Gunshots Sound Analysis,

Identification and Impact on Hearing

by

Bruno Tardif

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Electrical and Computer Engineering

Carleton University
Ottawa, Ontario

Abstract

Analyzing the sound resulting from a gunshot can help identify the type of gun used. This can help determine the appropriate public safety actions when a gunshot sound is detected in a public space. Furthermore, analyzing the sound can lead to a better understanding of the impact of this sound on hearing. For example, exposure to gunshot sounds can cause hearing impairments. Measuring these sounds from various locations and analyzing them can improve the design of hearing protectors and help enact regulations to minimize the impact of gunshot sounds on hearing impairments.

In this thesis, acoustic data was collected from different guns and a mortar for analysis. To capture their sound including any non-symmetric sound propagation from the gunshots, 27 high dynamic range pressure 1/8" microphones were placed around the weapons forming a polar grid pattern. Audio signals from all microphones were captured at 204.8 kHz sampling rate synchronously to preserve the fidelity of the impulse nature of the gunshots.

This thesis entails two types of analysis. The first is an image-based analysis method to take advantage of the recent advancement of image recognition techniques. In this thesis, we used Continuous Wavelet Transform (CWT) and Short-Time Fourier Transform (STFT) to obtain the time-frequency spectrogram of the gunshot recordings as images. We then employed a two-stage machine-learning classifier to process the images to identify the specific gun or mortar the gunshot was originated from. Classification accuracy of 98.1% was achieved in this example. A second type of analysis was conducted using a novel method for characterizing acoustic attenuation properties that respects the

spectral sensitivity of the human ear and preserves the energy in its octave bands at the same time.

Acknowledgements

I would like to express my gratitude to Carleton University, the Quality Engineering Test Establishment and the Munitions and Experimental Test Centre at the Canadian Department of National Defence for their support in the writing of this thesis and for their competence and skill demonstrated through all these years of operations.

Particularly, I would like to thank Prof. Rafik Goubran for his support, guidance and enlightenment through all these years and Dr. David Lo for being a wonderful colleague, supervisor and friend. Without him, the thesis would not have been possible. My thanks also go to Master Corporal Alex Saumure; recording the data was simple with him taking charge of all the logistics and his ensuring that we had all the equipment.

Finally, I would like to thank my love, Chantal Boutin. I am so lucky to have you by my side.

Statement of Originality

This thesis describes the research work that was done by the author at Carleton University for completion of the Master of Applied Science (M.A.Sc.) program in the Department of Systems and Computer Engineering (Faculty of Engineering and Design). A portion of this research has been reported in the following publications [1] [2]:

1. Bruno Tardif, David Lo, and Rafik Goubran, "**Gunshot Sound Measurement and Analysis**", *in Proceedings of 2021 IEEE Sensors Applications Symposium (SAS)*, Sundsvall, Sweden, Virtual Conference, August 2021.
2. Bruno Tardif, David Lo, and Rafik Goubran, "**Measurement and Characterization of Hearing Protection Devices in the Presence of Impulse Sound**", *in Proceedings of 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Neuchâtel, Switzerland, Virtual Conference, June 2021.

The results reported in the above-mentioned papers are reflected throughout this thesis. The author performed data analysis, prepared the documents for publication, and made all necessary changes based on feedback from co-authors and paper reviewers.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Analyzing the sound resulting from a gunshot can be used to identify the type of gun used and can also provide a better understanding of the impact of this sound on hearing. The first relates to public safety. Gun identification has proved to be crucial in determining the level of risk posed to the public and to execute a proper response. Among the various means available, using acoustic signals from gunshots is one of the most popular methods [3] [4] [5]. In the patrols of an armed conflict zone or during a forensic investigation, gunshot noise often provides key evidence for identifying the gun used [6].

The second area relates to health. A 2018 report published by the World Health Organization (WHO) notes that approximately 466 million people have some sort of hearing loss. That number may increase to more than 900 million people by 2050 [7]. Furthermore, the most common work-related illness in U.S. is occupational hearing loss [8]. For those regularly exposed to high impulse noise, such as members of military forces and police forces, roughly a third of them retire with chronic hearing problems [9]. These numbers show a need to improve how we currently protect our hearing health, and that includes protection against impulse noises.

1.1 Problem Statement

The measure of gunshot sounds, given their impulse nature, has been a challenging area of research for decades [10]. The evaluation of gunshot sounds requires careful attention [11].

Contrary to typical continuous noise in an industrial environment, audio signals from gunshots are impulsive and have a duration of only a few milliseconds [12]. Measuring impulsive sound correctly can also be challenging due to the intense sound pressure from gunshots that can reach the saturation levels of microphones, causing clipping in the recorded data. Consequently, the proper choice and installation of the equipment as well as the analysis of the audio recordings all require great care [13].

When trying to identify a gun from its audio recording, some studies have based their analysis on the acoustic characteristics of the mechanism of the firearms [14], while others have used time domain features [15], as well as features from frequency domain [16]. To take advantage of recent advancements in image recognition techniques, this study instead developed a novel image-based approach to the analysis of gunshot sound. To investigate the effectiveness of this image-based approach, we chose the problem of gun identification as a test bench. We applied the Continuous Wavelet Transform (CWT) and the Short-Time Fourier Transform (STFT) to obtain spectrogram images of the gunshot recordings. After a two-stage machine-learning classifier was trained, the spectrogram images were presented to the classifier for gun identification.

Unlike measurements for steady-state noises, which regulatory agencies commonly use to quantify the maximum daily exposure, measurements of the impact on hearing have no specific exposure limits when applied to impulsive noises such as gunshots. To demonstrate this point, the newest military standard MIL-STD-1474E [17] bases the daily maximum allowable gunshots on two different metrics without specifying which one to use. The first one, $L_{IAeq100ms}$, is an equal-energy model approach that is calculated over a 100-ms interval. The second one, the Auditory Risk Unit (ARU), is calculated with the

online available Auditory Hazard Assessment Algorithm for Humans (AHAAH) model [18], which is an electroacoustic model of the human auditory system. In all the technical choices, the question of properly evaluating an attenuation arises. Choosing a proper hearing protection device (HPD) or a suppressor installed on a rifle may become difficult.

1.2 Contributions

This thesis led to three contributions.

The first contribution is a method for accurately measuring impulse gunshot sounds using a grid of pressure microphones and a high-speed acquisition system. This contribution has been published in the following paper [1]:

- Bruno Tardif, David Lo, and Rafik Goubran, "Gunshot Sound Measurement and Analysis", *in Proceedings of 2021 IEEE Sensors Applications Symposium (SAS)*, Sundsvall, Sweden, Virtual Conference, August 2021.

The second contribution is a method for improving the identification of gunshots with modern image classification, combined with machine-learning processes, to accurately identify the type of gun based on the gunshot impulse sound. This contribution is presented in Chapter 4 and has been published in the following paper [2]:

- Bruno Tardif, David Lo, and Rafik Goubran, "Measurement and Characterization of Hearing Protection Devices in the Presence of Impulse Sound", *in Proceedings of 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Neuchâtel, Switzerland, Virtual Conference, June 2021.

The third contribution is a method to evaluate the attenuation impacting our hearing protection with a novel energy-preserving method for estimating the octave band insertion

loss using the continuous wavelet transform. This contribution is presented in Chapter 4. It relies on research previously published by the author in [2]:

- Bruno Tardif, David Lo, and Rafik Goubran, "Measurement and Characterization of Hearing Protection Devices in the Presence of Impulse Sound", *in Proceedings of 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Neuchâtel, Switzerland, Virtual Conference, June 2021.

1.3 Thesis Outline

Chapter 2 of this thesis presents a background literature review of the following topics:

- Spectral analysis of impulsive noise
- Impact of impulsive noise on hearing
- Acoustic sound insertion loss quantification

The literature review discusses the advantages and disadvantages of the prior art to help understand the new methods proposed in this thesis.

Chapter 3 discusses the data collection setup and methodology. This section is crucial for giving rare environment-controlled high-definition impulse sounds. The discussion is also important for validating the analysis used in our study, both for the tests of guns identification, based on impulse sounds and on the attenuation evaluation.

Chapter 4 presents the proposed methods for gun identification. First, it uses guns with dissimilar ammunition calibers in perfect conditions, then by testing the identification process with similar ammunition calibers under a simulated reverberant environment.

Chapter 5 evaluates sound attenuation by using a novel energy-preserving method for estimating the octave band insertion loss with the continuous wavelet transform. We

compare this method to existing ones, highlighting its originality and potential for enhanced precision.

Chapter 2: Background Review

2.1 Gunshot Sound Analysis

This section presents an overview of the spectral analysis methods that were used to analyze gunshot sounds.

2.1.1 The Short-Time Fourier Transform (STFT)

We can view the first spectral analysis method, the STFT, as a projection operator that projects signal energy onto the time-frequency plane [19]. It comes from the correlation at the instant t and frequency ω of the signal with a time-shifted window:

$$STFT(t, \omega) = \int s(\tau) h(\tau - t) e^{-i\omega\tau} \, d\tau \quad (1)$$

Where $s(t)$ is the signal and $h(t)$, the window. This study uses the Hamming window.

While the STFT has been used for years for analyzing non-stationary signals, one of its disadvantages is that it requires a fixed tiling in the time-frequency plane making it difficult to obtain sufficient time-resolution for the entire analysis [20].

2.1.2 The Continuous Wavelet Transform (CWT)

The second spectral analysis method, the CWT, is defined with the following formula [21]:

$$W_{\psi}(t, s) = \int_{-\infty}^{\infty} \frac{1}{s^n} \psi^* \left(\frac{\tau - t}{s} \right) x(\tau) d\tau \quad (2)$$

Where $\psi^*(\tau)$ is the complex conjugate of the wavelet, s is the scaling parameter and t, the

translation for which the Fourier transform is defined by [21]:

$$\Psi_{\beta,\gamma}(\omega) = \int_{-\infty}^{\infty} \psi_{\beta,\gamma}(t) e^{-i\omega t} dt = U(\omega) a_{\beta,\gamma} \omega^{\beta} e^{-\omega\gamma} \quad \textbf{(3)}$$

Where $a_{\beta,\gamma}$ is a normalization constant, $U(\omega)$ is the unit step function and β and γ are two parameters to control the shape of the wavelet. The value $P^2 = \beta\gamma$ is called the time-bandwidth product. Morse wavelets are the best function available when a good tradeoff between time and frequency precision is desired [22]. The values chosen for this thesis were $\gamma = 3$ and $P^2 = 60$. They provided a symmetric wavelet which in-turns resulted in a neutral time-frequency precision tradeoff [23].

Since the temporal length of the wavelets used in CWT is frequency dependent, the tiling in the time-frequency plane can be structured such that adequate time resolution for signal localizations can be obtained over the full acoustic spectrum when compared to the STFT. Therefore, the CWT resolution can be qualified as adjustable. The local resolution is known as the Heisenberg box [22]. Temporal and frequency resolutions are defined by the following two equations [24]:

$$\sigma_t^2(\gamma) = \int_{-\infty}^{\infty} (t - \mu_\gamma)^2 \left|\psi_\gamma(t)\right|^2 dt \quad \textbf{(4)}$$

$$\sigma_\omega^2(\gamma) = \int_{-\infty}^{\infty} (\omega - \xi_\gamma)^2 \left|\Psi_\gamma(\omega)\right|^2 d\omega \quad \textbf{(5)}$$

where $\mu_\gamma = \frac{1}{\|\psi_\gamma(t)\|^2} \int_{-\infty}^{\infty} t \left|\psi_\gamma(t)\right|^2 dt$

and $\xi_\gamma = \frac{1}{2\pi\|\psi_\gamma(t)\|^2} \int_{-\infty}^{\infty} \omega \left|\Psi_\gamma(\omega)\right|^2 d\omega$. $\Psi_\gamma(\omega)$ being the Fourier transform of the wavelet $\psi_\gamma(t)$.

2.2 Impact on Hearing Quantification

The following are the primary methods for quantifying the impact on hearing from gunshot sounds:

- MIL-STD 1474D
- Pfander effective level and Smoorenburg effective level
- $L_{Aeq,8h}$
- MIL-STD 1474E
- AHAAH model
- Kurtosis

2.2.1 MIL-STD 1474D

MIL-STD 1474D, also known as Committee on Hearing and Bio-Acoustics (CHABA) criteria, derives from research conducted 45 years ago and known for being defective [25]. The calculation leading for the maximum allowable exposure is based on B-duration, as shown in Figure 1. In practice, getting the B-duration is not an easy task and depends on several criteria.

The formula for determining the allowable number of impulses per day when wearing ear plugs or earmuff is:

$$\boldsymbol{N_1 = 10^x} \quad \textbf{(6)}$$

Where:

$$\boldsymbol{x = \frac{1}{5}\left[177 - L + 6.64\, log_{10}\frac{200}{T}\right]} \quad \textbf{(7)}$$

And $\boldsymbol{L}$ is the measured peak sound level in dB and $\boldsymbol{T}$ is the B-duration in ms. When wearing ear plugs and ear muff, the equation is:

$$\boldsymbol{N_2 = 20\, N_1} \tag{8}$$

This standard is now considered obsolete.

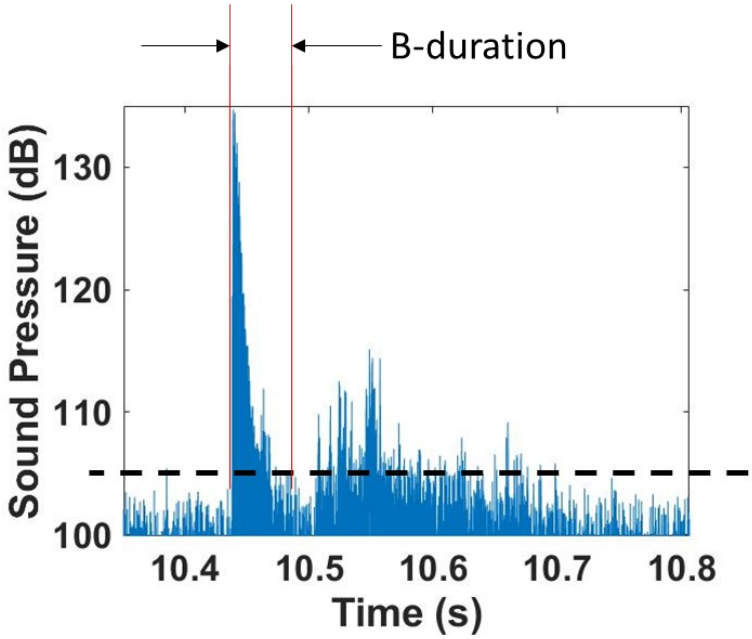


Figure 1: B-duration example

2.2.2 Pfander and Smoorenburg

With the objective to enhance the CHABA method, Pfander [26] and Smoorenburg [27] suggested these equations to evaluate the effective exposure level:

$$\boldsymbol{L_p = L_{peak} + 10\ \log T_C + 10 \log N} \tag{9}$$

$$\boldsymbol{L_S = L_{peak} + 10\ \log T_D + 10 \log N} \tag{10}$$

Where $\boldsymbol{L_P}$ and $\boldsymbol{L_S}$ are the effective levels for Pfander and Smoorenburg respectively, $\boldsymbol{L_{peak}}$ is the measured peak pressure, $\boldsymbol{T_C}$ and $\boldsymbol{T_D}$ are C-duration and D-duration, and $\boldsymbol{N}$ is the

number of impulses. T_C and T_D are variations of the B-duration from Figure 1, used in Germany only for Pfander. Neither method accounts for the spectral characteristic of the impulses, and in Chan study [28], both show a lower accuracy than the $L_{Aeq,8hr}$ method.

2.2.3 $L_{Aeq,8hr}$

$L_{Aeq,8hr}$ is an energy-equivalent method normalized to an eight-hour working day. Exposure to N identical impulsive events, such as weapon rounds fired, is expressed with [29]:

$$L_{Aeq,8hr} = 10 \log\left(\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\frac{p_A^2(t)}{p_0^2}dt\right) + 10 \log\frac{t_2-t_1}{T_{8hr}} + 10 \log N \quad (11)$$

Where $t_2 - t_1$ is the period over which the average is taken (the impulsive event in this case), $p_A(t)$ is the A-weighted pressure signal as a function of time, T_{8hr} is the 8-hours normalized working day (28 800 s) and N, is the number of impulsive events. For example, with the Canadian federal maximum exposure limit of 87 dB_A for a given working day ($L_{Aeq,8hr}$), we can calculate the maximum of impulsive event having a waveform $p_A(t)$ that an individual can be exposed [30].

The attractiveness of $L_{Aeq,8hr}$ is the simplicity in its application and the fact that it can be used for both continuous and impulsive noise [29].

2.2.4 MIL-STD 1474E

This standard offers two methods for evaluating the hazards associated with impulsive noises [17]. The first uses the model AHAAH as describe in section 2.2.5. The second is similar to $L_{Aeq,8hr}$. But a correction factor is applied to account for the duration of the impulse [31]. The formula is:

$$L_{Aeq100ms} = log\left\{\frac{1}{p_0^2 T}\int_0^T p_A^2(t)dt\right\} \quad (12)$$

Where T is 100 ms, p_0 is the reference pressure of 20 μPa and $p_A(t)$ is the A-weighted pressure signal as a function of time. There is correction factor when the impulsive noise is longer than 100 ms. The next step is then to apply some correction factors to convert the value to $L_{Aeq,8hr}$.

No indications exist to determine which of the two methods is preferable. However, $L_{aeq100ms}$ is indicated as simpler to use in the standard, and both are identified as superior to the method in MIL-STD 1474D.

2.2.5 AHAAH Model

Among the proposed methods for evaluating the hazards associated with impulse-noise exposure, the Auditory Hazard Assessment Algorithm for Humans (AHAAH) differs from the others as the proposed method is an electroacoustic model of the ear [32]. Instead of trying to find a correlation between a measurement method of sound pressure and hearing loss, the US Army Research Laboratory has tried to understand the physiological processes associated with the hearing loss and then to find a measure to predict it. Doing so, they came up with a circuit diagram of the electroacoustic model.

Using a model approach for predicting hazards from high impulsive noises has the advantage of addressing the two aspects of this kind of signal: its frequency and time components. Frequency modeling of the ear depicted by ear-response filters, such as the C-weighted and A-weighted filters, implemented years ago has proven the importance of the frequency spectrum of an acoustic signal regarding the auditory system susceptibility. However, the temporal pattern is also important and needs to be taken into account [33],

primarily because the auditory system is recovering or resting between exposures. Neglecting them may overestimate the hazards associated with impulsive exposures.

The method comes with a software calculating an Auditory Risk Unit (ARU), which is used to estimate the allowable number of exposures (ANE) that an individual can be exposed to. The model is now included as one of the proposed methods in the new standard MIL-STD 1474E.

Although this method is promising, studies have identified the following issues with its results and use [25] [29] [31] [34]:

- The model is not fully developed and validated.
- The model produced unsatisfactory results for several exposure conditions, leading to increased risk of hearing impairment when the peak-pressure level decreases, for example:
    I. The model gives two possible conditions, warned and unwarned. These conditions refer to the middle ear muscle contractions protecting the ear. AHAAH model uses it to calculate the recommended maximum noise levels. However, this option is debated and the warned condition seems to not be present in enough people to be applied as a valid form.
    II. Due to the software itself, the modeling of the acoustic hazard from a complex military environment with continuous noise from various different machinery and weapons produced simultaneously or in a reverberant environment is not possible.

2.2.5.1 AHAAH Results from Simulated Signals

To test the limits of AHAAH, we have analyzed the model by injecting simulated signal of the form:

$$\boldsymbol{y(t) = Ae^{-\lambda t}\, cos(\omega t + \varphi)} \quad \textbf{(13)}$$

Where $\boldsymbol{A}$ is the peak amplitude in dB, $\boldsymbol{\lambda}$ is the decay constant and $\boldsymbol{\omega}$, the frequency as illustrated in Figure 2.

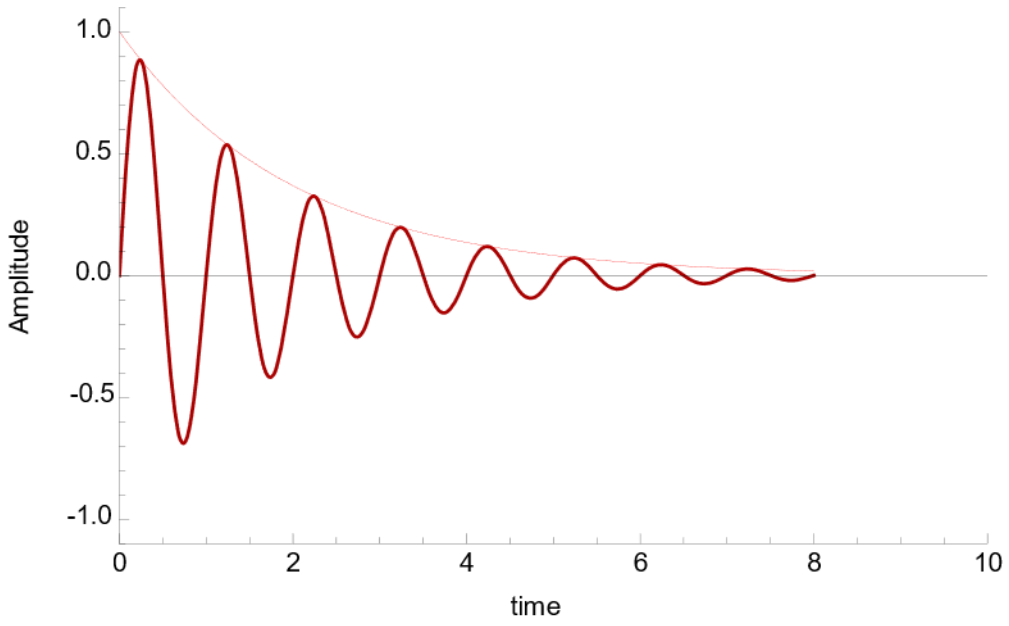


Figure 2: Simulated signal injected in AHAAH

Table 1 to Table 3 give the ANE results from the model. For comparison, the value from equation 11 is also calculated and inserted in the table and labelled with A-weighted identical impulsive events $\boldsymbol{N}$. Figure 3 to Figure 5 show the results for a fixed peak value $\boldsymbol{A}$ and decay time $\boldsymbol{\lambda}$.

Table 1: Allowable Number of Exposures (ANE) computed by AHAAH from simulated injected signals, λ = 25

		Decay Constant (λ) = 25			
Frequency (ω)	**Peak value (dB_A)**	**120**	**140**	**160**	**180**
16	**A-weighted N**	11500000	113000	1115	11
	Warned	121000000	979000	8822	198
	Unwarned	26427	3346	591	86
31.5	**A-weighted N**	2500000	24800	245	2
	Warned	25200000	218800	1828	48
	Unwarned	25300	2296	272	27
63	**A-weighted N**	283600	2791	28	0
	Warned	2090000	25800	273	11
	Unwarned	28700	605	29	4
125	**A-weighted N**	35300	347	3	0
	Warned	90579	1102	43	4
	Unwarned	1549	76	10	2
250	**A-weighted N**	6640	65	1	0
	Warned	8663	97	3	0
	Unwarned	216	10	1	0
500	**A-weighted N**	1940	19	0	0
	Warned	2170	27	1	0
	Unwarned	58	4	1	0
1000	**A-weighted N**	924	9	0	0
	Warned	234	4	0	0
	Unwarned	11	1	0	0
2000	**A-weighted N**	703	7	0	0
	Warned	78	1	0	0
	Unwarned	10	0	0	0
4000	**A-weighted N**	744	7	0	0
	Warned	27	0	0	0
	Unwarned	8	0	0	0
8000	**A-weighted N**	909	9	0	0
	Warned	2389	24	0	0
	Unwarned	1274	13	0	0
16000	**A-weighted N**	3285	32	0	0
	Warned	868000	8724	122	8
	Unwarned	95300	976	22	2

Table 2: Allowable Number of Exposures (ANE) computed by AHAAH from simulated injected signals, λ = 100

		Decay Constant (λ) = 100			
Frequency (ω)	**Peak value (dB_A)**	**120**	**140**	**160**	**180**
16	**A-weighted N**	2900000	28810	284	3
	Warned	15400000	139544	3772	455
	Unwarned	21002	2001	352	107
31.5	**A-weighted N**	1600000	16000	157	2
	Warned	7700000	77700	2695	150
	Unwarned	24268	1712	275	55
63	**A-weighted N**	497600	4898	48	0
	Warned	1800000	24792	697	37
	Unwarned	20460	658	41	6
125	**A-weighted N**	104400	1027	10	0
	Warned	220000	2816	113	12
	Unwarned	2242	105	14	3
250	**A-weighted N**	23700	233	2	0
	Warned	31200	354	13	2
	Unwarned	349	16	2	1
500	**A-weighted N**	7403	73	1	0
	Warned	7891	100	5	1
	Unwarned	91	6	1	0
1000	**A-weighted N**	3621	36	0	0
	Warned	908	14	1	0
	Unwarned	19	1	0	0
2000	**A-weighted N**	2783	14	0	0
	Warned	306	2	0	0
	Unwarned	20	0	0	0
4000	**A-weighted N**	2965	29	0	0
	Warned	106	1	0	0
	Unwarned	21	0	0	0
8000	**A-weighted N**	3641	36	0	0
	Warned	9594	97	1	0
	Unwarned	3805	39	0	0
16000	**A-weighted N**	13148	129	1	0
	Warned	865270	8701	121	8
	Unwarned	95038	973	22	2

Table 3: Allowable Number of Exposures (ANE) computed by AHAAH from simulated injected signals, λ = 200

		Decay Constant (λ) = 200			
Frequency (ω)	**Peak value (dB_A)**	**120**	**140**	**160**	**180**
16	**A-weighted N**	1060000	10400	102	1
	Warned	3700000	41275	1956	299
	Unwarned	19489	1329	235	96
31.5	**A-weighted N**	828000	8148	80	1
	Warned	2700000	32318	1727	241
	Unwarned	21387	1242	219	80
63	**A-weighted N**	435000	4282	42	0
	Warned	1200000	16619	990	67
	Unwarned	13830	826	62	9
125	**A-weighted N**	144090	1418	14	0
	Warned	270185	3651	187	20
	Unwarned	2964	159	18	4
250	**A-weighted N**	40718	401	4	0
	Warned	54412	629	24	3
	Unwarned	538	23	3	1
500	**A-weighted N**	13898	137	1	0
	Warned	13992	180	9	1
	Unwarned	141	9	2	0
1000	**A-weighted N**	7015	69	1	0
	Warned	1743	27	2	0
	Unwarned	31	2	0	0
2000	**A-weighted N**	5472	54	1	0
	Warned	599	9	1	0
	Unwarned	36	1	0	0
4000	**A-weighted N**	5878	58	1	0
	Warned	213	3	0	0
	Unwarned	40	1	0	0
8000	**A-weighted N**	7264	72	1	0
	Warned	19159	194	2	0
	Unwarned	7011	73	1	0
16000	**A-weighted N**	26216	258	3	0
	Warned	861421	8662	121	8
	Unwarned	94622	969	22	2

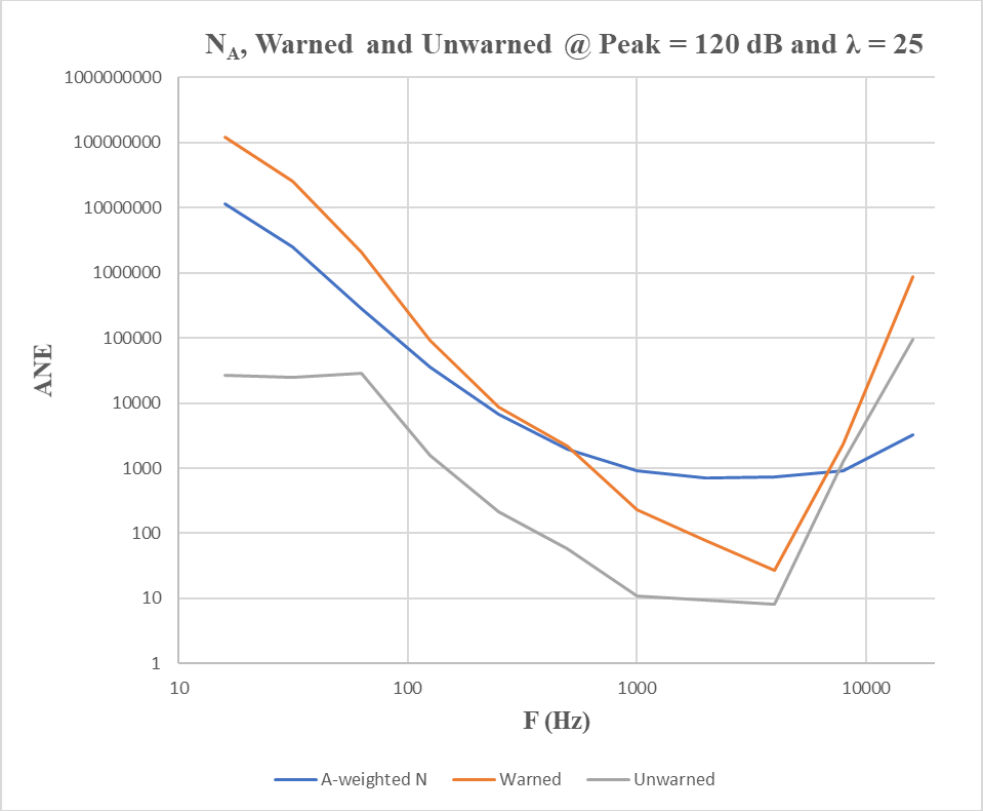


Figure 3: AHAAH model results from simulated injected signals, λ=25 and A=120dB

Figure 4: AHAAH model results from simulated injected signals, λ=100 and A=120dB

Figure 5: AHAAH model results from simulated injected signals, λ=200 and A=120dB

From the data, we can observe the following:

- The relation with the decay time is linear, and doubling λ (decreasing the length of the signal) is doubling mostly the allowed number of rounds for the 3 conditions (equivalent energy, warned and unwarned).
- The relation with the peak value is inversely proportional with the equivalent energy method. However, it is just inversely proportional with the warned and unwarned conditions.
- The effect of using warned and unwarned conditions leads to enormous differences.

2.2.5.2 AHAAH Results from Live Recorded Signals

In the objectives of finding unsatisfactory results as mentioned in the literature, recorded live signals have been injected in the model. Figure 6 shows these live recordings. Coming from various distances, the peak values are normalized to fixed values, as shown in Table 4. The results indicate that in the salvo effect, the ANE is not linearly related with the number of gunshots when using AHAAH, unlike the equivalent energy model (N_A), especially for the unwarned state.

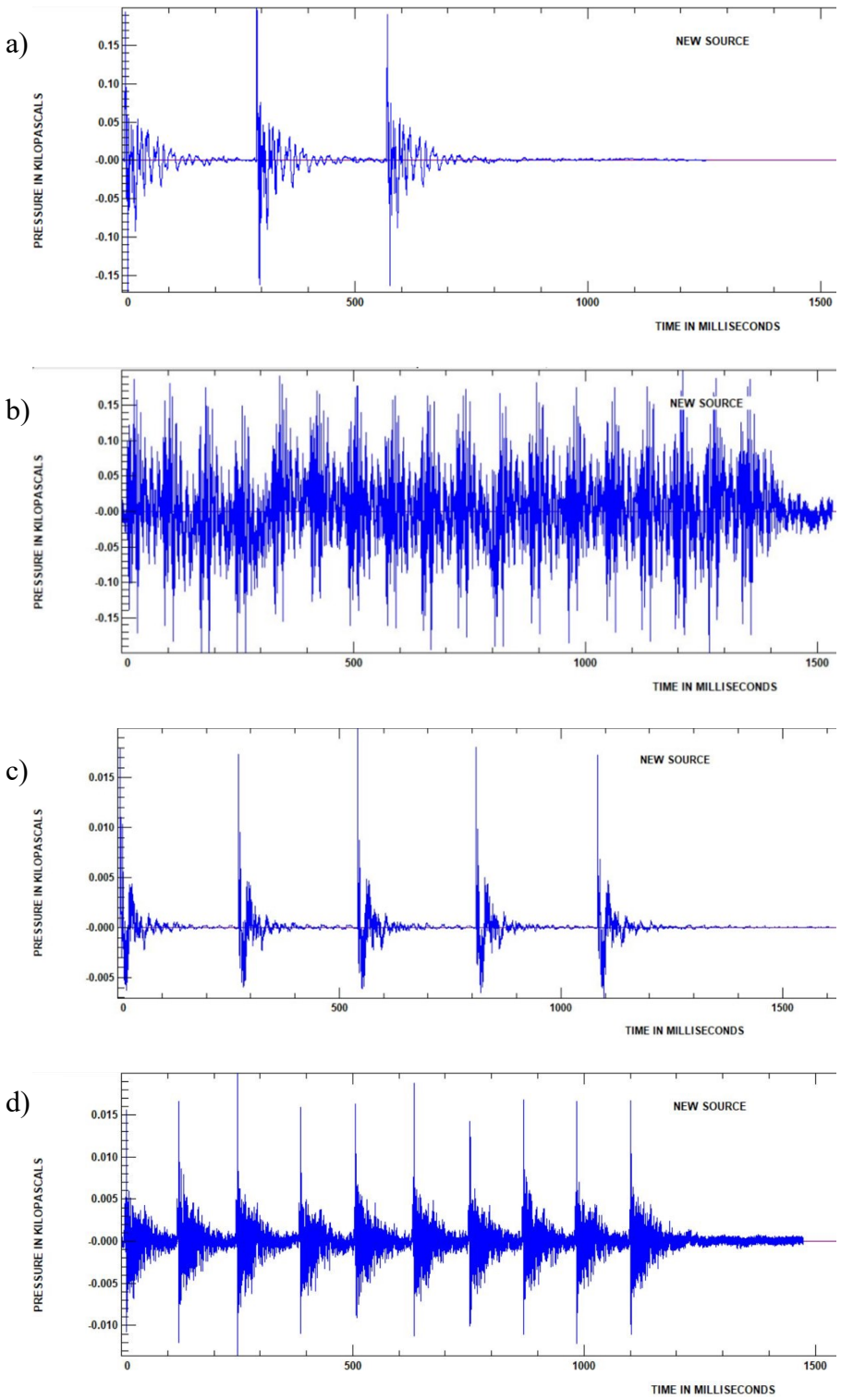


Figure 6: Live recorded signals from a) vehicle-mounted 25 mm single-barrel cannon, b) vehicle-mounted 7.62mm coaxial machine gun, c) 57mm naval artillery gun and d) 112.7mm (0.5”) naval machine gun

Table 4: AHAAH Allowable Number of Exposures (ANE) for live recorded signals

		1st peak			Salvo			
	Pk	**N_A**	**Warned**	**Unwarned**	**Nb**	**N_A**	**Warned**	**Unwarned**
LAV 6.0 25mm	120	29979	18913	636	3	10658	7363	653
	140	295	232	27		105	90	24
	160	2.9	9	2.3		1	3.1	1.5
LAV 6.0 C6	120	6658	1524	146	18	397	107	75
	140	66	19	4.1		3.9	1.3	1.1
	160	0.6	1	0.4		0.04	0.1	0.1
Naval 57mm	120	47946	10817	900	5	11381	2925	923
	140	472	139	35		112	37	23
	160	4.6	4.7	2.1		1.1	1.1	0.9
Naval 0.50	120	12035	1466	199	10	1660	163	123
	140	119	17	4.8		16	1.8	1.5
	160	1.2	0.5	0.3		0.2	0.1	0.1

2.2.6 Kurtosis

Kurtosis is often depicted as the "peakedness" of a signal [29] [31], which is not necessarily true [35]. The equation is [36]:

$$\boldsymbol{\beta} = \sum_{i=1}^{N} \left(\frac{X_i - \bar{X}}{S}\right)^2 \quad \textbf{(14)}$$

Where $\bar{X}$ is the mean and S, is the standard deviation. The kurtosis of a Gaussian noise is 3 and it represents the fourth moment of a statistical distribution. By comparison, the mean is the first moment, variance the second, and skew the third.

Earlier studies considered the kurtosis [37] [36] as a means to evaluate impairments from impulsive noises. However, no current study has concluded that this approach gives

better results than other methods [38]. Instead, the metric $L_{Aeq,8hr}$, the equivalent noise exposure level normalized to a nominal eight-hour working day [38] is preferable.

2.3 Insertion Loss Metrics

We used three methods to quantify insertion loss on acoustic sound signals:

- Impulse Peak Loss (IPIL)
- Impulsive Spectral Insertion Loss (ISIL)
- Octave Band Impulse Peak Insertion Loss (OBIPIL)

2.3.1 Impulse Peak Insertion Loss

IPIL is a single number representing the overall peak acoustic attenuation [39]. In our case, the equation is:

$$IPIL(dB) = Pk_{normal}(dB) - Pk_{w/attenuator}(dB) \quad \textbf{(15)}$$

Where Pk_{normal} is the peak sound pressure level without attenuator and $Pk_{w/attenuator}$ is the peak sound pressure level with an attenuator. The value has no frequency-related component and is purely a time-domain value. In this study, we subtracted the peak average sound level with attenuator from the peak average sound level without the attenuator of all the gunshots recorded.

2.3.2 Impulsive Spectral Insertion Loss (ISIL)

A contrast to IPIL, ISIL's goal is to explicitly quantify the frequency dependency of attenuation, using either a hearing protection device (HPD) or, as in this study, an attenuator such as a suppressor. Figure 7 shows the decomposition in band process using a

filter bank. Then we calculate the energy by evaluating the equivalent continuous sound pressure level in each 1/3 octave band [39]. In this study, the energy is calculated with the following equation:

$$\boldsymbol{L_{eq}(dB) = 10\ \log\left[\frac{1}{t_2 - t_1}\int_{t_1}^{t_2}\frac{\rho^2(t)}{\rho_0^2}dt\right]} \quad \textbf{(16)}$$

Where $\boldsymbol{\rho}$ is the sound pressure level, $\boldsymbol{\rho_0}$, the reference sound pressure level of 20 µPa and $\boldsymbol{t_2}$ and $\boldsymbol{t_1}$ are the chosen time interval. In our case, $\boldsymbol{t_1}$ is set to 10 ms before the peak sound pressure level of a gunshot and $\boldsymbol{t_2}$, to 20 ms after. As used in Fackler et al. [39], the filter bank used is FFT-based. Therefore, as seen in section 2.1.1, its time-frequency resolution cannot be traded-off and adjusted based on needs. Consequently, the precision in time and frequency may not be optimized through the full acoustic spectrum.

2.3.3 Octave Band Impulse Peak Insertion Loss

Recently, Sarray et al. [40] introduced OBIPIL, which is the IPIL but calculated in each 1/3 octave band instead of a single number. The result is a curve covering the frequency audio range. The use of a filter bank preserving the peak value in each 1/3 octave band is fundamental to maintain the integrity of the method. The process is the same as the one used for ISIL and is shown in Figure 7. We used a filter bank compliant with the standard ANSI S1.11-2004 [41]. Again, to maintain the analysis on the gunshot, we used a window of 10 ms before and 20 ms after the impulse.

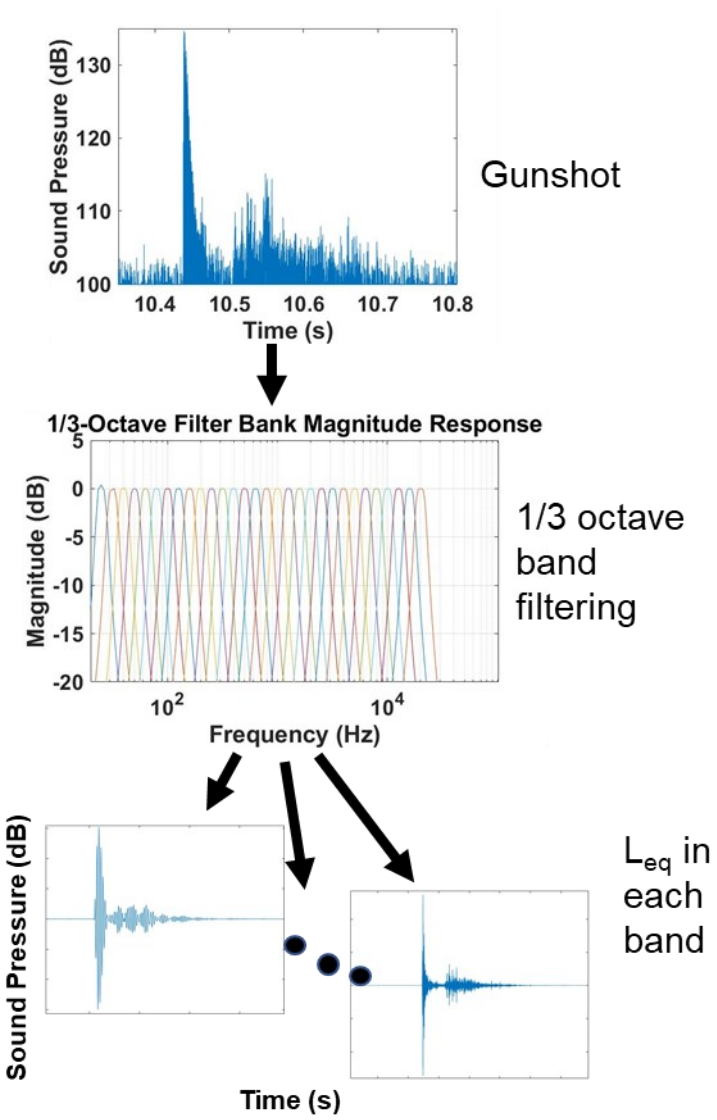


Figure 7: Gunshot decomposition in 1/3 octave bands

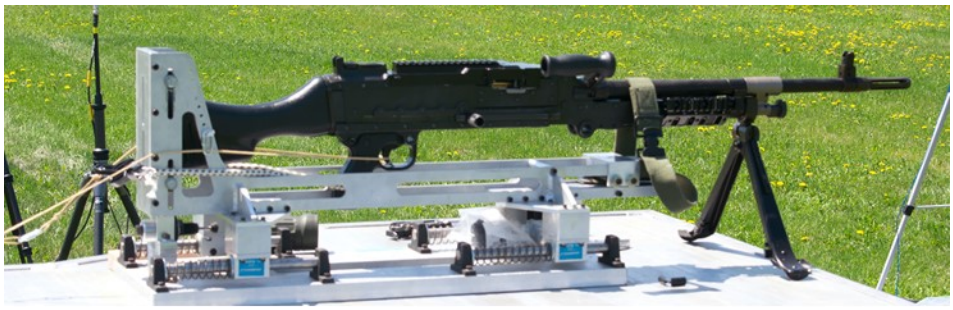
Chapter 3: Experimental Setup

3.1 Introduction

The purpose of this section is to present the methodologies used in designing the setup for gunshot acoustic data capture and the equipment chosen. A brief description of the guns and mortar used will be given.

3.2 Weapons Selection

We selected the guns and mortar for the gunshot identification according to ideal conditions, as indicated in Figure 8. Besides being able to be used in automatic mode, the guns were operated in a single-shot mode for the test with the shooter standing up. The mortar—with its much bigger 60mm-caliber—had an operation and function different from those of the two guns.



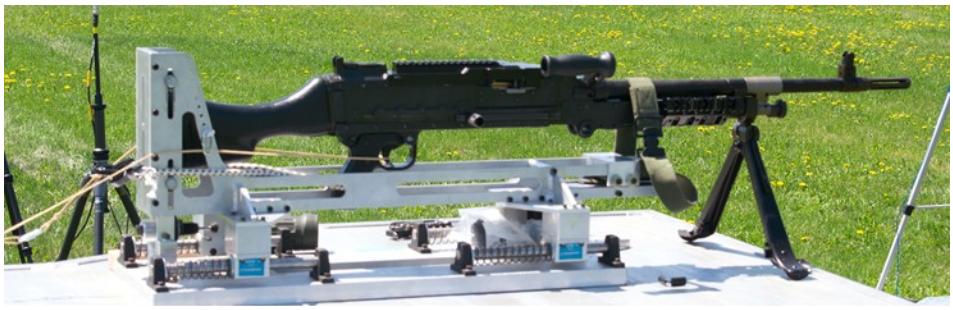
(a)

(b)

(c)

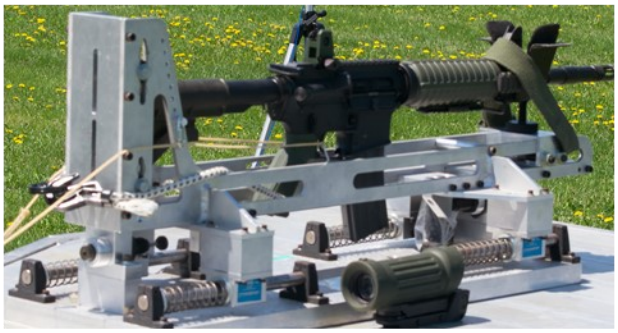
Figure 8: a) Gun #1, b) Gun #2 and c) Mortar used for the test

For the tests under reverberant environment, we used four similar guns. Gun #1 had a caliber of 7.62 mm and the other ones, 5.56 mm. For security reasons, the exact models are not named. They are simply being referred to as Gun #1, Gun #2, Gun #3, and Gun #4. Since the caliber and the operating mechanism of these guns are similar, comparable impulse noises are expected. Figure 9 shows the guns selected for the gunshot identification in a reverberant environment.

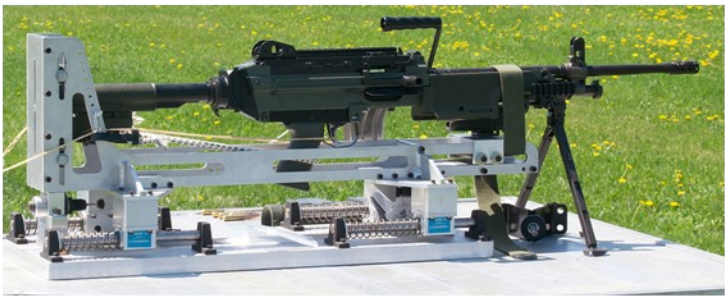


(a)

(b)



(c)



(d)

Figure 9: a) Gun #1, b) Gun #2, c) Gun #3, and d) Gun #4 used for the test

3.3 Microphones Selection

This study used G.R.A.S. 1/8" pressure microphones, model 46DD-FV, which can sustain up to 175 dB, reference 20µPa, sound pressure level. They also have a bandwidth of 10 Hz to 25 kHz with ± 1 dB flatness and 6.5 Hz to 140 kHz with ± 3 dB flatness. These parameters exceed those specified in the military standard MIL-STD-1474E [17] and those expressed by Brinkmann [13] and Rasmussen et al. [42].

Figure 10: Pressure Microphone used for the test

3.4 Data Acquisition Units

We used two Siemens LMS SCADAS SCR202 data acquisition units, 16 channels each, for the tests. In the master-slave mode, the units can synchronously record up to 32 channels of data with 204.8 kHz sampling rate. Their portability and 24 bits resolution on every channel provided reliable and precise measurements. Again, these parameters exceed the specifications of the military standard MIL-STD-1474E [17].

Figure 11: Acquisition Unit used for the test

3.5 Microphone Configurations

In compliance with the military standard [17], we installed microphones perpendicular to the incident pressure waves and placed them 1.6 m above ground with tripods for the guns and at 1.24 m for the mortar to limit reflections from the ground. Figure 12 and Figure 13 show the location of the microphones around the guns and mortar.

The height difference from guns to mortar is due to the operator's position for firing, standing up for the guns and kneeling for the mortar. The microphones were positioned in three concentrating arcs in sufficient numbers to capture the sound so that any asymmetry of sound propagations would be revealed. In addition, the selected distances captured the near-field effect and the far-field values, while avoiding saturations from the loud impulse signals. Finally, due to safety concerns and surrounding constraints, it was not possible to install microphones at 2 m 60°, 2 m 300°, 8 m 60° and 8 m 300° locations on the grid for

the guns test. They have been installed at 2m behind the guns to get as much recordings as possible.

Figure 14 shows a part of the field installation for the mortar test.

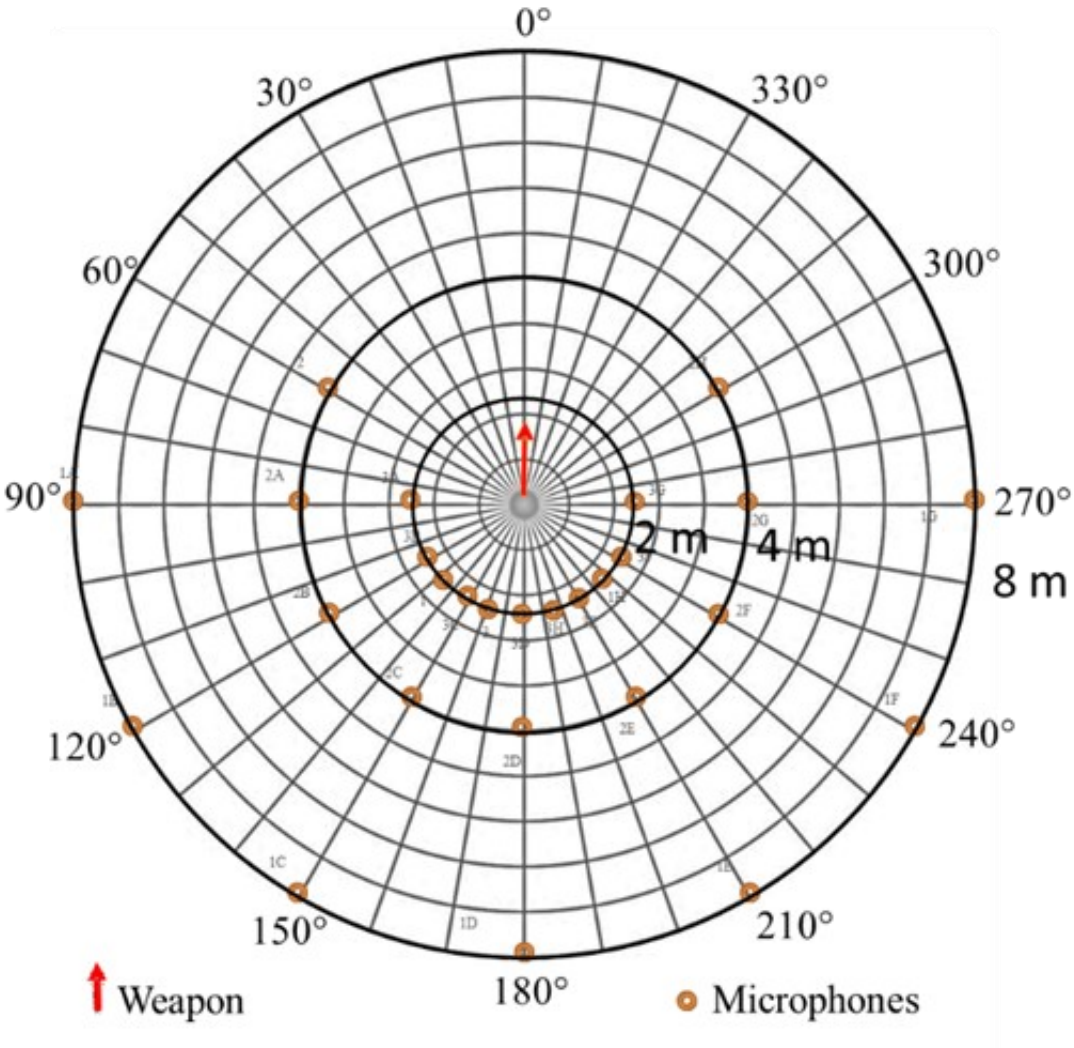


Figure 12: Microphones positions used for the gun test

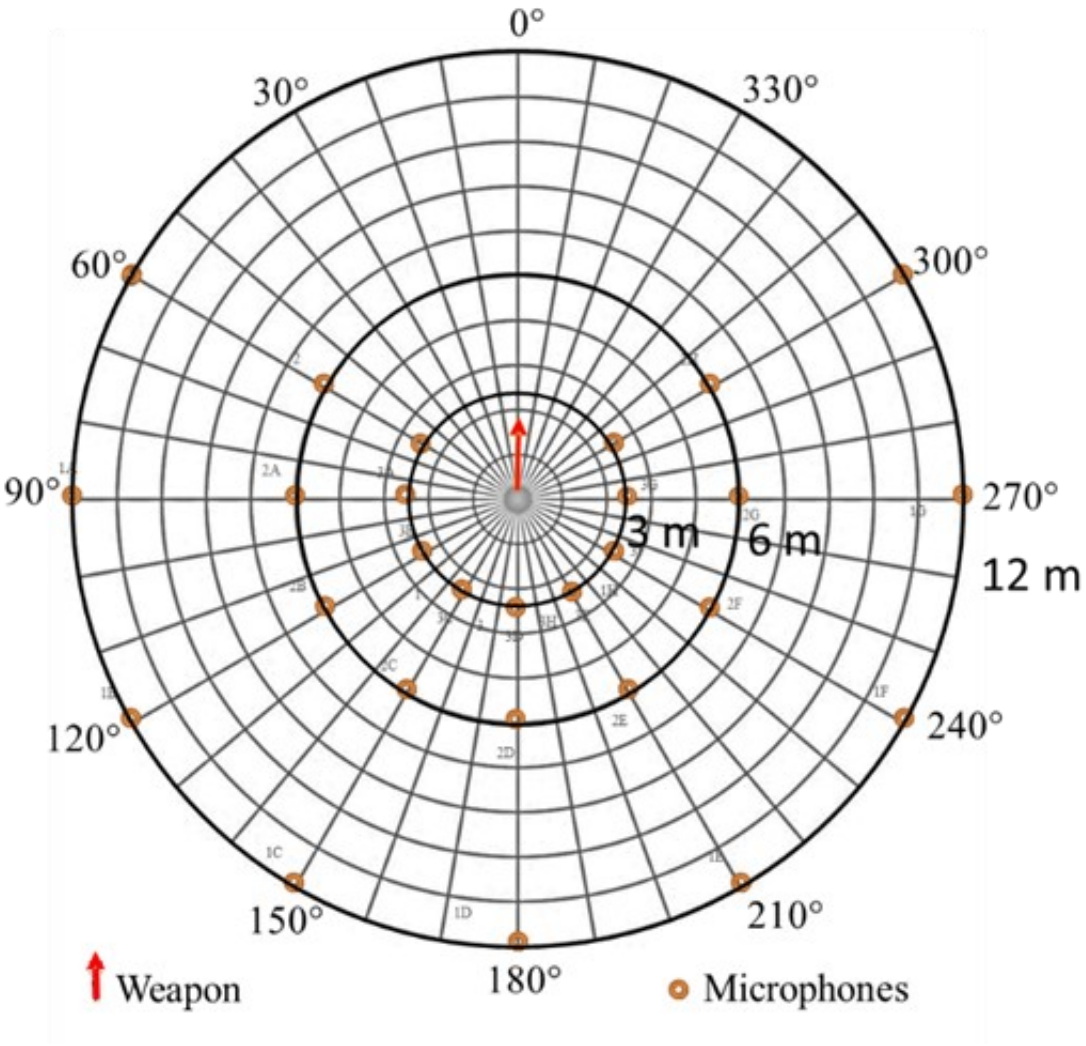


Figure 13: Microphones positions used for the mortar test

Figure 14: Microphones installation for the mortar test

3.6 Weather Conditions

The military standard [17] notes that the weather conditions affect the performance of the microphones, thus deserving consideration. The temperature, humidity, wind speed and direction, atmospheric pressure can all bias the results. Consequently, measurements must be taken with a wind speed below 20 km/h. We also need a day without fog, rain or snow. Our study took all the measurements between 10 and 15°C, well within the -20 to 55°C operating limitations of the microphones and data acquisition system. Finally, we used a calibrator regularly during the tests to ensure that the measurements by the microphones, coupled with the data acquisition unit, were valid and adjusted for the weather conditions at the moment of the tests.

Chapter 4: Gunshots Sound Identification

4.1 Gun Identification Under Ideal Conditions

The purpose of this section is to present the methodologies and results obtained when identifying the signal coming from two guns and a mortar under ideal conditions.

4.1.1 Introduction

Analyzing gunshot sounds can help identify the type of gun used. This is important for determining the appropriate public safety actions when a gunshot sound is detected in a public space. We collected acoustic data from two different guns and a mortar for analysis. Five shots were fired from each gun and the mortar. To capture their sound, including any non-symmetric sound propagation from the gunshots, we placed 27 high dynamic range pressure 1/8" microphones around the guns at 2, 4, and 8 m, forming a polar grid pattern discussed in Chapter 3. A similar set-up was used for the mortar with distances of 3, 6 and 12 m. Audio signals from all microphones were captured at 204.8 kHz sampling rate synchronously to preserve the fidelity of the impulse nature of the gunshots. An image-based analysis method was developed to take advantage of the recent advancement of image recognition techniques. In this thesis, we applied the Continuous Wavelet Transform (CWT) to the time-frequency gunshot recordings and obtained the spectrograms as images. A two-stage machine-learning classifier then processed the images to identify the specific gun or mortar that the gunshot originated from.

4.1.2 Peak Sound Pressure Levels

The positions in distance and propagation angle are related to Figure 12 for the guns and Figure 13 for the mortar. Peak pressure is the preferred metric in this study to avoid the band-limiting effect, as it would be while using A-weighting, or any energy averaging such as ***Leq*** or ***SEL*** (sound exposure level) [29]. Raw waveform examples from the gunshots are presented in Figure 18. Gunshots can easily be distinguished in the time domain between the guns and the mortar but not between the 2 guns. See Table 5 to Table 7, as well as Figure 15 to Figure 17, for the peak sound pressure levels measured in dB.

When we doubled the distance, the mean attenuation was 5.4 dB for gun #1, 5.5 dB for gun #2, and 6.0 dB for the mortar. The directivity of the sound propagation was also greater for the two guns than for the mortar, especially noticeable when the peak pressure level behind the guns, at 180°, as shown in Figure 15 and Figure 16.

Table 5: Peak sound pressure level as a function of distance and angle of propagation, gun #1, 7.62mm

	2 m	4 m	8 m
60°		158.7	
90°	159.1	154.1	148.0
120°	153.3	148.5	143.7
135°	149.7		
150°	147.9	144.3	138.6
165°	145.4		
180°	147.5	140.7	133.8
195°	147.5		
210°	150.0	144.3	138.6
225°	149.9		
240°	153.0	149.3	143.9
270°	160.0	154.5	148.4
300°		160.9	

Table 6: Peak sound pressure level as a function of distance and angle of propagation, gun #2, 5.52 mm

	2 m	4 m	8 m
60°		158.0	
90°	156.8	151.5	145.9
120°	153.3	146.4	141.4
135°	150.3		
150°	148.0	143.7	139.1
165°	144.1		
180°	144.0	139.6	133.8
195°	147.5		
210°	150.4	144.2	138.5
225°	150.3		
240°	152.4	146.2	140.0
270°	156.4	151.7	145.5
300°		157.0	

Table 7: Peak sound pressure level as a function of distance and angle of propagation, Mortar charge 3

	3 m	6 m	12 m
60°	167.8	162.5	156.8
90°	168.0	162.3	155.3
120°	167.9	163.2	154.9
150°	166.2	159.6	154.4
180°	165.8	159.5	154.3
210°	163.5	159.1	154.8
240°	166.1	160.2	153.5
270°	168.7	160.0	154.8
300°	169.4	162.0	157.3

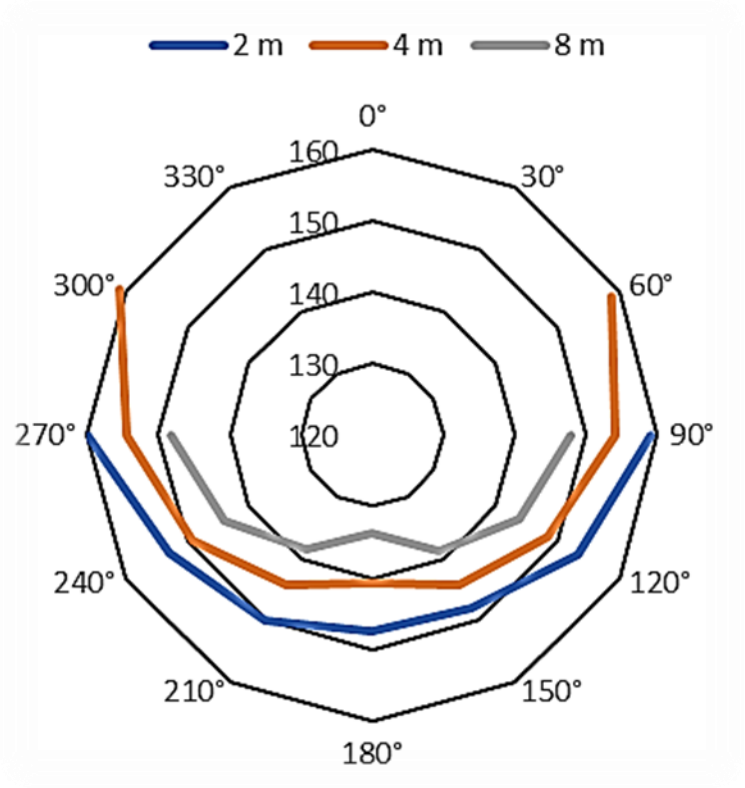


Figure 15: Peak sound pressure levels as a function of distance and angle of propagation for gun #1

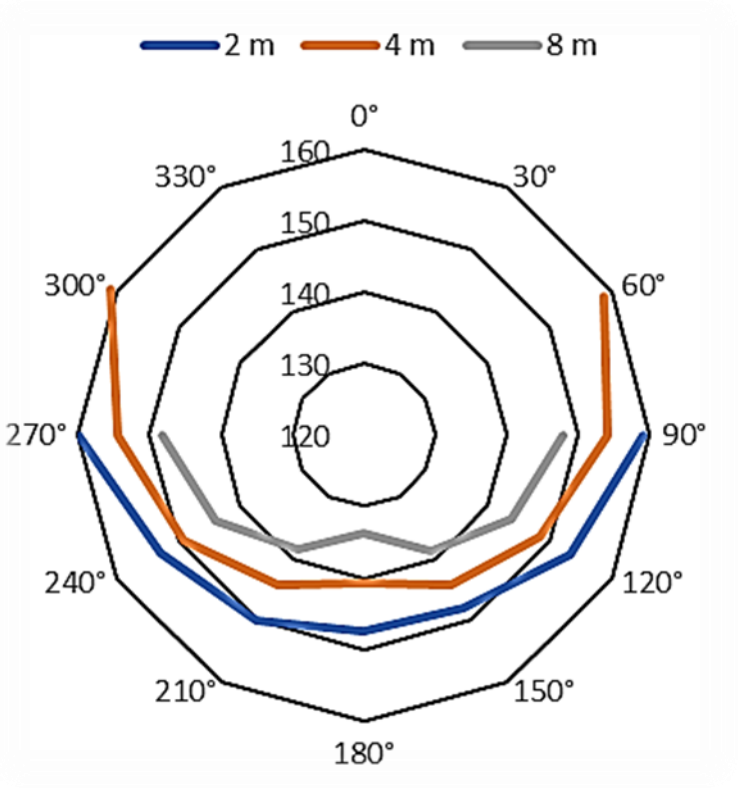


Figure 16: Peak sound pressure levels as a function of distance and angle of propagation for gun #2

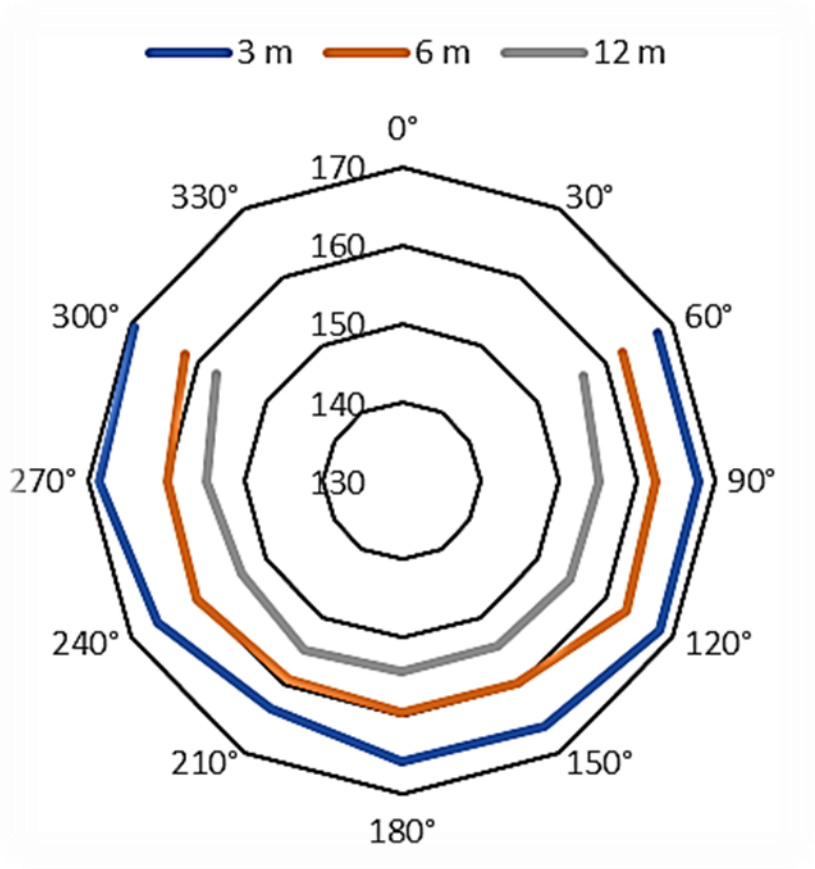


Figure 17: Peak sound pressure levels as a function of distance and angle of propagation for the mortar

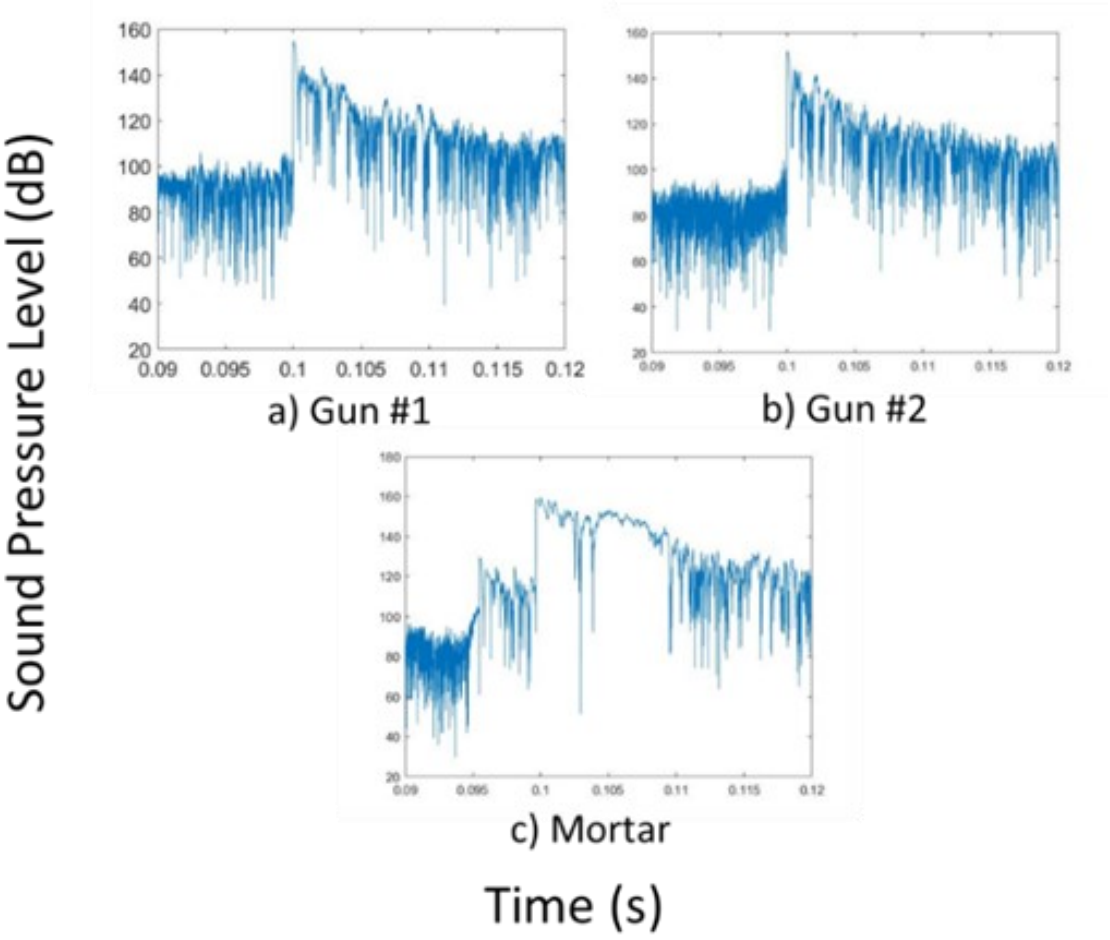


Figure 18: Gun sound pressure level at 4 m, 270° for gun #1 (a) and gun #2 (b); and at 6 m, 270° for mortar (c)

4.1.3 Spectral Analysis and Propagation

The spectrograms varied more with the distance of propagation for the two guns and more with the angle of propagation for the mortar. Figure 19, Figure 20 and Figure 21 illustrate the various spectrograms for the two guns and the mortar.

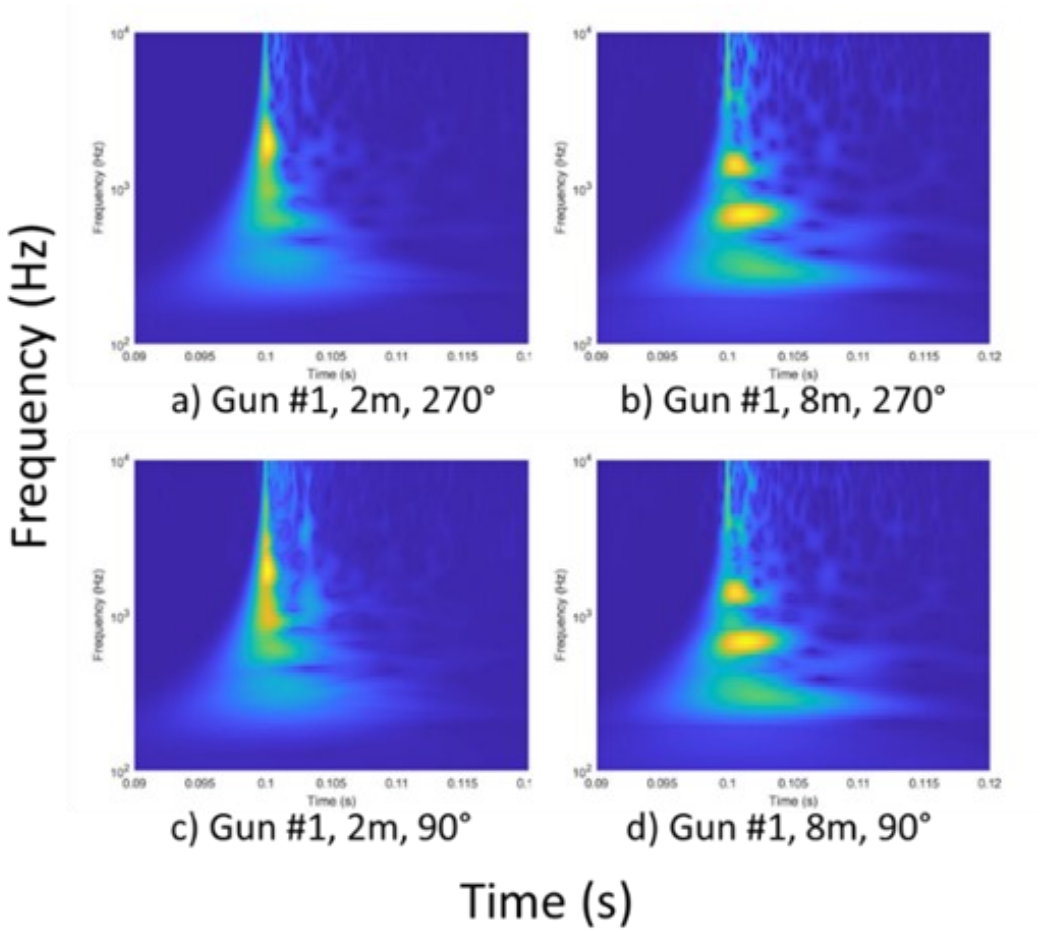


Figure 19: Gun #1 CWT spectrograms at 2 m, 270° (a), 8 m, 270° (b), 2 m, 90° (c) and 8 m, 90° (d)

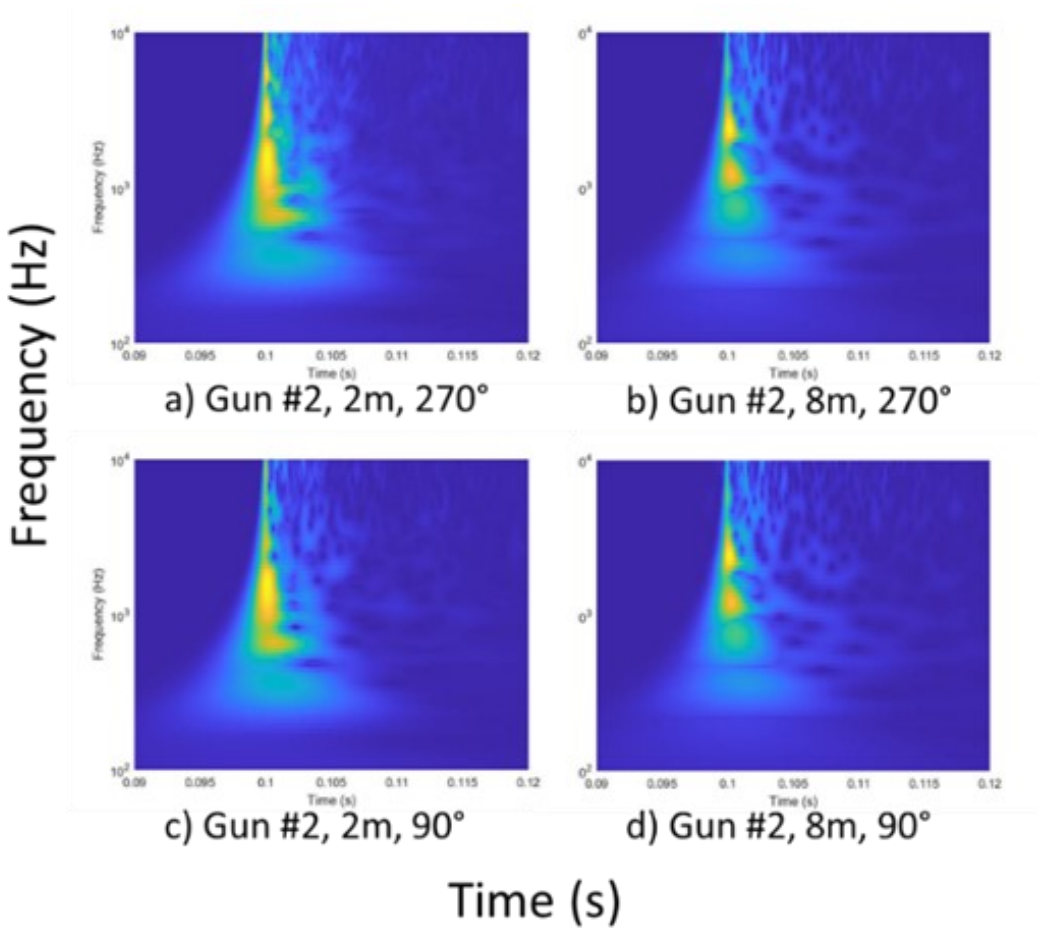


Figure 20: Gun #2 CWT spectrograms at 2 m, 270° (a), 8 m, 270° (b), 2 m, 90° (c) and 8 m, 90° (d)

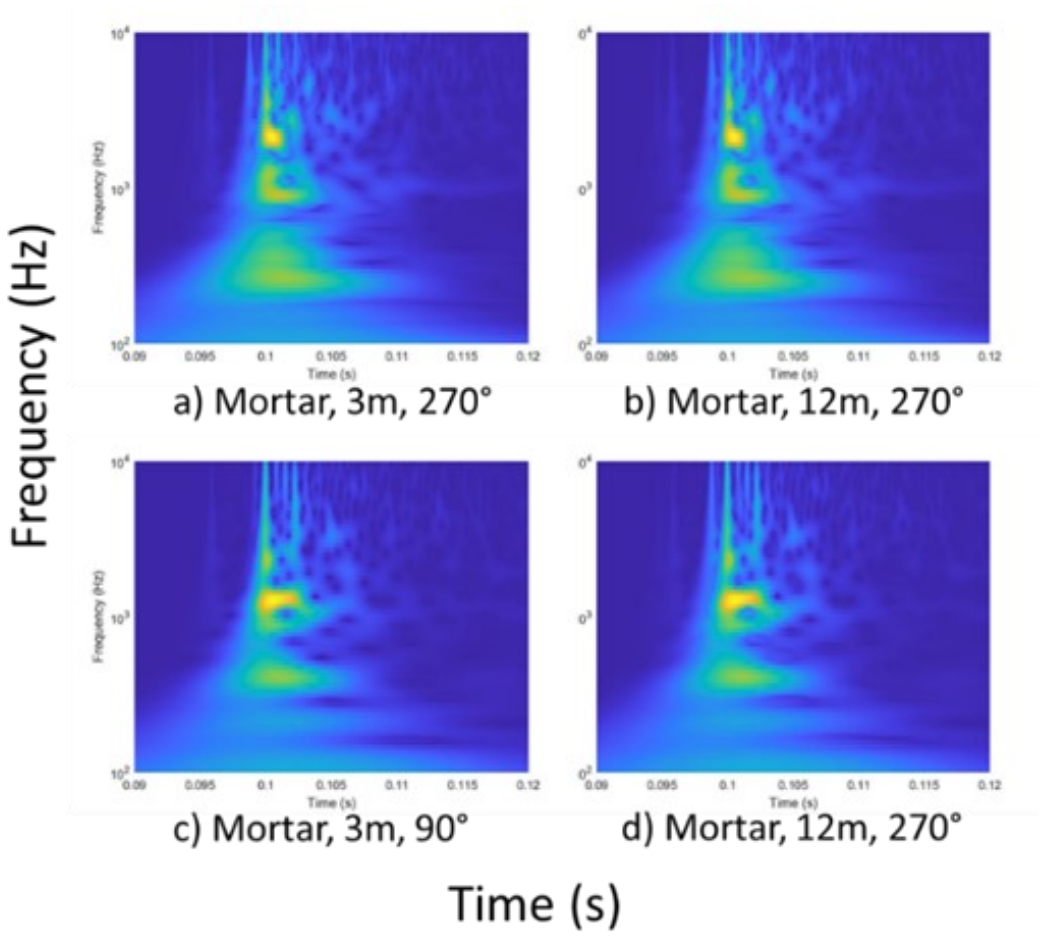


Figure 21: Mortar CWT spectrograms at 3 m, 270° (a),12 m, 270° (b), 3 m, 90° (c) and 12 m, 90° (d)

4.1.4 Machine-Learning Classifier Image Recognition

The image-recognition process, shown in Figure 22, is similar to a novel approach used in the medical domain [43]. The bank of image for the first stage came from the five shootings of each gun recorded with the setup-up of 27 microphones, as shown in Figure 12. Consequently, each gun had a bank of 135 images for the learning and testing process. Similarly, we used the setup of 27 microphones shown in Figure 13 to record five shots with the mortar, giving also a bank of 135 images. To avoid any saturation and to simplify the analysis, we used the number three propellant rings, also called charge #3, for the

mortar. Consequently, 405 images were used in the machine-learning process to identify the gunshots during the second phase.

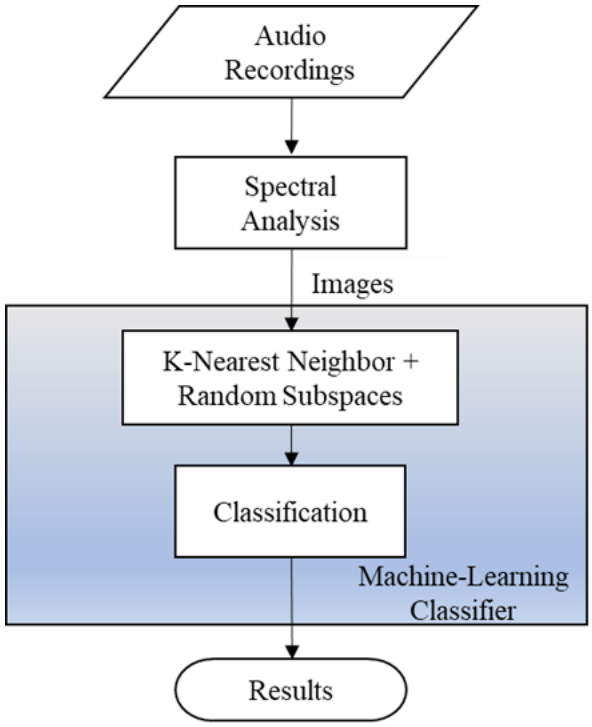


Figure 22: Process flowchart of the method

To select the machine learning classifier, twenty-two methods were tried. Each one was computed five times to evaluate its mean (μ) and standard deviation (σ). The k-nearest neighbor (KNN) with random subspaces (Subspace KNN) showed the highest accuracy and the lowest standard deviation. Because a low standard deviation indicates a higher repeatability, we selected this method. Table 8 shows the accuracy obtained with various machine learning models during the learning stage.

KNN, a simple classifier, has been known for years [44] and the use of random subspaces can enhance its precision [45]. In this study, we started the image processing using the speed up robust features (SURF) algorithm [46] to extract feature points from a

given image. Afterward, the extracted feature points are clustered using the K-means algorithm to give an image-specific bag of features which was then used to train the KNN classifier with a subspace dimension of 125 randomly selected elements from the 250 features selected with SURF. After that, 30 random nearest neighbor learners were created and the final decision or classification was based on the most popular of them.

Given the similarity of their ammunition caliber and the similarity of their mechanism, the guns spectrograms are similar. Figure 23 shows a confusion matrix obtained during one of the learning iterations. Figure 19 and Figure 20 also showed a similar result.

This study used 80% of the images for training the classifier, the rest for testing the accuracy. In view of the limited number of images, the process of training and testing was repeated 50 times with randomly selected image banks. The resulting average is the value that was reported.

Table 8: Tested Machine Learning Models

Machine Learning Model	μ	σ
Fine Tree	91.1	2.8
Medium Tree	91.1	2.8
Coarse Tree	85.9	4.4
Linear Discriminant	93.5	2.8
Kernel Naïve Bayes	77.5	4.5
Linear SVM	97.4	0.8
Quadratic SVM	97.6	0.8
Cubic SVM	97.0	1.4
Fine Gaussian SVM	84.4	0.6
Medium Gaussian SVM	97.4	0.9
Coarse Gaussian SVM	75.2	2.1
Fine KNN	96.6	1.3
Medium KNN	91.5	2.1
Coarse KNN	67.8	1.7
Cosine KNN	90.5	0.5
Cubic KNN	92.1	1.5
Weighted KNN	94.7	0.9
Boosted Trees	35.1	4.8
Bagged Trees	86.7	1.8
Subspace Discriminant	96.8	0.7
Subspace KNN	98.5	0.4
RUSBoosted Trees	35.1	4.8

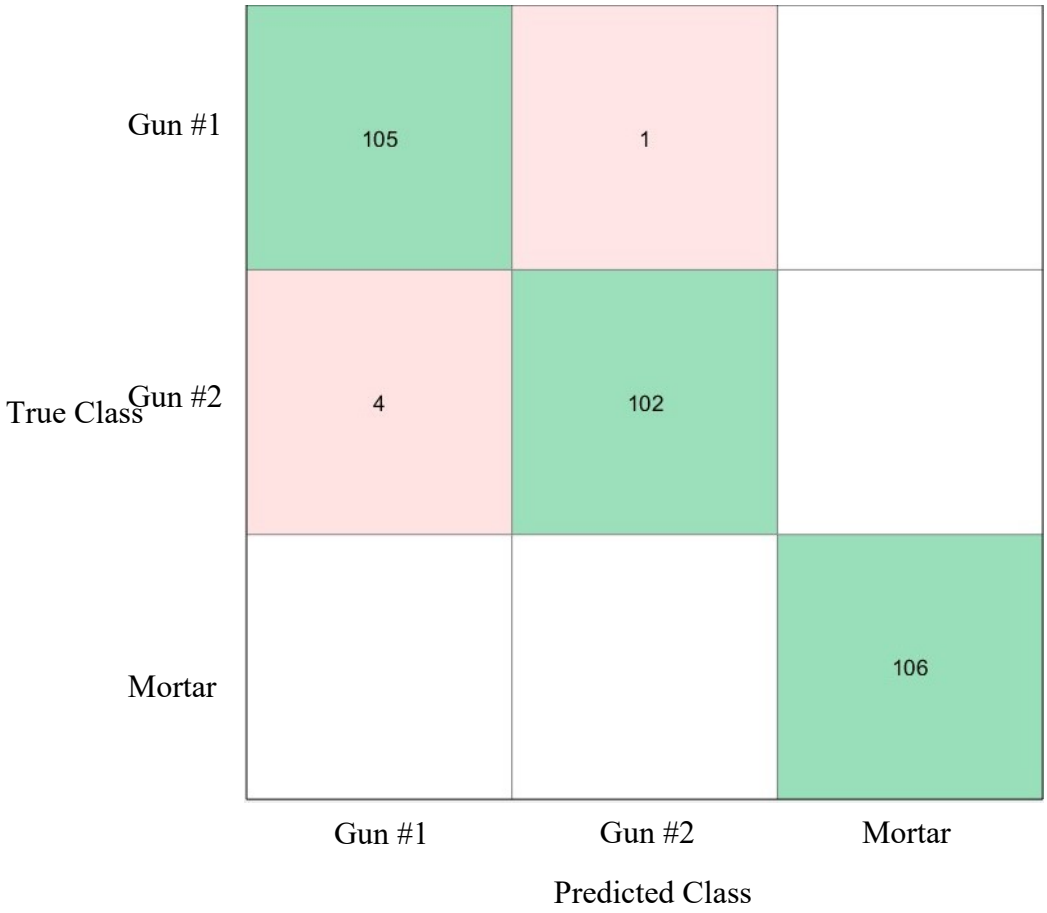


Figure 23: Confusion Matrix for the two guns and the mortar, Subspace KNN

4.1.5 Gunshot Identification

With the process discussed in section **4.1.4**, the accuracy obtained with the machine learning identification reached 98.1% when asked to identify correctly whether a spectrogram is generated from gun #1, gun #2 or the mortar. As noticed in section 4.1.3 and despite having a right-handed shooter, the peak sound pressure levels gave symmetric results from left to right and showed that the CWT spectrograms differ more with distance than with what is observed from the propagation angle. Therefore, using the CWT spectral analysis, the accuracy increased from 98.1 % to 99.8 % when using results from only 4 and 8 m in distance for the guns.

4.1.6 Discussion

As discussed is section 4.1.2, we used the peak non-weighted sound pressure level to avoid frequency attenuation that comes when applying A-weighting filter [47]. Also, unlike background noises [48], shorter recording times are preferred to avoid losing resolution at the impulse noise when microcomputer signal processing resources are limited. Consequently, we didn't applied any weighting nor energy-averaging method, giving the values in Table 5 to Table 7, Figure 15, as well as the spectrograms of Figure 19, Figure 20 and Figure 21.

Extracted from the field measurements, the peak sound pressure level decreased with the distance from the weapons. Depending on the surroundings and weather conditions, the level of attenuation should have decreased by 6 dB when the measuring distance doubles [12]. This was the value we got for the mortar. For the guns, we got 5.0 dB and 5.4 dB for gun #1 and #2 respectively when the distance increased from 2 m to 4 m. The near-field effect might have been a factor, especially for data recorded at the microphones located at 2 m. For example, Figure 19 and Figure 20 show the signal had a peak value at 800 Hz. At this frequency, using 343 m/s as the average speed of sound in the air, the sound wavelength is 0.43 m, and hence, will be defined as the near-field region [49]. In the near-field region, the sound pressure and acoustic-particle velocity are not in phase [49] causing non-linearity.

Due to the presence the shooter's body, the signal attenuation shown in Figure 15 and Figure 16 is at the propagation angle of 180°. This observation had been measured before [6] [50]. In short, the results indicate that gunshot sounds from some weapons can

have a non-linear propagation pattern. In the case of the mortar, the main factor affecting the amplitude and spectral analysis results from the audio signals was the angle of propagation. But for the two guns, the distance affected the analysis results the most.

Finally, the accuracy increased from 98.1 % to 99.8 % when spectrograms from the 2 m location were removed. At this range of identification accuracy, the proposed method gives a solid foundation for applications such as threat evaluation and forensic identification. Therefore, reference databases built from audio signals can be implemented in surveillance systems and onboard police or military vehicles to get real-time emergency systems that, on top of simply detecting a threat such as a gunshot, can also evaluate its seriousness and guide the appropriate response. Another application is the identification for forensic purposes. In this case, only known and different weapon models have been studied. More studies would be needed to validate the effectiveness when using a specific weapon model. With the proliferation of audio sensors in surveillance systems [51], more and more audio recordings of events exist. They are potentially useful in legal investigations and as pieces of evidence in court.

4.1.7 Conclusion

We have presented the methodology for designing an experimental setup including the placement of microphones and the choice of data acquisition system specifically chosen for the measurement of high-level impulse audio signal from gunshots. The setup recorded the gunshot sounds from two guns and a mortar at shooting ranges. Taking advantage of the recent advancements in image recognition, we developed a new image-based approach to acoustic signal analysis. To investigate the effectiveness of this image-based analysis

approach for gunshot sounds, we chose the problem of gun identification as a test bench.

We have demonstrated that an image-based signal analysis method can be effective in analyzing gunshot sounds. In this example, we obtained a classification accuracy of 98.1%.

The CWT and the machine-learning classifier chosen are examples of possible analysis techniques. Further studies are needed to investigate the effectiveness of other analysis techniques. Furthermore, the robustness of the measurement and the analysis methods we have developed require evaluation for more complex environmental conditions, such as a reverberant environment and guns using similar ammunitions. Also, evaluation in real life situations like military trainings and operations would improve the knowledge and validate the method.

4.2 Gun Identification Under Reverberant Environment

The objective of this section is to present the methodologies and results obtained when identifying the signal coming from four guns under reverberant environment and to compare the STFT and the CWT.

4.2.1 Introduction

In real-life applications, gunshots may originate in urban environments surrounded by buildings. Therefore, an identification method that is not sensitive to reverberations is preferred [52] [53]. To evaluate the effect of reverberations on the gun identifier used in this study, we simulated reverberations to add controlled amount of reverberation into the acoustic signal.

We applied the method presented in 4.1, but with two spectral analyses that are compared with controlled amount of reverberation added onto the acoustic signals. We recorded gunshots sounds from four different guns using the same ammunition for three of them and a slightly bigger version for the fourth. We then applied spectral analysis to obtain the time-frequency breakdown of each gunshot signal that could be represented as an image. By using image-recognition techniques, we analyzed and classified the result.

The two spectral analysis methods we used are the Short-Time Fourier Transform (STFT) and the Continuous Wavelet Transform (CWT). The STFT is a classical method for spectral analysis of impulse noises such as gunshots [24]. On the other hand, the CWT has shown advantages when analyzing signals of this kind [24].

From the spectral analysis, a machine-learning based image recognition method is applied to identify which gun a particular gunshot originated from. Different models have been tried and the k-nearest neighbor (KNN) with random subspaces was chosen. As discussed in section 4.1, among modern image-recognition techniques, the k-nearest neighbor classifier is one of the most widely used classifiers [54] due to its good generalization and easy implementation [44]. When combined with random subspaces, its accuracy can be further improved [45].

4.2.2 Audio Recordings Spectral Analysis

For this analysis, we performed these process steps:

1. Detected the gunshots signal peaks in the audio recordings.
2. Applied a simple peak detector with a threshold of 120dBA to the audio files, since the level of the ambient noises was much lower when compared to the gunshot

signals.

3. Selected a window of 0.2 s around the detected peaks.
4. Calculated the spectral analysis for each one.

Figure 24 shows the flow of this process.

As applied in section 4.1, we developed a novel image-based analysis method to perform gunshot sound analysis. Our approach differs from the traditional one, in which gunshot sound identification is performed in the time-frequency domain. That approach first analyses the known acoustic signature of the various guns followed by applying pattern recognition techniques that look for these known acoustic signatures [16] [14].

Similar to the system used by Kim et al. [55], our system consists of two stages. In the first, the audio signal is converted to images. In the second, the image-recognition technique analyzes the images and identifies any recognizable spectrogram. To test the effectiveness of this image-based approach on gunshot sound analysis, we computed the Short-Time Fourier Transform (STFT) and the Continuous Wavelet Transform (CWT), converting the gunshot sounds recording to spectral images. Afterward, a machine-learning classifier was first trained with known data and then used to distinguish the weapon a particular gunshot originated from. When evaluating the effect of simulated reverberation, the training data is still the same set of images that contain no reverberation. Simulated reverberations are then added to 20% of the images used for testing.

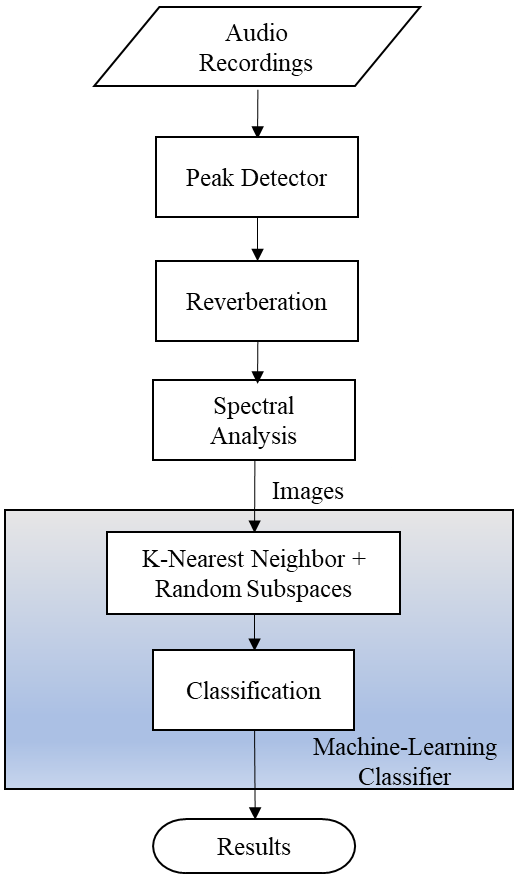


Figure 24: Process flowchart of the method used in this chapter

4.2.3 Machine Learning and Classification

To set the baseline for the whole study in this chapter, we used the bank of images resulting from the spectral analysis without reverberation for machine learning and image recognition. Each gun was fired five times and recordings were done using 27 microphones. That gave a database of 135 images for each gun, 540 images without reverberation in total.

Section 4.1 noted various machine learning models have been tested. Table 9 gives the results after five iterations. Because of the accuracy reached when identifying the gunshots, we chose the k-nearest neighbor with random subspaces (Subspace KNN) model. It gave the highest accuracy with the STFT and the CWT with a low standard deviation.

Based on the results, we selected the same machine learning method as the one used in section 4.1.

Table 9: Tested Machine Learning Models with the STFT and the CWT for the four guns

Machine Learning Model	STFT		CWT	
	μ	σ	μ	σ
Fine Tree	89.8	3.4	81.2	3.4
Medium Tree	89.7	3.4	78.8	2.5
Coarse Tree	64.5	3.4	56.2	5.7
Linear Discriminant	93.2	0.8	85.6	0.4
Kernel Naïve Bayes	73.3	2.9	45.1	2.4
Linear SVM	95.9	1.0	88.3	1.9
Quadratic SVM	98.3	0.4	92.7	1.4
Cubic SVM	97.3	0.7	92.5	1.0
Fine Gaussian SVM	90.1	1.2	85.3	1.4
Medium Gaussian SVM	96.8	0.9	90.4	1.2
Coarse Gaussian SVM	67.6	3.7	41.3	5.9
Fine KNN	97.3	0.7	92.0	0.9
Medium KNN	86.6	0.9	81.4	2.6
Coarse KNN	49.4	6.3	42.2	2.3
Cosine KNN	89.8	1.6	83.4	2.3
Cubic KNN	85.7	1.9	82.4	4.0
Weighted KNN	95.7	0.3	88.9	1.0
Boosted Trees	94.3	2.0	87.6	2.6
Bagged Trees	90.3	2.2	77.2	2.6
Subspace Discriminant	94.8	1.1	86.9	1.0
Subspace KNN	99.0	0.4	93.5	1.2
RUSBoosted Trees	88.7	4.5	78.4	3.0

4.2.4 Reverberation Model

Audio data obtained in the field often contains reverberations due to reflections from the surrounding surfaces [56]. The amplitude, delay and decay in the reverberation vary with the specific environmental condition at the moment of the recording [57]. So, it is important to use model that can control such parameters. The reverberator function in MATLAB [58] was used to introduce a controlled amount of reverberation into the audio signal. Figure 25 shows the reverberation model used in the reverberator function. The parameter values chosen are a Delay of 0, a Decay Factor of 0.5, a High Frequency Damping of 0.0005, and a Diffusion of 0.5. Furthermore, to test the robustness of the STFT and CWT as the spectral analysis methods chosen in the gun identifier, we computed the resulting identification accuracies with varying levels of reverberation. We did this by adjusting the reverberations to the peak-direct-sound ratio in the reverberator ranged from 0 to 0.5.

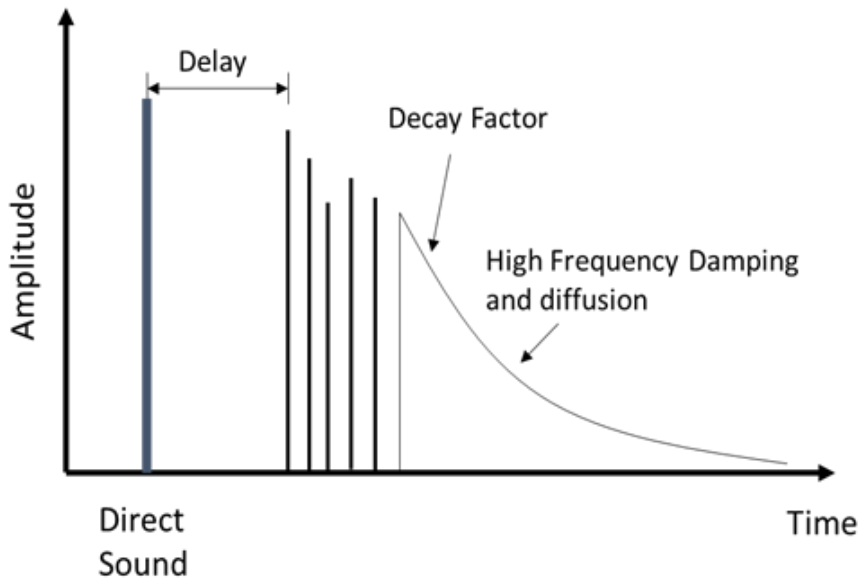


Figure 25: Reverberation model [58]

4.2.5 Spectral Analysis

Examples of the CWT and STFT spectral analysis from Gun #1 without and with reverberation are shown in Figure 26 to Figure 41. These are selected images from which the bags of features are extracted.

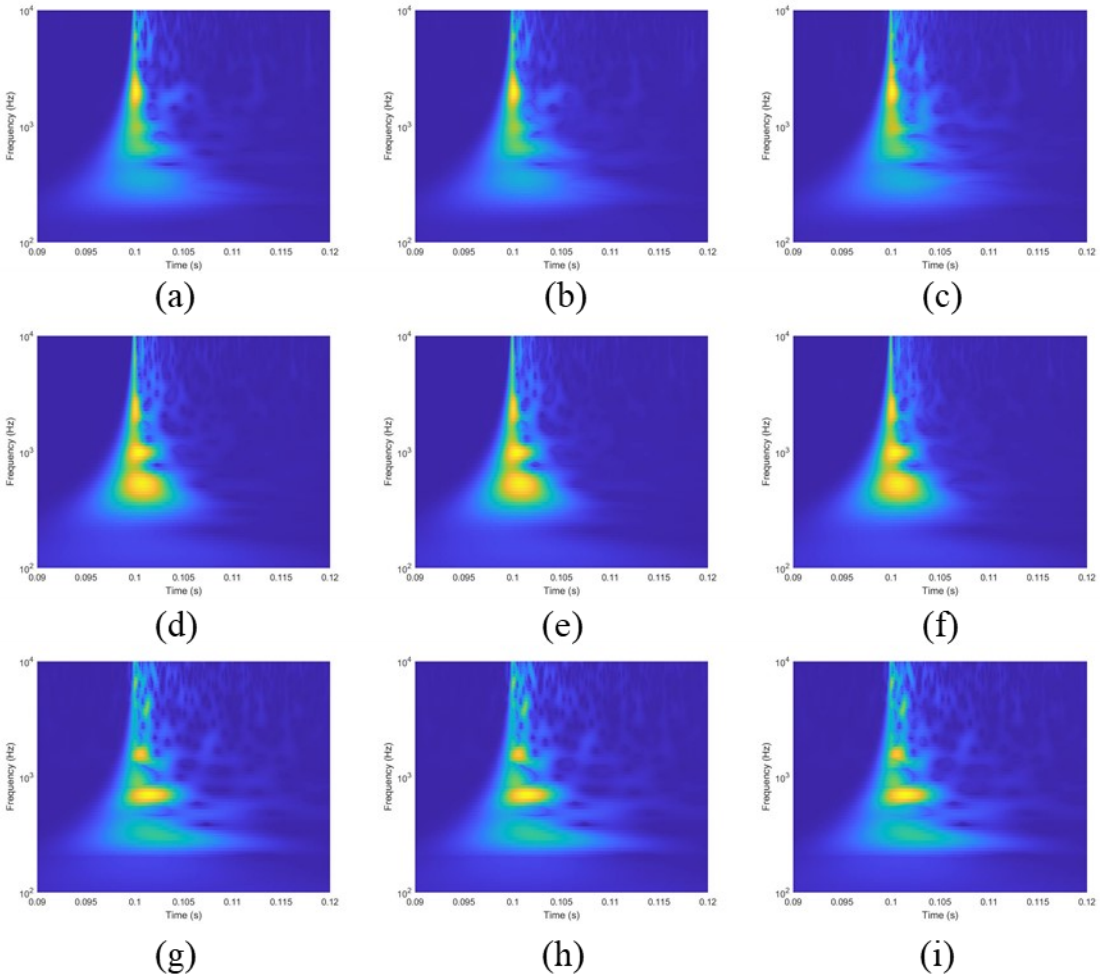


Figure 26: Gun #1 CWT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

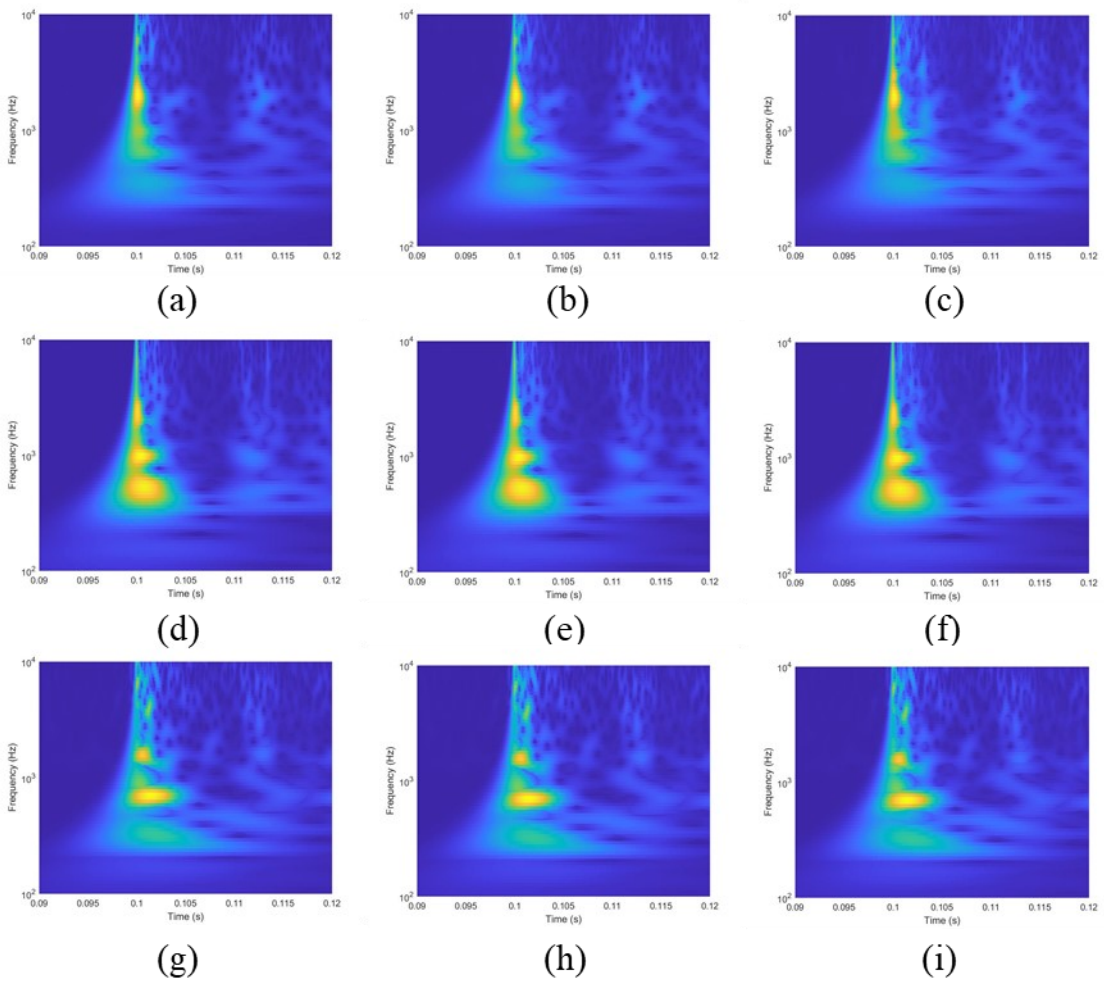


Figure 27: Gun #1 CWT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

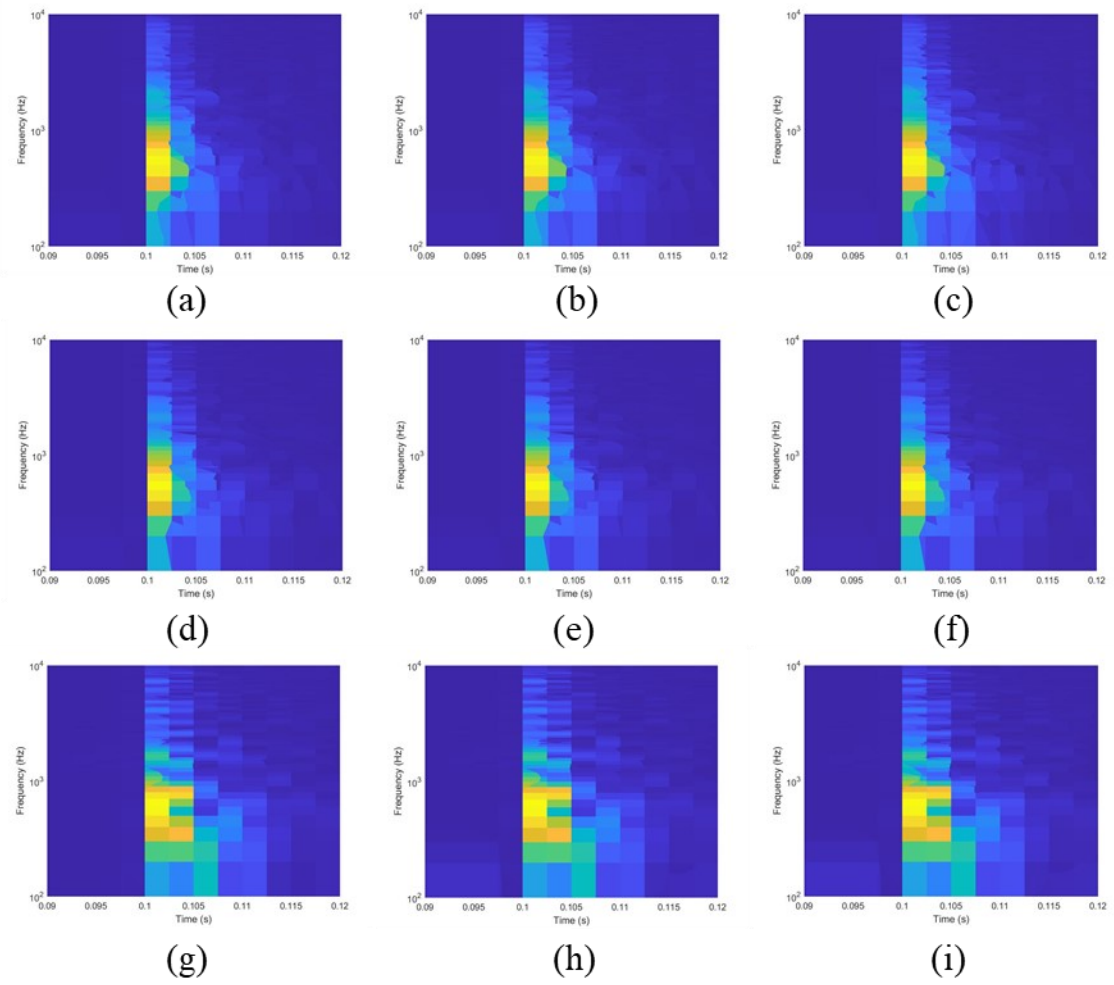


Figure 28: Gun #1 STFT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

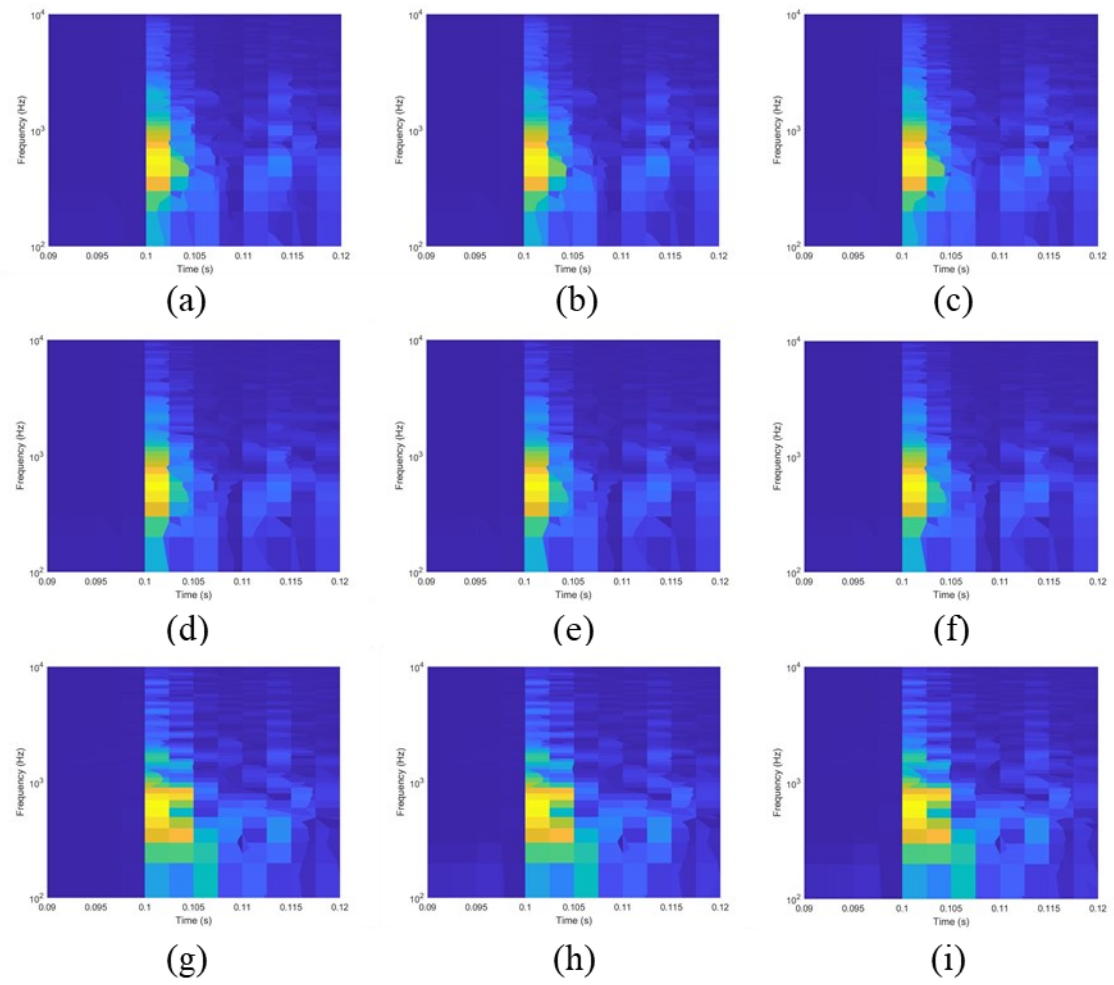


Figure 29: Gun #1 STFT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

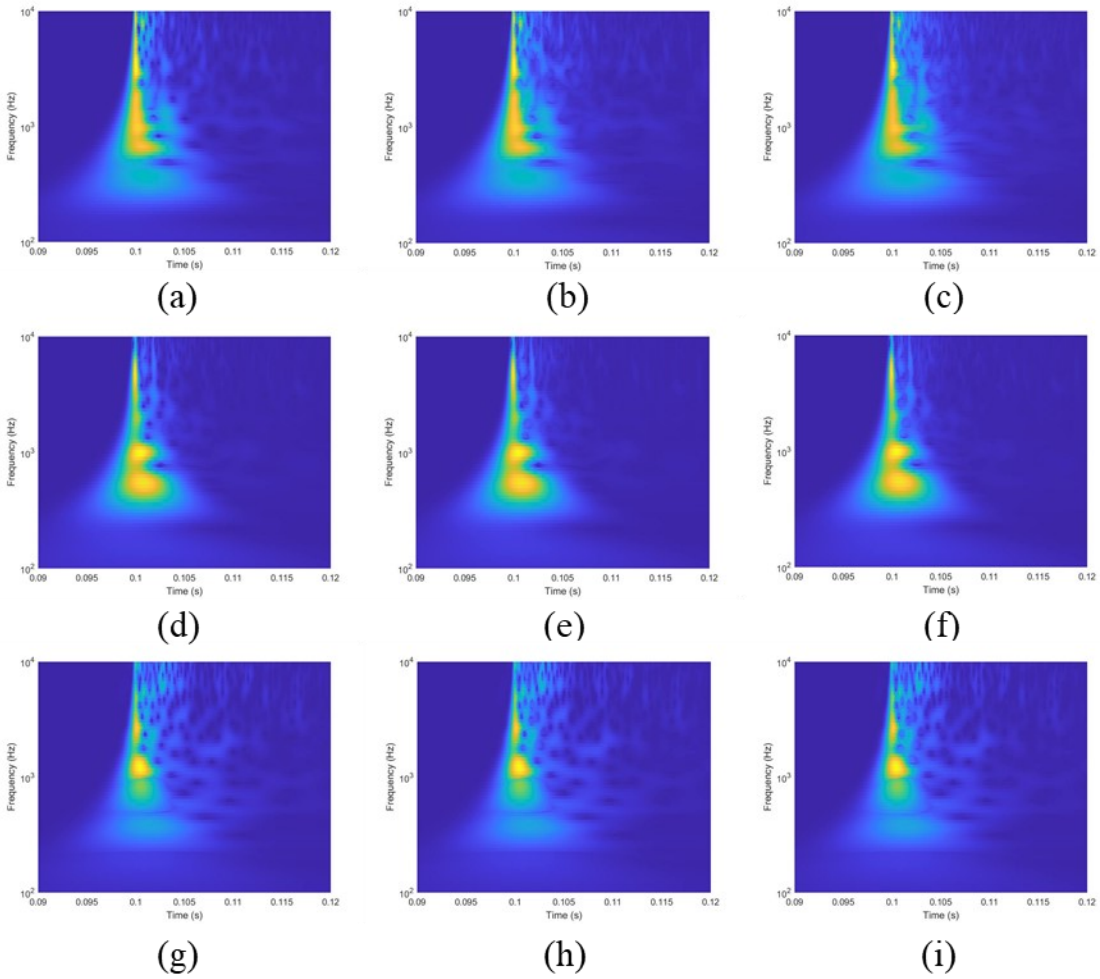


Figure 30: Gun #2 CWT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

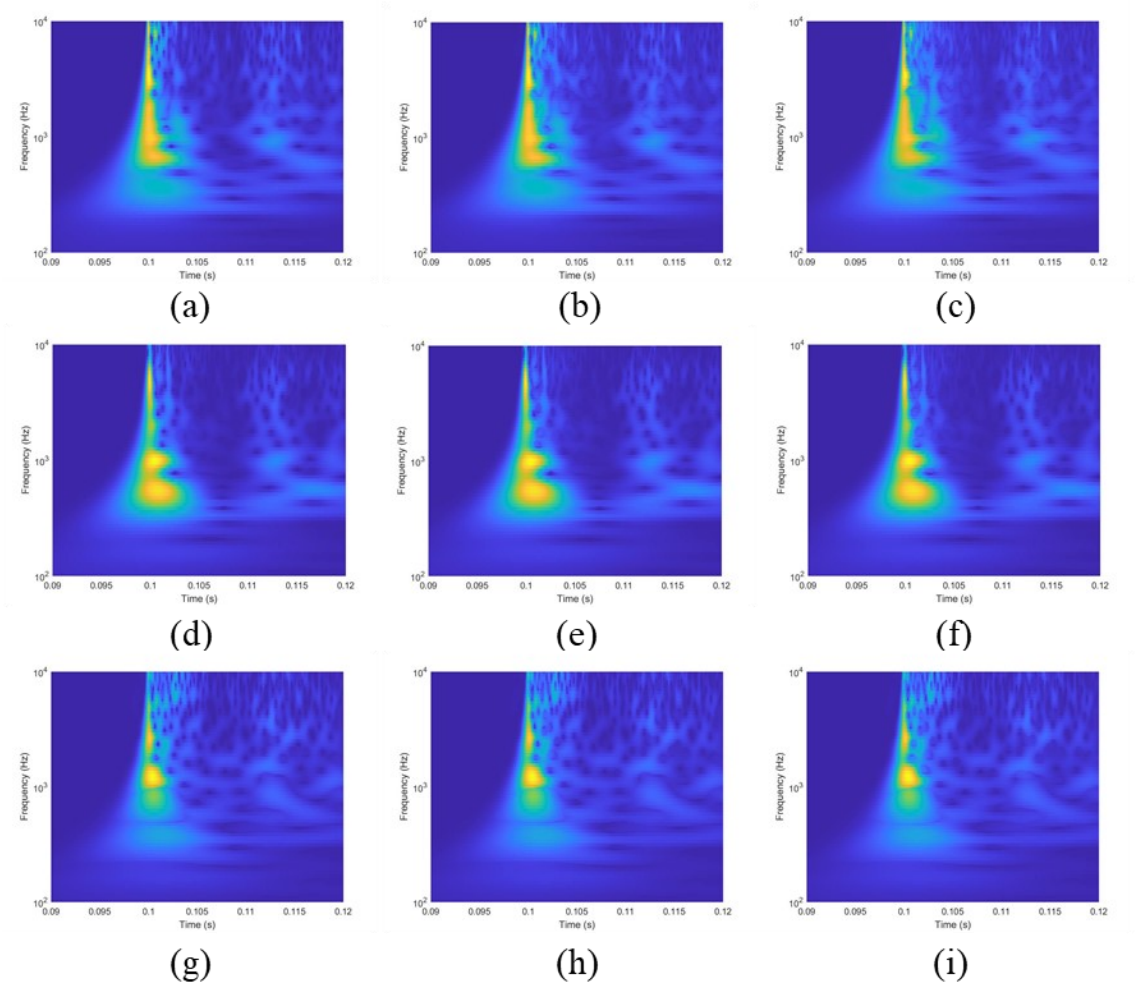


Figure 31: Gun #2 CWT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

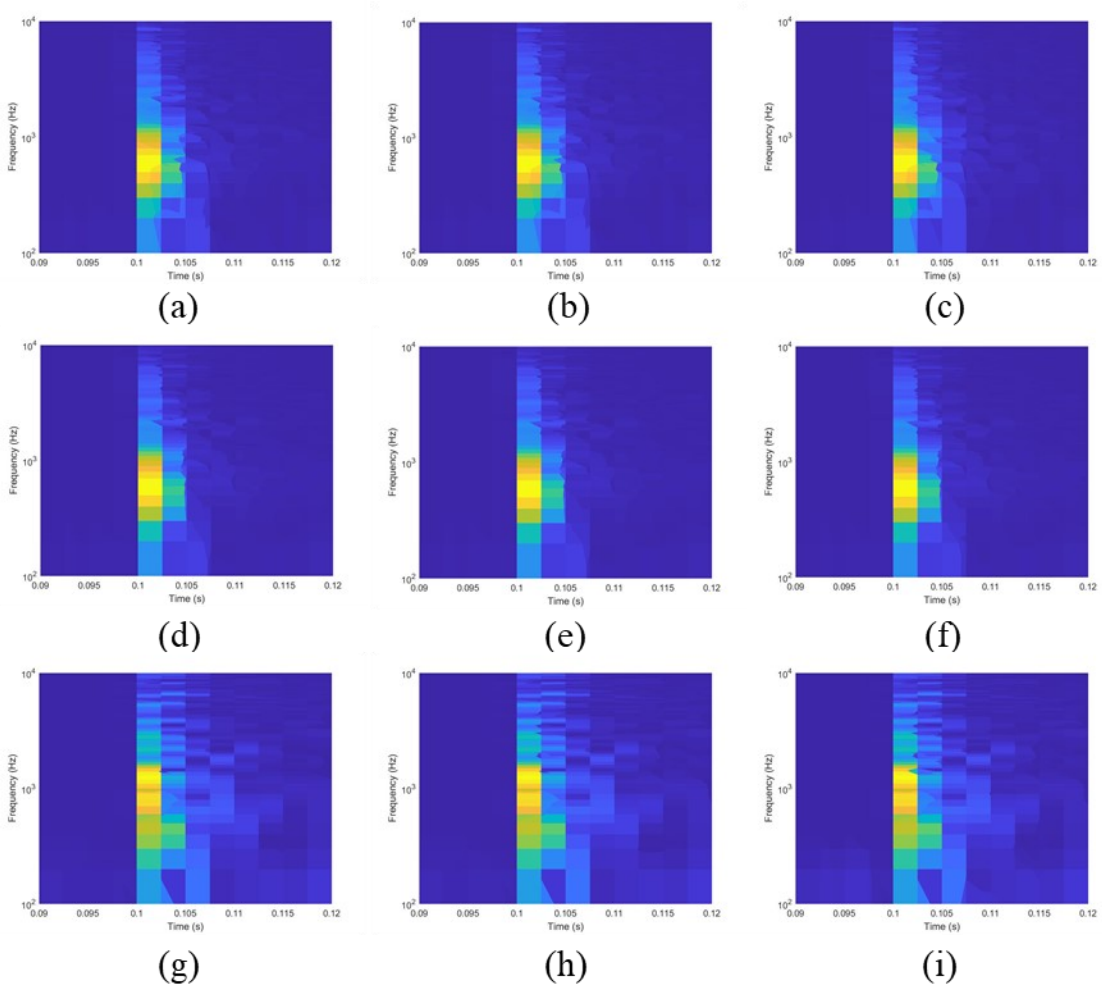


Figure 32: Gun #2 STFT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

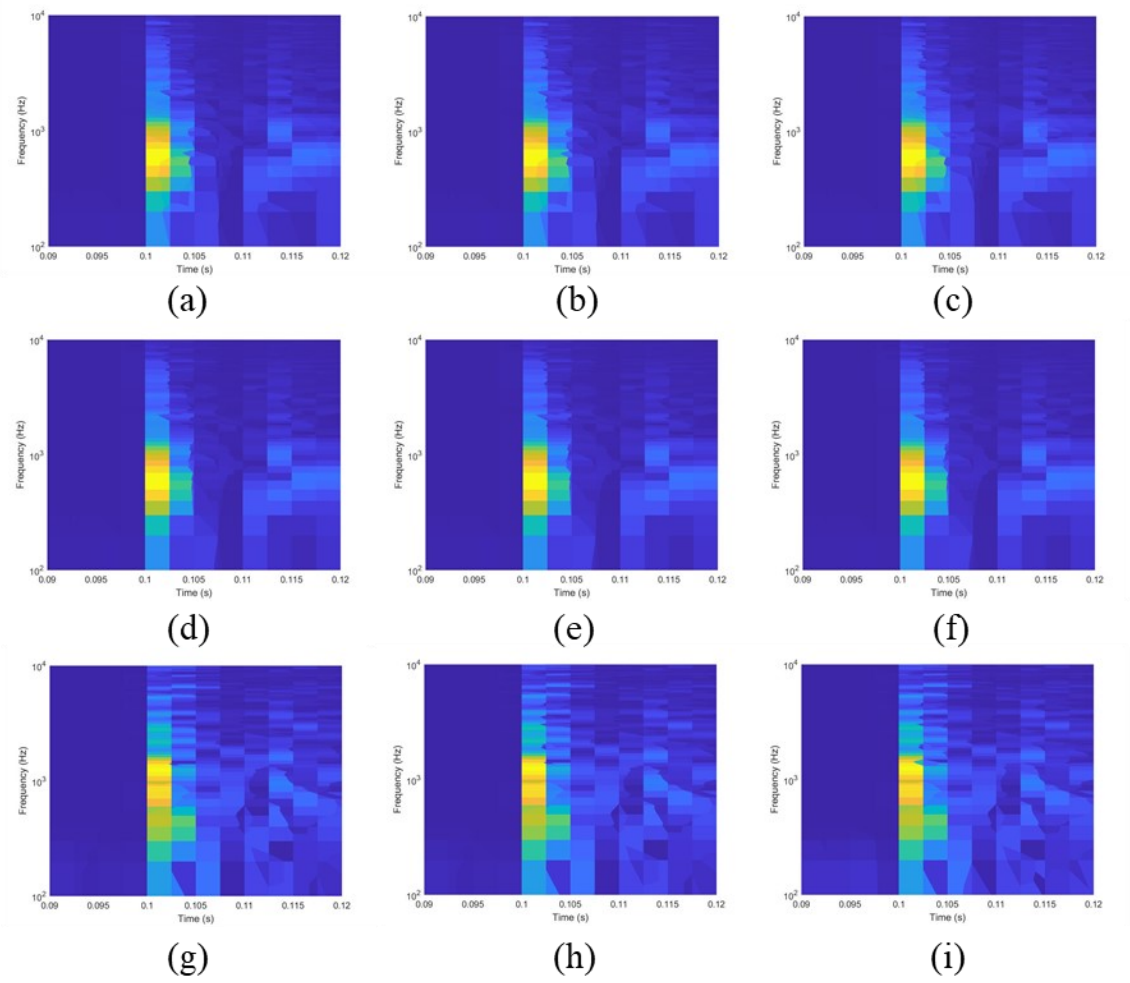


Figure 33: Gun #2 STFT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

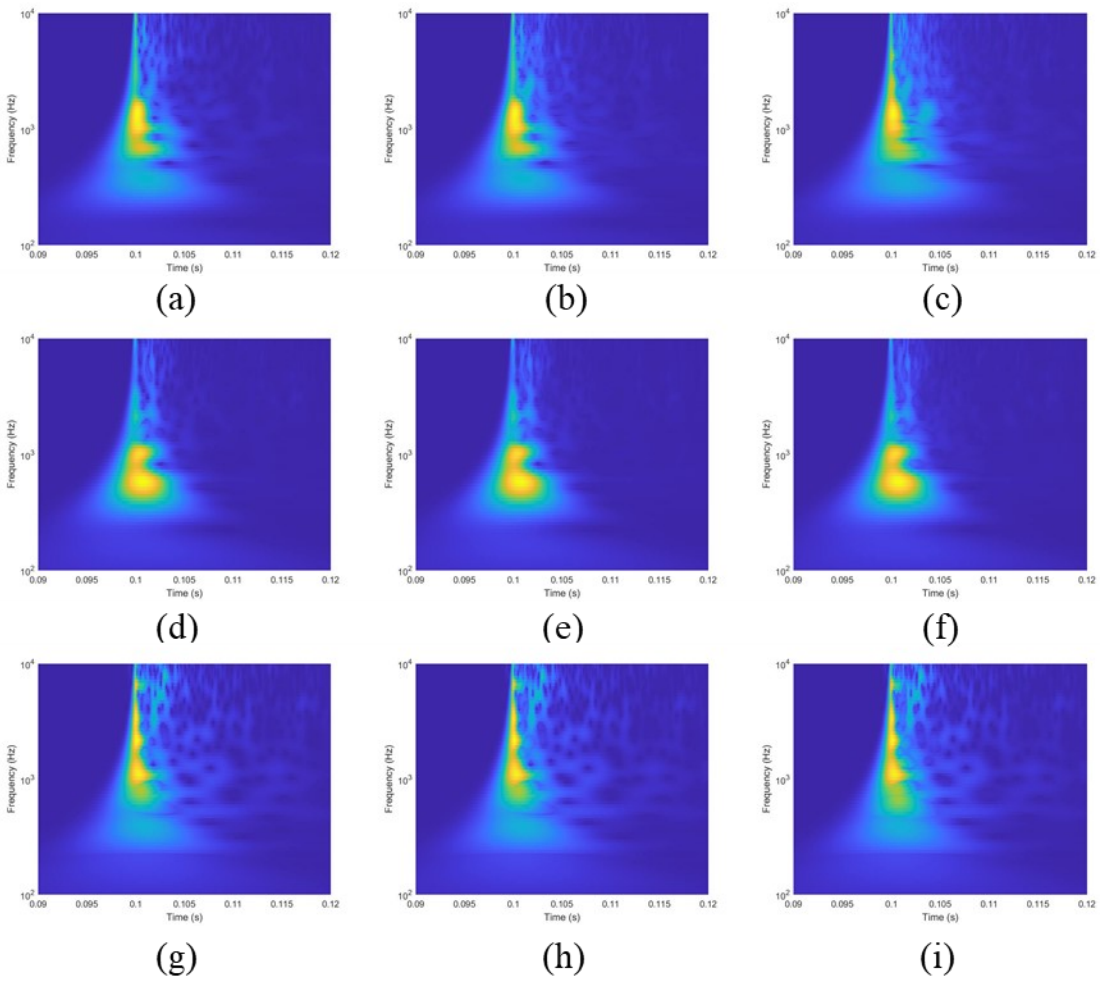


Figure 34: Gun #3 CWT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

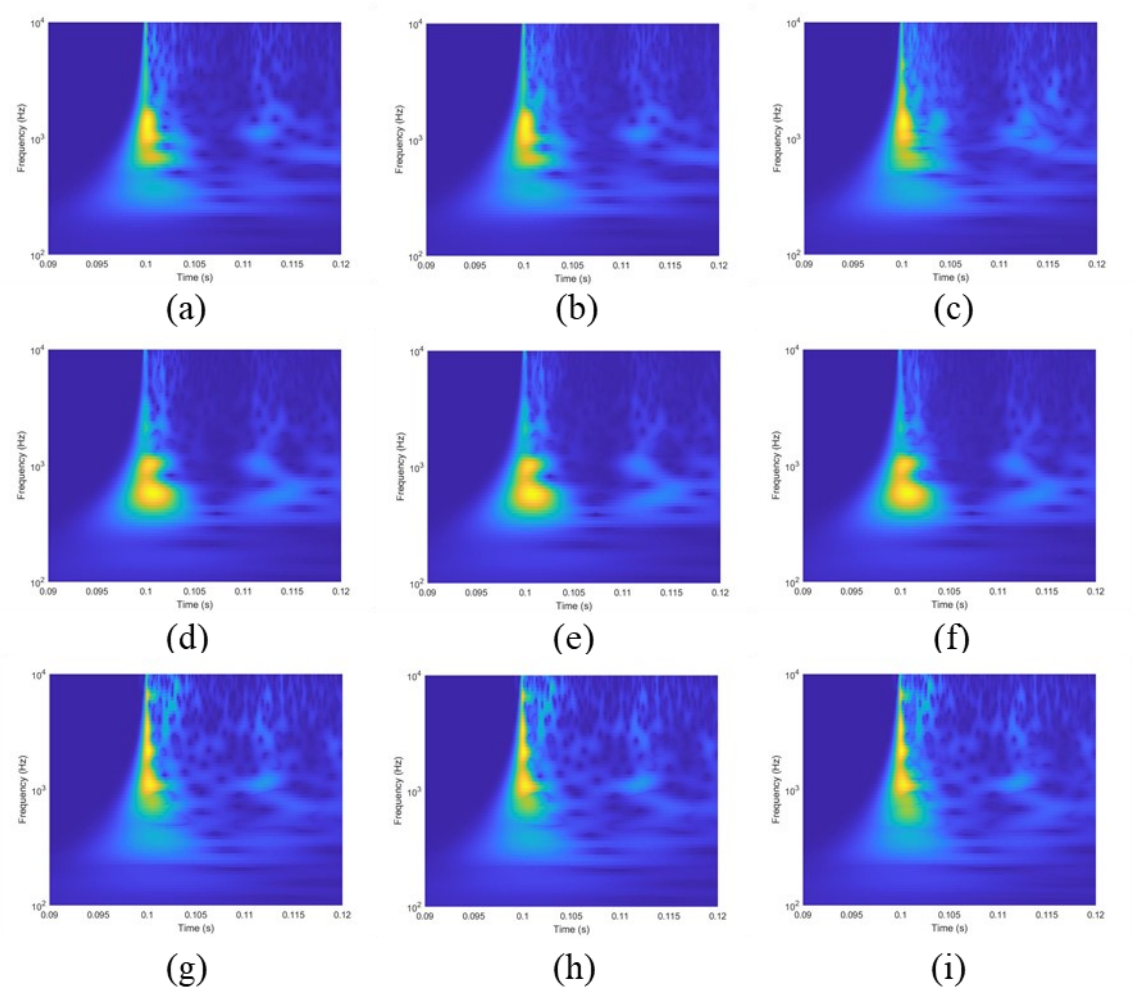


Figure 35: Gun #3 CWT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

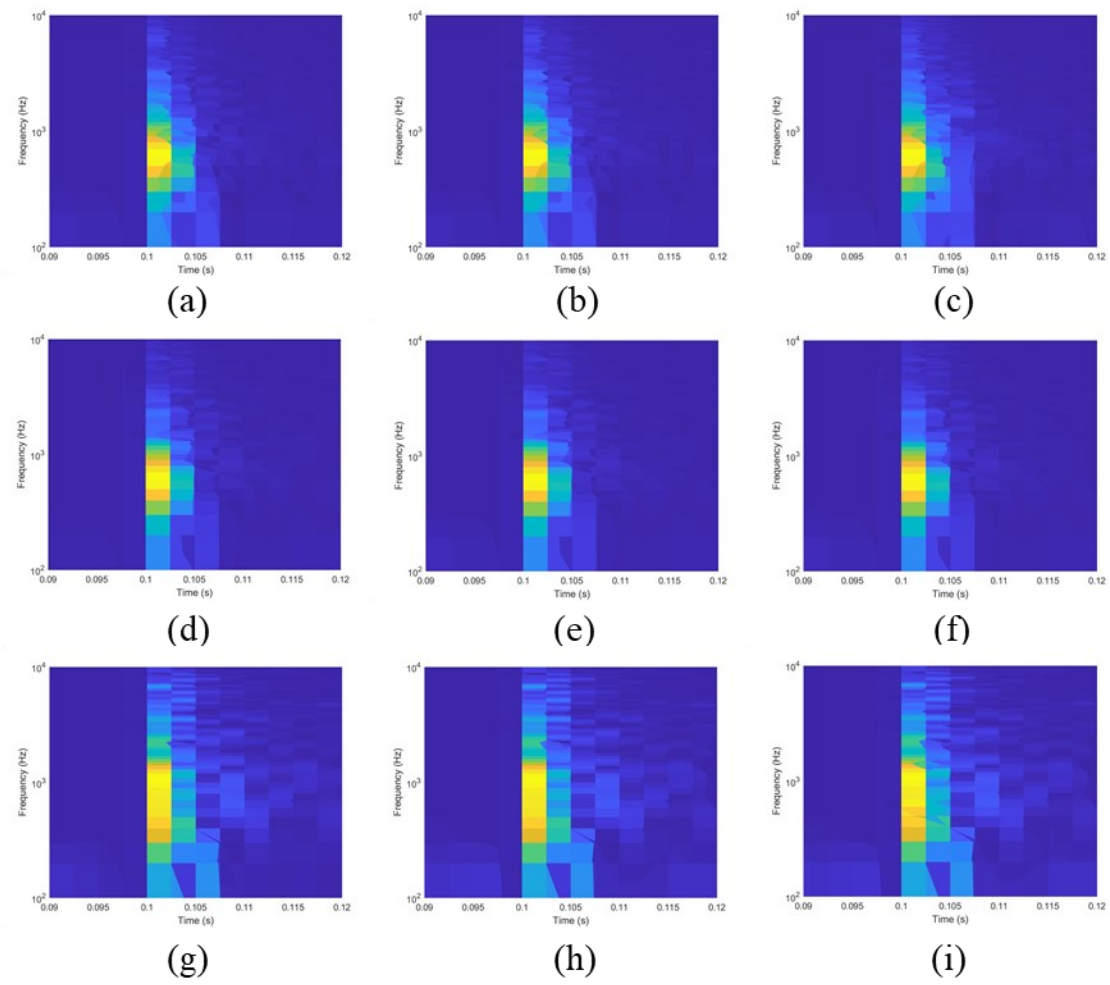


Figure 36: Gun #3 STFT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

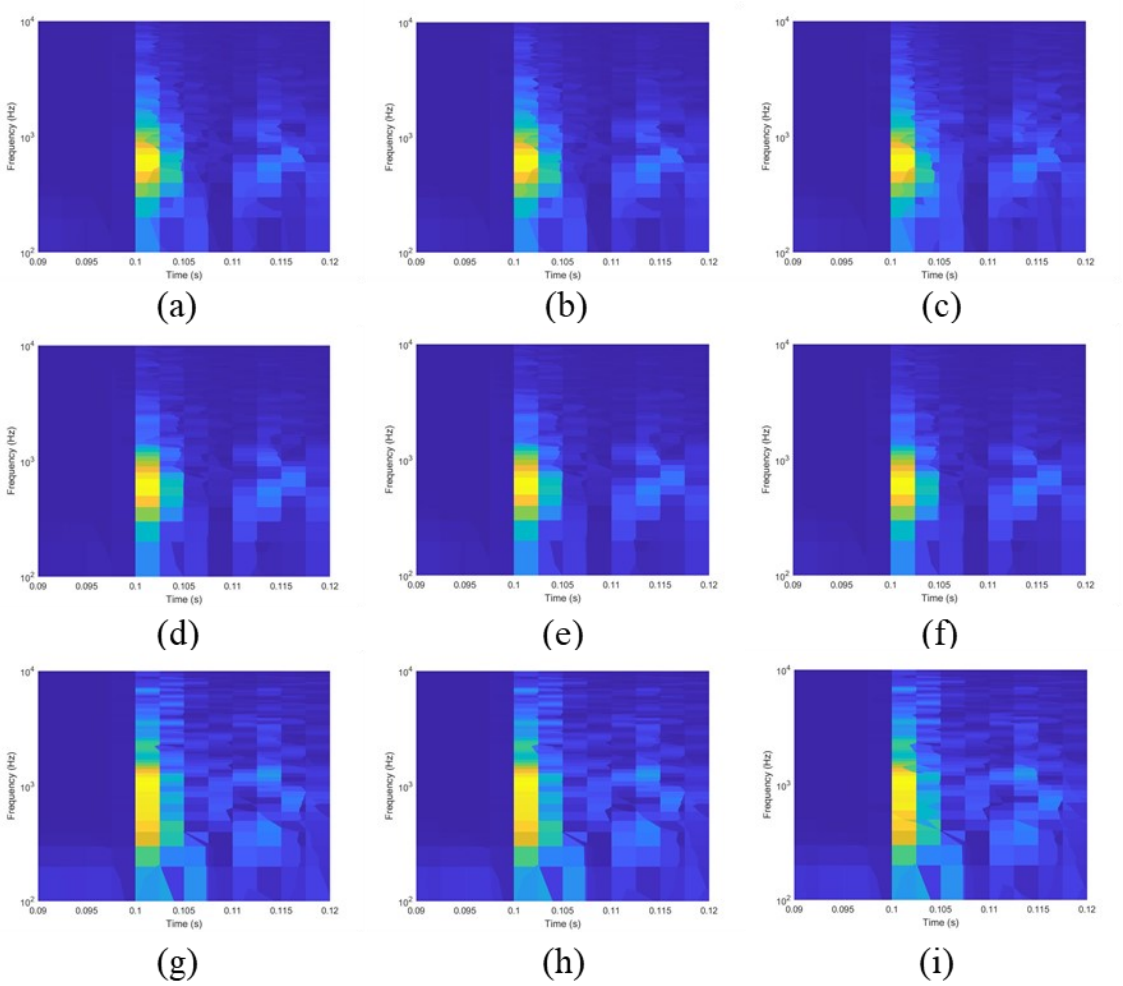


Figure 37: Gun #3 STFT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

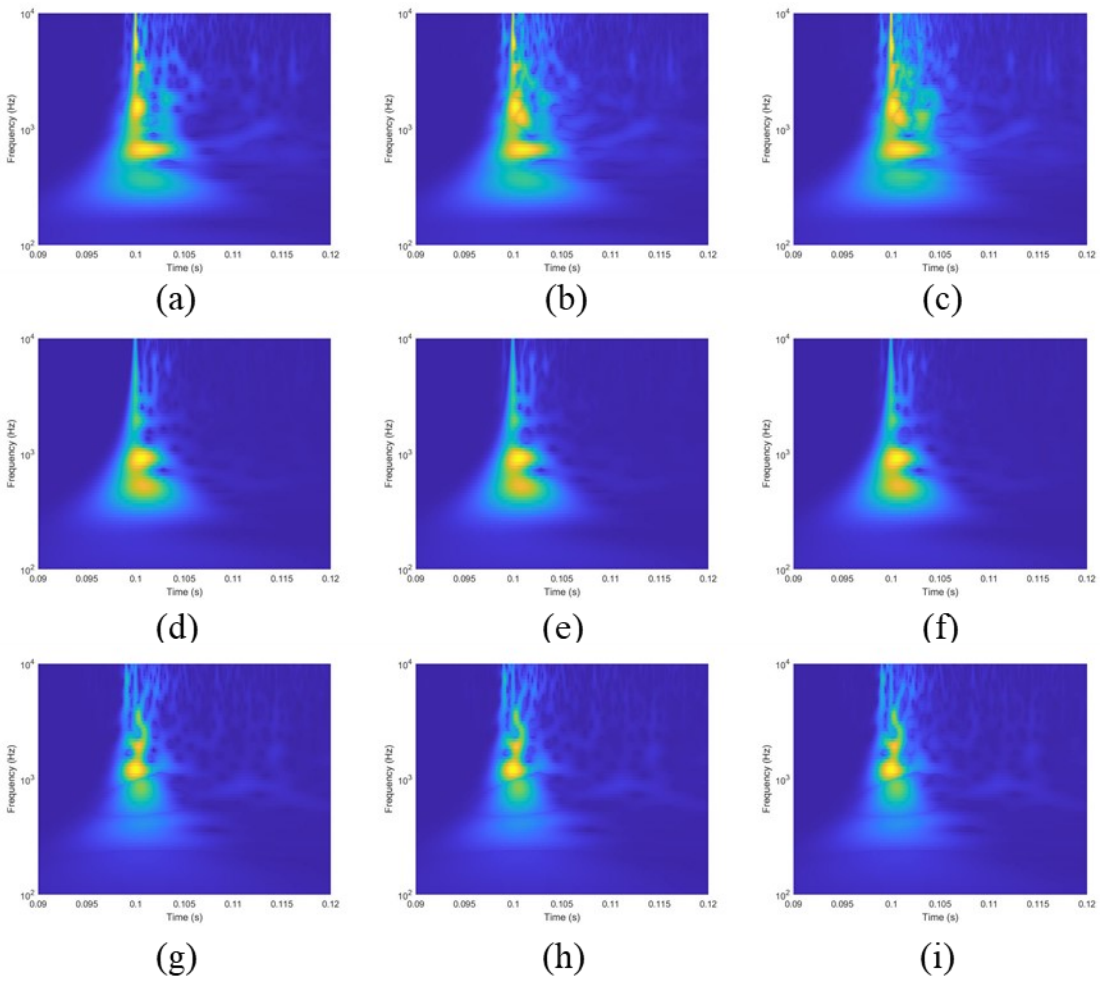


Figure 38: Gun #4 CWT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

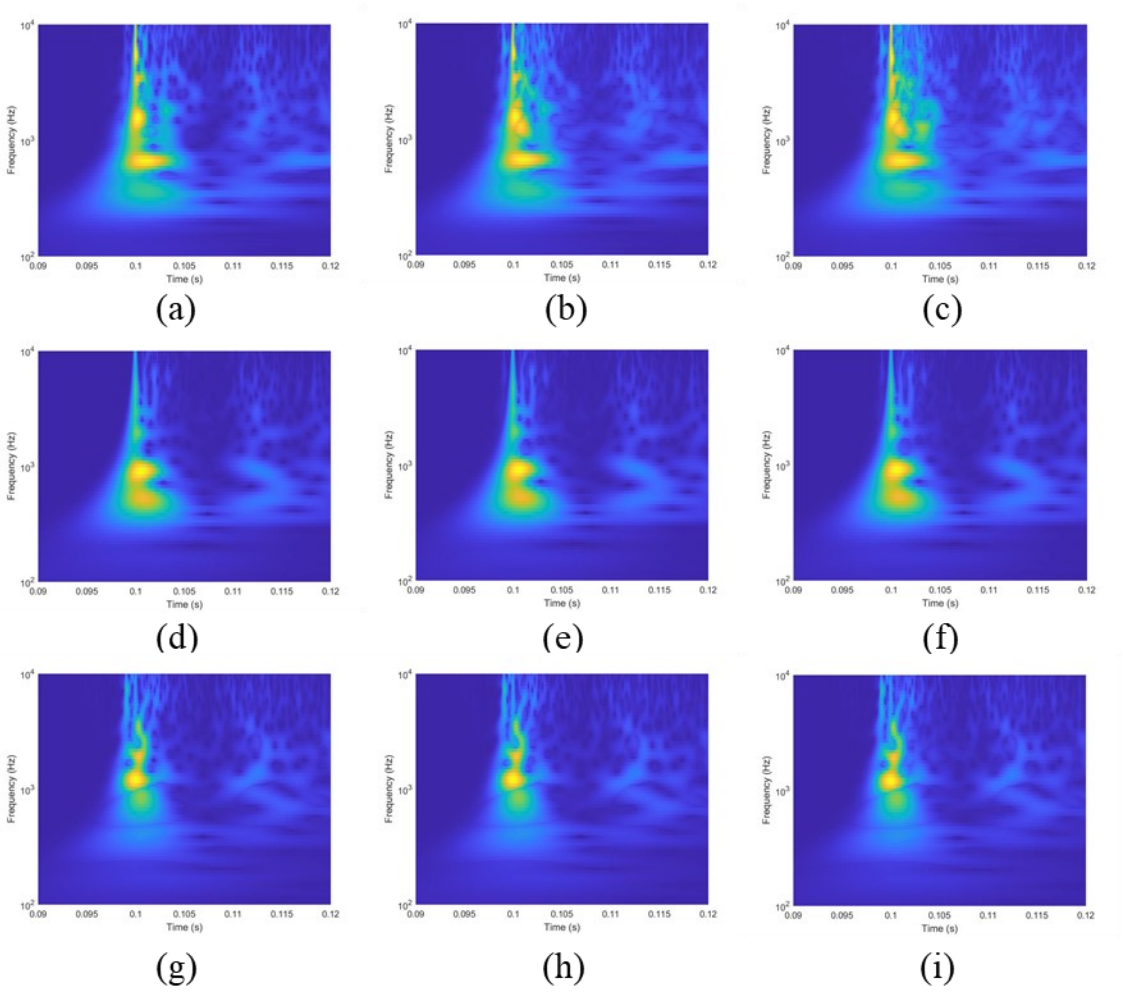


Figure 39: Gun #4 CWT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

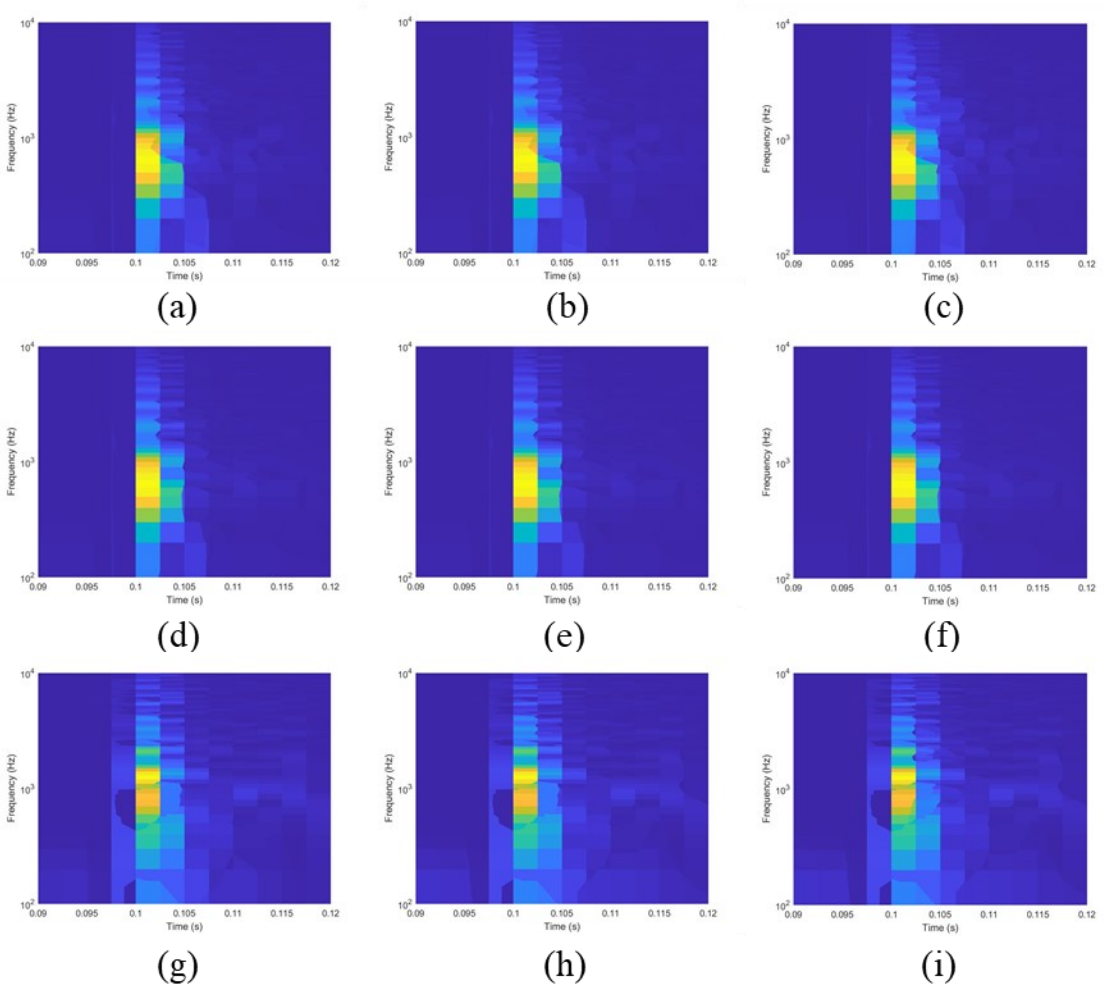


Figure 40: Gun #4 STFT spectrogram, no reverberation, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

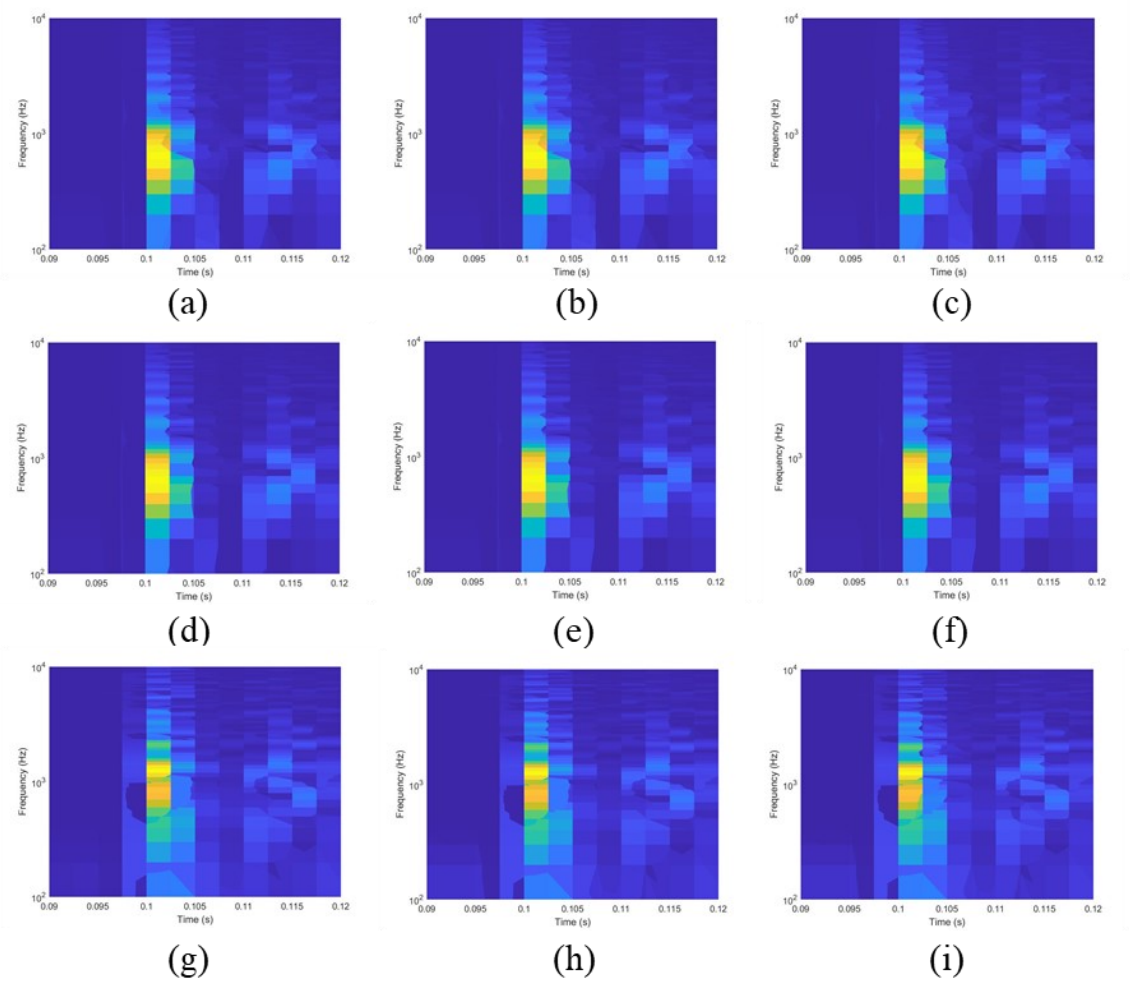


Figure 41: Gun #4 STFT spectrogram, reverberation ratio of 0.5, signal from a) 90° 2m, b) 180° 2m, c) 270° 2m, d) 90° 4m, e) 180° 4m, f) 270° 4m, g) 90° 8m, h) 180° 8m and i) 270°, 8m

4.2.6 STFT Parameters

Many parameters can shape how a spectrogram is produced by the STFT. The most common are the window sample size, the overlap between two adjacent windows, and the number of elements over which the STFT is calculated. Among them, the window sample size determines the resolution in time and frequency and, consequently, a trade-off must be found. To get high precision in time and frequency at the same time is not possible due to the Heisenberg uncertainty principle. In this study, we used a window size of 1024 with

512 overlapped samples and 2048 frequency points to optimize the classification accuracy.

While each test showed a very good accuracy with no reverberation, Case 1 in Table 10 did best when the reverberation was added. This table shows the tests done in the attempt to find the best fit in accuracy and robustness when simulated reverberation was added to the acoustic signal.

The accuracy can get very low when the window size is wider and the reverberation gets higher. This can be a consequence of having reverberation signal that are computed with the signal coming from the direct sound impulse. If we look at Figure 42 c), we see that the reverberation signal appears around the x-axis value of 0.11s. Reported in Figure 42 d), it is right at the junction of the calculated adjacent windows. Consequently, the reverberation signal may affect directly the primary signal from the direct sound impulse. Which is not the case with smaller windows.

Table 10: STFT Parameters Optimization

	No reverberation		**Reverberation amplitude ratio at 0.5**		**STFT parameters**		
Case	**Accuracy (%)**	**σ (%)**	**Accuracy (%)**	**σ (%)**	**Window**	**Overlap**	**NFFT**
1	99.3	0.6	88.7	7.1	1024	512	2048
2	99.2	1.0	83.3	7	1024	512	4096
3	98.9	0.9	37.4	4.2	2048	512	2048
4	99.4	0.9	75.5	10.6	512	256	2048
5	98.2	1.0	78.7	5.6	512	256	4096

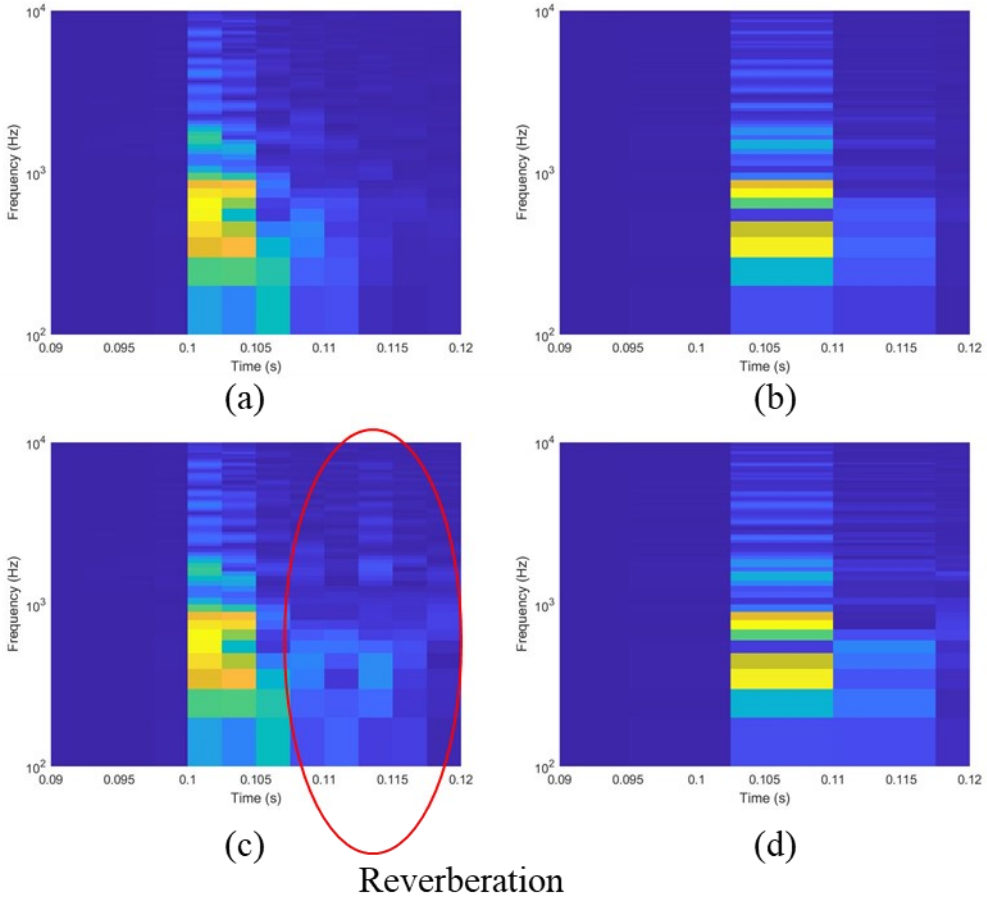


Figure 42: Table 10 STFT spectrogram for Gun #1, signal from 90° 2m for a) Case 1, no reverberation, b) Case 3, no reverberation, c) Case 1, reverberation ratio of 0.5 and d) Case 3, reverberation ratio of 0.5

From these results, we can trace a comparison between the spectrum analysis coming from the best scenario for the STFT and the CWT. Table 11 and Figure 43 show the results. The process using the STFT showed a better classification accuracy with no reverberation. The accuracy obtained from the STFT did not decrease as much then that obtained from CWT.

Table 11 Reverberation effect on classification accuracy

Amplitude ratio	STFT accuracy	CWT accuracy
0.0	99.0 %	94.7 %
0.3	98.8 %	91.5 %
0.4	91.8 %	72.2 %
0.5	88.7 %	54.2 %
0.6	78.3 %	32.5 %

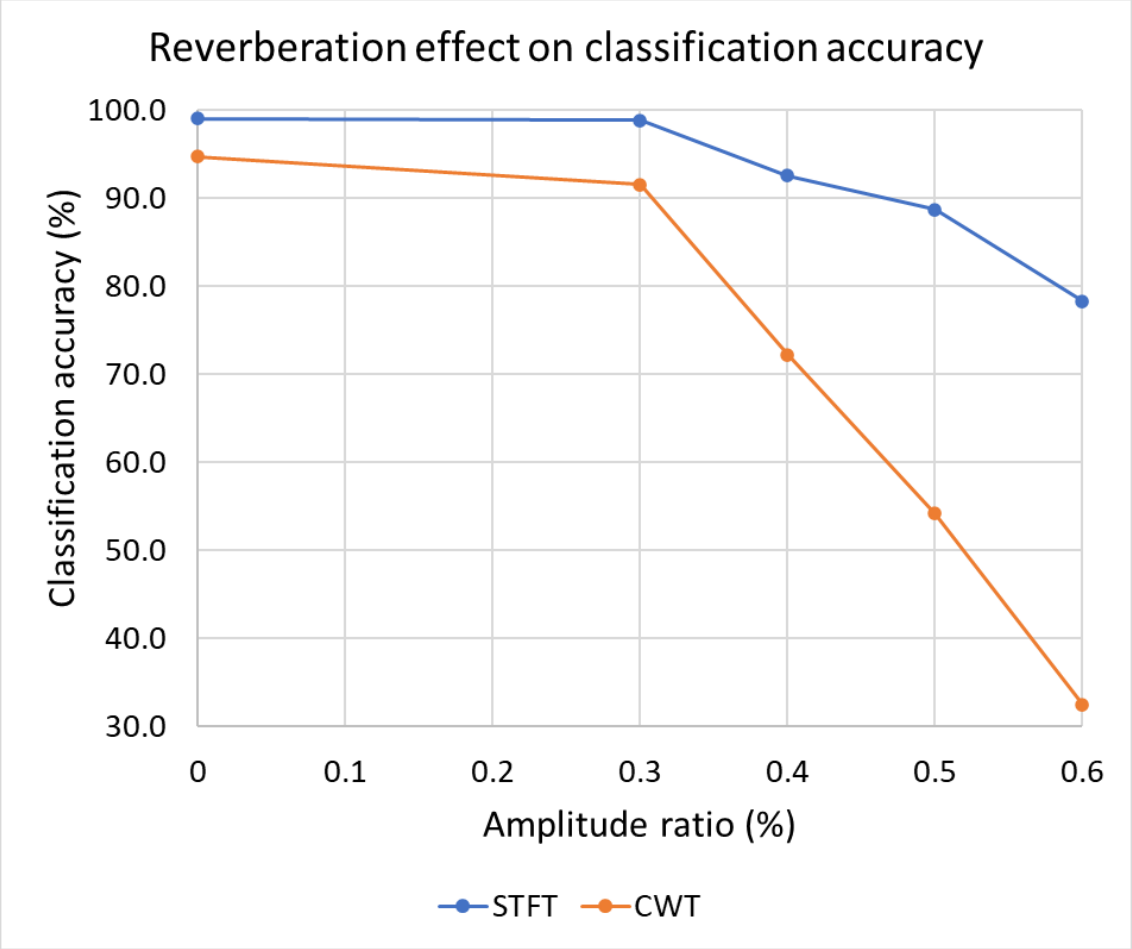


Figure 43: Effect of the reverberation ratio (%) on the classification accuracy

4.2.7 Discussion

Public authorities use acoustic gunshots identification for threat evaluation. We used a hybrid approach of combining the use of modern mathematical methods with machine-learning image-recognition tools as a classifier. We compared two spectral analysis methods when submitted to reverberation conditions.

Both STFT and CWT bring their own uncertainties over time and frequency domains,. A sampling rate of 204.8 kHz was used to record the audio signals. We chose the highest possible sampling rate allowed in our data acquisition system to get the maximum accuracy over the waveforms of the impulse noises. For the same reason, we used the window size around the peak of the impulse noises and the range displayed in the image when performing the spectral analysis. The window peak, however, did not need to be too large; it consumed too much computing time without improving the precision and resolution around the peaks. The range needed to be zoomed enough to show the dissimilarities between the guns. With a longer window area after the pulse, the spectral analysis captured the acoustic reverberations from the gun. The same is also true in the frequency domain. For the purpose of the analysis, it was unnecessary to keep any data with a frequency below 100 Hz and over 10 kHz; for the purpose of this study, no useful information was shown outside these frequencies.

For the machine learning process, the trials with the k-nearest neighbors showed that a value greater than 1 for k degraded the accuracy. Using random subspaces permitted an increase in accuracy of around 3%. Consequently, using the nearest neighbor with a random-subspaces algorithm gave us an almost perfect classifier. To avoid too much

fluctuation in the accuracy of the results, we repeated and averaged over 50 times the process of learning and classifying.

The STFT was more accurate compared to the CWT without reverberation and more robust when submitted to reverberation, as Table 11 and Figure 43 illustrate. At an amplitude ratio over 0.3, the accuracy for identifying one of the four guns did not deteriorate as much when compared to using the CWT. By observing Figure 44 and Figure 45, reverberation affects mostly higher frequency-time area of the spectrogram. The selected parameters leading to a coarser spectral-time resolution of the STFT results in a less sensitive response to the subtle changes produced by the reverberation signal.

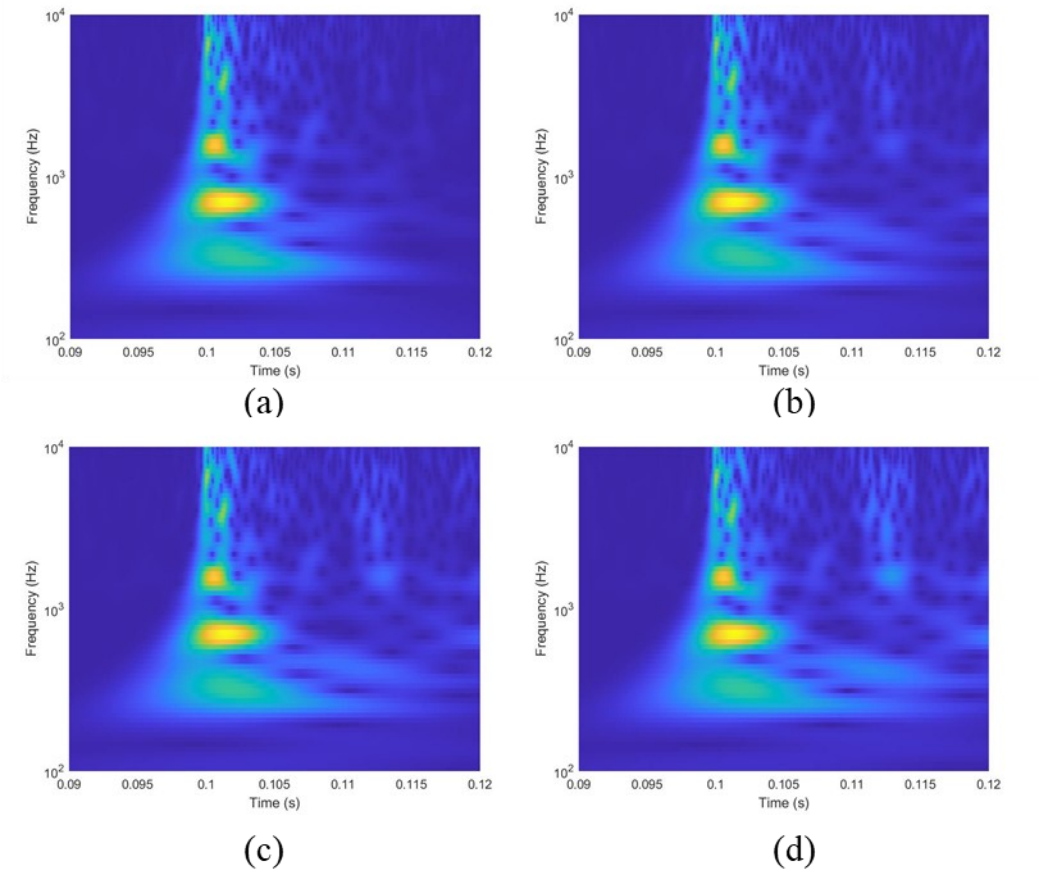


Figure 44: Table 11 CWT spectrogram for Gun #1, signal from 135° 2m with a) no reverberation, b) reverberation ratio 04, c) reverberation ratio 0.5 and d) reverberation ratio 0.6

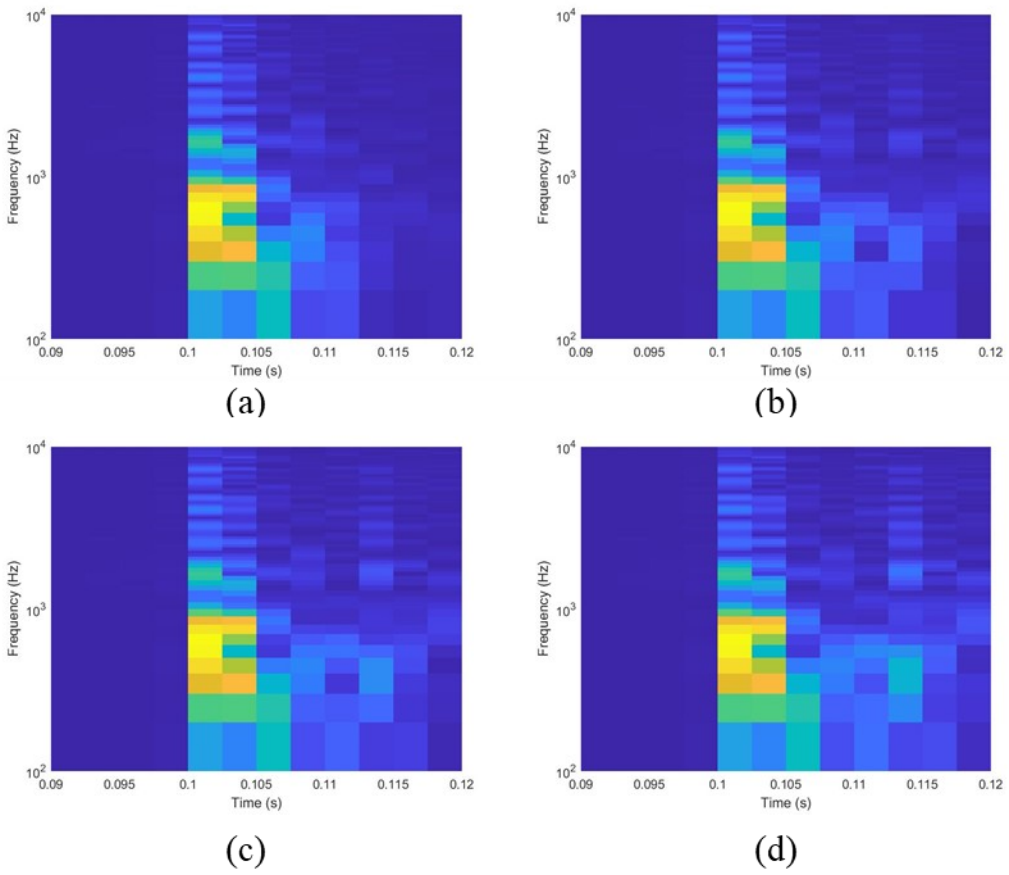


Figure 45: Table 11 STFT spectrogram for Gun #1, signal from 135° 2m with a) no reverberation, b) reverberation ratio 04, c) reverberation ratio 0.5 and d) reverberation ratio 0.6

4.2.8 Conclusion

To construct a gun identifier, we used two different spectral analysis methods combined with a machine-learning based identifier. We investigated the impact on gun identification accuracy by introducing controlled amounts of reverberation into the audio recordings of gunshots. We also collected audio data from four similar guns. To perform gun identification, audio recordings of gunshots were first converted to spectral images using either STFT or CWT.

A machine-learning image processing method then determined the level of identification accuracy. The results that the STFT provided a better level of identification accuracy than the CWT, and remained even when reverberation was introduced into the audio signals. In addition, the STFT maintained its better identification accuracy as the overall accuracy deteriorated with an increasing level of reverberations. STFT and CWT are just two examples of spectral analysis methods; further studies are needed to investigate the use of other spectral analysis methods and the ways they perform in a reverberant environment.

Chapter 5: Impact on Hearing

5.1 Introduction

While regulatory agencies commonly measure steady-state noises with the equivalent energy model to quantify the maximum daily exposure, other methods have been proposed for impulsive noises such as gunshots. (We discussed these in section **2.2**.) The newest military standard MIL-STD-1474E [17] bases the daily maximum allowable gunshots on two different metrics, yet without specifying which one to use. The first is $\boldsymbol{L_{IAeq100ms}}$, an equal-energy model calculated over a 100-ms interval. The second is the Auditory Risk Unit (ARU), calculated with the online available Auditory Hazard Assessment Algorithm for Humans (AHAAH) model [18]. In a study by Murphy et al. [29], $\boldsymbol{L_{Aeq,8hr}}$ presented itself as the best model to predict impairment over the AHAAH model.

Out of the various technical choices arises the question of properly evaluating an attenuation. Properly choosing a hearing protection device (HPD) or a suppressor to be installed on a rifle may become difficult. Two studies [47] [59] have proposed a method to evaluate the allowable number of exposures (ANE) during a day with the following equation [59]:

$$\boldsymbol{ANE = 10^{((EEL - L_{Aeq,8hr} - IPIL)/10)}} \quad \textbf{(17)}$$

Where $\boldsymbol{EEL}$ is the Equal Energy Limit which is 85 or 87 dBA in many jurisdictions and $\boldsymbol{IPIL}$ is the impulse peak insertion loss. An energy metric, the $\boldsymbol{L_{Aeq,8hr}}$, is subtracted with a

peak attenuation value, the IPIL.

This chapter examines the replacement of the attenuation factor, IPIL, with something more compatible with the energy terms EEL and $\boldsymbol{L_{Aeq,8hr}}$. IPIL represents a single number unique to the entire audio frequency range. Studies show that the organ of hearing is more sensitive to certain frequencies resulting in potential hearing damage varying with the audio spectrum [60] [61] [62]. Hearing protection devices (HPD) also give attenuations that also are frequency-dependent [39]. Consequently, we need a way to replace the model in equation **(17)** that uses IPIL for a metric with one that takes into account the energy and the frequency dependency.

In 2017, Fackler et al. [39] proposed a model to meet these two characteristics with the Impulsive Spectral Insertion Loss (ISIL). ISIL evaluates the energy attenuation in 1/3 octave frequency bands with a filter bank. However, the filter bank derives from the Fast Fourier Transform (FFT) and its time-frequency resolution cannot be traded-off and adjusted based on needs [20]. Consequently, the precision in time and frequency may not be optimized through the full acoustic spectrum. This chapter presents a method to overcome this issue using orthogonal wavelets as filters which have the property of conserving energy [63] and optimizing the resolution by its adjustable time-frequency resolution capability.

The metrics are applied to real data coming from the use of a suppressor installed on the C20 semi-automatic sniper rifle.

5.2 Wavelets Octave Band Insertion Loss

Some medical applications employ the Continuous Wavelet Transform (CWT). This method de-noises ultrasound images [64], estimates with precision the oxygen level [65], and analyzes spectral data for posturographic signal processing [66]. The Continuous Wavelet Transform (CWT), is defined with the following equation [21]:

$$\boldsymbol{W_{\psi}(t,s) = \int_{-\infty}^{\infty} \frac{1}{s^n} \psi^* \left(\frac{\tau - t}{s}\right) x(\tau) d\tau} \quad \textbf{(18)}$$

Where $\boldsymbol{\psi^*(\tau)}$ is the complex conjugate of the wavelet, $\boldsymbol{s}$ is the scaling parameter and $\boldsymbol{t}$, the translation.

WOBIL uses the same process as ISIL and OBIPIL, shown in Figure 7. The filter bank discriminating the various 1/3 octave bands used in this study is built with the orthogonal Generalized Morse Wavelet (GMW) superfamily for which the Fourier transform is defined by [21]:

$$\boldsymbol{\Psi_{\beta,\gamma}(\omega) = \int_{-\infty}^{\infty} \psi_{\beta,\gamma}(t) e^{-i\omega t} dt = U(\omega) a_{\beta,\gamma} \omega^{\beta} e^{-\omega\gamma}} \quad \textbf{(19)}$$

Where $\boldsymbol{a_{\beta,\gamma}}$ is a normalization constant, $\boldsymbol{U(\omega)}$ is the unit step function and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are two parameters to control the shape of the wavelet. The value is called the time-bandwidth product. As mentioned in the previous chapters, Morse wavelets give the best function available when a good tradeoff between time and frequency precision is desirable [22]. The values chosen for this paper were $\boldsymbol{\gamma}$ = 3 and $\boldsymbol{P_2}$ = 60, providing symmetric wavelets which, in turn, resulted in a neutral time-frequency precision tradeoff [23].

An orthonormal wavelet, by its finite energy nature, preserves the energy of a signal during its analysis [63]. It also provides a better time-frequency analysis over the full acoustic bandwidth due to its ability to adjust its time-frequency resolution analysis

capability compared to the Short-Time Fourier Transform (STFT) [67]. The ISIL equation (17) calculates the energy in each band with the same time window.

In its final application for evaluating the ANE during a day, WOBIL is inserted in this equation (15):

$$\boldsymbol{ANE = 10^{((EEL - L_{Aeq,8hr} - WOBIL)/10)}} \quad \textbf{(20)}$$

5.3 Rifle and Acoustic Impulse Measurements

The rifle we used was a semi-automatic sniper weapon, the C20 with 7.62 X 51 mm ammunition, shown in Figure 46. The first series of gunshots was done without a suppressor, followed by one with the suppressor HX-QD-762 from OSS.

We used a G.R.A.S. 1/8" pressure microphone model 46DD-FV for the tests, which can sustain up to 175 dB, reference 20μPa, for the sound pressure level. It has a bandwidth of 10 Hz to 25 kHz with ± 1 dB flatness and 6.5 Hz to 140 kHz with ± 3 dB flatness. We installed it at 1.6 m above ground, 4 m away on the right side of the shooter, aligned with the hammer of the rifle, as Figure 46 illustrates.

A Siemens LMS SCADAS SCR202, 16 channels, was used for the tests. This unit synchronously records up to 16 channels of data with 204.8 kHz sampling rate and a resolution of 24 bits.

a) Rifle C20

b) Suppressor installed on the muzzle

C) Microphone position

Figure 46: Equipment and set-up a) C20 rifle used for the test b) Suppressor on the muzzle
c) microphone position for recording

5.4 Results and Discussion

Table 12 and Figure 47 show the result from the average of 10 firings with the suppressor and 10 firings without it. It confirms that the curve for WOBIL differs from ISIL and OBIPIL. The same time window of 30 ms used for ISIL and OBIPIL has been applied.

IPIL gives a single value, independent of the frequencies. It is purely an indication of the peak sound level attenuation observed, without any consideration for the frequency content. In this study, the value was an attenuation of 21.7 dB.

By contrast, ISIL is a curve varying with frequencies, thus giving a more detailed analysis of the observed attenuation over the frequencies. The attenuation is energy-based with the calculus of $\boldsymbol{L_{eq}}$ in each band. However, the filtering in each 1/3 octave band is based on the FFT. As presented in section 2.1, it doesn't have the capability of offering an adjustable analysis and the precision in time and frequency may not be optimized through the full acoustic spectrum.

OBIPIL, like IPIL, is based on the peak-sound pressure level. But unlike it, OBIPIL is calculated in each 1/3 octave band. It requires a proper filter that maintains the peak level in each band. However, the energy itself may not be preserved and the incongruity of equation (15) remains because energy variables such as ***LAeq,8hr*** and EEL are not the same as the peak attenuation measured with OBIPIL.

The introduction of energy attenuation with wavelets, WOBIL, seeks to overcome the incongruities and deficiencies of IPIL, ISIL and OBIPIL, taking into account the frequency dependence of the attenuation coming from suppressors and HPDs. But it also gives a more accurate version by preserving the energy in each 1/3 octave band and by using an adjustable analysis. That way, if an attenuation device observes an energy shift, the model will still give an accurate measurement. The analysis for evaluating the hearing damage risk will be more in line with the real outcomes thus avoiding overestimation or underestimation.

Figure 47 shows that ISIL and OBIPIL gave similar results for up to 2.5 kHz when we analyze the attenuation given by the suppressor. By contrast, WOBIL gave more attenuation below 500 Hz and less above 1 kHz.

Table 12 Average Attenuation Results by ⅓ Octave Bands

1/3 Octave bands (Hz)	OBIPIL (dB)	WOBIL (dB)	ISIL (dB)
25	8.4	9.8	8.2
32	8.9	10.4	8.8
40	9.6	11.9	9.7
50	10.3	12.9	10.2
63	11.1	13.7	10.9
79	11.8	14.8	11.9
100	12.5	16.5	12.3
126	13.2	17.8	13.3
158	14.6	18.9	14.5
200	16.4	19.1	15.9
251	16.9	20.6	17.1
316	18.0	21.0	17.8
398	19.4	21.9	19.4
501	21.5	21.3	21.9
631	22.9	22.5	23.2
794	20.9	23.3	22.2
1000	21.8	21.1	21.4
1259	21.3	19.1	21.7
1585	20.4	17.7	21.2
1995	20.7	17.0	20.2
2512	19.4	15.6	19.3
3162	21.9	14.0	18.6
3981	19.5	13.2	17.8
5012	18.2	12.3	16.4
6310	16.5	11.7	14.6
7943	17.6	11.0	13.6
10000	17.1	10.2	13.0
12589	15.8	9.5	11.7
15849	16.6	8.8	10.3
19953	16.8	8.1	9.8

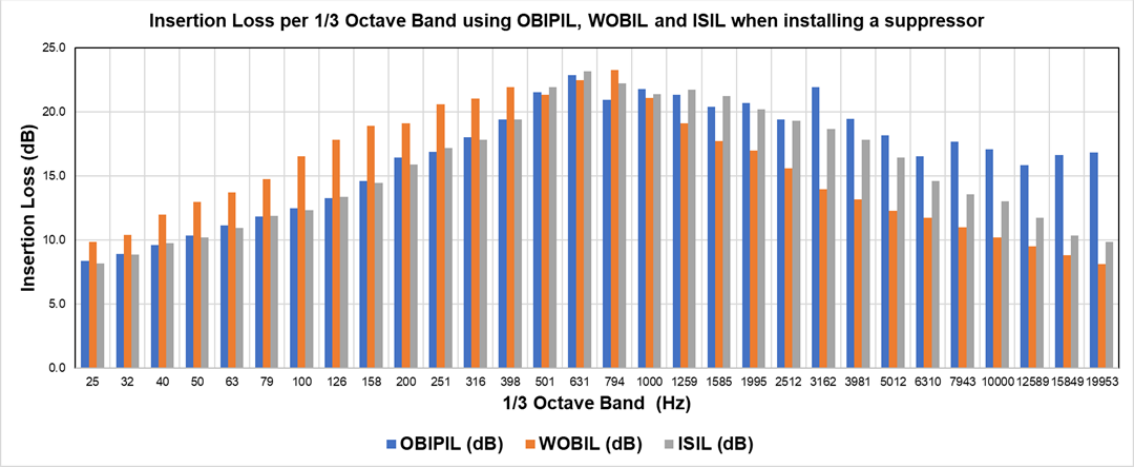


Figure 47: Attenuation results by 1/3 octave bands

5.5 Conclusion

In this chapter, we have illustrated an innovative method to quantify attenuation of sound pressure levels based on wavelet analysis to preserve the 1/3 octave band energy content of a gunshot acoustic signal. We applied our method to gunshot signals recorded from a sniper rifle and evaluated the attenuation when installing a suppressor. When compared to previous methods, the wavelet analysis showed an attenuation curve different from those other methods, in giving higher attenuation at low frequencies and lower attenuation at higher frequencies. Further studies are needed to better understand the effect of this new method and to determine whether the precision of hearing impairment can be more accurately predicted.

Chapter 6: Conclusion

This study has developed a gunshot sound analysis to reach two goals: a) gun identification and b) impact on hearing. Gun identification is important in detecting and mitigating the risks posed to the public. Better understanding and the estimation of the impact of gunshot sound on hearing is needed to improve how we currently protect our hearing health from impulse noises.

We collected acoustic data from four different guns and a mortar. To capture their sound, including any non-symmetric sound propagation, we placed 27 high-dynamic range-pressure microphones around the guns and mortar, forming a polar grid pattern. Audio signals were captured at 204.8 kHz sampling rate synchronously to preserve the fidelity of the impulse nature of the gunshots. We developed an image-based analysis method to take advantage of the recent advancement of image recognition techniques. This study applied two spectral analysis methods—Short Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT)—to get the spectrogram of the gunshot audio signal. To classify these spectrograms and identify which gun the particular gunshot originated from, we employed machine learning using the k-nearest neighbor and random subspaces. And to test the robustness of the gunshot identification, we added simulated reverberation to the recorded signals.

Impulsive sounds can cause severe hearing damage and even hearing loss. Sound-protection devices are widely used to attenuate impulsive sounds and reduce their impact

on hearing. Properly measuring and characterizing the sound attenuation is essential when choosing a specific hearing-protection device. Such measurement can help in enacting safety regulations. To increase the precision in evaluating insertion loss, this study has proposed the use of a novel energy preserving method for estimating the 1/3 octave band insertion loss using the continuous wavelet transform. This precision is important for the evaluation of suppressors, for example, and to evaluate the effectiveness of a hearing-protection device. Currently, such devices are often characterized using the Impulse Peak Insertion Loss (IPIL), which measures the total attenuation across all frequency bands. But IPIL does not provide any information about the spectral attenuation of the device.

Human hearing being spectrally sensitive, and the risk of noise-induced hearing damage being frequency-dependent, the characterization of hearing protection devices therefore has to be done for both the peak and the full audible frequency spectrum from 20 Hz to 20k Hz. The proposed energy-preserving method was applied to real recorded data. To do so, we collected gunshot audio sounds from firing a sniper rifle and evaluated the sound attenuation effect of adding a sound protection device (or sound suppressor) to the rifle. We have compared the method that we called Wavelets Octave Band Insertion Loss (WOBIL) to existing methods such as the IPIL, the Impulsive Spectral Insertion Loss (ISIL) and the recently published Octave Band Impulse Peak Insertion Loss (OBIPIL).

6.1 Future Research

Further studies should be done to better understand the effect of the proposed methods in this thesis, both in the gunshot identification process and the impact on hearing.

6.1.1 Research in Gunshot Identification

The following future research efforts have the potential of improving our knowledge and robustness of the gunshot identification:

- Record more guns, mortars and weapons gunshots to increase the audio database
- Test the learning process used in this study with more items than three or four classification categories
- Use wavelets other than the Morse wavelets or try the Gabor transform to compare with our results
- Use recordings containing noise or with multiple guns shooting in close range to simulate combat zone environment
- Use recordings from real data recorded during military operations and trainings.

6.1.2 Research in Impact on Hearing

The following efforts can improve our knowledge of the impact of energy preservation among octave bands when analyzing attenuation:

- Measure and analyze attenuation from hearing protection devices to quantify the octave band attenuation and to see if significant dispersion is observed among protection having similar rating

- Measure the real impact on hearing with people wearing hearing protection devices having dissimilar WOBIL

References

[1] B. Tardif, D. Lo and R. Goubran, "Gunshot Sound Measurement and Analysis," in *IEEE Sensors Applications Symposium (SAS)*, Sundsvall, Sweden, 2021.

[2] B. Tardif, D. Lo and R. Goubran, "Measurement and Characterization of Hearing Protection Devices in the Presence of Impulse Sound," in *2021 IEEE International Symposium on Medical Measurements and Applications*, Neuchâtel, Switzerland, 2021.

[3] E. Kiktova, M. Lojka, M. Pleva, J. Juhar and A. Cizmar, "Gun type recognition from gunshot audio recordings," in *3rd International Workshop on Biometrics and Forensics (IWBF 2015)*, Gjøvik, Norway, 2015.

[4] S. Ntalampiras, I. Potamitis and N. Fakotakis, "On acoustic surveillance of hazardous situations," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, 2009.

[5] A. Pikrakis, T. Giannakopoulos and S. Theodoridis, "Gunshot detection in audio streams from movies by means of dynamic programming and Bayesian networks," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, 2008.

[6] R. Maher and S. Shaw, "Directional Aspects of Forensic Gunshot Recordings," in *AES 39th International Conference*, Hillerød, Denmark, 2010.

[7] H. O. Durmuş, E. Çetin and B. Karaböce, "An Evaluation of Performance Tests of In-Ear Hearing Aids," in *2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Istanbul, Turkey, Turkey, 2019.

[8] E. A. Masterson, P. T. Bushnell, C. L. Themann and T. C. Morata, "Hearing Impairment Among Noise-Exposed Workers — United States, 2003–2012," Centers for Disease Control and Prevention, Atlanta, GA, USA, 2016.

[9] J. M. Thompson, L. VanTil, K. Feder, J. Sweet, M. Boswall, C. Courchesne, G. A. Banta, P. Lamontagne, L. Bogaert, K. McKinnon and A. Poirier, "Prevalence of hearing problems among Canadian Armed Forces Veterans: Life After Service Studies," *Journal of military, veteran and family health,* vol. 2, no. 2, pp. 62-72, 2016.

[10] A. Nakashima and R. Farinaccio, "Review of weapon noise measurement and damage risk criteria: considerations for auditory protection and performance," *Military medicine,* vol. 180, no. 4, pp. 402-408, 2015.

[11] C. Liguori, A. Ruggiero, D. Russo and P. Sommella, "Task-Based Measurements for Estimating the workers' exposure to acoustic noise," in *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Houston, Texas, 2018.

[12] S. D. Beck, H. Nakasone and K. W. Marr, "Variations in recorded acoustic gunshot waveforms generated by small firearms," *The Journal of the Acoustical Society of America,* vol. 129, no. 4, pp. 1748-1759, 2011.

[13] H. Brinkmann, "Techniques and procedures for the measuremente of impulse noise," NATO, Meppen, 2000.

[14] P. Giverts, S. Sofer, Y. Solewicz and B. Varer, "Firearms identification by the acoustic signals of their mechanisms," *Forensic Science International,* vol. 306, January 2020.

[15] M. Hrabina and M. Sigmund, "Gunshot recognition using low level features in the time domain," in *2018 28th International Conference Radioelektronika (RADIOELEKTRONIKA)*, Prague, 2018.

[16] P. Thumwarin, N. Wakayaphattaramanus, T. Matsuura and K. Yakoompai, "Audio forensics from gunshot for firearm identification," in *The 4th Joint International Conference on Information and Communication Technology, Electronic and Electrical Engineering (JICTEE)*, Chiang Rai, Thailand, 2014.

[17] MIL-STD 1474E, "Design Criteria Standard Noise Limits," US Department of Desense, 2015.

[18] Army Research Laboratory, "Auditory Hazard Assessment Algorithm for Humans (AHAAH)," U.S. Army Combat Capabilities development Command, [Online]. Available: https://www.arl.army.mil/who-we-are/directorates/hred/ahaah/. [Accessed 12 01 2020].

[19] N. A. Khan, M. N. Jafri and S. A. Qazi, "Improved resolution short time Fourier transform," in *2011 7th International Conference on Emerging Technologies*, 2011.

[20] Mathworks, "Time-Frequency Analysis and Continuous Wavelet Transform," Mathworks, 06 01 2021. [Online]. Available: https://www.mathworks.com/help/wavelet/ug/time-frequency-analysis-and-continuous-wavelet-transform.html. [Accessed 06 01 2021].

[21] J. M. Lilly and S. C. Olhede, "Generalized Morse Wavelets as a Superfamily of Analytic Wavelets," *IEEE Transactions on Signal Processing,* pp. 6036-6041, November 2012.

[22] L. Aguiar-Conraria and M. J. Soares, "The Continuous Wavelt Transform: moving beyond uni and bivariate analysis," *Journal of Economic Surveys,* vol. 28, no. 2, pp. 344-375, 2014.

[23] Mathworks, "Morse Wavelets," Mathworks, 2019. [Online]. Available: http://mathworks.com/help/wavelet/ug/morse-wavelts.html. [Accessed 10 October 2019].

[24] P. Sun and J. Qin, "Analysis of Impulse Noise Based on Wavelet Transform for Military Applications," in *2014 IEEE AUTOTEST*, St. Louis, Missouri, USA, September 2014.

[25] F. L. Wightman, G. Flamme, A. J. Campanella, G. A. Luz, T. Rosenthal and S. Gallo, "American Institute of Biological Sciences Peer review of Injury Prevention and Reduction Research Task Area Impulse Noise Injury Models," American Institute of Biological Sciences, Frederick, MD, 2010.

[26] F. Pfander, H. Bongartz, H. Brinkmann and H. Kietz, "Danger of auditory impairment from impulse noise: A comparative study of the CHABA damage-risk criteria and those of the Federal Republic of Germany," *Journal of the Acoustical Society of America,* pp. 67:628-633, 1980.

[27] G. Smoorenburg, "Damage risk criteria for impulse noise," in *New Perspectives on Noise-Induced Hearing Loss*, New-York, Raven Press, 1982, pp. 471-490.

[28] P. Chan, K. Ho, K. KAN, J. Stuhmiller and M. Mayorga, "Evaluation of impulse noise criteria using human volunterr data," *Journal of Acoustical Society of America,* pp. 110:1967-1975, 2001.

[29] W. J. Murphy and C. A. Kardous, "A case for using A-weighted equivalent energy as a damage risk criterion," NIOSH, 2012.

[30] Government of Canada, "Levels of Sound - Canada Occupational Health and Safety Regulations, Part VII - IPG-074," April 2010. [Online]. Available: https://www.canada.ca/en/employment-social-development/programs/laws-regulations/labour/interpretations-policies/074.html. [Accessed 29 07 2021].

[31] A. Nakasima, "A comparison of metrics for impulse noise exposure," Defence Research and Development Canada, Toronto, 2015.

[32] R. G. Price and J. T. Kalb, "The Philosophy, Theorical Bases, and Implementation of the AHAAH Model for Evaluation of Hazard from Exposure to Intense Sounds," US Army Research Laboratory, Aberdeen, 2018.

[33] R. Danielson, M. Henderson, L. Bianchi and R. Salvi, "The importance of "temporal pattern" in traumatic impulse noise exposures," *The Joournal of the Acoustical Society of america,* p. 209, 1991.

[34] W. Tielemans, R. Angus, G. Smoorenburg, H. Brinkmann, R. Crabtree, A. Dancer, M. Forrest, M. . Mayorga, F. P. Mérat, G. Price, G. Stevins and O. Woxen, "Reconsidaration of the Effects of Impulse Noise," NATO, 2003.

[35] P. H. Westfall, "Kurtosis as Peakedness, 1905–2014. R.I.P," *The American Statistician,* pp. 68:3, 191-195, 2014.

[36] D. Henderson and R. P. Hamernik, "The Use of Kurtosis Measurement in the Assessment of Potential Noise Trauma," in *Noise-Induced Hearing Loss - Scientific Advances*, New York, NY, Springer, 2012, pp. 41-55.

[37] R. R. Davis and O. Clavier, "Impulsive noise: a brief review," *Hearing research,* vol. 349, pp. 34-36, 2017.

[38] J. de la Rosa, A. Moreno-Munoz, A. Gallego, R. Piotrkowski and E. Castro, "Spectral Kurtosis based system for transients' detection: Application to termite targeting," in *2009 IEEE Sensors Applications Symposium*, New Orleans, LA, USA, 2009.

[39] C. J. Fackler, E. H. Berger, W. J. Murphy and M. E. Stergar, "Spectral analysis of hearing protector impulsive insertion loss," *International Journal of Audiology,* vol. 56, no. Sup1, pp. 13-21, 2017.

[40] S. Sarray, A. Nakashima, H. R. Dajani, M. Bouchard, D. Lo and S. ginet, "Octave band impulse peak insertion loss: a method to characterize hearing protection devices when firing with small arms," in *26th International congress on Sound and Vibration*, Montreal, 2019.

[41] American National Standard, *Specification for Octave-Band and Fractional-Octave-Band Analog and Digital Filters,* American National Standards Institute, 2009.

[42] P. Rasmussen, G. Flamme, M. Stewart, D. Meinke and J. Lankford, "Measuring recreational firearm noise," *The Journal of the Acoustical Society of America,* vol. 127, no. 3, pp. 1794-1794, 2010.

[43] S. A. Singh and S. M. a. M. Mishra, "Classification of short unsegmented heart sound based on deep learning," in *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Auckland, New Zealand, 2019.

[44] N. García-Pedrajas and D. Ortiz-Boyer, "Boosting k-nearest neighbor classifier by means of input space projection," *Expert Systems With Applications,* vol. 36, no. 7, pp. 10570-10582, 2009.

[45] T. K. Ho, "Nearest neighbors in random subspaces," in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, Beijing, 1998.

[46] H. Bay and T. Tuytelaars, "SURF: Speeded Up Robust Features," in *Lecture Notes in Computer Science, Computer Vision – ECCV 2006: 9th European Conference on Computer Vision*, Graz, Austria, 2006.

[47] C. Giguère and C. Laroche, "Hearing loss prevention program in the military environment," *Canadian Acoustics,* vol. 33, no. 4, pp. 21-30, 2005.

[48] C. Liguori, A. Ruggiero, D. Russo and P. Sommella, "Influence of measurement time on acoustic noise uncertainty," in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Torino, Italy, 2017.

[49] C. H. Hansen, "World Health Organization," [Online]. Available: https://www.who.int/occupational_health/publications/noise1.pdf. [Accessed 28 10 2020].

[50] R. D. Rasband, A. T. Wall, K. L. Gee, S. H. Swift, C. M. Wagner, W. J. Murphy and C. A. Kardous, "Impulse noise measurements of M16 rifles at Marine Base Quantico," in *Proceedings of Meetings on Acoustics*, Minneapolis, Minnesota, 2018.

[51] P. Atrey, N. Maddage and M. Kankanhalli, "Audio based event detection for multimedia surveillance," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing*, Toulouse, France, 2006.

[52] A. M. C. R. Borzino, J. A. Apolinario and M. L. R. de Campos, "Robust DOA estimation of heavily noisy gunshut signals," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Bisbane, Australia, 2015.

[53] T. Ahmed, M. Uppal and A. Muhammad, "Improving efficiency and reliability of gunshot detection systems," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, 2013.

[54] M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot and C.-A. Deledalle, "Machine learning in acoustics: Theory and applications," *The Journal of the Acoustical Society of America,* vol. 146, no. 5, pp. 3590-3628, 2019.

[55] J. Kim, C. Park, J. Ahn, Y. Ko, J. Park and J. C. Gallagher, "Real-time UAV sound detection and analysis system," in *2017 IEEE Sensors Applications Symposium (SAS)*, Glassboro, NJ, USA, 2017.

[56] R. MacLeod, "Modeling of active reverberation by time delay estimation," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, Seattle, WA, USA, 1998.

[57] R. Patole and P. Rege, "A comparative analysis of classifiers and feature sets for acoustic environment classification," *The Journal of the Audio Engineering Society,* vol. 67, no. 12, pp. 939-952, 2019.

[58] MatLab, "Reverberator," Mathworks, 2020. [Online]. Available: https://www.mathworks.com/help/audio/ref/reverberator-system-object.html. [Accessed 17 10 2020].

[59] A. Nakashima, "A comparison of metrics for impulse noise exposure," DRDC Toronto Research Center, Toronto, 2015.

[60] N. Fink, H. Z. Pikkel, A. Eisenkraft and G. A. Banta, "Hearing protection devices and methods used for their evaluation: A military perspective," *Journal of military, veteran and family health,* vol. 5, no. 1, pp. 141-147, 2019.

[61] W. Ahroon, R. Hamernik and R. Davis, "Complex noise exposure: an energy analysis," *Journal of Acoustical Society of America,* vol. 93, no. 2, pp. 997-1006, 1993.

[62] M. Roberto, R. P. Hamernik, R. J. Salvi, D. Henderson and R. Milone, "Impact noise and the equal energy hypothesis," *The Journal of the Acoustical Society of America,* vol. 77, no. 4, pp. 1514-1520, 1985.

[63] S. C. Olhede and A. T. Walden, "Generalized Morse Wavelets," *IEEE Transactions on Signal Processing,* vol. 50, no. 11, pp. 2661-2670, 2002.

[64] G. Andria, F. Attivissimo, G. Cavone and A. Lanzolla, "Selection of wavelet functions and thresholding parameteres in ultrasound image denoising," in *2013*

IEEE International Symposium on Medical Measurements and Applications (MeMeA), Gatineau, QC, Canada, 2013.

[65] M. Raghuram, K. V. Madhav, E. H. Krishna, N. R. Komalla, K. Sivani and K. A. Reddy, "Dual-tree complex wavelet transform for motion artifact reduction of PPG signals," in *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings*, Budapest, Hungary, 2012.

[66] L. Iuppariello, G. D'Addio, G. Pagano, A. Biancardi, M. Romano, P. Bifulco and M. Cesarelli, "Effects of wavelets analysis on power spectral distributions in posturographic signal processing," in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Benevento, Italy, 2016.

[67] X. Zhu and J. Kim, "Application of analytic wavelet transform to analysis of highly impulsive noises," *journal of sound and vibration,* vol. 294, no. 4-5, pp. 814-855, 2006.

[68] F. J. Harris, "On the use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE,* vol. 66, no. 1, pp. 51-83, 1978.

[69] K. R. Borisagar, R. M. Thanki and B. S. Sedani, "Fourier Transform, Short-Time fourier Transform, and Wavelet Transform," in *Speech Enhancement Techniques for Digital Hearing Aids*, Cham, Switzerland, Springer, 2018, pp. 63-74.