Effects of Varying Levels of Background Noise on Room Acoustic Parameters, Measured with ESS and MLS Methods

Nikolaos M. Papadakis 1,2,*, Smaro Antoniadou 1 and Georgios E. Stavroulakis 1

1 Institute of Computational Mechanics and Optimization (Co.Mec.O), School of Production Engineering and Management, Technical University of Crete, 73100 Chania, Greece; jd.smaro@gmail.com (S.A.); gestavroulakis@tuc.gr (G.E.S.)
2 Department of Music Technology and Acoustics, Hellenic Mediterranean University, 74100 Rethymno, Greece
* Correspondence: nikpapadakis@tuc.gr

Abstract: Typically, background noise of different types and levels is present during the measurement of the impulse response in spaces. The two methods that are, in practice, most frequently used in the measurement of the impulse response, are the exponential sine sweep (ESS), and the maximum length sequence (MLS). This study’s objective was to estimate the impact of background noise (white noise, tonal noise) on the acoustic parameters (\(T_{30}\), EDT, \(C_{80}\), and \(D_{50}\)) for ESS and MLS measurements, by introducing artificial background noise, employing an external sound source. For this purpose, measurements were performed with varying levels of external noise (in steps of 2 dB), and the effect was assessed, using the relative error compared to measurements without artificial background noise. According to the findings for white noise (as background noise), in the case of \(T_{30}\) and EDT, the difference between the two methods, as well as the relative error, for the initial levels of added background noise, was small. However, for higher levels of added background noise, there was a sharp increase in the relative error, which was greater for the ESS method, both for \(T_{30}\) and EDT. Regarding \(C_{80}\) and \(D_{50}\), while initially the differences between the ESS and MLS methods were small, cumulatively, as the background noise increased, the relative error increased for both methods, with the ESS method showing the largest error. In the case of tonal noise (as background noise), the results were consistent with those observed in the case of white noise. The study’s findings contribute to a better understanding of the ESS and MLS methods, and suggest the expected relative error of acoustic parameters when various types and levels of background noise are present. Additionally, the study suggests, based on background noise and level, the optimum method to conduct impulse response measurements.

Keywords: acoustic measurement; exponential sine sweep; maximum length sequence; room acoustics; ISO 3382-1:2009; impulse response; MLS; ESS; swept-sine; architectural acoustics

1. Introduction

As we are all, consciously or unconsciously, aware of the sounds and acoustic environment we are exposed to every day, satisfactory acoustics are crucial for quality of life. Modeling techniques, in both enclosed [1,2] and open spaces [3,4], as well as related acoustic theories for the accurate calculation of the sound field, have been the subject of numerous significant recent developments. However, sound measurement techniques in existing spaces will always be relevant, as they provide an accurate description of the acoustic field, as well as the information needed for its assessment and optimization.

The impulse response can be used to obtain most of the acoustic parameters that define a space, which in turn are useful for providing guidance for possible improvements. ISO 3382-1 [5] addresses the proper measurement of the impulse response and acoustic parameters in performance spaces. Other standards, such as ISO 3382-2 [6], cover the measurement of ordinary rooms, and ISO 3382-3 [7] of open plan offices. The measurement
of impulse responses is useful in other standards, such as ISO 354:2003 [8], which specifies a method of measuring the sound absorption coefficient, or the equivalent sound absorption area, in a reverberation room. In various domains, including virtual and augmented reality, impulse response measurements are also required for the estimation of head-related transfer functions (HRTFs), which are crucial for auralization applications [9].

In order to simulate the omnidirectional radiation characteristics of a monopole as a sound source for acoustic measurements [10], a dodecahedron speaker is frequently employed, with other sources being available [11]. Alternative excitation signals are used in practice, because impulsive signals applied directly to dodecahedron speakers cannot produce enough acoustic output. The use of these signals ensures that the speaker can produce enough acoustic energy for an extended period of time, without exceeding its peak power limit. Postprocessing can be applied to obtain the impulse responses with high time resolution. In addition, the repeatability of the measurements is improved by the accurate reproduction of these deterministic excitation signals. Some of the most well-known signals are the maximum length sequence (MLS), exponential sine sweep (ESS), inverse repeated sequence (IRS), and time-stretched pulses.

The MLS method was introduced by Schroeder [12]. The method is based on the excitation of the acoustic space by a “periodic pseudorandom signal having almost the same stochastic properties as a pure white noise” [13], and also demonstrating excellent immunity to distortion [14]. It has been demonstrated that the signal-to-noise ratio for the technique increases by 3 dB when the period length of the MLS sequence is doubled [15]. The MLS signal was quite popular in the 1990s, with the use of the first laptops [16]. Some of the method’s practical facets are detailed in [17,18]. With the advantage of lowering distortion peaks, the IRS method was created as an alternative approach to the MLS [19]. The method’s drawback is the lengthier computation time for the deconvolution employing high-order FFT and IFFT filters [20]. Another approach for the measurement of impulse responses is the time-stretched pulses [21]. With the intention of reducing peak distortions, this approach seeks to raise the sound-to-noise ratio. Last but not least, the ESS method [22,23] is based on the notion of using an exponential time-growing frequency sweep, which enables the simultaneous deconvolution of the system’s linear impulse response, and selective separation of each impulse response corresponding to the harmonic distortion orders taken into account. With an ideal excitation signal for the quick measurement of an acoustical impulse response, and even without averaging, the ESS approach aims to solve the majority of the shortcomings in conventional measurement techniques.

Of the above methods, the most commonly used in practice are the MLS and ESS, which are likewise included in Annex A and B of ISO 18233 [24]. Various studies have attempted to compare these two methods. In a study by Stan et al. [13], it is demonstrated that the MLS (or IRS) approach is susceptible to producing superior results than the other methods (ESS) in a (nonrandom) noisy environment. In the study, it was also found that, contrary to expectations, in the case of impulsive noise, the results favored the ESS method. The ESS approach appears to be the most suitable in a quiet environment. In a study by Guidorzi et al. [25], differences and advantages among impulse response measurements by the MLS and ESS methods were thoroughly analyzed and discussed. In addition, measurements were performed using the ESS and MLS methods in the presence of a hum, an impulsive noise, and with ideal conditions. The reverberation times ($T_{10}$, $T_{15}$, $T_{20}$, $T_{30}$) were measured and compared for MLS and ESS. Some of the conclusions were that the ESS method has a weak rejection for steady tonal disturbances, and that impulsive noise was the cause of differences in the reverberation times, especially $T_{10}$ and $T_{15}$. In a study by Torras-Rosell et al. [26] comparing the measurement of impulse responses between pseudorandom sequences (MLS, IRS) and sweep signals, it was stated that for MLS, background noise contaminates the measured signal along the time domain in a more-or-less uniform manner, because circular cross-correlation between the background noise and the excitation signal spreads the noise artifacts across the measured impulse response. However, the noise artifacts are not evenly distributed in the retrieved impulse response when using the sweep
technique. All of the frequency components of the impulsive noise above the sweep’s instantaneous frequency are pushed into the impulse response’s non-causal part by the sweep technique, whilst all of the frequency components below the sweep’s instantaneous frequency are drawn into the causal part. However, it is stated in the study [26] that: “unlike what is claimed in the literature, sweep signals cannot reject all distortion artifacts from the causal part of the estimated impulse response”. Other studies including comparisons of the methods can be found in [20,27,28].

This study focused on further extending the current knowledge of ESS and MLS measurements, by performing the methods in the presence of artificial background noise (white noise, tonal noise) of various levels. Therefore, the aim of this study was to estimate how background noise affects acoustic parameters, by comparing the relative error of measurements with and without artificial background noise.

For this purpose, this paper has been structured as follows: the methodology used in this study is described in Section 2. The results of this study are presented in Section 3. In Section 4, the data are analyzed, the research question is addressed, and the research’s limitations, and areas for future research, are noted. The conclusion provides a concise overview, and places the work in context.

2. Methods
2.1. Acoustic Parameters

The impulse response can be used to obtain most of the acoustic parameters that define a space. A room’s acoustical parameters are defined in ISO 3382-1 [5], along with methods for estimating these parameters from the measured impulse response. For this research, some of the most important and widely used acoustic parameters were selected: reverberation time (T), early decay time (EDT), clarity (C80), and definition (D50).

2.1.1. Reverberation Time (T)

Reverberation time (T) is the most significant and well-known room acoustic parameter. According to ISO 3382-1, it is defined as [5], “the duration required for the space-averaged sound energy density in an enclosure to decrease by 60 dB after the source emission has stopped”. For the estimation of T, the decay curve must be measured, which is defined, according to ISO 354 [8], as the “graphical representation of the decay of the sound pressure level in a room as a function of time after the sound source has stopped”. Typically, the decay curve, or r(t), is an irregular curve that can be roughly approximated by a linear decay (Figure 1).

![Figure 1. Typical decay curve.](image-url)
The decay curve can be calculated with the interrupted noise method (the direct recording of the decay of the sound-pressure level after exciting a room with broadband or band-limited noise), or with the integrated impulse response method (the reverse-time integration of the squared impulse responses) \cite{5}. For this study, as the goal was the measurement of the impulse responses using the ESS and MLS methods, the integrated impulse response method was utilized. The generation of the decay curve for each octave band can be obtained from the backward integrated squared impulse response \(h(t)\) \cite{12}:

\[
r(t) \approx \int_t^\infty h^2(\tau)d\tau
\]

This expression can be applied in more practical, normalized logarithmic form:

\[
10 \log r(t) = 10 \log \left( \frac{\int_t^\infty h^2(t)dt}{\int_0^\infty h^2(t)dt} \right)
\]

In this expression, the denominator represents the total energy. Again, as stated in \cite{5}, “\(T\) can be evaluated based on a smaller dynamic range than 60 dB and extrapolated to a decay time of 60 dB”. For example, if decay values from 5 dB to 35 dB below the initial level are used, it is labelled \(T_{30}\).

2.1.2. Early Decay Time (EDT)

The early decay time (EDT) is defined as \cite{29}, “the time interval required for the sound energy level to decay 10 dB after the excitation has stopped”. It is evaluated from the slope of the integrated impulse response curves (as the conventional reverberation time). The slope of the decay curve is determined from the slope of the best-fit linear regression line of the initial 10 dB (between 0 dB and \(-10\) dB) of the decay \cite{5}. The outcome is multiplied by a factor of six, to allow for direct comparison with the \(T\). The human perception of a room’s reverberation is more closely correlated with EDT than with \(T\) \cite{30}.

2.1.3. Clarity (\(C_{80}\))

The balance between early- and late-arriving energy is important for room acoustics. An early-to-late-arriving sound-energy ratio, “can be calculated for either a 50 ms or an 80 ms early time limit, depending on whether the results are intended to relate to conditions for speech or music, respectively” \cite{5}, using Equation (4). \(C_{80}\) is usually defined as clarity \cite{5}. A low clarity value denotes an unclear, highly reverberant sound. On the contrary, a high clarity value indicates a significant portion of early energy, which is equivalent to a subjective sensation of clarity.

\[
C_{80} = 10 \log \left( \frac{\int_0^{80 \text{ ms}} h^2(t)dt}{\int_{80 \text{ ms}}^\infty h^2(t)dt} \right)
\]

2.1.4. Definition (\(D_{50}\))

The acoustic parameter of “definition” (\(D_{50}\)) is used for speech conditions as per Equation (5), and is a measure of the speech definition \cite{31}. It can be used as a measurement of the early-to-total sound-energy ratio.

\[
D_{50} = 100 \frac{\int_0^{50 \text{ ms}} h^2(t)dt}{\int_0^\infty h^2(t)dt} (%)
\]

2.2. Measurements

For the measurements, an amphitheater at the Hellenic Mediterranean University, Department of Music Technology and Acoustics, Crete, Greece was used. Impulse responses
were measured with the use of a dodecahedron loudspeaker (Type DO12, 01 dB-Stell). The dodecahedron loudspeaker’s directivity characteristics meet ISO 3382-1 requirements. Before the measurements were conducted, the loudspeaker was evaluated by the authors to ensure that everything was in order with its operation [5]. Although a dodecahedron speaker is typically employed, alternative acoustic sources are also accessible for a variety of reasons (e.g., high cost, transportation difficulties, availability) [32–34]. For both methods, the same source and microphone positions were used, according to ISO 3382-1 [5]. In order to compare the properties of the impulse response objectively, the sound levels of the ESS and MLS signals in the measurement positions were set to approximately 84 dB for both methods. This level was preferred because it corresponds with a value between the optimum levels for the MLS and ESS signals, as proposed by Stan et al. [13]. Precautions were taken in order that the background-sound level would be approximately the same for every measurement. For the measurements, the 63 Hz octave band was excluded, due to variations observed in the background-noise levels. The measurements of the impulse response had a sampling frequency of 44.1 kHz. According to the anticipated $T$, an acceptable sequence length and time constant for the methods was selected. A single iteration was performed for each of the measurement points for each method. Consequently, no stepwise rotation of a dodecahedron sound source was used to increase the precision of the acoustic measurements in the room [35]. An omnidirectional microphone (Type 4190, Earthworks) was used for each of the measurements. The dodecahedron loudspeaker was placed in the center of the amphitheater stage (Figure 2). For all the measurements, as a noise compensation method, according to ISO 3382, the truncation method was used, which truncates (removes) the part of the impulse response tail that is close to, or below, the noise level.

Figure 2. (a) Dodecahedral loudspeaker placed in the center of the stage of the amphitheater. (b) Dodecahedral loudspeaker, additional sound source, omnidirectional microphone, sound-level meter.
With the aid of a second sound source (Tower V8, JWS, Falkenberg, Sweden), measurements were conducted with background-noise levels that varied in 2 dB steps. The sound source was directed toward the wall, in order to achieve the diffusion of the artificial background noise. Sound level measurements were made at the microphone position, using a sound-level meter (01dB-Steel SdB02). The maximum acceptable deviation from the expected sound level was 0.1 dB. Tonal noise and white noise were the two types of background noise that were employed.

The acoustic parameters that were calculated from the impulse responses were $T_{30}$, early decay time (EDT), clarity ($C_{80}$), and definition ($D_{50}$). In order to form an objective comparison of the measuring methods, all the room acoustics parameters were computed using the ARTA software (version 1.9.5, ARTALABS, Croatia). In each case, the mean relative error was used for the comparison of the results. In general, relative error is defined as the ratio of the absolute error of a measurement to the actual measurement. For this study, the actual measurement was considered the measurement without artificial background noise. The absolute error was considered the absolute difference between the measurement with the artificial background noise (for each case of background noise, and each measurement method), and the measurement without artificial background noise (the actual measurement). The mean relative error for all the octave bands was considered the average of all relative errors for the octave bands that were used.

3. Results

Impulse response measurements were performed as described in the methods section. Initially, the measurements were made without the addition of noise, in order to measure the impulse response and the acoustic parameters in the room.

Then, using an extra sound source and varied background-noise levels, impulse response measurements were made. The increments for the experiments were 2 dB. The acoustic parameters that were calculated from the impulse responses were $T_{30}$, EDT, $C_{80}$, and $D_{50}$, and are presented in Figure 3. The acoustic parameters were calculated for eight octave bands, from 125 to 8 KHz. Figure 3 presents the mean relative error for each acoustic parameter, and for each step of additional varying background noise. Figure 4 presents the results for the relative error for each octave band in the case of white noise as background noise, for $T_{30}$, EDT, $C_{80}$, and $D_{50}$ (in the graphs, there is a common y-axis limit per acoustic parameter).

A second set of measurements was performed, in which tonal noise was added as background noise. With the aid of a second sound source, impulse response measurements were carried out with different background-noise levels, and steps of 2 dB. The acoustic parameters that were calculated from the impulse responses were again $T_{30}$, EDT, $C_{80}$, and $D_{50}$, and are presented in Figure 5. The acoustic parameters were calculated for eight octave bands from 125 to 8 KHz. Figure 5 presents the mean relative error for each acoustic parameter, and for each step of additional varying background noise.
Figure 3. The mean relative error for all octave bands (M.R.E.O.B.) (125–8000 Hz) in the case of white noise as external background noise for $T_{30}$, EDT, $C_{80}$, and $D_{50}$.
Figure 4. Relative error for each octave band (R.E.O.B.) (125–8000 Hz) in the case of white noise as background noise for $T_{30}$, EDT, $C_{80}$, and $D_{50}$ (in the graphs, there is a common $y$-axis limit per acoustic parameter).
Figure 5. Mean relative error for all octave bands (M.R.E.O.B.) (125–8000 Hz) in the case of tonal noise (1000 Hz) as external background noise for $T_{30}$, EDT, $C_{80}$, and $D_{50}$.

4. Discussion

In the results section, according to the type of artificial background sound added in the space (white noise, tonal noise), the relative error for the four acoustic parameters ($T_{30}$, EDT, $C_{80}$, and $D_{50}$), and for the ESS and MLS methods, was measured and presented. In the case of white noise (Figure 3) as background noise, the results suggest that, in the case of $T_{30}$ and EDT, the difference between the two methods, as well as the mean relative error for the initial levels of added background noise is relatively small. However, for higher levels of added background noise, it seems that there is a sharp increase in the mean relative error, which is greater for the ESS method, both for $T_{30}$ and EDT. This is likely due to sharp changes in the distribution of energy in the energy decay curve, as the methods could not possibly effectively handle the excessive levels of the added noise. Regarding $C_{80}$ and $D_{50}$, while initially the differences between the ESS and MLS methods are again small, cumulatively, as the background noise increases, the error increases, with the ESS method seeming to have a greater error. Since both $C_{80}$ and $D_{50}$ (Equations (4) and (5), respectively) express a ratio of the impulse response’s energy before time $t_e$, and its energy after time $t_e$ ($t_e = 80$ ms, and $t_e = 0$ ms, respectively), cumulatively, as the background noise increases, the denominator increases (and the numerator, also, to a lesser extent), thus increasing the relative error.

In addition, the mean relative error for each individual octave band in the case of white noise as background noise for $T_{30}$, EDT, $C_{80}$, and $D_{50}$ is presented in Figure 4. In the case of $T_{30}$ and EDT, it seems that the sharp changes occur in more octave bands for the ESS method, compared to the MLS method, while they also occur at a lower added-background-noise level. Regarding $C_{80}$ and $D_{50}$, cumulatively, as the background noise increases, the error increases for the octave bands, with the ESS method presenting a greater error in comparison to the MLS method. Again, for the ESS method, the relative error seems to increase noticeably at lower background-noise levels.
In the case of tonal noise as background noise, the results are consistent with those observed in the general case of white noise. Again, regarding \( C_{80} \) and \( D_{50} \), it seems that cumulatively, as the background noise increases, the error increases, with the ESS method presenting a greater error. Regarding \( T_{30} \) and EDT for higher levels of added background noise, it seems that there is a sharp increase in the relative error, with the MLS method being less sensitive to these changes.

Our findings appear to be continent with the expected results in the cases of white and tonal background noise, according to previous studies. As presented in the introduction, in a previous study by Stan et al. [13], it is stated that, “in a noisy environment the MLS method is subject to giving better results than the ESS, since the MLS method possess the ability of randomizing the phase of any component in the recorded signal that is not correlated to the input signal emitted in the acoustical space”. Thus, any additional noise (white or even impulsive) will be distributed uniformly along the deconvolved impulse response. It is also stated in the study by Guidorzi et al. [25], for tonal disturbances, that the ESS method has a weak rejection. Moreover, in the study by Torras-Rosell et al. [26], it is mentioned that the noise artifacts are not evenly distributed in the retrieved impulse response when using the sweep technique.

In general, the overall direction of the results shows that in the cases of white and tonal noise, the MLS method performs better than the ESS method. Therefore, the MLS is to be preferred in noisy environments indoors, as well as outdoors (e.g., measurements of noise barriers in situ [36]). However, the ESS method is generally more suitable than the MLS method for use in architectural acoustics [25], and the perfect and complete rejection of the harmonic distortions prior to the impulse response, their individual measurement, and the excellent signal-to-noise ratio make it the best impulse response measurement technique in an unoccupied and quiet room [13]. Furthermore, developments are constantly emerging for the improvement of ESS-method measurements in the presence of non-stationary noise [37].

Limitations

We acknowledge that our research may have some limitations. Some of those limitations stem from the fact that there are many variations on how to perform acoustic measurements with the different methods.

The first limitation is that this study has only investigated the performance of the ESS and MLS signal qualities in the case where the signals in the measurement position had the same excitation level. However, optimum signal levels have been proposed for the ESS and MLS [13], which are different for each method. Consequently, additional measurements are needed in the case of optimum signal levels for both methods.

Secondly, we used a single noise compensation method (the truncation method), as presented in the methods section. It has been stated that, “when these (noise compensation) methods are used to suppress noise effects, their performances differ significantly” [38].

In addition, we used a single software implementation, as presented in the methods section. Again, it has been shown that there are some (relatively small) differences in the results among software implementations [39]. However, the differences for measuring the \( T_{30} \) from the impulse response were found to increase for smaller peak-to-noise ratios [39].

5. Conclusions

In this study, we assessed the effect of background noise (white noise, tonal noise) on acoustic parameters (\( T_{30} \), EDT, \( C_{80} \), and \( D_{50} \)) for ESS and MLS measurements, by introducing artificial background noise, employing an external sound source. In the case of white noise (as background noise), the results suggest that, in the case of \( T_{30} \) and EDT, the difference between the two methods, as well as the mean relative error for the initial levels of added background noise, is small. However, for higher levels of added background noise, there is a sharp increase in the relative error, which is greater for the ESS method, both for \( T_{30} \) and EDT. Regarding \( C_{80} \) and \( D_{50} \), while initially the differences between the
ESS and MLS methods are again small, cumulatively, as the background noise increases, the error increases, with the ESS method presenting a greater error. In the case of tonal noise, the results are consistent with those observed in the general case of white noise.

The findings presented in this study make several contributions to the current literature, by adding to our understanding of the ESS and MLS methods for impulse response measurements. The findings support the idea that the MLS method for white and tonal background noise can provide better results in cases of high background-noise levels. However, in cases of low background noise, the results of the two methods are similar. These results are in agreement with the literature. This work also contributes to existing knowledge by suggesting the expected relative error of acoustic parameters when various types and levels of background noise are present. Finally, the study suggests, based on background noise and level, the optimum method to conduct impulse response measurements.

Author Contributions: Conceptualization, N.M.P. and S.A.; methodology, N.M.P. and S.A.; software, N.M.P. and S.A.; formal analysis, N.M.P. and S.A.; investigation, N.M.P. and S.A.; resources, N.M.P.; data curation, N.M.P., S.A. and G.E.S.; writing—original draft preparation, N.M.P.; writing—review and editing, G.E.S.; visualization, N.M.P.; supervision, N.M.P. and G.E.S.; project administration, N.M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We would like to thank everyone who helped in the measurement sessions. We would also like to thank the anonymous reviewers for their thoughtful observations and comments.

Conflicts of Interest: The authors declare no conflict of interest.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.