. Álvarez-Morales, L., Galindo, M., Girón, S., Zamarreño, T., Cibrián, R.M.: Acoustic characterisation by using different room acoustics software tools: a comparative study. Acta Acust. United Acust. **102** (2016) (in press)
16. Jacobsen, F.: A note on acoustic decay measurements. J. Sound Vib. **115**, 163–170 (1987)
17. Kob, M., Vorländer, M.: Band filters and short reverberation times. Acta Acust. United Acust. **86**, 350–357 (2000)
18. International Organization for Standardization: ISO 3382–1 Acoustics-Measurement of Room Acoustic Parameters Part 1. Geneva (2009)
19. International Electrotechnical Commission: IEC 61260 Ed. 1, Electroacoustics-Octave-band and fractional-octave-band filters, International Electrotechnical Commission, Geneva (1995)
20. International Electrotechnical Commission: IEC 61260-1 Ed. 1, Electroacoustics-Octave-band and fractional-octave-band filters-Part 1: Specifications, International Electrotechnical Commission, Geneva (2014)
21. Venturi, A., Farina, A., Tronchin, L.: On the effects of preprocessing of impulse responses in the evaluation of acoustic parameters on room acoustics. J. Acoust. Soc. Am. **133**, 3224 (2013). (POMA 19, 015006)
22. Faiget, L., Legros, C., Ruiz, R.: Optimization of the impulse response length: application to noisy and highly reverberant rooms. J. Audio Eng. Soc. **46**, 741–750 (1998)
23. Xiang, N.: Evaluation of reverberation times using a nonlinear regression approach. J. Acoust. Soc. Am. **98**, 2112–2121 (1995)
24. ITA-Toolbox: A MATLAB Toolbox for the Needs of Acousticians. Institute of Technical Acoustics, RWTH Aachen University, Aachen (2013) http://www.ita-toolbox.org/
25. Cabrera, D., Jimenez, D., Martens, W.L.: Audio and Acoustical Response Analysis Environment (AARAE): a tool to support education and research in acoustics. In: Proceedings of Internoise, Melbourne (2015)
26. Cabrera, D., Lee, D., Leembruggen, G., Jimenez, D.: Increasing robustness in the calculation of the speech transmission index from impulse responses. Build. Acoust. **21**, 181–198 (2014)
27. International Electrotechnical Commission: IEC 50268-16 Ed. 4, Sound System Equipment-Part 16: objective Rating of Speech Intelligibility by Speech Transmission Index. International Electrotechnical Commission, Geneva (2011)
28. International Organization for Standardization: ISO 3382 Acoustics-Measurement of the reverberation time of rooms with reference to other acoustical parameters. Geneva (1997)