EVALUATING ADAPTIVE REVERBERATION TIME ESTIMATION FOR NON-STANDARD CONDITIONS

Aron Karsai[1], Leon Kaiser[1], Matthias Frank[1], Benjamin Pohler[1], Alois Sontacchi[1]

[1] *Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria*

Abstract

The practices of measuring the impulse response (IR) of a room and obtaining the reverberation time (RT) are standardized in ISO 3382-2. It specifies the use of an approximately omnidirectional sound source and multiple microphone positions (and, where appropriate, multiple source positions) to obtain spatially averaged room acoustic parameters. The standard also specifies minimum distances between the sound source, microphones, and reflecting surfaces to avoid near-field effects and measurement bias. This ensures that the direct sound and the non-diffuse decay of early reflections do not affect the RT measurement.
However, these criteria may be difficult to fulfill in practical scenarios, e.g., when measuring with directional loudspeakers or in smaller rooms. To enable RT measurement under such non-standard circumstances, this study investigates approaches that modify the evaluation range of the linear regression used to calculate the RT from the energy decay curve. We inspect modifications based on the degree of non-linearity of the regression as well as adapting the range based on a diffuse field assumption utilizing the loudspeaker-microphone distance, room volume and an estimate of the RT.
An evaluation across six acoustic scenarios indicates that these approaches can reduce the error under certain configurations. Specifically, the distance based approach improves the measurement more often than not.

Keywords: *reverberation time, adaptive estimation, room acoustics*

1. Introduction

The RT is a fundamental parameter for evaluating room acoustics, yet standard estimation methodologies face real-world challenges. Theoretical predictive models often require rigorous boundary calibration to align with empirical data [1]. Furthermore, physical measurements are highly vulnerable to environmental noise. Capturing valid room impulse responses requires advanced techniques to filter out transient disturbances [2], while standard post-processing relies on complex, automated truncation to prevent stationary noise from compromising decay integration [3]. Also, the physical dimensions of the room can pose challenges to set up the loudspeakers and microphones for the measurement in a standard-conform manner. It may also occur that the equipment available for the measurement does not comply with the standard, while measurements still need to be performed using that equipment. Collectively, these vulnerabilities highlight the need for a modified, robust approach to standard RT estimation, and the fact that it is a topic of active research.
This paper focuses on the problems that are caused by inadequate loudspeaker-microphone distances. We provide solutions to these problems via different methods, which are then evaluated by measurements. The last one of these methods also provides a possible solution to the problem that is caused by directed loudspeakers. However, the evaluation of the solution to this problem is not the topic of this paper.
The remainder of this paper is organized as follows. First, we formulate the problem that is the main focus of this paper in Section 2. Afterwards, we provide a brief review of RT measurement according to the standard in Section 3. We then introduce our novel methods for adaptive RT estimation in non-standard cases in Section 4. Finally, we evaluate the proposed methods using measurements described in Section 5, and present and discuss the results in Section 6.

2. Problem Formulation

According to the standard [4], the RT of a room is measured using the impulse response (IR) of the room. In the time domain, the IR can be described as consisting of three parts: the direct sound, the early reflections, and the late reverberation. The mixing time is the

Corresponding author: *a.karsai@kug.ac.at.*

point in time that separates the deterministic early reflections from the stochastic late reverberation of the specific room IR [5]. To measure the RT, the standard makes the specific assumption about the direct sound range of the IR that it contains a limited amount of energy compared to the rest of the IR. This is specified in more detail in Subsection 3.1.

The impact of the direct sound can be characterized by the distance-based direct-to-reverberant ratio (DRR). This is for a point source – using the theoretical energy of a diffuse sound field – in dB is [5]:

$$\mathrm{DRR}(d, Q, T_{60}, V) = 10\log_{10}\left(\frac{Q/4\pi d^2}{cT_{60}/13.8V}\right), \quad (1)$$

where Q is the directivity factor of the loudspeaker, d is the loudspeaker-microphone distance, c is the velocity of sound in air[1], T_{60} is the RT and V is the room volume.

The DRR expresses how much of the energy at the microphone is the result of the direct sound compared to the reverberant, diffuse decaying sound field. From Equation (1) we can see, that increasing Q leads to and increased DRR. To this end, the standard requires to have a loudspeaker with an omnidirectional characteristic, as this corresponds to the minimum $Q = 1$ value. Furthermore, in the same fashion we can also observe that decreasing d also results in a higher DRR. Therefore, the standard requires to have a minimum distance between the loudspeaker and the microphone $d_{\min}$, defined as:

$$d_{\min} = 2\sqrt{\frac{V}{c\hat{T}}}, \quad (2)$$

where $\hat{T}$ is some approximation of the RT. Since the RT is generally frequency-dependent, $\hat{T}$ is usually some aggregated value of it. In this study, $\hat{T}$ is estimated by the average of the RT values in the octave bands with center frequencies 500 Hz and 1 kHz. We specifically focus on the non-standard measurement setups, where the criterion of $d_{\min}$ is not fulfilled.

3. Standard RT Measurement Methods

3.1. Preprocessing of the IR

Let $h[k]$ be the sampled IR for $k = 0, 1, ..., K-1$. According to [4], the RT in each octave band of interest is obtained by filtering the IR using third-order Butterworth bandpass filters, with center frequencies corresponding to the respective octave bands and a -3 dB bandwidth of one octave. Then, the samples of the energy decay curve (EDC) are calculated in dB from each filtered IR using Schröder's backwards integration [6]. It is also normalized so that at the start of the IR at $k = 0$ the remaining energy is 0 dB.

$$\mathrm{EDC}[k] = 10\log_{10}\left(\frac{\sum_{l=k}^{K-1} h^2[l]}{\sum_{l=0}^{K-1} h^2[l]}\right) \quad (3)$$

[1]We approximate the velocity of sound by $c \approx 343\ \frac{\mathrm{m}}{\mathrm{s}}$.

As the methods adapting the regression range themselves are independent of frequency band, we provide their descriptions based on a hypothetical general EDC in Subsection 3.2 as well as in Section 4. They can then be applied to all actual EDCs corresponding to each frequency band of interest.

The RT of a room in a given frequency band is defined as the duration between turning off the loudspeaker and the point in time at which the spatially averaged sound energy density has decreased by 60 dB. Measuring the RT over a full 60 dB decay range requires a SNR of 60 dB, which is rarely achievable in practice. To address this limitation, standard evaluation methods analyze a smaller dynamic range of the EDC and extrapolate the measured slope to determine the final RT. These extrapolated measurements are called T_{20} if extrapolated from the range where $\mathrm{EDC}[k] \in [-25, -5]$ and T_{30} if extrapolated from the range where $\mathrm{EDC}[k] \in [-35, -5]$, respectively. Here, the conditions leading to a diffuse sound field from at least -5 dB in the EDC are assumed, i.e. the mixing time has to be shorter than the time it takes for the EDC to reach -5 dB. The dB scale transforms the exponential decay of the diffuse IR to linear decay on the EDC. Additionally, the standard states that if these extrapolation techniques are to be used, there should be a 10 dB buffer between the lower border of extrapolation range and the noise floor. Since the SNR is the negated noise floor, the SNR should be at least 35 dB and 45 dB for T_{20} and T_{30}, respectively.

In Section 4, we focus on approaches based on 20 dB ranges, as transferring these concepts to a 30 dB range is trivial. Furthermore, we loosen the rather strict criterion of the 10 dB buffer to 3 dB in our modified approaches.

3.2. Standard T_{20} Measurement

Before introducing our novel algorithms for adaptive T_{20} estimation, we provide a short summary of the T_{20} measurement method standardized in [4].

To select the specified region for T_{20}, the ordered index set $I = \{k \mid -5 \geq \mathrm{EDC}[k] \geq -25\} = \{k_1, k_2, ..., k_n\}$ is constructed, where $n = |I|$. Then, the samples L_i from this theoretically linear region of the EDC for $i = 1, 2, ..., n$ are selected:

$$L_i = \mathrm{EDC}[k_i]. \quad (4)$$

Since this region of the EDC is only perfectly linear in theory, linear regression is used to estimate its slope. Formally, we want the estimate

$$\hat{L}_i = a + bk_i, \quad (5)$$

where

$$b = \frac{\sum_{i=1}^{n}(k_i L_i) - n\overline{k}\overline{L}}{\sum_{i=1}^{n}(k_i^2) - n\overline{k}^2}, \qquad a = \overline{L} - b\overline{k}, \quad (6)$$

and $\overline{L} = 1/n \sum_{i=1}^{n} L_i$ and $\overline{k} = 1/n \sum_{i=1}^{n} k_i$. Then, the extrapolated RT measurement T_{20} is given by:

$$T_{20} = \frac{-60}{b}. \tag{7}$$

To quantify the quality of the linear regression and therefore the T_{20} measurement, the standard also provides the degree of nonlinearity ξ. Its value is derived from the correlation coefficient r^2, defined as

$$r^2 = \frac{\sum_{i=1}^{n}(\hat{L}_i - \overline{L})^2}{\sum_{i=1}^{n}(L_i - \overline{L})^2}. \tag{8}$$

The value of the correlation coefficient lies between 0 and 1, where $r^2 = 1$ corresponds to a perfect linear fit. Then, degree of nonlinearity is defined as the difference to a perfect linear fit in per mille:

$$\xi = (1 - r^2) \cdot 1000. \tag{9}$$

This relatively simple method of T_{20} measurement is assumed to work well when the conditions for the microphone and loudspeaker positions as well as $Q = 1$ for the loudspeaker directivity specified in the standard are met.

4. Adaptive RT Measurement Methods

In the case of an increased direct sound energy, it might happen that at -5 dB the EDC is still showing the impact of the direct sound instead of diffuse decay. This then in turn increases ξ, so the quality of the linear regression and therefore the T_{20} measurement diminishes. This effect is illustrated in Figure 1. In the following, we present novel algorithms to modify the range $[-25, -5]$ of the linear regression to calculate T_{20} in order to avoid the this influence of the direct sound.

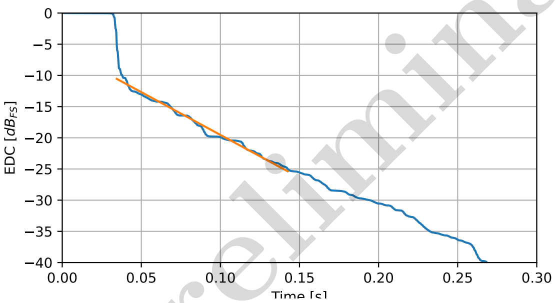


Figure 1: EDC with the direct sound affecting the linear regression in the range of [-25, -5] dB.

4.1. Range Selection Based on the Grid-Search Optimal Linear Fit

As the first approach, an intuitive idea to determine a modified range for the linear regression is to examine the value of ξ. and estimate the RT from the range of the EDC, where – after multiple estimates were obtained from a grid of different ranges – ξ is the lowest. The resulting RT estimate is then called T_{uw}, where the subscript 'uw' stands for unweighted linear regression.

Let $I_j = \{k \mid -5 - 0.5j \geq \mathrm{EDC}[k] \geq -25 - 0.5j\} = \{k_{1,j}, k_{2,j}, ..., k_{n_j,j}\}$ with $n_j = |I_j|$ and $j \in \{0, 1, ..., J\}$ be now a series of index sets, and the series of samples $L_{i,j}$ selected from the theoretically linear region of the EDC for $i = 1, 2, ..., n_j$ be selected as:

$$L_{i,j} = \mathrm{EDC}[k_{i,j}]. \tag{10}$$

This results in a series of ranges selected from the EDC, where $L_{i,0} = L_i$ and starting with $L_{i,1}$ onward the upper and lower borders of the ranges are shifted down by 0.5 dB $j = J$. The value of J is determined from the 3 dB buffer and the SNR, such that the lowest of the lower borders of the ranges is not lower than the noise floor +3 dB or -40 dB. Hence:

$$J = \lfloor -50 - 2\min(-\mathrm{SNR} + 3, -40) \rfloor. \tag{11}$$

Afterwards, linear regression is applied to each $L_{i,j}$ using equations (5)–(9), and the values $T_{\mathrm{uw},j}$ and $\xi_{\mathrm{uw},j}$ are obtained. Then, T_{uw} is chosen from the series of $T_{\mathrm{uw},j}$ values where the corresponding $\xi_{\mathrm{uw},j}$ is the lowest, i.e.:

$$T_{\mathrm{uw}} = \arg\min_{T_{\mathrm{uw},j}} \ \xi_{\mathrm{uw},j}(T_{\mathrm{uw},j}). \tag{12}$$

4.2. Range Selection Based on the Grid-Search Optimal Weighted Linear Fit

The second approach is almost identical to the first one. However, greater emphasis is placed on EDC values near the upper boundary of the linear regression range. This ensures that deviations from linearity caused by increased direct sound energy are penalized more strongly by the degree of nonlinearity parameter. To achieve this, the conventional linear regression is replaced by a weighted linear regression, assigning higher weights to samples closer to the upper boundary of the range. The result is then named T_{w} where the subscript 'w' refers to the weighted linear regression. As we also did in the previous method, we construct the series of index sets I_j for $j = 0, 1, ..., J$ where J is given by (11) and select the series of samples from the EDC $L_{i,j}$ according to (10). However, the linear regression and, consequently, all expressions derived from it have to be adapted to account for the weighting. We define the weights $w_{i,j}$ for each EDC sample $L_{i,j}$ in each range j to be the sample value $L_{i,j}$ in linear units. I.e., $w_{i,j} = 10^{L_{i,j}/10}$, thereby ensuring that samples towards the start of the range have higher weights than the ones towards the end. Then, the linear regression model is:

$$\hat{L}_{\mathrm{w},i,j} = a_{\mathrm{w},j} + b_{\mathrm{w},j} k_{i,j}, \tag{13}$$

where

$$b_{\mathrm{w},j} = \frac{\sum_{i=1}^{n_j}(w_{i,j} k_{i,j} L_{i,j}) - W_j \overline{k}_{\mathrm{w},j} \overline{L}_{\mathrm{w},j}}{\sum_{i=1}^{n_j}(w_{i,j} k_{i,j}^2) - W_j \overline{k}_{\mathrm{w},j}^2},$$
$$a_{\mathrm{w}_j} = \overline{L}_{\mathrm{w},j} - b_{\mathrm{w},j} \overline{k}_{\mathrm{w},j}, \tag{14}$$

and $W_j = \sum_{i=1}^{n_j} w_{i,j}$ and $\overline{L}_{\mathrm{w},j} = {}^1/_{W_j} \sum_{i=1}^{n_j} w_{i,j} L_{\mathrm{w},i,j}$ and $\overline{k}_{\mathrm{w},j} = {}^1/_{W_j} \sum_{i=1}^{n_j} k_{i,j}$. Then, $T_{\mathrm{w},j}$, $r^2_{\mathrm{w},j}$ and $\xi_{\mathrm{w},j}$ are calculated as:

$$\begin{aligned} T_{\mathrm{w},j} &= \frac{-60}{b_{\mathrm{w},j}}, \qquad \xi_{\mathrm{w},j} = (1 - r^2_{\mathrm{w},j}) \cdot 1000, \\ r^2_{\mathrm{w},j} &= \frac{\sum_{i=1}^{n_j} w_{i,j} (\hat{L}_{\mathrm{w},i,j} - \overline{L}_{\mathrm{w},j})^2}{\sum_{i=1}^{n_j} w_{i,j} (L_{i,j} - \overline{L}_{\mathrm{w},j})^2}. \end{aligned} \tag{15}$$

Afterwards, T_w is chosen from the series of $T_{\mathrm{w},j}$ values according to (12), swapping the subscripts 'uw' appropriately to 'w'.

4.3. Range Selection Based on Microphone-Loudspeaker Distance

For this last approach, instead of trying to minimize the non-linearity of the linear regression, we utilize the physical parameters of the measurement to provide a better estimate for T_{20} in non-standard conditions. Specifically, we utilize the microphone-loudspeaker distance-based DRR to modify the borders of the linear regression range.

Since we want to counteract the increased direct sound energy possibly interfering with the linear regression around -5 dB in the EDC, we shift the evaluation range of the regression by how much the direct sound energy theoretically increased. To this end, we define

$$\Delta(d, d_{\min}, Q, \hat{T}, V) = \begin{cases} \mathrm{DRR}(d, Q, \hat{T}, V) - \\ \mathrm{DRR}(d_{\min}, 1, \hat{T}, V), & \text{if } d < d_{\min}, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

which corresponds to the increase in the direct sound energy if the measurement is made at a distance d shorter than $d_{\min}$. Here, the first term is the DRR of the actual measurement setup, using the actual loudspeaker-microphone distance and the actual directivity factor of the loudspeaker. The second term is the theoretical DRR of the perfectly standard-conform measurement setup at the minimum distance and using $Q = 1$. Note that $\hat{T}$ is used, which is also used in the estimation of $d_{\min}$ according to (2). Since $\hat{T}$ is unknown before the RT measurement, we use the T_{20} measurement from the standard method as an estimate of the RT in the 500 Hz and 1 kHz octave bands to calculate $\hat{T}$.

The modification of the linear regression range is then straightforward. Both bounds of the linear regression range are lowered by the estimated increase in direct sound energy. In this way, we can be sure that linear regression is applied in a region of the EDC where the direct sound energy has already vanished, and there is a diffuse sound field. We define now the ordered index set $I_\mathrm{d} = \{k \mid -5 - \Delta \geq \mathrm{EDC}[k] \geq -25 - \Delta\} = \{k_{\mathrm{d},1}, k_{\mathrm{d},2}, ..., k_{\mathrm{d},n_\mathrm{d}}\}$ where $n_\mathrm{d} = |I_\mathrm{d}|$. Then, the samples $L_{i,d}$ from the theoretically linear region of the EDC for $i = 1, 2, ..., n_d$ are selected as:

$$L_{\mathrm{d},i} = \mathrm{EDC}[k_{\mathrm{d},i}]. \tag{17}$$

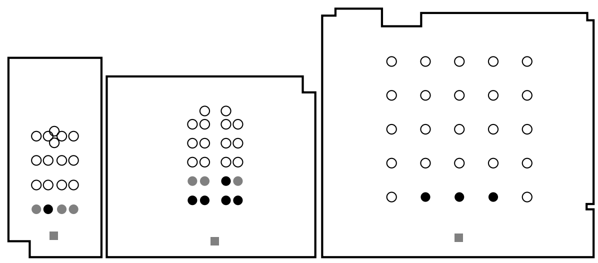
Afterwards, the linear regression is applied once again according to equations (5)–(9), and T_d is obtained.

5. Measurement Setup and Evaluation

5.1. Acoustic Scenarios and IR Measurement

For the evaluation of the proposed methods, we measured the IRs of 3 different rooms. Each room was measured with and without acoustic treatment, resulting in a total of 6 different acoustic scenarios. The 3 rooms had varying sizes and volumes. *Room 1* was a small room with a volume of 54 m³, *Room 2* was a medium-sized room with a volume of 116 m³, and *Room 3* was a large room with a volume of 184 m³. The IRs were obtained using the exponential sine sweep method described in [7]. The measurements were conducted in the 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz octave bands. Below 500 Hz, the assumption of a homogeneous sound field can no longer be made, which makes a comparison between different positions infeasible. In each measurement, we used a 1/7/1 loudspeaker as a source, described in [8]. The loudspeaker has a radiation characteristic which is well approximated by an omnidirectional characteristic in the frequency bands of interest; therefore $Q = 1$ for all measurements. An array of NTi Class 1 omnidirectional microphones (M2230) was also placed in each room. The loudspeaker and the microphones were positioned at the typical seated ear height of 122 cm.

First, the loudspeaker was placed in the front of each room and the microphone array was set up. This is depicted in Figure 2, where the shapes and shades indicate which microphones were placed according to the standard.



Room 1 **Room 2** **Room 3**

Figure 2: Measurement setups with loudspeaker in front position. Grey square: loudspeaker. Hollow circles: microphones placed further than $d_{\min}$ in both the treated and the untreated case. Grey circles: microphones placed further than $d_{\min}$ in only the untreated case. Black circles: microphones placed closer than $d_{\min}$ for both the treated and the untreated case.

Afterwards, the loudspeaker was placed in the middle of the microphone array to obtain more measurements with distances less than $d_{\min}$. For *Rooms 2* and *3*, 11 of these measurements were conducted. For *Room 1*, there was only one measurement with the center loudspeaker position. The setup is depicted in Figure 3, where once again each microphone is distinguished by shade to indicate the relation of their distance to the loudspeaker compared to d_{min}.

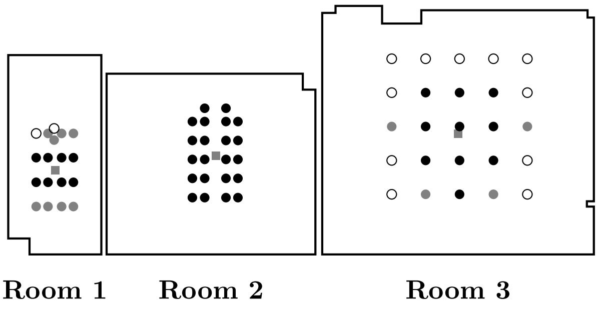


Figure 3: Measurement setups with loudspeaker in center position. Grey square: loudspeaker. Hollow circles: microphones placed further than d_{min} in both the treated and the untreated case. Grey circles: microphones placed further than d_{min} in only the untreated case. Black circles: microphones placed closer than d_{min} for both the treated and the untreated case.

5.2. Evaluation of the RT Estimation Methods

After acquiring the IRs, the EDCs and T_{20} values were calculated from each IR for each band of interest. If a T_{20} estimation could not be conducted because of the lack of the 10 dB buffer, the measurement was excluded. Then, the data was split into two sets. The first one called the ground truth (GT) set contains all the measurements from the setup of Figure 2, where the microphone was further from the loudspeaker than d_{min}. The second set called the non-standard set contains all the measurements from the setups of Figures 2 and 3, where the microphone was closer to the loudspeaker than d_{min}. The GT set is used to evaluate the performance of the modified RT estimation methods for non-standard circumstances. Both sets were then subdivided into measurements from each frequency band f and acoustic scenario a. First, the assessment of the GT set was performed. For each frequency band, the T_{20} measurements arising from the same acoustic scenario showed variance. Upon observing their histograms [9] and Q-Q plots [10], we assumed the GT values to be normally distributed for each band and acoustic scenario. Figure 4 shows the histogram of GT T_{20} values for the 500 Hz octave band in *Room 2* in the treated case.

To verify the assumption of normality, we conducted Shapiro-Wilk tests [11] in each band for each acoustic scenario. For the chosen significance level of $\alpha = 0.05$, the results of the tests showed that the null hypothesis of normality could not be rejected for most cases. However, in the cases of untreated *Room 2* and treated *Room 3* in the 500 Hz octave band, as well as of untreated *Room 1* and treated *Room 2* in the 1 kHz octave band, the p-values were below α,

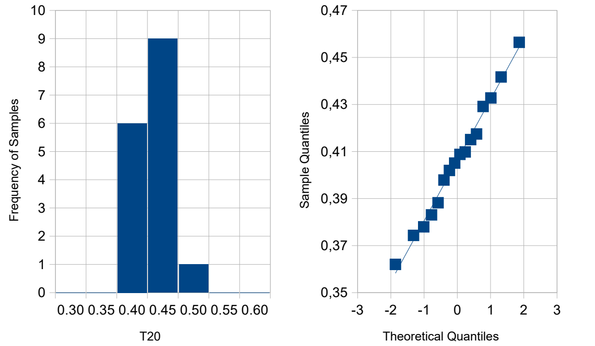


Figure 4: Histogram of GT T_{20} values for the 500 Hz octave band in Room 2 in the treated case.

indicating that the null hypothesis of normality should be rejected. Therefore, we excluded these cases from the evaluation of the modified RT estimation methods. Afterwards, the values T_{uw}, T_{w}, and T_{d} were calculated additionally for each measurement in the non-standard set. Measurements for which the required values could not be calculated due to the missing 3 dB buffer were excluded. To evaluate the consistency of these values and T_{20} with the GT distributions across octave bands and acoustic scenarios, we calculated the log-likelihood of each subset of non-standard measurements under the corresponding GT distribution, defined by its mean and variance. This was done separately for the values T_{20}, T_{uw}, T_{w}, and T_{d} for each subset. The log-likelihood of a specific method $\log L_{\mathrm{method},a,f}$ for acoustic scenario a and frequency band f was calculated as follows:

$$\log L = \sum_{s=1}^{S_{a,f}} \log \mathcal{N}(T_{\mathrm{method},s} | \mu_{a,f}, \sigma^2_{a,f}), \qquad (18)$$

where $S_{a,f}$ is the number of measurements in the subset of non-standard measurements for acoustic scenario a and frequency band f, $T_{\mathrm{method},s}$ is the value of the method under consideration for measurement s, and $\mu_{a,f} = {}^1/_{S_{a,f}} \sum_{s=1}^{S_{a,f}} T_{\mathrm{GT},s}$ and $\sigma^2_{a,f} = {}^1/_{(S_{a,f}-1)} \sum_{s=1}^{S_{a,f}} (T_{\mathrm{GT},s} - \mu_{a,f})^2$ are the sample mean and variance of the GT values for acoustic scenario a and frequency band f. The function $\mathcal{N}(x|\mu, \sigma^2) = {}^1/\sqrt{2\pi\sigma^2} e^{-((x-\mu)^2/2\sigma^2)}$ is the probability density function (PDF) of the normal distribution.

6. Results and Conclusion

After carrying out the measurements and evaluations as described in the previous section, the results of the log-likelihood calculations for each method, acoustic scenario, and frequency band are shown in Table 2. There is no frequency band where any of the modified RT estimation methods outperforms T_{20} in all acoustic scenarios. However, a tendency can be observed, where the higher the frequency band, the more times T_{uw} and T_{w} outperform T_{20}. Still, T_{d} shows the best performance overall, outperforming or closely match-

500 Hz	T_{20}	T_{w}	T_{uw}	T_{d}
Room 1 – Treated	-28.5	-34.2	-34.4	-30.1
Room 1 – Untreated	-16.7	-20.9	-21.7	**-12.9**
Room 3 – Treated	excluded	excluded	excluded	excluded
Room 3 – Untreated	-148.0	-180.7	-155.0	**-134.0**
Room 2 – Treated	-376.7	-674.6	-586.9	-504.5
Room 2 – Untreated	excluded	excluded	excluded	excluded
Sum	-569.9	-910.4	-798.1	-681.4
1000 Hz	T_{20}	T_{w}	T_{uw}	T_{d}
Room 1 – Treated	-51.3	-88.1	-90.7	**-40.9**
Room 1 – Untreated	excluded	excluded	excluded	excluded
Room 3 – Treated	-195.3	-228.0	-225.6	**-184.4**
Room 3 – Untreated	-218.5	**-190.6**	**-186.1**	**-145.8**
Room 2 – Treated	excluded	excluded	excluded	excluded
Room 2 – Untreated	-285.7	-312.4	-295.3	**-253.1**
Sum	-750.7	-819.1	-797.7	**-624.1**
2000 Hz	T_{20}	T_{w}	T_{uw}	T_{d}
Room 1 – Treated	-50.0	-86.0	-70.5	-51.5
Room 1 – Untreated	-6.5	-7.2	**-6.4**	-7.1
Room 3 – Treated	-243.5	-345.6	-332.1	**-219.8**
Room 3 – Untreated	-357.0	**-331.4**	**-305.5**	**-350.4**
Room 2 – Treated	-610.7	-625.9	**-548.2**	**-445.9**
Room 2 – Untreated	-307.5	-350.4	-340.0	**-264.8**
Sum	-1575.1	-1746.4	-1602.7	**-1339.5**
4000 Hz	T_{20}	T_{w}	T_{uw}	T_{d}
Room 1 – Treated	-35.4	-43.2	-42.5	**-30.0**
Room 1 – Untreated	-9.0	-8.9	**-11.7**	-9.8
Room 3 – Treated	-291.9	**-267.0**	**-256.9**	**-224.3**
Room 3 – Untreated	-167.1	-187.5	-237.5	-398.9
Room 2 – Treated	-461.0	**-424.8**	**-389.1**	**-360.2**
Room 2 – Untreated	-281.5	-308.4	-316.2	**-266.1**
Sum	-1245.9	**-1239.8**	-1253.9	-1289.2
8000 Hz	T_{20}	T_{w}	T_{uw}	T_{d}
Room 1 – Treated	-41.8	-50.4	-53.1	-45.2
Room 1 – Untreated	-18.7	-49.1	-37.0	**-17.8**
Room 3 – Treated	-784.7	**-397.6**	**-375.9**	**-249.8**
Room 3 – Untreated	-280.4	**-166.5**	**-148.4**	**-141.0**
Room 2 – Treated	-930.6	**-769.0**	**-794.3**	**-487.6**
Room 2 – Untreated	-1099.1	**-631.7**	**-682.6**	**-506.9**
Sum	-3155.2	**-2064.3**	**-2091.2**	**-1448.1**

Table 2: Log-likelihood values of the different RT estimation methods for the non-standard measurements given the distributions of the GT values. Separate for each frequency band and acoustic scenario. Gray background: Standard T_{20}. Bold font: Better alignment with GT than T_{20}.

ing T_{20} in all frequency bands above 500 Hz. The only outlier is in the 4 kHz band in the case of untreated *Room 3*. In the 500 Hz band, all modified methods perform worse than T_{20}, except for T_{d} in untreated *Room 1* and *Room 3*. This is likely due to room modes influencing the EDCs in this lower frequency range, so the assumption of a homogeneous sound field is partially violated even here.

Overall, the results indicate that the modified RT estimation methods can provide better estimates of T_{20} in non-standard measurement conditions, especially at higher frequencies. T_{d}, in particular, shows promise as a robust alternative to the standard T_{20} estimation method in the scenarios where the microphone is placed closer to the loudspeaker than the recommended d_{min}. Studying the outliers and the specific conditions under which the modified methods fail could provide further insights into their limitations and potential improvements.

Another possible future research direction could be to investigate the performance of the modified RT estimation methods in other non-standard measurement conditions, such as having a directed loudspeaker, as the opportunity to include the loudspeaker directivity is already present in the method for T_{d}. However, in this case a shift in the mixing time due to an altered reflection density depending on the radiation characteristics is also expected.

7. Acknowledgments

Financial support by the Austrian Research Promotion Agency (FFG) (Grant No. 66369593 - Room Imagineer) and the cooperation with BETTERMEETINGROOMS GmbH are gratefully acknowledged.

References

[1] K. Prawda, S. J. Schlecht, and V. Välimäki, “Calibrating the sabine and eyring formulas,” *The Journal of the Acoustical Society of America*, vol. 152, no. 2, pp. 1158–1169, 2022.

[2] K. Prawda, S. Schlecht, and V. Välimäki, “Rule of two: Methods to detect and localize nonstationary noise in sweep measurements,” *methods*, vol. 2, p. 5, 2023.

[3] M. Janković, D. G. Ćirić, and A. Pantić, “Automated estimation of the truncation of room impulse response by applying a nonlinear decay model,” *The Journal of the Acoustical Society of America*, vol. 139, no. 3, pp. 1047–1057, 2016.

[4] *Acoustics – measurement of room acoustic parameters*, ISO, 3382-2, 2009.

[5] S. Weinzierl, *Handbuch der Audiotechnik*. Springer Science & Business Media, 2008.

[6] *Acoustics – measurement of room acoustic parameters*, ISO, 3382-1, 2009.

[7] A. Farina, “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” *Journal of The Audio Engineering Society*, 2000. [Online]. Available: `https://api.semanticscholar.org/CorpusID:9614437`.

[8] J. Moling, “Compact spherical 171 loudspeaker array with increased vertical directivity,” Toningenieurprojekt, Universität für Musik und darstellende Kunst Graz, 2024. [Online]. Available: `https://phaidra.kug.ac.at/o:134675`.

[9] R. Srivastava, “Karl pearson and “applied” statistics,” *Resonance*, vol. 28, pp. 183–189, Feb. 2023. DOI: `10.1007/s12045-023-1542-3`.

[10] M. B. Wilk and R. Gnanadesikan, “Probability plotting methods for the analysis of data,” *Biometrika*, vol. 55, no. 1, pp. 1–17, 1968, ISSN: 00063444, 14643510.

[11] S. S. Shapiro and M. B. Wilk, “An analysis of variance test for normality (complete samples),” *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965, ISSN: 00063444, 14643510.