INVESTIGATION OF THE EFFECTS OF SIGNAL AMPLITUDE ON DISCRETE FOURIER TRANSFORM (DFT) AND DFT RESOLUTION APPROACHES

Supported by tetfund

Olatilewa Raphael Abolade[1] and Alice Oluwafunke Oke[2]

[1]Department of Computer Engineering, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria

[2]Department of Computer Engineering, Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

Corresponding author email: abolade.raphael@oouagoiwoye.edu.ng

Received: June 17, 2024 Accepted: August 28, 2024

Abstract: In this work, five audio signal datasets: Speech On.wav (SO), Speech Off.wav (Sof), Speech Disambiguation.wav (SD), Speech Misrecognition.wav (SM) and Window Navigation Start.wav (WNS), were acquired from Microsoft Windows Sound Notification Collection to investigate the effects of signal amplitude on two frequency estimation methods, namely, Discrete Fourier Transform (DFT) and DFT resolution approaches. The signals were analyzed using MATLAB version R2018b. The DFT and DFT resolution estimates of the fundamental frequency of each sample was determined. The Power spectrum analysis of the audio signals was carried out to extract data about the signals. The amplitude of each of the signals was recorded. The DFT estimate values for the five audio signal datasets were 703.4180Hz, 349.2124Hz, 1176Hz, 347.0099Hz and 988.434Hz respectively, and their DFT Resolution estimates were 700.8324Hz, 349.1250Hz, 1177Hz, 343Hz, and 998.0526Hz while the amplitudes were 0.1236m, 0.1635m, 0.0222m, 0.1399m, and 0.0264m respectively. The DFT Resolution method outperformed DFT method with SO, Sof, SM by 2.5856Hz, 0.0874HZ and 4.0099Hz, respectively while DFT outperformed DFT Resolution with SD and WNS by 1Hz and 9.6186Hz respectively. The relationship between the amplitude of the signals and the performance of the estimation methods was examined by plotting a graph of amplitude against the improvement in DFT estimation methods. The application of DFT Resolution to the signals demonstrates a constant, albeit more pronounced, inverse relationship between amplitude and the estimated frequency. However, DFT methods performed better when analyzing signals with very low amplitudes as seen in the analysis of SD and WNS.

Keywords: Amplitude, frequency, energy, signal, waveforms, bins.

Introduction

In many fields of study, a signal's frequency is crucial. Since signals convey information while they move from one place to another, they can be thought of as functions that transport data made up of various frequency components, and because they carry information, the study of signals and their properties is of interest in many areas of research. The amplitude and frequency of signals are of particular interest. However, the process of determining the frequency has known weaknesses. The process involves converting a signal from the time domain to the frequency domain in order to determine the frequency. The value of the resulting frequency itself is approximated in this process. The use of the resulting frequency in an application could impact results negatively.

To further analyze frequency, Fourier analysis is used to describe the complex waveforms of signals. The signal would be broken into elemental units to make it possible to specify it. According to Madisetti (2018), a mathematical technique that aids in signal encoding and signal performance analysis is the Fourier transform. It expands signals into a spectrum of sinusoidal components. The Fourier series is a method of analyzing periodic functions in the time domain. First, the function is expressed as the sum of a sequence of sines and cosines, or sinusoids. Each sinusoid is specified by its frequency, which is then added together. Baron Jean Fourier demonstrated that the linear sum of harmonically linked sine and cosine waves with specified frequencies, amplitudes, and phases can be used to represent a periodic waveform of unlimited complexity. A sound wave's amplitude is determined by the energy required to produce it, whereas frequency indicates the occurrence. The Fourier transform is used to obtain a discrete form of a signal, which

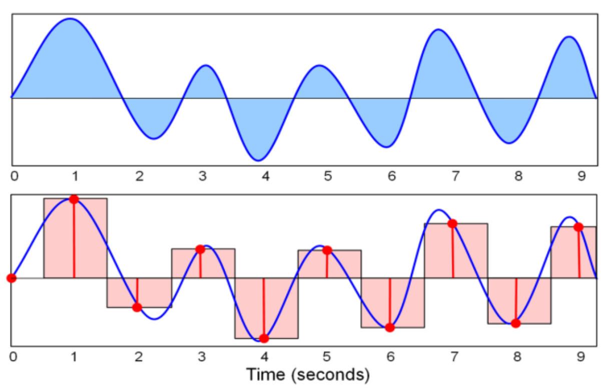

Figure 1: Discrete Fourier Transform approximation of a signal. (Source: Muller, 2015).

results in the frequency domain. The Discrete Fourier Transform approximation of a signal is represented in Figure 1. The frequency of sampling a signal is the same as the inverse of the time domain's period in the frequency domain, but sampled signals are uniformly distributed in the time domain, and the period represents the separation between samples (Trauth, 2020).

It is well known that the assumptions made by the Fourier Transform technique has some inadequacies when converting a signal from the time domain to the frequency domain. For instance, the Fourier transform algorithm that is used in converting signals from time domain to frequency domain operates on the assumption that a signal's time record is exactly repeated throughout and that, as a result, the signals inside it are periodic at intervals that match the time record's duration. Because of this insufficiency, spectral leakage—a phenomenon in which a single-toned signal becomes dispersed across multiple frequencies following the application of the discrete Fourier transform—occurs, making it difficult to determine the signal's true fundamental frequency and resulting in imprecise fundamental frequency estimation (Deraemaeker and Worden, 2012).

The amplitude and loudness of a sound wave are determined by the energy required by the source to produce the sound. Evans *et al.* (2024) report a strong correlation between loudness and wave amplitude. This suggests that a wave's energy in the time domain will be transformed into a different form in the frequency domain. A wave's energy is directly correlated with the square of its amplitude. Andrews (2023) states that amplitude and frequency have an inverse connection, which may be mathematically stated as $E \propto A^2$, where E represents the wave's energy and A, its amplitude. In this work, the impact of a signal's amplitude on fundamental frequency estimation was investigated. This paper is presented in sections thus: section one; introduction, section two; materials and methods, section 3; results and discussion, and section 4; conclusion.

Materials and Methods

The study employed five signals that were sourced from the Microsoft Windows Sound library. The signals are Speech On.wav (SO), Speech Off.wav (Sof), Speech Disambiguation.wav (SD), Speech Misrecognition.wav (SM) and Window Navigation Start.wav (WNS). First, the amplitudes of the sample signals were determined by time-domain analysis.

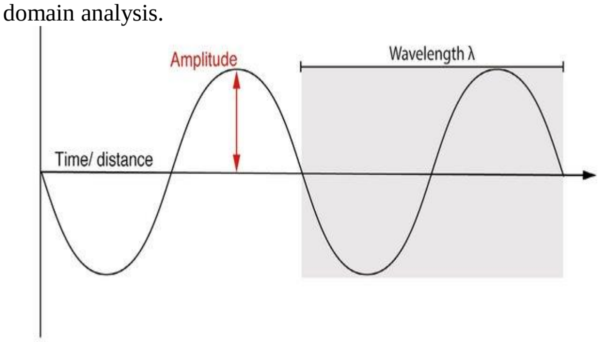

Figure 2: Graph of time (T) against amplitude (A) (Source: Browning *et al.*, 2017)

The plot of amplitude vs. time is displayed in Figure 2. The sinusoidal waveforms were then examined in MATLAB. Figure 2 illustrates how the components of a signal can be displayed graphically in the time domain. A and T stand for the signal's peak amplitude and period, respectively. The maximum movement of points on a wave is represented by the amplitude, which is the maximum displacement of a sinusoidal oscillation from the equilibrium position. According to Clark and Dunn (2022), the amplitude of a sound wave is the amount of movement of air particles; the more the displacement, the higher the wave's energy.

Frequency is expressed as $F_s = \frac{1}{T_s}$, where T_s is the sample interval and F_s is the sampling frequency. Shmaliy (2010), expressed the frequency of sampling as $F_s = \frac{1}{T_s}$, where T_S is the period between samples and the frequency of sampling is F_s. The frequency of sampling is the inverse of period between samples.

Analysis of a sinusoidal signal in the time domain.

The time waveforms of the input sinusoidal signals were analyzed in MATLAB. The sample signals were read using the MATLAB function $wavread()$; the function took the file as input, and a vector was defined to represent the time from zero to a maximum value possible. A plot of time versus the amplitude was done to give the time domain characteristics of the waveform using MATLAB's plot function.

Conversion of sinusoidal signal in time domain to frequency domain.

The Fourier transform algorithm is used to convert signals from the time domain to the frequency domain (McManus *et al.*, 2020). The frequency was defined thus; from $-fs/2$ to $+fs/2$, the frequency domain represents f as an increment of $f/(n-1)$. The signals' Fourier transform (FFT) was derived by plotting f versus the signal's absolute value, also known as the complex modulus or ABS.

Table 1: Simulation Parameters

S/N	Parameter	Description
1	X	Sample data size
2	T_S	Period between samples
3	F_S	Frequency of sampling
4	N	Number of samples
5	T_{max}	Time of last sample
6	Δf	Occurrence Interval between Bins

Table 1. shows the simulation parameters used in DFT and DFT resolution methods. The separation between bins is the frequency resolution. Each bin is separated by Δf. This implies that the analysis of the magnitude spectrum will only be accurate within Δf. A smaller value of Δf would yield better accuracy. A better resolution of the signal can be obtained by increasing N. (Kramer and Eden, 2016). Increasing the number of samples (N) decreases Δf. Where a higher resolution is required, a longer record must be obtained. In this case, a larger size of the sample was taken to give a small Δf sample with a better resolution. The frequency resolution is the distance between bins, and equation 1 expresses this distance which divides each bin apart. This suggests that only within Δf would the magnitude spectrum analysis be accurate. A lower value would result in more precision. Increasing N will result in a

higher signal resolution. (Kramer and Eden, 2016). Δf is reduced as the number of samples (N) increases. It is necessary to have a longer record when a greater resolution is needed (Qaisar *et al.*, 2023). To provide a sample with a better resolution in this instance, a larger sample size was taken such that $\Delta f = \frac{F_s}{N}$ (Eq1)

Through simulations, values for amplitude, DFT, and DFT resolution were obtained, and the results from DFT and DFT resolution were compared for each signal. The plots of the signals in time and frequency domains were done using the MATLAB R2018B simulator. The maximum amplitude of each signal was plotted, and the DFT and DFT resolution results were compared for each of the signals. To study the relationship between amplitude and the performance of the frequency estimation methods, the maximum amplitude was plotted against the expected improvement in the DFT resolution method.

Results and Discussion

A wave's energy is directly correlated with the square of its amplitude. This relationship can be expressed mathematically as $E \propto A^2$, where E is the energy of the wave and A is its amplitude. DFT was used to generate the signal's power spectrum and realize the frequencies of the sample signals. Through DFT resolution, a more accurate approximation of the fundamental frequency was obtained. The power spectrum was used to extract data regarding the frequency peak and its related bins, which revealed information about the fundamental frequency. A frequency domain analysis of the signals was necessary to show how the energy of the signals is distributed throughout a range of frequencies.

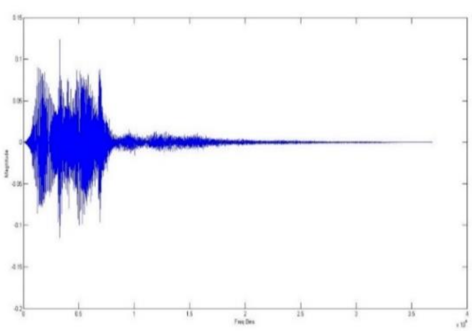

Figure 2: Time domain plot of Speech On.wav

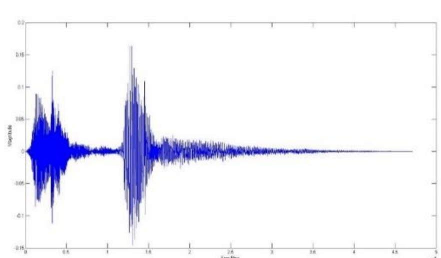

Figure 3: Time domain plot of Speech Off.wav

The five sample signals' time and magnitude are represented by the Y and X axes in the time domain graphs. Figures 2 and 3 respectively, display the time domain plots for Speech On.wav and Speech Off.wav. When magnified, the result in every instance resembled a sine wave. It was noted that the amplitude changed quickly at first and then steadily as the experiment went on. A sound wave's amplitude and loudness are determined by the energy needed for the source to produce the sound. This implies that a wave's energy in the time domain will be transformed in the frequency domain.

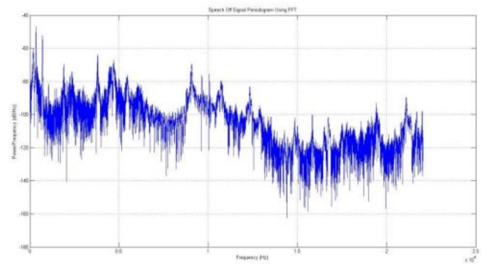

Figure 4: Power spectrum of Speech On .wav

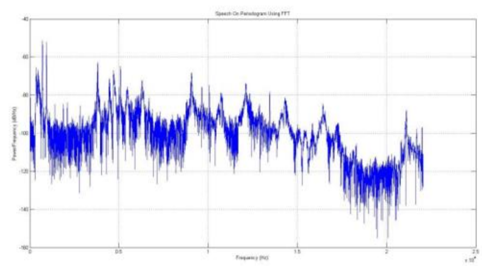

Figure 5: Power spectrum of Speech Off .wav

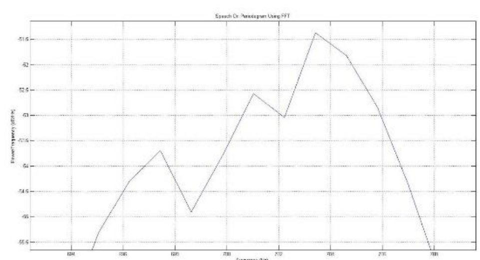

Figure 6: Signal Power spectrum of Speech On signal zoomed

The power spectrum of a continuous-time signal in signal processing represents the power distribution among frequency components making up the signal. Figures 4 and 5 show the sampled frequencies of Speech On.wav and Speech Off.wav, respectively, providing an indication of each frequency's relative contribution to the overall power of the signal in the power spectrum. The X-axis represents the frequency of the signal in Hz, while the Y-axis represents the power/frequency measured in dB/Hz.

In Figure 6, $-53.6978dB$, $-52.5742dB$, $-51.3706dB$ represent the powers of the peaks and $697.4366Hz$, $701.0254Hz$, $703.4180Hz$, are the corresponding frequency bins, while the resolution is 1.1963 in the case of Speech On signal. The estimated fundamental frequency obtained was 700.5980 Hz. The power spectrum of the Speech On signal is zoomed in in Figure 6.

Table 2: Amplitude, DFT and DFT Resolution Values

S/N	Input Signal	Max Amplitude	DFT f_{0est}	DFT f_{0res}	DFT f_{0est} – DFT f_{0res} (Hz)
1	Speech On.wav	0.1236	703.4180	700.8324	2.5856
2	Speech Off.wav	0.1635	349.2124	349.1250	0.0874
3	Speech Disambiguation.wav	0.0222	1176	1177	-1
4	Speech Misrecognition.wav	0.1399	347.0099	343	4.0099
5	Windows Navigation Start	0.0264	988.4340	998.0526	-9.6186

Each sample's maximum amplitude, DFT, DFT resolution, and the expected improvement in DFT resolution results are presented. In Table 2, Speech Disambiguation.wav and Window Navigation Start.wav having amplitudes of about 10 percent compared to others have slightly better results using the DFT method than with the DFT Resolution method, whereas for larger amplitude values, the DFT Resolution performed better than DFT as expected. In Table 2, the maximum amplitude, DFT, DFT resolution and the difference in DFT and DFT resolution methods results are presented. Comparing the DFT and DFT resolution of the input signal, it was observed that Speech Disambiguation.wav and Window Navigation Start.wav, having much lower amplitudes of 0.0222m and 0.0264m, respectively, performed better with DFT than with DFT resolution, contrary to expectation.

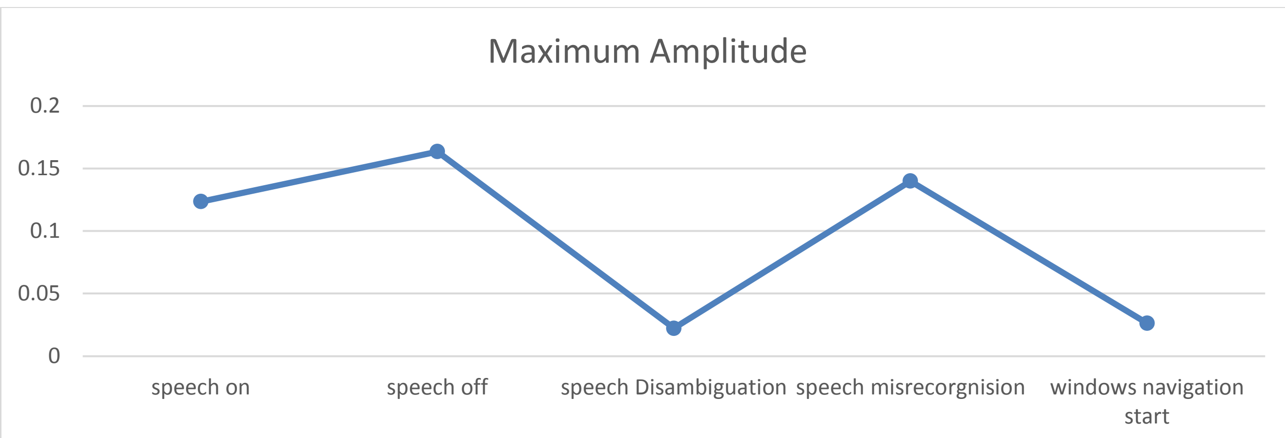

Figure 7: Maximum amplitude of samples

In Figure 7, a plot of the maximum amplitude against the corresponding signal is presented. Speech Off signal has the largest amplitude at 0.1635m, while Speech Disambiguation has the smallest amplitude at 0.0222m. This indicates that Speech Off signal possesses the highest energy of all the signals, while Speech Disambiguation signal possesses the least energy, according to Clark and Dunn (2022) the amplitude of a sound wave is the amount of movement of air particles, the more the displacement, the higher the wave's energy. According to Andrews (2023), there is an inverse relationship between amplitude and frequency. The impact of magnitude of amplitude on the inverse relationship between amplitude and frequency was investigated in relation to the methods used in DFT and DFT resolution of signals.

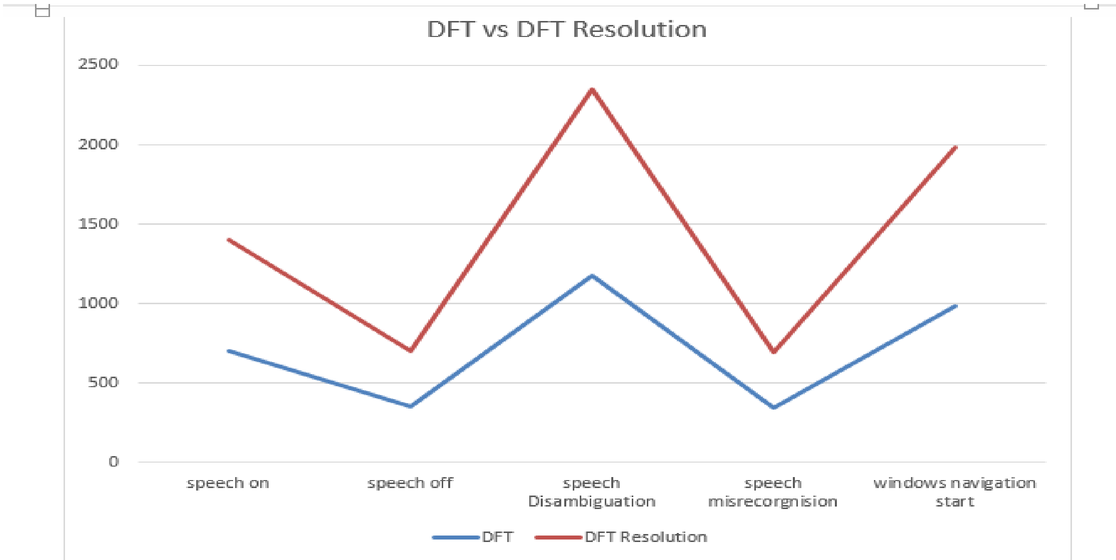

Figure 8: DFT and DFT Resolution

In Figure 8, the inverse relationship between amplitude and frequency was observed, and when compared to Figure 7, it was also observed that the inverse proportionality was more, and consistent with what was observed when the DFT resolution method was used compared to the DFT method.

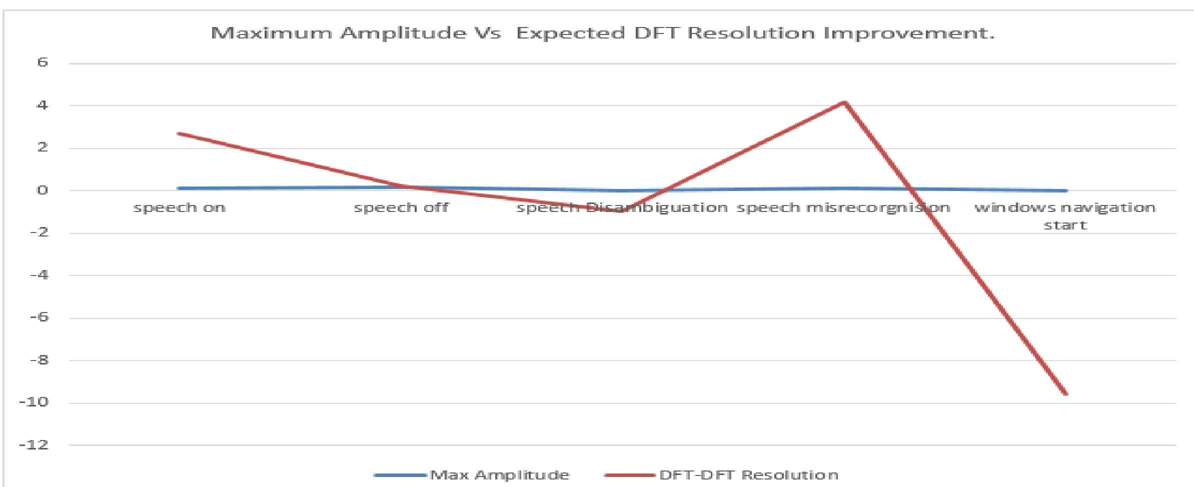

Figure 9. Maximum Amplitude and Expected DFT Resolution Improvement.

A graph of the expected improvement from DFT and DFT resolution against the maximum amplitude of each signal is plotted in Figure 9. The plot shows that Speech Misrecognition has the largest margin of improvement in result when comparing the result from both methods, with Speech Misrecognition having a value of 4.0099Hz, while the result for Windows Navigation Start was poor with a value of -9.6186.

Conclusion.

When compared to the DFT estimates of the fundamental frequency of the signals, the application of DFT resolution to the signals demonstrates a consistent, albeit more pronounced, inverse relationship between amplitude and frequency regardless of the signal's size. However, DFT outperformed the DFT resolution method with Speech Disambiguation.wav and Window Navigation Start.wav, which were noted to have much lower amplitude compared to other signals. In the conversion of signals from time domain to frequency domain using DFT, the DFT resolution method may not necessarily yield a better result with signals having small amplitudes.

Acknowledgement

I wish to thank Professor Oke A.O. for her support and guidance towards the success of this work.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Clark, F. E., & Dunn, J. C. (2022). From soundwave to soundscape: A guide to acoustic research in captive animal environments. Frontiers in Veterinary Science, 9, 889117.

Deraemaeker, A., and Worden, K. (Eds.). (2012). *New trends in vibration based structural health monitoring* (520). Springer Science & Business Media.

Evans, J. P., Lin, K., & Savostyanov, A. N. (2024). Effects of fundamental frequency changes on spoken sound loudness. Perception, 03010066241249451.

Andrews, S. S. (2023). Light and Waves: A Conceptual Exploration of Physics. Germany: Springer International Publishing.

McManus, L., De Vito, G., and Lowery, M. M. (2020). Analysis and biophysics of surface EMG for physiotherapists and kinesiologists: toward a common language with rehabilitation engineers. Frontiers in neurology, 11.

Browning, E., Gibb, R., Glover-Kapfer, P., and Jones, K. E. (2017). Passive acoustic monitoring in ecology and conservation.

Shmaliy, Y. (2010). Continuous-Time Signals. Netherlands: Springer Netherlands.

Trauth, M. H. (2020). MATLAB® Recipes for Earth Sciences. Germany: Springer International Publishing.

Muller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications.* Springer.

Madisetti, V. K., (2018). *The Digital Signal Processing Handbook-3 Volume Set*. CRC press.