Acoustic analysis of snoring sounds towards testing relationships with other respiratory signals and health outcomes

By

Melody Lu

Master of Engineering (Biomedical)

Thesis
Submitted to Flinders University
for the degree of

Master of Engineering (Biomedical)

College of Science and Engineering
21/12/2022

Declaration

I certify that this thesis:

does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university

and the research within will not be submitted for any other future degree or diploma without the permission of Flinders University; and

to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

Signature of student..

Print name of student...............MELODY LU.............

Date...............17/10/2022...

Acknowledgments

I would like to express my gratitude to my primary supervisor, Kristy Hansen, who guided me throughout this project. I would also like to thank my co-supervisors, Peter Catcheside and Bastien Lechat who supported me and offered deep insight into the study. This endeavour would not have been possible without their editing help, late-night feedback sessions, and research assistance. I wish to extend my special thanks to Duc Phuc Nguyen for his assistance in developing MATLAB programmes for this study. Lastly, I would be remiss in not mentioning my friends and family, especially my parents. Their belief in me has kept my spirits and motivation high during this process.

Abstract

Snoring has been a prevalent sleeping disorder among the general population in Australia with around 20%-25% of Australians regularly snoring on most nights. Even though a lot of studies have been conducted to develop more accurate acoustic methods to identify and classify snoring signals, further investigation is still needed to advance snoring assessment methods. Moreover, the consensus for the objective definition of snoring is still lacking. This study aims to explore the information embedded in the acoustic signals of snoring and to examine the performance of a snoring detection algorithm. Another objective of this study is to advance the acoustic analytic techniques for future development that would allow for a more comprehensive differential diagnosis and evaluation of snoring and obstructive sleep apnoea in a home setting. A total of 2330 20-second audio segments from 6 participants were used for the acoustic analysis of snoring sounds. The performance of a snoring detection algorithm was evaluated by a confusion matrix, ROC curve, and other performance evaluation metrics. This study that indicated snoring happens in the lower frequency range at about 200 Hz. There were no noticeable differences between the snoring and non-snoring episodes at frequencies between around 200 Hz and 1000 Hz. For this snoring dataset, 0.62 was the best-performing cut-off relative power. Interestingly, the snoring and non-snoring power spectra appeared to fluctuate more in those participants who had a higher number of snoring events recorded. The results of this study may help in the development of an objective definition of snoring based on its acoustic characteristics. In addition, this work can be beneficial for the development of snoring detection algorithms. Future studies to support the conclusions of this study can be carried out by adding more human raters and participants.

Contents

Table of Figures

Table of Tables

Table of Equations

1. Introduction

1.1 Background of the study

Snoring is prevalent among the general population with around 20%-25% of Australians regularly snoring on most nights (Better Health Channel, 2014; Sleep Disorders Australia, 2020) and up to 60% of men and 40% of women snoring to at least some degree. Moreover, snoring becomes more common as people get older (American Academy of Sleep Medicine, 2014).

Snoring is a breathing sound that typically occurs during inhalation in sleep but can also occur during exhalation (American Academy of Sleep Medicine, 2014). Acoustically, snoring occurs when the patency of the upper airway is diminished, and the soft-tissue walls of the oropharynx begin to vibrate to produce snoring sounds (Dalmasso & Prota, 1996), which can range from light minimally intrusive sounds to very heavy and disruptive snorting (Godman, 2022). The harshness of snoring sounds is thought to largely reflect flutter of the soft palate (Pevernagie et al., 2010), although other soft tissues structures may also contribute, and snoring loudness also depends on breathing effort.

In general, occasional snoring is considered to be a largely innocuous phenomenon with no significant negative effects on health (Better Health Channel, 2014). However, habitual loud snoring can be highly problematic, particularly for bed-partners of snorers. The American Academy of Sleep Medicine describe snoring as a respiratory sound that “occurs without episodes of apnoea, hypopnea, respiratory effort related arousals (RERAs) or hypoventilation” (American Academy of Sleep Medicine, 2014). Thus, loud or frequent bouts of snoring are widely considered to be a strong sign of significant sleep-disordered breathing and obstructive sleep apnoea (OSA) (Sogebi et al., 2011). People with occasional snoring are very likely to develop obstructive sleep apnoea with aging or weight gain (American Academy of Sleep Medicine, 2014), so snoring assessments may provide useful prognostic markers of future health risks.

Assessments on snoring are highly dependent on the definition of snoring. Dalmasso and Prota (1996) emphasised that snoring should be defined based on the method of measurement used for sound and noise signals.

Early work mainly focused on acoustic investigations of snoring mechanisms, loudness, and upper airway cross-sectional area (CSA) and sites of obstruction (Dalmasso & Prota, 1996). Dalmasso and Prota (1996) reported that Leq-equivalent continuous sound level, power spectrum, and linear prediction code for CSA are the primary methods to analyse and measure snoring.

In recent years, the acoustic study of snoring for the diagnosis of sleep disorders has gained a lot of popularity in sleep medicine. Many researchers have established different acoustic analysis techniques to distinguish between normal snoring and obstructive sleep apnoea. However, a major limitation of existing work to date is that current classification systems used to define sleep disordered breathing and obstructive sleep apnoea rely on human scoring of the most obvious transient complete (apnoea) or partial (hypopnoea) upper airway obstruction events that last at least 10 seconds and are associated with oxygen desaturation or brief arousals from sleep (American Academy of Sleep Medicine, 2014). This approach ignores individual breath-by-breath level data more relevant to snoring, including extended periods of snoring that do not meet current definitions for obstructed breathing events. Nevertheless, snoring assessments are likely to be useful for simplifying the detection of classically defined obstructive sleep apnoea and Wang and Peng (2017) point out that the performance of snoring detection and acoustic analysis can be improved by using advanced signal processing techniques such as wavelet transform, independent component analysis, and Hilbert-Huang Transform (HHT).

Although many studies have been carried out to establish more accurate acoustic methods of identifying and classifying snore signals, further work remains warranted to improve snoring assessment techniques. In addition to an ongoing lack of information on acoustic features of snoring associated with particular disorders (Dalmasso & Prota, 1996), very little is known regarding likely relationships between snoring and the degree of upper airway obstruction, or the specific sites of upper airway collapse likely associated with different acoustic features. More research on the relationship between the physiological characteristics of obstructed breathing and relationships with snoring-related acoustic signals is required towards more effective and comprehensive assessments of obstructed breathing during sleep.

1.1 Project Focus

The objectives of this project were to explore information embedded in snoring acoustics and to evaluate the performance of a snoring detection algorithm. It was expected that more information on the acoustic features of snoring could provide improved methods for assessing sleep disordered breathing. Furthermore, this project aimed to advance methods towards future development of acoustic analysis amenable to more comprehensive differential diagnosis and assessment of snoring and obstructive sleep apnoea in a home setting.

1.2 Project Value

The initiative of this project arose from the need to extract more comprehensive information regarding relationships between snoring, the severity of obstructed breathing during sleep and clinical outcomes in sleep related breathing disorders. Moreover, the project also sought to generate more comprehensive definitions of snoring than current methods that do not adequately consider the periodic and highly variable nature of upper airway tissue vibrations underlying snoring.

Since the current acoustic analyses are not sufficient to correlate the acoustic features of snore signals to pathological indications, this project examined existing algorithms for snoring detection towards exploring the acoustic characteristics of snoring. In this regard, the results of this project have the potential to assist future acoustic approaches to identify underlying health problems arising from obstructed breathing in sleep and to help establish a more standardised method to record snoring.

1.3 Structure of the Thesis

This thesis is divided into five sections. The first section gives an introduction of the project which briefly describes the background of the study, project aim, and project value. The next section contains the literature review that presents the theoretical underpinnings of the research topic. A review of relevant works of literature on the definition of snoring, detection and classification of snoring acoustics, and the correlation between snoring and physiological features are covered.

The third section details materials and methods used to obtain the necessary information, which consists of database properties and data preparation, ranking of snore events, and the selection of acoustic methods for snoring detection and classification. Then, the results obtained from the procedures are presented in the fourth section, which aims to provide an extensive analysis of the results.

In the last section, the implications and applications of the results are explored and discussed for the purpose of enhancing the field of diagnostic acoustic analysis for snoring. Conclusions, weaknesses of the study, and recommendations for future work on snoring are also provided.

2 Literature Review

2.1 Snoring Definition

In general, snoring is a noisy breathing sound generated by airway tissue vibrations that predominantly emanate from the upper airway tissues including the soft palate, tongue, uvula, and tonsils during sleep. Snoring is regarded as a social nuisance not only to the snorers themselves but particularly to bed partners. There have been different definitions for snoring according to different scholars and sleep specialists. Rohrmeier et al. (2013) stated that “Although snoring is a common problem, no unequivocal definition yet exists for this acoustic phenomenon”.

Early research had defined a healthy simple snore event as “a power spectrum of snoring signal with a harmonic structure and a fundamental frequency that ranged 110-190 Hz” (Fiz et al., 1996). However, this definition only characterised snoring signals on the basis of the sound frequency and the presence of harmonics. Other acoustic features, especially of snoring itself, had not been considered.

In 1996, Dalmasso & Prota (1996) described snoring as “a typical inspiratory sound, even though a small expiratory component can be heard or recorded (especially in OSA patients) with different spectral features”. The authors then emphasised the importance of taking into account the type of acoustic measurements when it comes to defining snoring sounds (Dalmasso & Prota, 1996).

More recently, Pevernagie et al. (2010) discussed the mechanism of snoring as a “vibration of anatomical structures in the pharyngeal airway. Flutter of the soft palate accounts for the harsh aspect of the snoring sound”. The authors found that “the pitch of the snoring sound is in the low-frequency range (<500 Hz) and corresponds to a fundamental frequency with associated harmonics. The pitch of snoring is determined by vibration of the soft palate, while nonpalatal snoring is more ‘noise-like’, and has scattered energy content in the higher spectral sub-bands (>500 Hz)”. Pevernagie et al.’s research approached snoring sounds by employing speech analysis as the mechanisms of sound generation are similar. The definition had only taken account of the changes in snoring pitch and anatomic site of snoring.

Swarnkar et al. 2017 defined snoring as “a ‘Breath Episode’ containing periodic packets of energy, even for a small portion of the episode”. In the article, a ‘Breath Episode’ is defined as “the sound originated from the patient from the start of an inspiration to the corresponding end of expiration” (Swarnkar et al., 2017). However, snoring can sometimes start from expiration and can be non-consecutive and irregular across breaths so this definition by Swarnkar et al. (2017), like others, remains somewhat arbitrarily focussed on inspiration. Nevertheless, it is worth noting that the flow dependence of snoring dictates that snoring is periodic in nature and can occur during either inspiration or expiration in the presence of at least some airflow.

Janott et al. conducted a study in 2019 with a focus on finding a more objective acoustic definition of snoring. In their research, the authors undertook a comprehensive acoustic and psychoacoustic analysis in an effort to arrive at a more robust acoustic differentiation between snoring and loud breathing. Although Janott et al's findings have accelerated the development of an objective definition of snoring, such a definition is still built on the basis of the subjective impressions of 25 raters.

Despite the fact that multiple research groups are relatively active in the study of the snoring acoustics, a clear definition that can be used objectively as a snoring acoustic measure is still lacking to date (Janott et al., 2019).

2.2 Acoustic Methods of Snoring Detection and Classification

Automatic snore detection is not a novel concept in sleep medicine. However, snoring detection algorithms still warrant improvements to achieve more meaningful assessments of upper airway obstruction severity. In this context, several investigations on acoustic analysis of snoring were carried out. Dalmasso & Prota (1996) highlighted three primary methods for the analysis of snoring which are Leq-equivalent continuous sound level, power spectrum (PS) and linear prediction code (LPC) for assessing upper airway cross-sectional area (CSA) changes. These methods have been assessed to associate snoring acoustics with anatomical and pathological implications. The focus of Dalmasso & Prota's paper was to distinguish nasal and oral snoring patterns. There is less information about the differentiations between snoring and other noises such as loud breathing and environmental noises.

A series of methods for snoring detection have been proposed and evaluated recently. Lee et al. (2013) approached the detection of snoring episodes from both acoustical and mechanical perspectives. In their study, a hidden Markov model-based method to detect snoring using a piezo snoring sensor was presented. The degree of vibrations related to snoring and non-snoring events were measured with the use of a piezo snoring sensor attached to the neck. However, this study was established on data that were subjectively selected and heartbeat and breathing noises were not taken into account in their snoring detection model.

Shin & Cho (2014) developed a snoring detection model that considered different types of noises in a standard bedroom. Shin & Cho extracted snoring parameters by characterising the recorded sounds. The authors then conducted a formant analysis, quadratic classification, and ten-fold cross-validation to identify snoring episodes. Both Lee et al. and Shin & Cho used formant analysis for snoring detection. Nevertheless, the lack of understanding of the formant frequencies and their significance remains a key unresolved issue (Shin & Cho, 2014). The ambiguities of the characteristics of snoring sound formants and acoustic properties of snoring are all due to the absence of a universally accepted definition of snoring.

Wang & Peng (2017) carried out an extensive review of acoustic analysis of snoring to evaluate recent acoustic methods for snoring detection and classification. For snoring detection, Wang & Peng (2017) identified that "Wavelet Transform is able to separate the fine details in a signal by using the multi-scale operation of scaling and translation, which is a very useful tool to analyse the instantaneous and time-varying non-stationary signals". A key problem with wavelet transforms is the difficulty of constructing a proper wavelet basis function to analyse snoring acoustics. This problem can be avoided by using Hilbert-Huang Transform (HHT) as HHT demonstrated higher time-frequency resolution than the time-frequency distribution (Wang & Peng, 2017). However, HHT also has some limitations such as the optimization of empirical mode decomposition algorithm, and boundary problems (Wang & Peng, 2017). Wang & Peng (2017) further reported that "snoring detection algorithms are mainly divided into two categories at present: one uses signal processing method including the short-time energy threshold method (Xu et al., 2013), double-threshold end-point detection (Liu et al., 2013), snoring enhancement method based on autocorrelation character (Ng et al., 2008) and so on". Depending on the snoring features, and category theory such as artificial neural network method, support vector machines method and Gaussian mixture model method, some complex algorithms are able to achieve higher precision for snoring detection (Wang & Peng, 2017).

As for snoring classification, Wang & Peng (2017) have reviewed different methods to classify snoring by identifying the different characteristics of snoring and snoring locations. According to Wang & Peng (2017), Bayesian classification, Gaussian mixture model (GMM), and K-nearest neighbour (KNN) are the typical classification approaches. The authors noted that GMM performs well in terms of specificity and sensitivity, but with a complicated algorithm and a slow real-time response (Wang & Peng, 2017). On the other hand, Bayesian classification is more straightforward and effective with good precision in the diagnosis of OSA and disease severity (Wang & Peng, 2017). Additionally, KNN is also a simple technique, but it requires a larger dataset to successfully classify snoring (Wang & Peng, 2017).

Sola-Soler et al. (2011) made efforts to classify snoring acoustics with the use of snore features in combination with apnoea-related information. Sola-Soler et al. proposed Bayes multi-group classification with kernel gaussian probability density function estimation to determine the severity of snoring related to sleep apnea hypopnea syndrome. Although Sola-Soler et al.'s work has advanced the research of snoring classification technology, further validation is still needed.

2.3 Correlation between Snoring Acoustics and Physiological Features

Snoring is viewed as a clinical sign and risk factor for several underlying medical conditions, most notably obstructive sleep apnoea and hypertension. Research into the clinical implications related to snoring has a long history. Dalmasso & Prota (1996) investigated the anatomical, pathological, and acoustical aspects of snoring and provided a detailed review that concentrated on correlating snoring signals with physiological features using acoustic methods. However, the characteristics of snoring and relationships with a range of clinical outcomes remain poorly understood and limited by a lack of detailed acoustic data in patient groups at risk of adverse clinical outcomes such as hypertension, strokes and myocardial infarction potentially associated with snoring (Dalmasso & Prota, 1996).

Furukawa et al. (2016) examined the correlation between snoring sound intensity and morning blood pressure. The authors found a relationship between morning blood pressure and tracheal sound intensity suggestive of a potential pathophysiological relationship. However, Furukawa et al.'s research did not consider the acoustic features of snoring related to OSA. Moreover, instead of using a clinical sample or conducting the study on the general population, their study was only focused on a worker population

A more recent study by Kayabekir & Yağanoğlu (2021) found an association between delta waves in EEG channel C3-M2 and snoring sound waves in polysomnography (PSG) evaluations of sleep. Kayabekir & Yağanoğlu's research only explored the properties of snoring sounds relationships with EEG signals. The research was based on people with primary snoring and with OSA using the apnoea-hypopnea index (AHI). Further studies in other sleep-breathing disorders involving hypoventilation and hypoxemia and relationships with snoring characteristics such as the severity and pitch period could potentially reveal clinically useful snoring assessment methods.

2.4 Research Gap

Due to the lack of a universally accepted definition of snoring, the interpretation of many sleep health studies is problematic. Previous work has largely failed to present a clear unequivocal definition of snoring that can be reproduced in other studies. A fundamental issue is that independent objective measures of upper airway obstruction are largely lacking from existing studies to date. Thus, differentiation between snoring acoustics and loud breathing with no or minimal airway tissue obstruction and tissue vibration is not yet reliably possible. Despite many theories that have been accumulated over the decades, the fundamental origins and characteristics of snoring are still not well understood. Some publications utilise specific acoustic features to characterise snoring, whereas others just use relative or absolute sound pressure level thresholds. In light of snoring definitions changing depending on the acoustic measures employed, several acoustic analyses of snoring detection have been formulated. The current snoring detection algorithms require further evaluation in order to make more effective and comprehensive use of snore data.

In addition, snoring acoustics likely contain a wealth of physiological information regarding the sites and severity of airway collapse. There have been numerous studies to develop and investigate acoustic analysis for snoring classification over the past few years. These techniques each have their own benefits and drawbacks. The selection of the appropriate acoustic method and characterisation are likely to be important for evaluating the mechanisms and severity of upper airway obstruction during sleep that can be helpful towards developing a consensus definition of snoring. Classification results are strongly depending on underlying methods and the information embedded in snoring acoustics may not be sufficiently utilised in existing methods to meaningfully evaluate airway obstruction. There has been a growing need for improved high-performance algorithms for evaluating airway obstruction during sleep for which snoring detection and assessment is likely to be especially useful compared to other more intrusive measurement techniques otherwise needed to assess airflow and pressure changes across the airway.

3 Materials and Methodology

3.1 Setting

The study was conducted in the Flinders University Adelaide Institute for Sleep Health sound attenuated sleep laboratory (overnight background noise levels 19dB(A)). Sleep and acoustic recordings were obtained as part of a study examining the effects of wind farm compared to traffic noise exposures during sleep. 68 study participants attended the laboratory for 7 consecutive nights. The first was an acclimation night, followed in random order by 6 different noise exposure nights including 20-sec or 3-min environmental noise exposures up to 48 dBA during sleep across the night, a quiet night (control), and windfarm noise exposure at 25 dB(A) for the full night, only during wake, or only during sleep (Liebich et al. 2022). Noise in the room was monitored using a PROSIG P8004 24-bit Data Acquisition System and a GRAS 40AZ microphone mounted ~1 meter above the participant's head to monitor and record acoustic data on each night. The microphone dynamic range was 17 to 132 dB and from 0.5 Hz to 20 kHz (frequency range ± 2dB). Each recording was obtained following microphone and recording system calibration against a 94 dB(A) noise source. Participants were also instrumented for polysomnographic sleep recordings, which primarily consisted of electroencephalographic channels (EEG: C3-M2, C4-M1), electrooculograms (EOG), submental electromyogram (EMG), nasal cannula pressure, thoracic and abdominal motion and oximetry signals, all acquired using Compumedics Grael hardware and Profusion sleep recording software according to recommended sleep signal acquisition guidelines (American Academy of Sleep Medicine 2007). For the purpose of this study, the primary signals of interest were the acoustic recordings obtained from studies in which prominent snoring was audibly present over and above environmental noise exposures. The acquired signals were digitalised as CSV files from which audio signal processing was performed using PyCharm CE and MATLAB_R2022a.

3.2 Data Acquisition

The dataset used for this analysis comprised of 6 acoustic signals obtained from 6 participants during sleep. Each audio was approximately 2 hours in length and was divided into 20-second segments using a Python algorithm developed for this study to provide a total of 2330 20-sec audio segments. The average duration of a respiratory sleep cycle is around 5 seconds. Cutting the audios into 20-second segments ensured that each segment contained complete respiratory cycles and snoring events could be captured from consecutive snoring breaths.

Further analysis was performed by randomly selecting 170 segments from each participant to generate a bench-mark data set from which both manual and automated scoring could be achieved within the practical time-constraints of this study. This random selection was achieved by developing a Python algorithm to extract the audio files. All segments were stored in wav format at 2048 Hz sampling rate, 24-bit resolution, and mono channel. In total, this final bench-marking dataset comprised of 1020 20 second duration data segments.

3.3 Data Selection

Segments were classified by 1 human rater. The time points of each snoring episode in a segment were manually labelled by using Audio Labeller APP in MATLAB shown in Figure 1. Each snoring event was scored and labelled based on the confidence level (definitely snoring, probably snoring or unsure Table 1) of one human rater. The weighted confidence scores were used to sort out 4 different levels of confidence in terms of identifying snoring sounds as shown in Table 1 (Warby et al., 2014). When the audio was considered to definitely be a snore event, the time points where the snoring starts and ends wre labelled as 100%. 75% would be labelled when the rater considered the audio is probably a snoring sound. If the rater was only guessing or unsure, the audio was labelled with 50% confidence. The unlabelled sections were considered as non-snoring events.

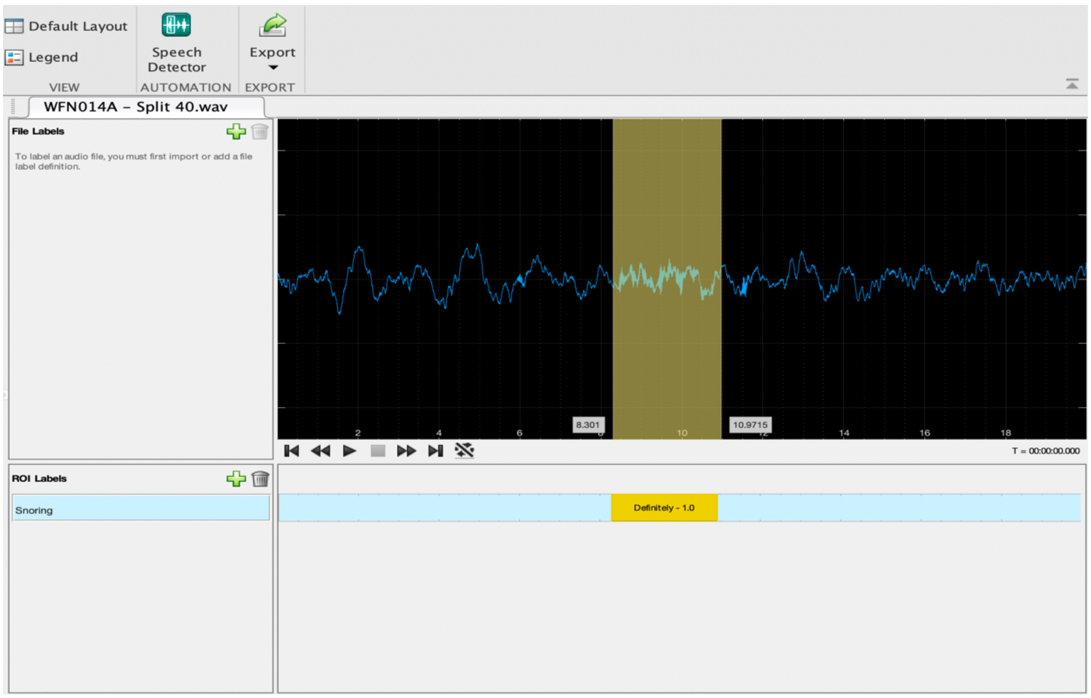

Figure 1 Audio labeler

Table 1 Scoring matrix

Scores	Confidence Level
100%	Definitely Snoring
75%	Probably Snoring
50%	Maybe/Guessing
0%	No Snoring

3.4 Methods of Characterisation

For characterisation, once snore events in each audio were collected and the benchmark dataset was created, the characteristics of the snore data were evaluated using a number of features. The features used in this study were: (1) event duration (seconds) (2) the number of snore events (3) sound frequency (Hz) (4) sound pressure level (dB).

The extraction and acoustic characterisation of the labelled snore episodes was performed using MATLAB software R2022a for 64-bit MacBook system. The snore event duration is described in seconds. The relationship between the snore event duration and the number of snore events are presented in a histogram.

The sound pressure level and frequency of the extracted snore episodes were characterised for each participant. The reference pressure level is 20 μPa as that is very close to what a normal human can hear at a sound frequency of 1000 Hz. The frequency resolution of power spectrum used for graph demonstration in this project is 20. The sound pressure level of snore events is presented as mean with the range of standard deviation for each participant. However, for participant WFN014A, the power spectrum also shows the sound pressure level of each snore event because only a few snore events were identified.

3.5 Evaluation of snoring detection algorithm

The evaluation of the snoring detection algorithm evaluated in this study was conducted by developing a code in MATLAB software R2022a for 64-bit MacBook system. The snoring detection method evaluated the relative power in the frequencies between 100 and 800 Hz. In 2012, Marshall et al. proposed that if the relative power surpassed 50% of the overall power, a snoring event would be classified (Marshall et al., 2012).

The relative power of a signal is defined as the ratio of the power of a frequency band to the total band power (Zhou et al., 2021). The formula to calculate the relative power in a power spectrum is:

$$\boldsymbol{Relative\ Power\ (RP)} = \frac{\boldsymbol{the\ area\ under\ the\ curve\ in\ a\ certain\ frequency\ band}}{\boldsymbol{the\ area\ under\ the\ curve\ in\ the\ total\ frequency\ band}} \qquad (1)$$

In this study, the area under the curve in the frequencies between 100 and 800 Hz was calculated as shown in pink in Figure 2. The total band power was obtained by computing the entire area under the curve as indicated in grey in Figure 2.

Based on this method, a MATLAB program was developed to calculate the relative power, from which confusion matrices were calculated at each relative power to determine the sensitivity and specificity of relative power to correctly classify snoring against human rater scoring at each cut-off. The plot of 1-specificity versus sensitivity was subsequently used to plot the receiver operating characteristic (ROC) curve of relative power to classify snoring. The area under the curve (AUC) is a useful measure of overall classifier performance, and from which the optimal cut-off for correct classification can be evaluated.

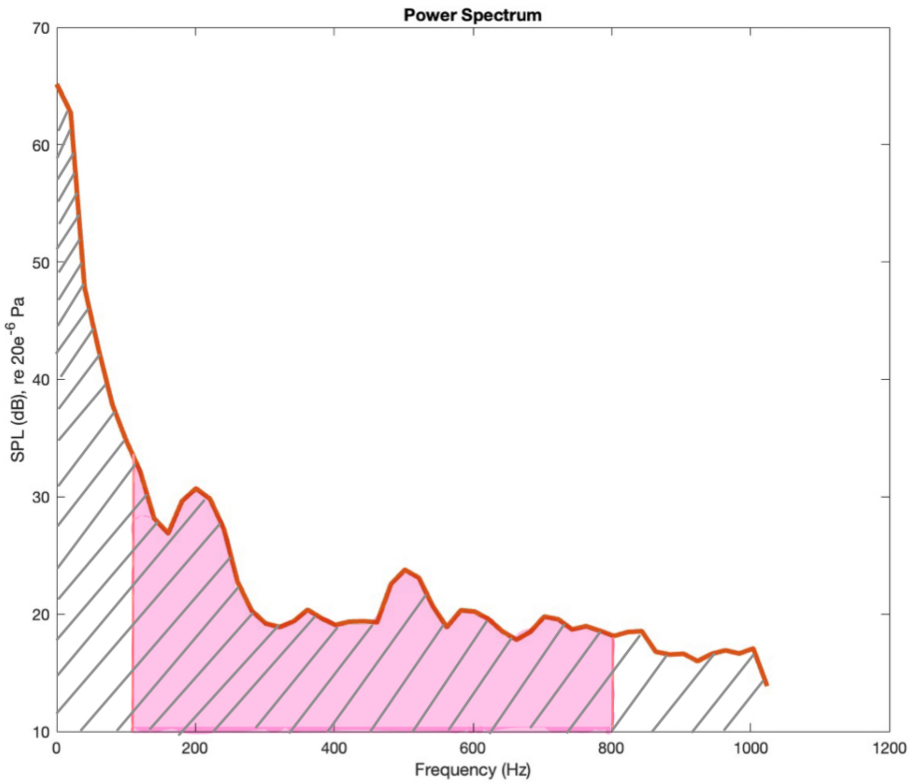

Figure 2 Power spectrum

Using a confusion matrix, the performance of a snoring detection algorithm can be visualised and summarised in a table (Singh et al., 2021). A confusion matrix is demonstrated in Table

2 where TN stands for True Negative, TP stands for True Positive, FP shows False Positive, and FN means False Negative.

Table 2 Confusion matrix

		Predicted	
Actual		Negative	Positive
	Negative	TN	FP
	Positive	FN	TP

Note. Adapted from *5 - Foundations of data imbalance and solutions for a data democracy* (pp. 83-106), by Kulkarni, A., Chong, D., & Batarseh, F. A., 2020, Academic Press. Copyright 2020 by Elsevier Inc.

The representation of the TN, FN, FP, and TP, are:

(1) True Positive (TP): TP represents the number of snoring events which have been properly detected by the snoring detection algorithm and the human rater.
(2) True Negative (TN): TN represents the number of non-snoring events that are correctly detected by the snoring detection algorithm and the human rater.
(3) False Positive (FP): FP represents the number of snoring events that were mis-detected by the snoring detection algorithm, but they are actually non-snoring events.
(4) False Negative (FN): FN represents the number of non-snoring events which were misclassified by the snoring detection algorithm, but they are actually considered as snoring events by the human rater.

Using the numbers of TP, TN, FP, and FN, the performance metrics such as accuracy, precision, recall, and F1 score can be computed (Singh et al., 2021).

The accuracy of an algorithm is defined as the ratio of correctly detected events (TP+TN) to the total number of events (TP+TN+FP+FN). The formula of accuracy is:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \tag{2}$$

The precision of an algorithm is measured by the ratio of correctly detected snoring events (TP) to the total number of correctly detected snoring and non-snoring events (TP+FP). The precision is denoted as:

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \tag{3}$$

The recall metric or sensitivity is represented as the ratio of correctly identified snoring events (TP) to the total number of the actual snoring events (TP+FN). The formula for recall metric or sensitivity is denoted as:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \tag{4}$$

The specificity is a representation of how many non-snoring events that were correctly detected. It is a ratio of correctly identified non-snoring events (TN) to the total number of the actual non-snoring events (TN+FP). The formula for specificity metric is expressed in:

$$\boldsymbol{Specificity} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}} \tag{5}$$

The F1 score of an algorithm represents the harmonic mean between recall and precision metrics. The formula to calculate the F1 score is represented as:

$$\boldsymbol{F1\ score} = \frac{\boldsymbol{2*Precision*Recall}}{\boldsymbol{Precision+Recall}} \tag{6}$$

4 Results

4.1 Results of Characterisation

The results of the data characterisation are presented in terms of the different acoustic features. Of 1020 20-second segments from 6 participants, a total of 3390 snoring events were identified. Overall, all six Participants evaluated snored at some point during the full-night sleep study. One participant (WFN029) snored almost the entire night with 574 identified snoring events from a total of 3390 identified snoring events. In contrast, WFN014A participant had mild snoring with only 8 snoring events identified from the entire 3390 identified snoring events.

4.1.1 Total Number of Snore Events and Snore Event Duration

The first characterisation was done by comparing the total number of snore events with the duration of snoring. Snore event duration distribution for the entire dataset is shown in Figure 3. As illustrated in Figure 3, the longest snore event that was detected was 18.4 seconds. The duration for most of the snore events was between 0.5 seconds and 2.2 seconds which can be seen more clearly in Figure 4. In addition, there were 40 snore events with a duration of around 1 second which was the largest number in this snore dataset.

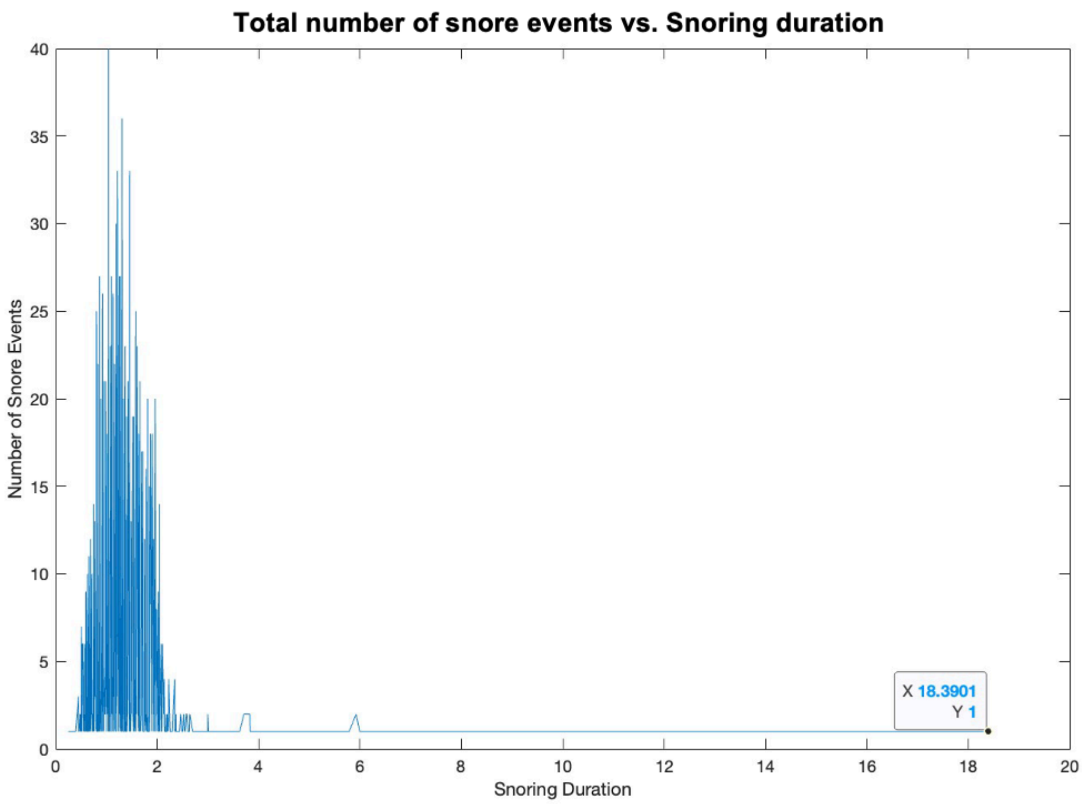

Figure 3 The number of snore events vs. event duration for the entire dataset

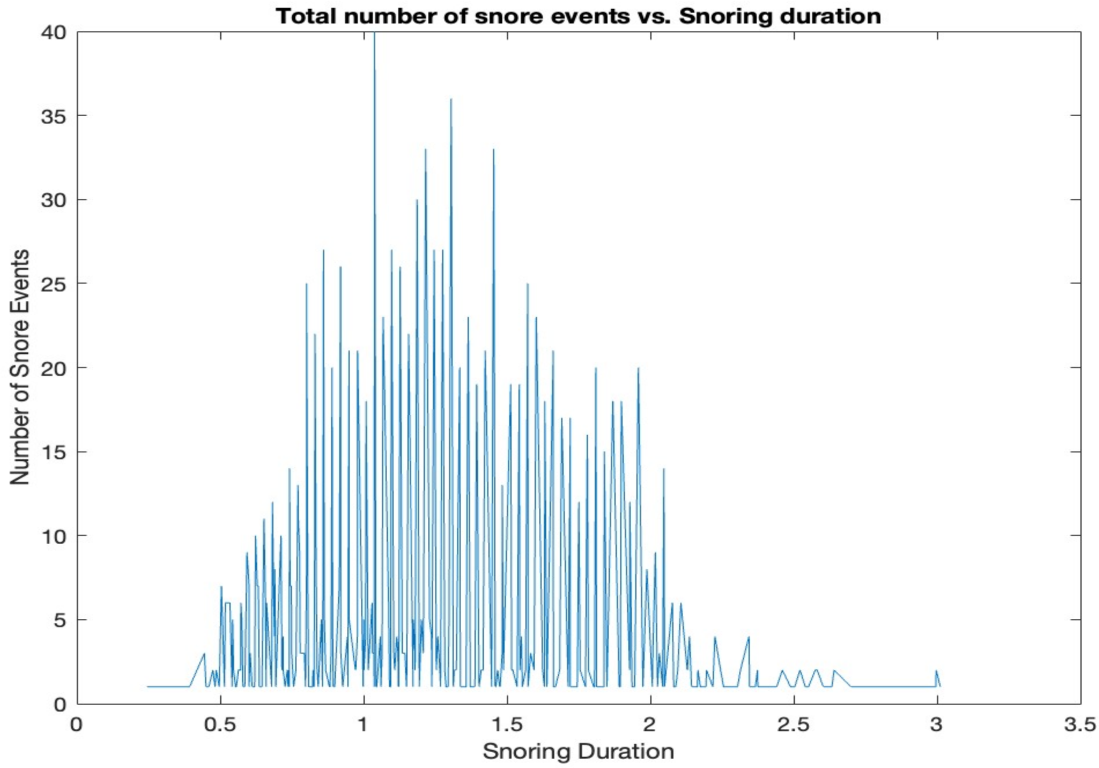

Figure 4 The number of snore events vs. snore event duration up to 3s

4.1.2 Power Spectrum

Snoring and non-snoring events can also be characterised by the power spectrum. In this study, snoring sound analysis was presented by power spectrum for each individual participant.

4.1.2.1 Participant WFN014A

Figure 5 illustrates the average sound pressure level of snores with the range of standard deviation across the entire frequency band for the participant WFN014A. As participant WFN014A only had 8 detected snore events, each power spectrum of a snoring event is also shown in Figure 5. As can be seen in Figure 5, the power spectra reveal the strongest energy of snore events in the lower frequency band. One snore event has the highest sound pressure level of 50 dB at around 100 Hz. Most of the highest sound pressure levels occur at the frequency of 200 Hz with a mean sound pressure level of 30 dB. A number of peaks can also be observed at approximately 370, 500, 600, 700, 830, and 1000 Hz with some harmonics.

The difference in power spectrum between non-snoring and snoring events is demonstrated in Figure 6. The most prominent differences can be seen in the frequency range from 180 to 290 Hz and from 370 to 830 Hz.

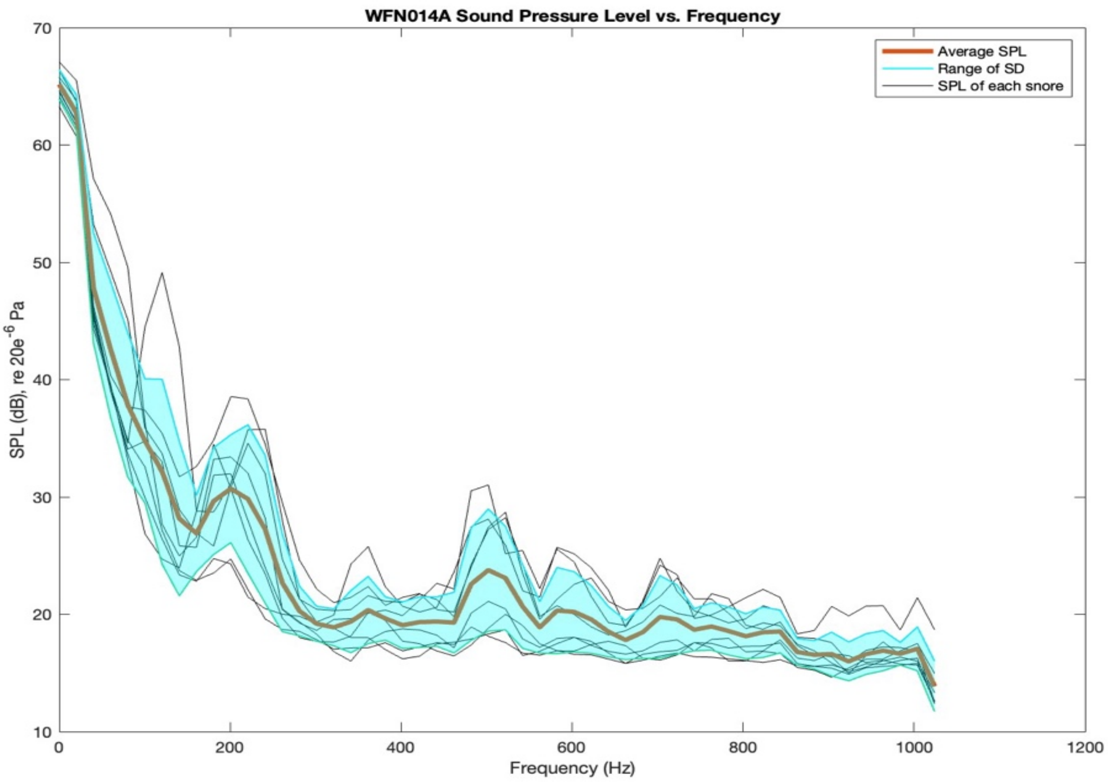

Figure 5 Average SPL with SD (WFN014A)

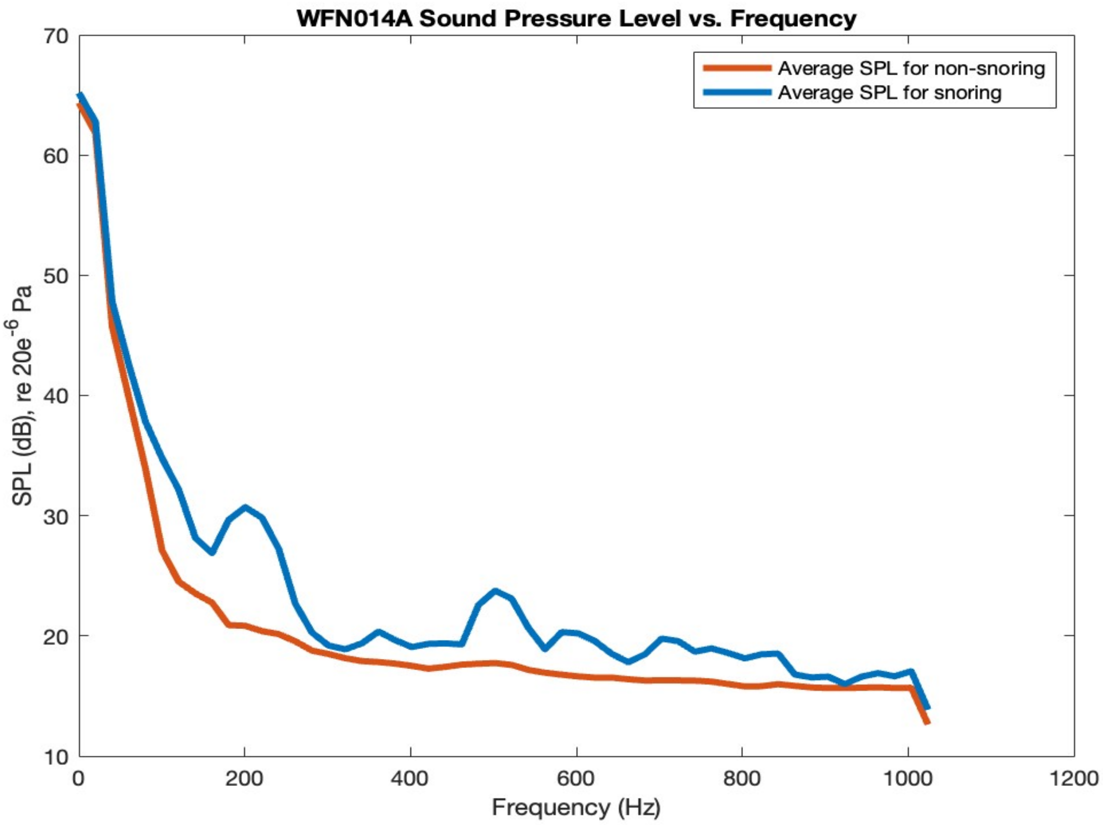

Figure 6 Average SPL for non-snoring and snoring events (WFN014A)

4.1.2.2 Participant WFN029

The average sound pressure level with a standard deviation of all the snore events identified from the participant WFN029 recording is demonstrated in Figure 7. It can be observed that the mean energy is greater than 20 dB. The strongest energy is found in the lower frequencies. The biggest peak has a sound pressure level of 45 dB at a frequency of about 200 Hz. After the big peak at 200 Hz, the average sound pressure levels fluctuate around 25 dB in the frequency range from 400 to 800 Hz. Moreover, the range of the standard deviation found in Figure 7 is larger compared to the standard deviation found in other participants, especially in the frequency band between 320 and 1020 Hz.

On the other hand, a comparison between the sound pressure levels of snoring and non-snoring events is shown in Figure 8. From the frequencies between 300 and 850 Hz, the difference in sound pressure level between snoring and non-snoring events is around 10 dB. Furthermore, the largest differences appeared to occur within the frequency ranging from 100 to 300 Hz and from 850 to 900 Hz.

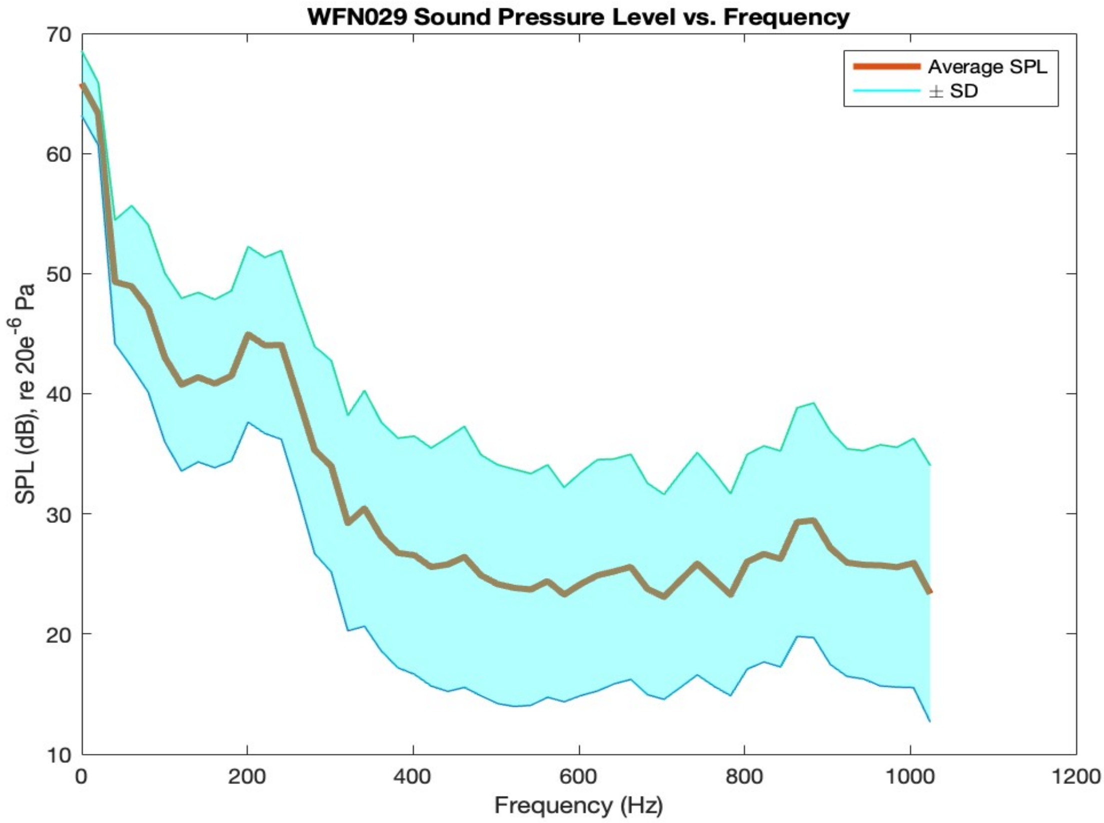

Figure 7 Average SPL with SD (WFN029)

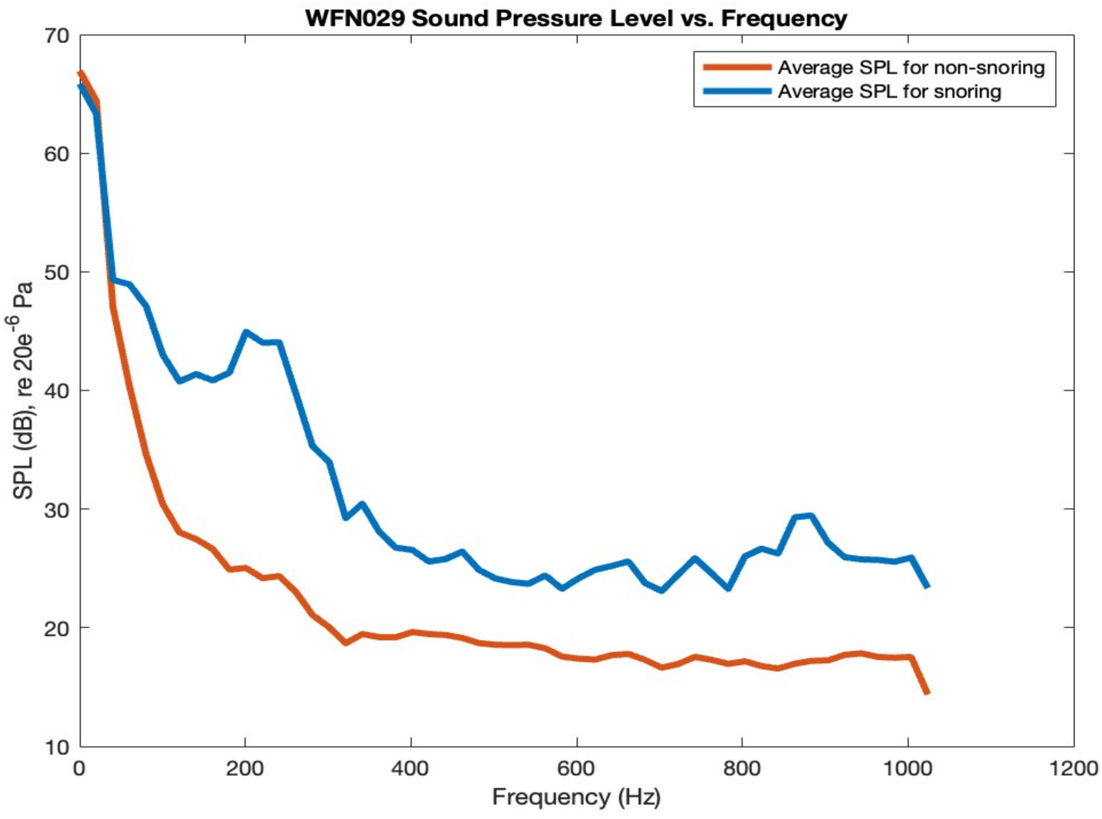

Figure 8 Average SPL for non-snoring and snoring events (WFN029)

4.1.2.3 *Participant WFN054A*

As can be seen in Figure 9, the average sound pressure level of snoring in participant WFN054A is larger than about 25 dB in the entire frequency band. This power spectrum shows that the strongest energy is found at 200 Hz with around 48 dB sound pressure level. There are several discernible peaks at 350, 440, 520, 620, 780, and 1000 Hz. The standard deviation is around 5 dB for all the frequencies.

The power spectra of the snoring and non-snoring events can be observed in Figure 10. The power spectrum for snoring events has several identifiable peaks whereas the power spectrum shows the most prominent peak at around 200 Hz.

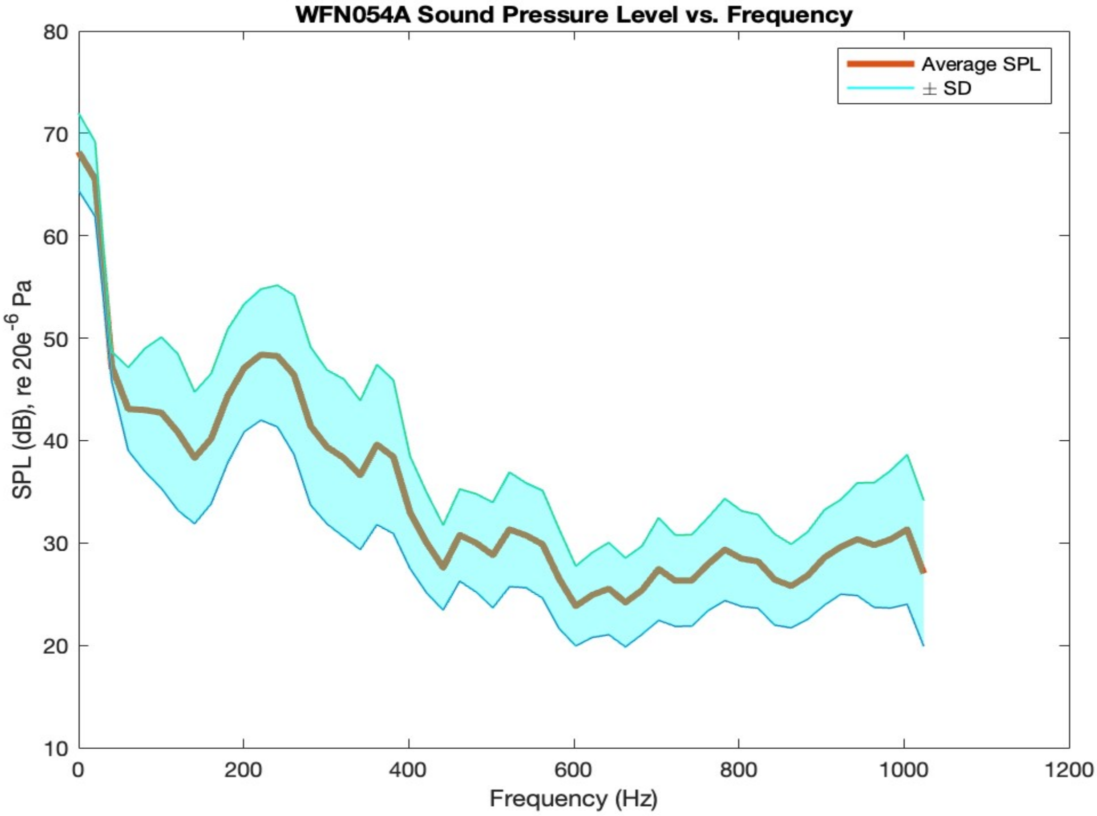

Figure 9 Average SPL with SD (WFN054A)

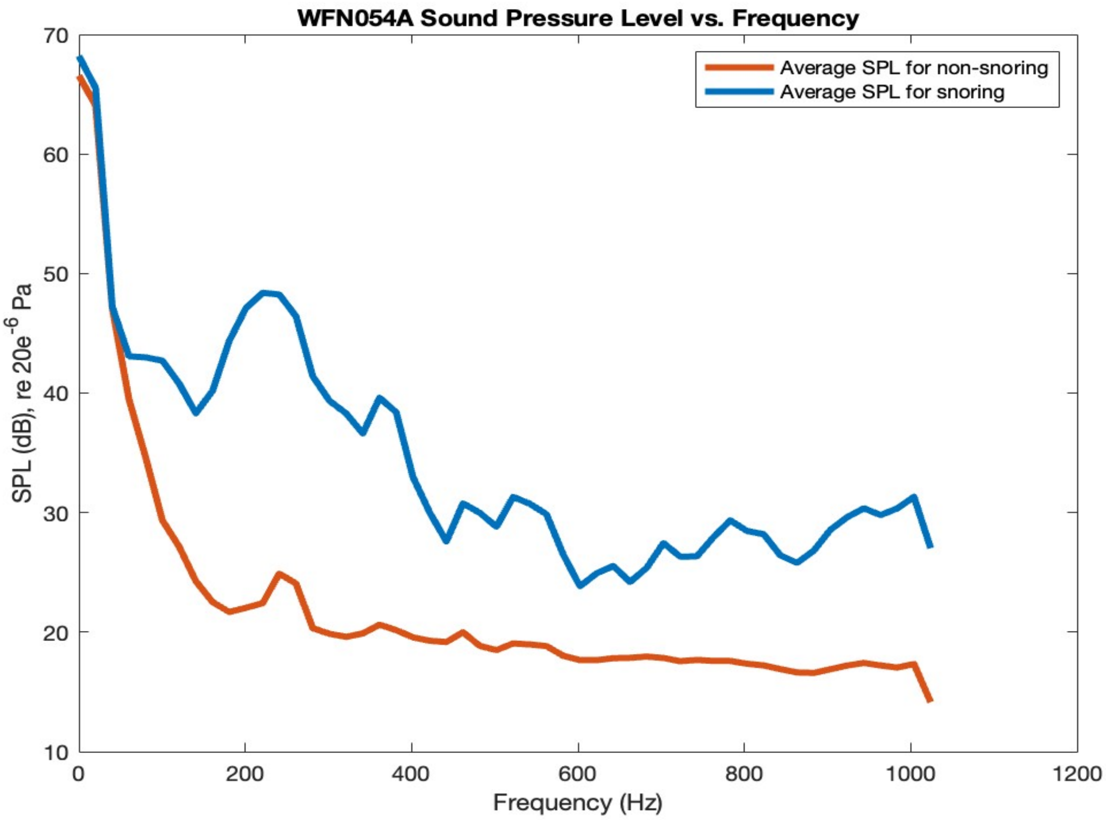

Figure 10 Average SPL for non-snoring and snoring events (WFN054A)

4.1.2.4 Participant WFN054D

Figure 11 shows the average sound pressure level of snoring events of participant WFN054D as well as the range of standard deviation. It is noticeable that the stronger energy was found in the lower frequency band. Similar to other participants, the highest peak that can be seen is at around 200 Hz with peak SPL around 50 dB, followed by several peaks at 350, 500, 700, 800, and 1000 Hz. Moreover, the average sound pressure level is larger than 20 dB for all frequencies.

In Figure 12, the most prominent differences between non-snoring and snoring were in the frequencies between 100 and 400 Hz, and between 850 and 1000 Hz. In the frequencies from 400 to 850 Hz, the difference in sound pressure level of snoring and non-snoring events was around 5 dB.

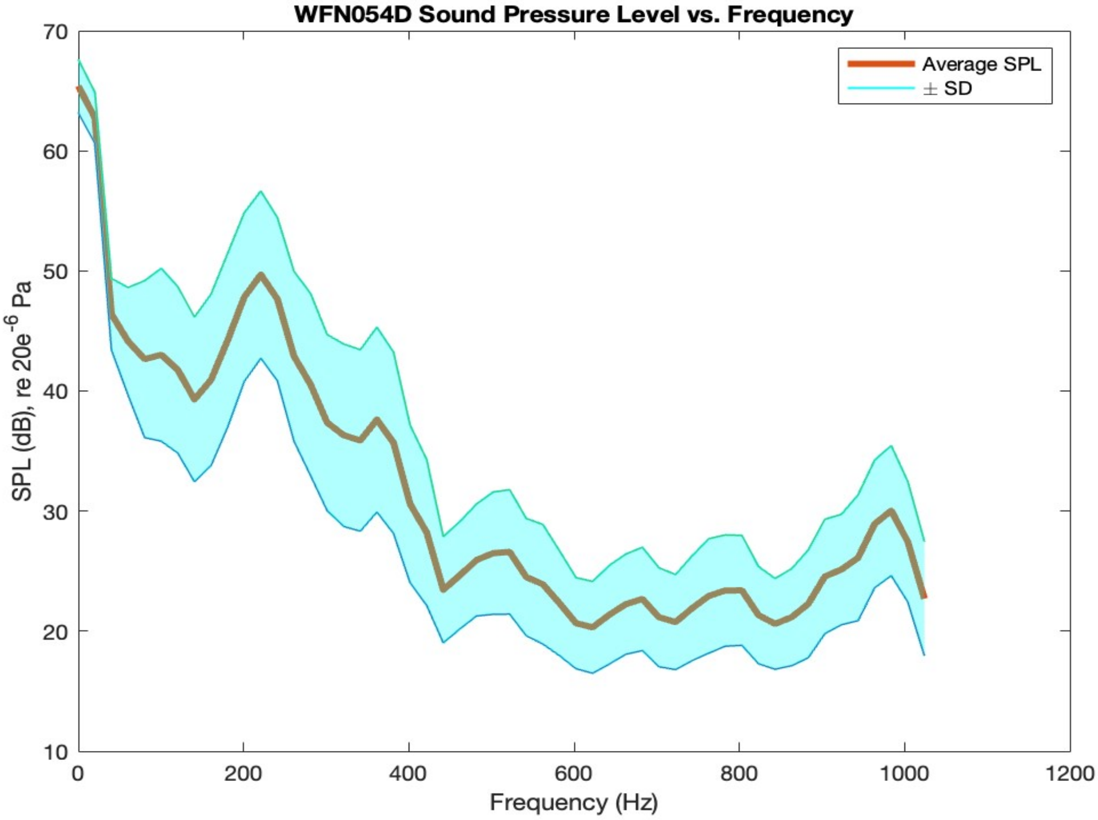

Figure 11 Average SPL with SD (WFN054D)

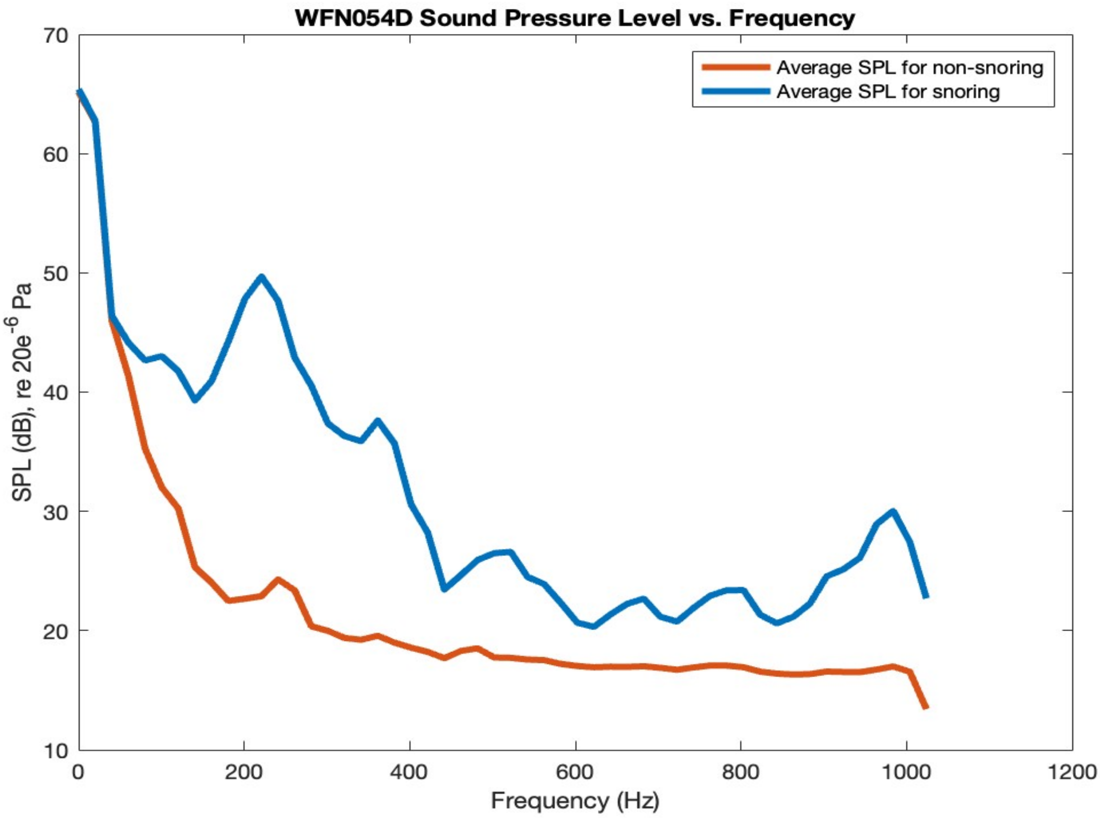

Figure 12 Average SPL for non-snoring and snoring events (WFN054D)

4.1.2.5 *Participant WFN077A*

Figure 13 demonstrates the average sound pressure level with the range of standard deviation of participant WFN077A. Most of the snore events were found to have a sound pressure level of 10 to 40 dB. However, the strongest energy sound observed was at 100 Hz, lower than in other participants in this study. The average sound pressure level peaks at 200, 450, 700 and 1000 Hz. Interestingly, a significant drop can be observed at the frequency of 800 Hz.

From Figure 14, the difference in sound pressure levels between snoring and non-snoring events can be seen. Compared to other participants, the difference in the sound pressure levels appears to be larger and nearly across the entire frequency band with around a 20 dB difference.

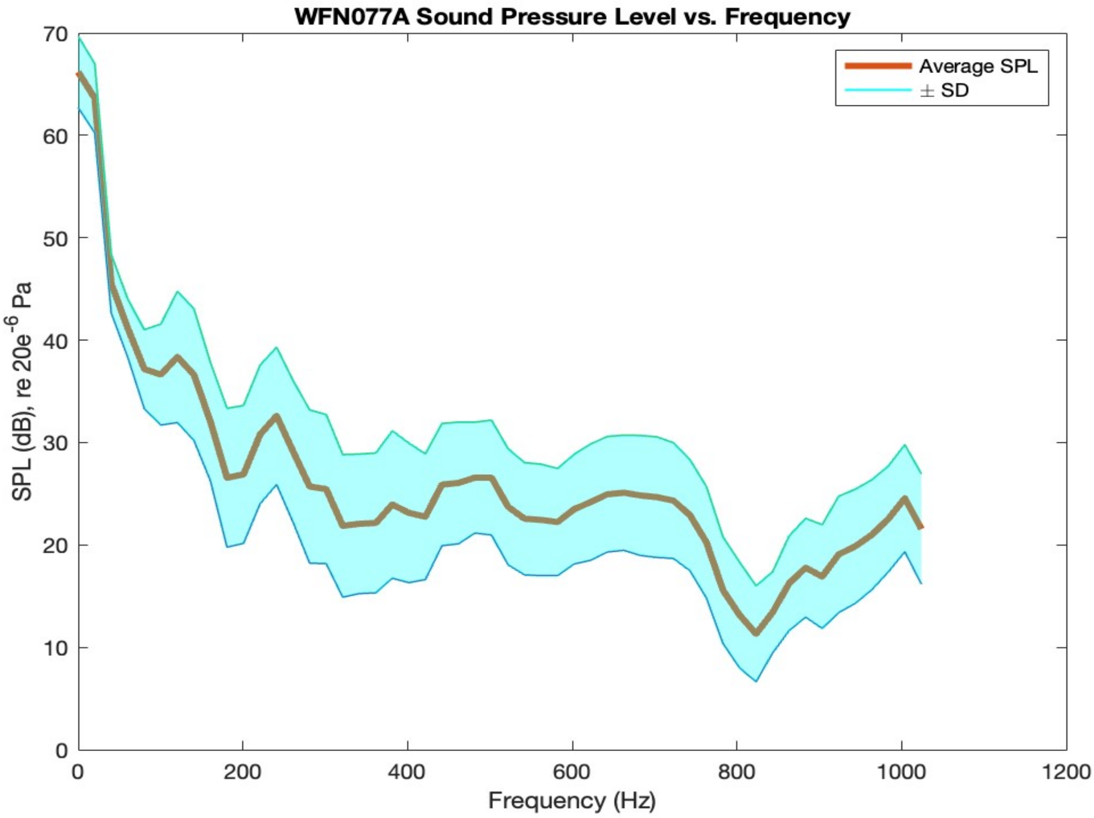

Figure 13 Average SPL with SD (WFN077A)

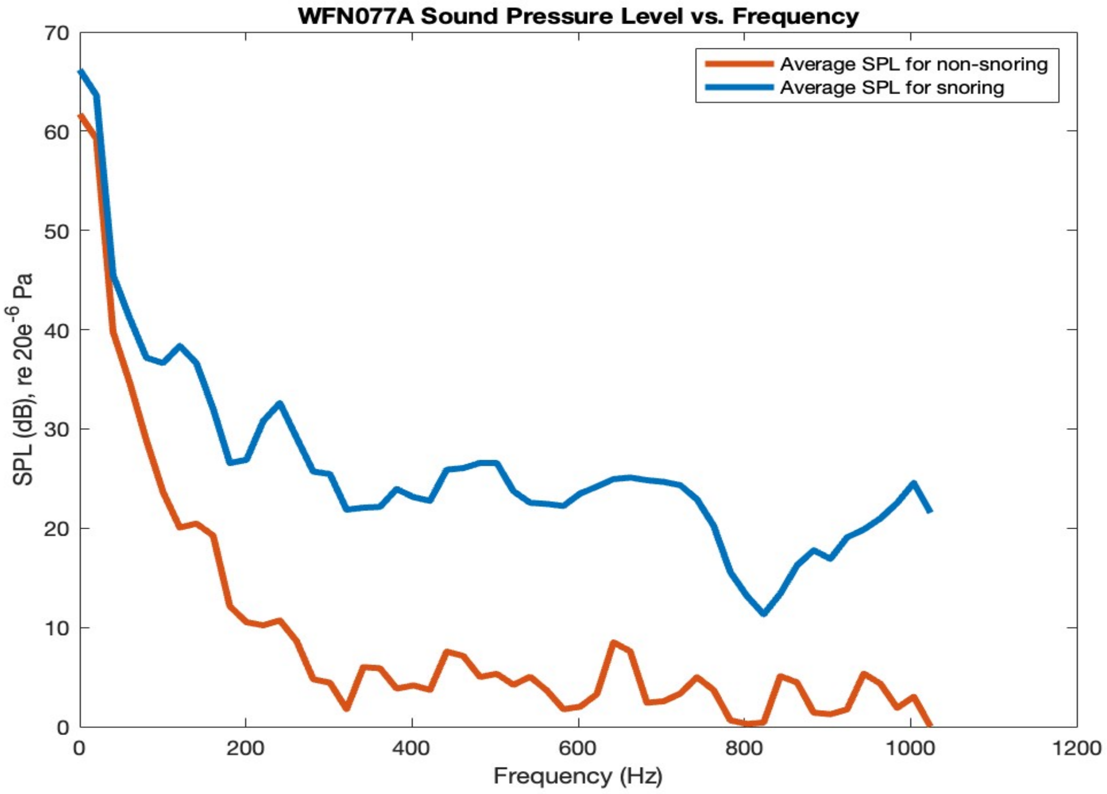

Figure 14 Average SPL for non-snoring and snoring events (WFN077A)

4.1.2.6 Participant WFN0151A

Figure 15 illustrates the average power spectrum of participant WFN0151A's snores. The most intense energy was found in the snoring spectrum between 100 and 400 Hz. In addition, there were a number of discernible energy peaks at 200, 450, 600, 780, and 950 Hz. For all frequencies, the energy of the snore was greater than 20 dB.

Figure 16 exhibited pronounced variance in sound pressure levels between snoring and non-snoring events. As demonstrated, the prominent difference was found at the frequency range from 100 to 380 Hz.

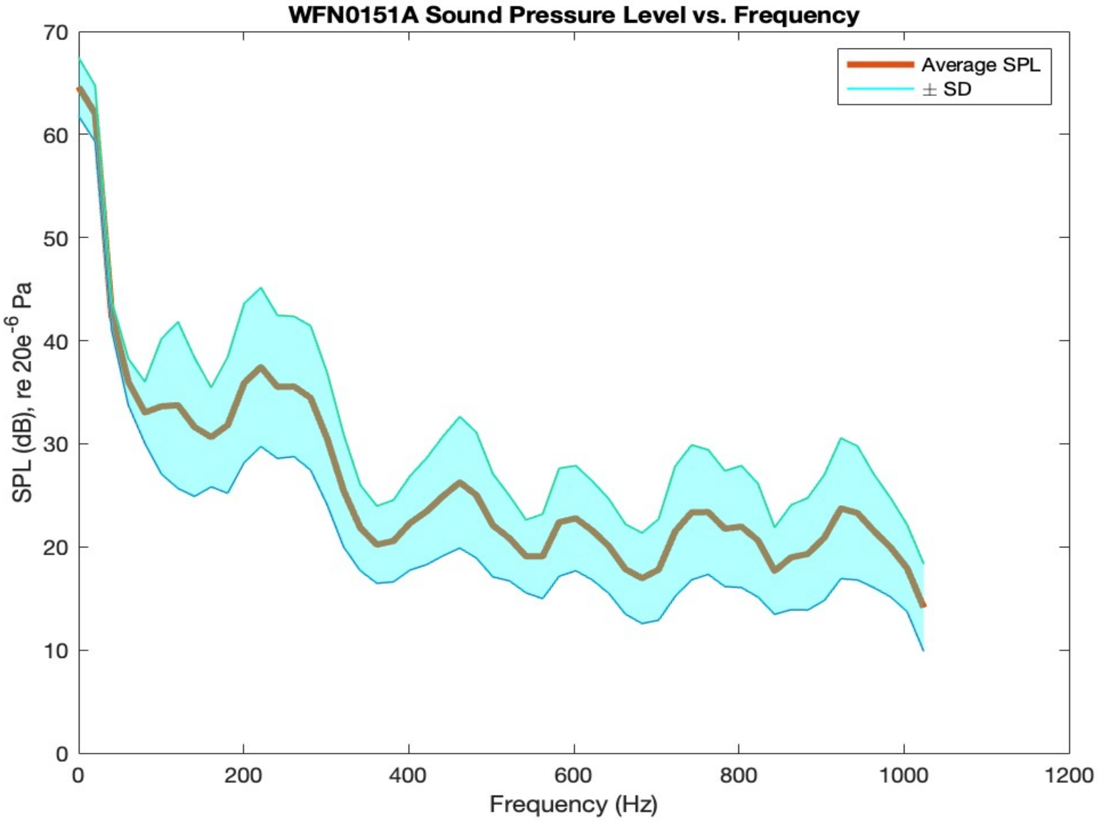

Figure 15 Average SPL with SD (WFN0151A)

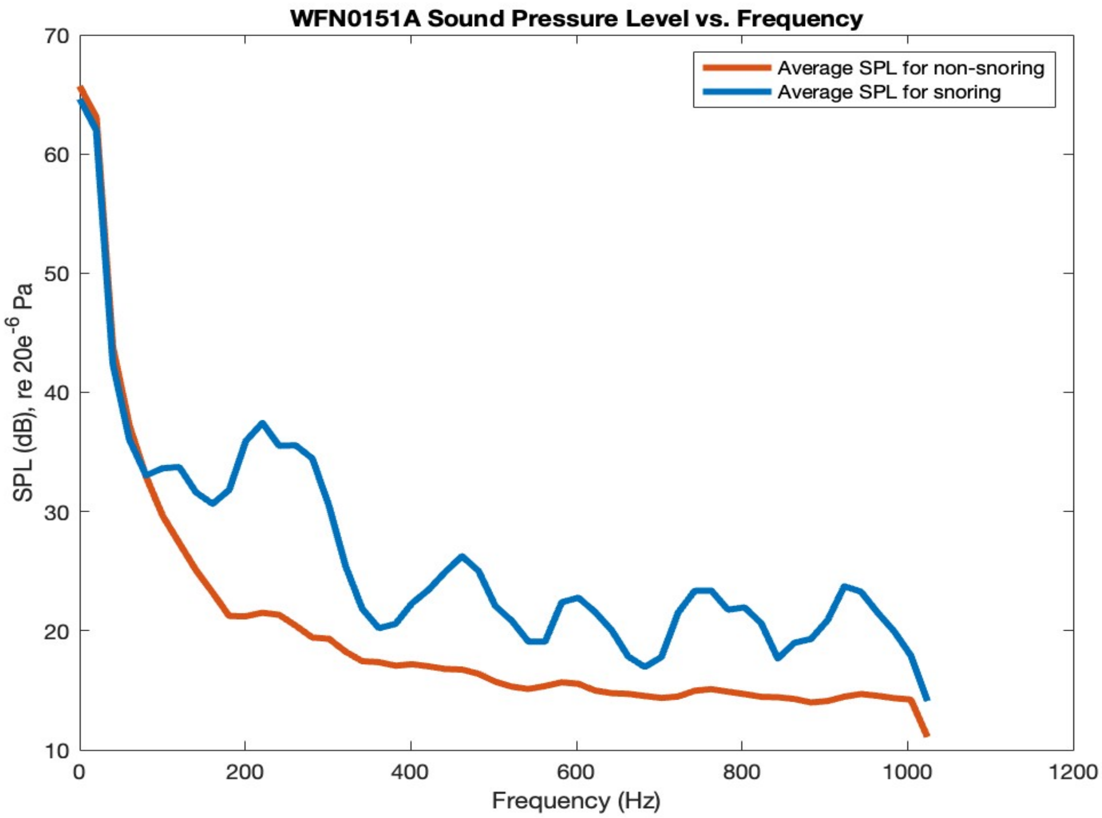

Figure 16 Average SPL for non-snoring and snoring events (WFN0151A)

4.2 Results of Snoring Detection Evaluation

4.2.1 ROC Curve

Initially, the evaluation was conducted using the relative power of 0.5 in the frequencies between 100 and 800 Hz and classifying the snoring sound if the relative power surpassed 50% of the overall power (Marshall et al., 2012). However, most of the values were greater than 0.5 (50%) due to the different sampling rates used in this study. Therefore, ROC analysis was employed to help evaluate if 0.5 or a somewhat different cut-off value may be optimal for this study.

As can be observed in Figure 17 and Table 3, the ROC curves of participants WFN054A, WFN054D, WFN077A, and WFN0151A demonstrate high sensitivity and specificity, with high AUC indicating that the snoring detection algorithm performs very well compared to human scoring in the snore dataset in this study.

Compared to the data of the other participants, the ROC curves of participants WFN014A and WFN29 show the performance of the snoring detection test was less optimal although specificity and sensitivity were still high above 90%.

Overall, the average values of the area under the curve (AUC), sensitivity and specificity were 0.942, 0.925, and 0.899 respectively as illustrated in Table 3. The optimal cut-off value (closest to the top-left corner of the ROC curve) indicated in the ROC analysis was 0.620. Therefore, the relative power used for snoring detection in this study was 0.620 instead of 0.5 as suggested by Marshall et al., 2012.

Table 3 ROC curve analysis

Participant ID	Area Under Curve	Specificity	1-Specificty	Sensitivity
WFN014A	0.927	0.864	0.136	0.875
WFN0151A	0.936	0.887	0.113	0.973
WFN029	0.901	0.857	0.143	0.889
WFN054A	0.965	0.924	0.076	0.946
WFN054D	0.987	0.969	0.031	0.939
WFN077A	0.934	0.893	0.107	0.928
Mean	0.942	0.899	0.101	0.925

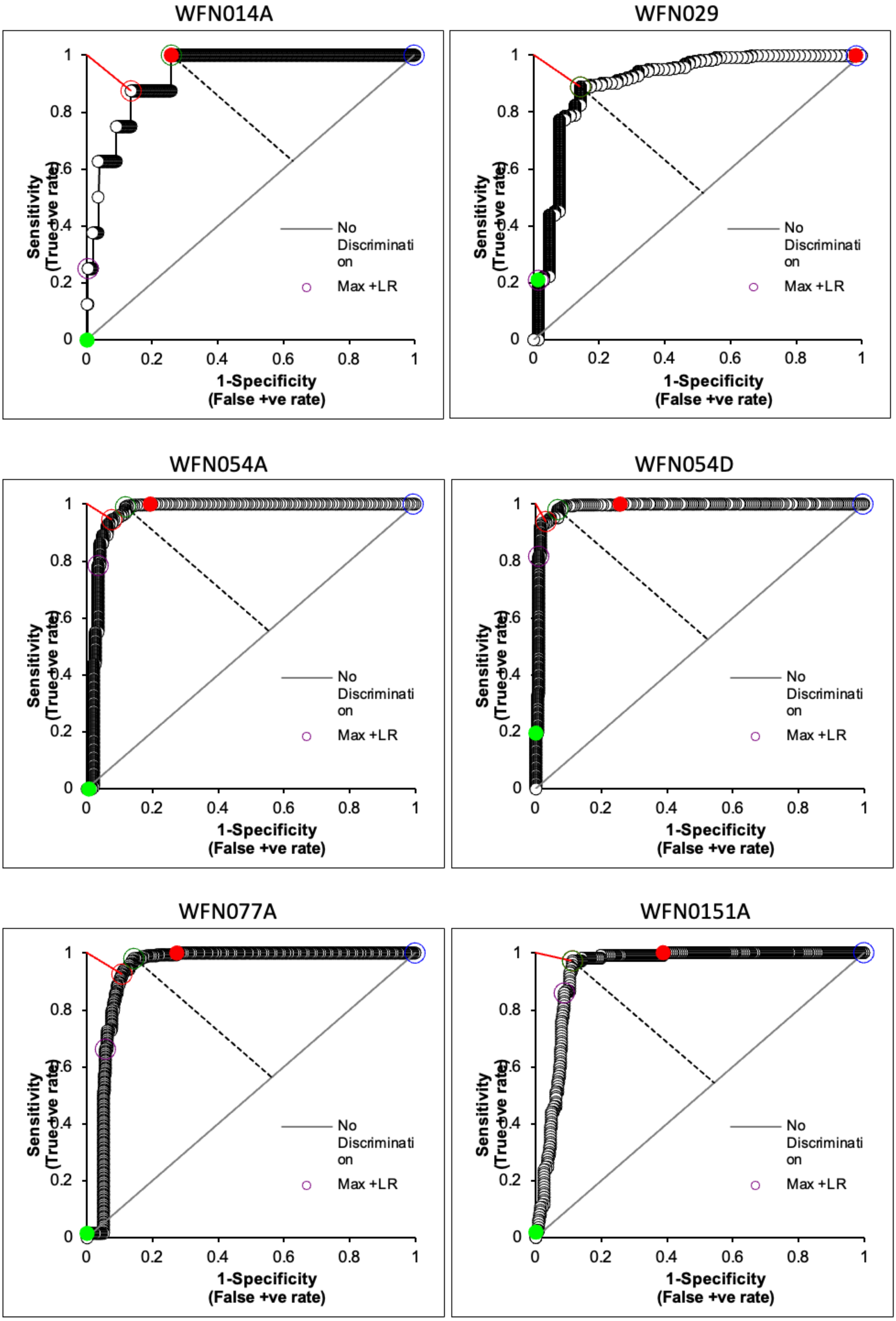

Figure 17 ROC curves of all participants

4.2.2 Confusion Matrix

The confusion matrix was utilised to compare the benchmark dataset with the results of deploying the snoring detection algorithm. With the application of the cut-off value of 0.62, the snoring detection algorithm shows great performance on the snore data as shown in Table 4 with a total true positive count value of 1535 and the true negative value of 1564.

Table 4 Confusion matrix of this study

		Predicted	
	Total samples (n) = 3390	Positive (snore)	Negative (non-snore)
Actual	Positive (snore)	TP = 1535	FN = 135
	Negative (non-snore)	FP = 156	TN = 1564

From the confusion matrix, the values of accuracy, precision, recall, and F1 score can be computed. The calculation can be seen as follows:

$$Accuracy = \frac{1535+1564}{1535+156+135+1564} = \frac{3099}{3390} = \mathbf{0.9142} \quad (7)$$

$$Precision = \frac{1535}{1535+156} = \frac{1535}{1691} = \mathbf{0.9077} \quad (8)$$

$$Recall = \frac{1535}{1535+135} = \frac{1535}{1670} = \mathbf{0.9192} \quad (9)$$

$$Specificity = \frac{1564}{1564+156} = \frac{1564}{1720} = \mathbf{0.9093} \quad (10)$$

$$F1\ score = \frac{2*0.9077*0.9192}{0.9077+0.9192} = \mathbf{0.9134} \quad (11)$$

Accuracy is a metric that can be used to assess the performance of the classification against a comparator gold-standard, in this case human scoring. In this study, the accuracy was 0.914 which indicates that snoring detection has high performance in classifying snoring and non-snoring data. Moreover, the recall metric calculated from the confusion matrix is 0.919 which shows that the snoring detection algorithm can correctly identify 91.9% of snore events.

In addition, the precision of the snoring detection algorithm was also high at 90.8% as was the F1 score (0.913) indicating high precision and recall.

5 Conclusion and Discussion

5.1 Discussion

This study was carried out with the aim of further investigating the acoustic features of snoring sounds and evaluating existing acoustic methods of snoring detection. Several techniques were used in this study to objectively evaluate the performance of a snoring detection algorithm. The definition of snoring is a crucial requirement to effectively develop algorithms that are able to reliably detect snoring episodes during sleep using acoustic methods. Nevertheless, no unequivocal definition yet exists for this acoustic phenomenon. Due to the lack of an unequivocal definition of snoring, different scholars have defined snoring sounds in various ways. In general, snores are regarded as relatively high frequency acoustic phenomena with sound intensity greater than a specific amplitude value. Some researchers only analysed a few characteristic snore events without looking at the full-study sleep cycle whereas others characterised snoring by the mechanisms of sound generation with the use of speech analysis.

The approach used in this study includes generating an objective benchmark dataset and using comprehensive sets of metrics to assess the performance of the snoring detection algorithm. It was found that the performance of the snoring detection algorithm may have been affected by the sampling rate when the relative power in the frequencies between 100 and 800 Hz was deployed since the best-performing cut-off relative power for this snore dataset was 0.62 instead of 0.5 which was stated in Marshall et al.'s research in 2012. The reason for the difference in the cut-off relative power value is that other recording differences such as different sampling ratees, type of microphones used, or recording setup could potentially have played a role in the detecting snoring events using this algorithm.

One of the findings in this study showed that snoring occurs in the lower frequency band, at around 200 Hz. When comparing the snoring and non-snoring events, prominent differences were revealed at frequencies between around 200 Hz and 1000 Hz. Interestingly, those participants with a larger number of snoring events identified appeared to have more fluctuations in both non-snoring and snoring power spectra. Furthermore, the sound pressure levels of non-snoring and snoring events also appeared to have larger differences in power spectra.

Another interesting finding was that small peaks were observed in the power spectra of non-snoring events for the participants WFN054A and WFN054D. These small peaks might suggest the possibility of the misclassification caused by the manual scoring.

5.1.1 Limitation

A clear limitation of this study is the small sample size. Although there were a total of 1020 20-second segments analysed for the experiments, these segments were extracted from only 6 participants. Consequently, a larger sample sizes of snoring and non-snoring participants are required in future studies to help evaluate the generalisability of the snoring detection algorithm in this study.

Another limitation of this study is that the method used to create a benchmark dataset remained subjective as it is based on the judgement of one human rater. Even though random extraction of the acoustic segments was applied to help ensure objectivity, the snoring identification for the benchmark dataset was still generated by listening to all the 20-second segments and based on the human rater's confidence level and perception of snoring. Furthermore, snoring classifications were determined by the assessment of a single rater. Increasing the number of raters would clearly be useful towards establishing consensus human scoring and for evaluating the level of algorithm versus human individual and consensus scoring agreement.

In addition, the placement of microphones is also a limitation of this study as the microphone position may affect the results of the snoring detection algorithm evaluation. The relative power is calculated from the sound pressure level which is determined by the distance and the position of the microphone to the sound source. As people may move during sleep, different sleep positions may skew results. However, this may be an issue regardless of microphone positioning, so standardisation applied in this study may ultimately be the most reasonable and practical approach.

5.1.2 Future Work

Further investigation of the snoring acoustics should be conducted with a larger sample size to allow a more accurate assessment of the snoring detection algorithm. The representativeness of larger sample size can be easier to evaluate and generalise the findings. Another work that can be made in the future to improve the study is to increase the number of human raters. Having more raters in the study allows the study to be more objective.

This study can also be improved by evaluating different snoring detection algorithms and comparing their performance in terms of classifying snoring and non-snoring events. Since there is no unequivocal definition of snoring to date, assessing the performance of different snoring detection algorithms will give more information about snoring acoustics.

5.2 Conclusion

In this study, an investigation of snoring acoustics was conducted by generating a benchtop dataset and evaluating the snoring detection algorithm. From the results obtained, the following conclusions can be drawn:

1. The duration of snoring events was around 0.5-2.2-second long.
2. Most non-snoring sounds are less than 30 dB.
3. The most prominent differences were revealed at a frequency of around 200 Hz.
4. Using the relative power of 0.5 in the frequencies between 100 and 800 Hz to detect snoring events can be affected by the sampling rate.
5. With the cut-off relative power value of 0.62, the performance of the snoring detection algorithm has a better performance on the snore dataset in this study.

The findings obtained in this study may facilitate finding an objective definition of snoring using the acoustic features of snoring. This work can also contribute towards the development of snoring detection algorithms. Further studies are required to support the findings by increasing the number of human raters and recruiting more participants.

6 References

American Academy of Sleep Medicine. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules. Terminology and Technical Specification. 1st ed.* (C. Iber, S. Ancoli-Israel, A. L. Chesson, & S. F. Quan, Eds.). American Academy of Sleep Medicine – Association for Sleep Clinicians and Researchers. https://aasm.org/clinical-resources/scoring-manual/

American Academy of Sleep Medicine. (2014). *The International Classification of Sleep Disorders - Third Edition*. Darien, IL : American Academy of Sleep Medicine.

Better Health Channel. (2014, August 30). *Snoring - Better Health Channel*. Www.betterhealth.vic.gov.au. https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/snoring

Çavuşoğlu, M., Poets, C. F., & Urschitz, M. S. (2017). Acoustics of snoring and automatic snore sound detection in children. *Physiological Measurement*, *38*(11), 1919–1938. https://doi.org/10.1088/1361-6579/aa8a39

Dalmasso, F., & Prota, R. (1996). Snoring: analysis, measurement, clinical implications and applications. *European Respiratory Journal*, *9*(1), 146–159. https://doi.org/10.1183/09031936.96.09010146

Fiz, J. A., Abad, J., Jane, R., Riera, M., Mananas, M. A., Caminal, P., Rodenstein, D., & Morera, J. (1996). Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea. *European Respiratory Journal*, *9*(11), 2365–2370. https://erj.ersjournals.com/content/9/11/2365.short

Furukawa, T., Nakano, H., Yoshihara, K., & Sudo, N. (2016). The Relationship between Snoring Sound Intensity and Morning Blood Pressure in Workers. *Journal of Clinical Sleep Medicine*, *12*(12), 1601–1606. https://doi.org/10.5664/jcsm.6340

Godman, H. (2022, February 11). *Is Snoring Bad?* U.S. News. https://health.usnews.com/wellness/articles/snoring

Janott, C., Rohrmeier, C., Schmitt, M., Hemmert, W., & Schuller, B. (2019, July 1). *Snoring - An Acoustic Definition*. IEEE Xplore. https://doi.org/10.1109/EMBC.2019.8856615

Kayabekir, M., & Yağanoğlu, M. (2021). The relationship between snoring sounds and EEG signals on polysomnography. *Sleep and Breathing*, *26*(3), 1219–1226. https://doi.org/10.1007/s11325-021-02516-8

Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). 5 - Foundations of data imbalance and solutions for a data democracy. In F. A. Batarseh & R. Yang (Eds.), *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering* (pp. 83–106). Academic Press. https://www.sciencedirect.com/science/article/pii/B9780128183663000058

Lee, H.-K., Lee, J., Kim, H., Ha, J.-Y., & Lee, K.-J. (2013). Snoring detection using a piezo snoring sensor based on hidden Markov models. *Physiological Measurement*, *34*(5), N41–N49. https://doi.org/10.1088/0967-3334/34/5/n41

Lee, J. (2022, March 26). *25 Alarming Sleep Apnea & Snoring Statistics Australia Needs to Know (2022) | WhatASleep*. WhatASleep! https://whatasleep.com.au/blog/sleep-apnea-snoring-statistics-australia/

Liebich, T., Lack, L., Hansen, K., Zajamsek, B., Micic, G., Lechat, B., Dunbar, C., Nguyen, D. P., Scott, H., & Catcheside, P. (2022). An experimental investigation on the impact of wind turbine noise on polysomnography-measured and sleep diary-determined sleep outcomes. *Sleep*, *45*(8), zsac085. https://doi.org/10.1093/sleep/zsac085

Markandeya, M. N., Abeyratne, U. R., & Hukins, C. (2018). Characterisation of upper airway obstructions using wide-band snoring sounds. *Biomedical Signal Processing and Control*, *46*, 201–211. https://doi.org/10.1016/j.bspc.2018.07.013

Marshall, N. S., Wong, K. K. H., Cullen, S. R. J., Knuiman, M. W., & Grunstein, R. R. (2012). Snoring Is Not Associated With All-Cause Mortality, Incident Cardiovascular

Disease, or Stroke in the Busselton Health Study. *Sleep*, *35*(9), 1235–1240. https://doi.org/10.5665/sleep.2076

Pevernagie, D., Aarts, R. M., & De Meyer, M. (2010). The acoustics of snoring. *Sleep Medicine Reviews*, *14*(2), 131–144. https://doi.org/10.1016/j.smrv.2009.06.002

Rohrmeier, C., Herzog, M., Ettl, T., & Kuehnel, T. S. (2013). Distinguishing snoring sounds from breath sounds: a straightforward matter? *Sleep and Breathing*, *18*(1), 169–176. https://doi.org/10.1007/s11325-013-0866-8

Section on Pediatric Pulmonology, & Subcommittee on Obstructive Sleep Apnea Syndrome. (2002). Clinical Practice Guideline: Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome. *PEDIATRICS*, *109*(4), 704–712. https://doi.org/10.1542/peds.109.4.704

Shin, H., & Cho, J. (2014). Unconstrained snoring detection using a smartphone during ordinary sleep. *BioMedical Engineering OnLine*, *13*(1), 116. https://doi.org/10.1186/1475-925x-13-116

Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Chapter 5 - Diagnosing of disease using machine learning. In K. K. Singh, M. Elhoseny, A. Singh, & A. A. Elngar (Eds.), *Machine Learning and the Internet of Medical Things in Healthcare* (pp. 89–111). Academic Press. https://www.sciencedirect.com/science/article/pii/B9780128212295000033

Sleep Disorders Australia. (2020). *SNORING WHAT IS SNORING?* https://www.sleepoz.org.au/_files/ugd/a1218b_523739e8468e45159e34c98f5187c4b3.pdf

Sogebi, O. A., Oyewole, E. A., & Olusoga-Peters, O. O. (2011). Sleep disordered breathing (SDB) experiences associated with snoring in adult Nigerians. *African Health Sciences*, *11*(3), 309–314.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261016/#:~:text=Habitual%20snoring%20is%20one%20of

Sola-Soler, J., Fiz, J. A., Morera, J., & Jane, R. (2011). Bayes classification of snoring subjects with and without Sleep Apnea Hypopnea Syndrome, using a Kernel method. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6071-6074. https://doi.org/10.1109/iembs.2011.6091500

Solà-Soler, J., Fiz, J. A., Morera, J., & Jané, R. (2012). Multiclass classification of subjects with sleep apnoea–hypopnoea syndrome through snoring analysis. *Medical Engineering & Physics*, *34*(9), 1213–1220. https://doi.org/10.1016/j.medengphy.2011.12.008

Stuck, B. A., Dreher, A., Heiser, C., Herzog, M., Kühnel, T., Maurer, J. T., Pistner, H., Sitter, H., Steffen, A., & Verse, T. (2013). S2k-Leitlinie „Diagnostik und Therapie des Schnarchens des Erwachsenen“. *HNO*, *61*(11), 944–957. https://doi.org/10.1007/s00106-013-2775-3

Swarnkar, V. R., Abeyratne, U. R., & Sharan, R. V. (2017). Automatic picking of snore events from overnight breath sound recordings. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2822-2825. https://doi.org/10.1109/embc.2017.8037444

Wang, C., & Peng, J. (2017). The Methods of Acoustical Analysis of Snoring for the Diagnosis of OSAHS. *J Sleep Med Disord*, *4*(2), 1080.

Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G. S., Carrillo, O., Sorensen, H. B. D., Jennum, P., Peppard, P. E., Perona, P., & Mignot, E. (2014). Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nature Methods*, *11*(4), 385–392. https://doi.org/10.1038/nmeth.2855

Wilson, K., Stoohs, R. A., Mulrooney, T. F., Johnson, L. J., Guilleminault, C., & Huang, Z. (1999). The Snoring Spectrum. *Chest*, *115*(3), 762–770. https://doi.org/10.1378/chest.115.3.762

Zhou, Q., Lin, J., Yao, L., Wang, Y., Han, Y., & Xu, K. (2021). Relative Power Correlates With the Decoding Performance of Motor Imagery Both Across Time and Subjects. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.701091

7 Appendices

7.1 Appendix 1: MATLAB Code

```
%% count the total number of snore events

clc
num_snore = 0;
for i = 1:170
    % check if it's an empty cell
    data_i = labelData.Labels.snoring{i,1};
    % if the cell contains data

    if ~isempty(data_i)
    ROI_i = labelData.Labels.Snoring{i,1}.Value; % extract the labelled audios

        for j = 1:length(ROI_i)
            if labelData.Labels.Snoring{i,1}.Value(j) == 'Definitely =1.0'
                num_snore = num_snore + 1;
            end

        end
    end
end

%% create a time array (duration) for each participant

clc
num_snore = 0;
WFN014A_time_array = [0 0];

% extract all the definite snore events
```

```
for i = 1:170
    % check if it's an empty cell
    data_i = WFN014A.Labels.Snoring{i,1};

    % if the cell contains data
    if ~isempty(data_i)
    % extract all the snore events
    ROI_i = WFN014A.Labels.Snoring{i,1}.Value;

        for j = 1:length(ROI_i)
            if  WFN014A.Labels.Snoring{i,1}.Value(j)  ==  'Definitely  =1.0'  %
extract definite snore events
                num_snore = num_snore + 1;
                % calculate duration
                duration    =    WFN014A.Labels.Snoring{i,1}.ROILimits(j,2)    -
WFN014A.Labels.Snoring{i,1}.ROILimits (j,1);
                WFN014A_time_array(num_snore) = duration;
            end

        end
    end
end

%% put all the time arrays together

WFN014A_time_array (9:(length(WFN0151A_time_array)+8)) = WFN0151A_time_array;
WFN014A_time_array (122:(length(WFN029_time_array)+121)) = WFN029_time_array;
WFN014A_time_array (696:(length(WFN054A_time_array)+695)) = WFN054A_time_array;
WFN014A_time_array (1107:(length(WFN054D_time_array)+1106)) = WFN054D_time_array;
WFN014A_time_array (1485:(length(WFN077A_time_array)+1484)) = WFN077A_time_array;

WFN014A -> WFN0151A -> WFN029 -> WFN054A -> WFN054D -> WFN077A
```

```
%% normal distribution for duration


pd = fitdist(time_array,'Normal');
x_value = (pd.mu-4*pd.std):0.0001:(pd.mu+4*pd.std);  % set ticks for x axis
y = pdf(pd,x_value);
hold on;
grid on;
% plot the bell curve
plot(x_value,y,'LineWidth',2);

% plot the empirical rule (3 sigma rule)
p1        =        plot([pd.mu+3*pd.std       pd.mu+3*pd.std],[0       0.0031],'--
','Color',"#4DBEEE",'LineWidth',1.2);
p2        =        plot([pd.mu-3*pd.std       pd.mu-3*pd.std],[0       0.0031],'--
','Color',"#4DBEEE",'LineWidth',1.2);
p3        =        plot([pd.mu+2*pd.std       pd.mu+2*pd.std],[0       0.0382],'--
','Color',"#77AC30",'LineWidth',1.2);
p4        =        plot([pd.mu-2*pd.std       pd.mu-2*pd.std],[0       0.0382],'--
','Color',"#77AC30",'LineWidth',1.2);
p5        =        plot([pd.mu+1*pd.std       pd.mu+1*pd.std],[0       0.1713],'--
','Color',"#7E2F8E",'LineWidth',1.2);
p6        =        plot([pd.mu-1*pd.std       pd.mu-1*pd.std],[0       0.1713],'--
','Color',"#7E2F8E",'LineWidth',1.2);
p7 = plot([pd.mu pd.mu],[0 0.2825],'--','Color',"#EDB120",'LineWidth',1.2);
title('Normal Distribution',FontSize=12)
xlabel('Duration',FontSize=12)
ylabel('Probability Density',FontSize=12)
legend([p7  p5  p3  p1],{'\mu','\mu  \pm  1\sigma','\mu  \pm  2\sigma','\mu  \pm
3\sigma'},'Location','northwest')

% to find y value from x:
% desiredY = interp1(x_value,y,pd.mu+3*pd.std)
```

```
%% create time array
time_array = xlsread ("time_array.xlsx");

%% Duration Plot
x_value = xlsread ("duration.xlsx",'Sheet2','A1:A325');
y_value = xlsread ("duration.xlsx",'Sheet2','B1:B325');
plot (x_value,y_value)
xlabel('Snoring Duration')
ylabel('Number of Snore Events')
title('Total number of snore events vs. Snoring duration')

%% Frequency vs. SPL (average)

clc
num_snore = 0;
for i = 1:170
    % check if it's an empty cell
    data_i = labelData.Labels.Snoring{i,1};
    % if the cell contains data

    if ~isempty(data_i)
    ROI_i = labelData.Labels.Snoring{i,1}.Value; % extract the labelled audios

        for j = 1:length(ROI_i)
            if labelData.Labels.Snoring{i,1}.Value(j) == 'Definitely =1.0'
                num_snore = num_snore + 1;
                duration = labelData.Labels.Snoring{i,1}.ROILimits(j,:); %
extract duration
                filename = labelData.Source{i, 1}; % extract corresponding file
name of the audio sample
                [y,fs] = audioread(filename); % audio signal: y is sampled data,
```

```
fs is sample rate

                start_point = floor(duration(1)*fs)+1;  % +1 due to index starts
at 1
                end_point = floor(duration(2)*fs);
                y_i = y(start_point:end_point); % extract selected segment

                % power spectrum
                [SPLn,fn] = f_PS(y_i,fs);  % call the f_PS function
                SPLn_all(:,num_snore) = SPLn;

%                 % plot power specturm of each snore event
%                 p4 = plot(fn, SPLn, '-k');
%                 xlabel('Frequency (Hz)')
%                 ylabel('SPL (dB), re 20e^{-6} Pa')
%                 hold on

            end

        end
    end
end

%% plot the average SPL
sz = size(SPLn_all);
mean_SPL(:,1) = mean(SPLn_all,2); % compute mean in each row
p2 = plot (fn,mean_SPL,'Color',"#D95319",LineWidth=3);
hold on

% compute SD
SD_SPL (:,1) = std (SPLn_all,0,2); % compute SD in each row
SD1 = mean_SPL-SD_SPL;
SD2 = mean_SPL+SD_SPL;
```

```
p1 = plot (fn,SD1,'-c',LineWidth=1);
p3 = plot (fn,SD2,'-c',LineWidth=1);

% shade SD in between area
shade(fn, SD1, fn, SD2, 'FillType', [1 2;2 1],'FillColor','c')
% FillType specifies the area between line 1 and line 2 should be filled,
% whether line 1 is above line 2 or vice versa
legend([p2 p1],{'Average SPL', '\pm SD'},'Location','northeast');


xlabel('Frequency (Hz)')
ylabel('SPL (dB), re 20e^{-6} Pa')
title('WFN029 Sound Pressure Level vs. Frequency')



%% SPL vs. Num of snore

clc
num_snore = 0;
for i = 1:170
    % check if it's an empty cell
    data_i = labelData.Labels.snoring{i,1};

    % if the cell contains data
    if ~isempty(data_i)
    ROI_i = labelData.Labels.snoring{i,1}.Value; % extract the labelled audios

        for j = 1:length(ROI_i)
            if labelData.Labels.snoring{i,1}.Value(j) == 'Definitely =1.0'
                num_snore = num_snore + 1;
                duration    =    labelData.Labels.snoring{i,1}.ROILimits(j,:);    %
extract duration
```

```
                filename = labelData.Source{i, 1}; % extract corresponding file
name of the audio sample
                [y,fs] = audioread(filename); % audio signal: y is sampled data,
fs is sample rate

                start_point = floor(duration(1)*fs)+1;  % +1 due to index starts
at 1
                end_point = floor(duration(2)*fs);
                y_i = y(start_point:end_point); % extract selected segment

                % power spectrum
                [SPLn,fn] = f_PS(y_i,fs);  % call the f_PS function
                SPLn_all(:,6) = SPLn;

            end

        end
    end
end

% sz = size(SPLn_all);
% SPLn_all (:,sz(2)+1) = mean(SPLn_all,2);
% p1 = plot (fn,SPLn_all(:,end),'Color',"#D95319",LineWidth=3);
% legend(p1,'Average SPL','Location','northeast');
% xlabel('Frequency (Hz)')
% ylabel('SPL (dB), re 20e^{-6} Pa')
% title('WFN014A Sound Pressure Level vs. Frequency')

clear data_i
clear duration
clear end_point
clear filename
clear fn
```

```
clear fs

clear i

clear j

clear labelData

clear ROI_i

clear SPLn

clear start_point

clear y

clear y_i


x_value = xlsread ("results.xlsx",'SPL','A1:A312');

y_value = xlsread ("results.xlsx",'SPL','C1:C312');

plot (x_value,y_value)

xlabel('Snoring Sound Power Level')

ylabel('Number of Snore Events')

title('Total number of snore events vs. Snoring sound power level')



%% create an array of relative power in the frequencies for 6 participants

clc

num_snore = 0;

for i = 1:170

    % check if it's an empty cell

    data_i = labelData.Labels.snoring{i,1};

    % if the cell contains data


    if ~isempty(data_i)

    ROI_i = labelData.Labels.snoring{i,1}.Value; % extract the labelled audios


        for j = 1:length(ROI_i)

            if  labelData.Labels.snoring{i,1}.Value(j)  ==  'Definitely  =1.0'  ||
labelData.Labels.snoring{i,1}.Value(j) == 'Probably = 0.75'

                num_snore = num_snore + 1;
```

```
                duration   =   labelData.Labels.snoring{i,1}.ROILimits(j,:);   %
extract duration
                filename = labelData.Source{i, 1}; % extract corresponding file
name of the audio sample
                [y,fs] = audioread(filename); % audio signal: y is sampled data,
fs is sample rate

                start_point = floor(duration(1)*fs)+1;  % +1 due to index starts
at 1
                end_point = floor(duration(2)*fs);
                y_i = y(start_point:end_point); % extract selected segment

                % power spectrum
                [SPLn,fn] = f_PS(y_i,fs);  % call the f_PS function

                % to calculate the area under a curve
                int_all = trapz(fn, SPLn);
                int_selected = trapz(fn(6:40,1),SPLn(6:40,1));

                % calculate the relative power in frequencies
                relative_power = int_selected/int_all;
                relative_power_all(num_snore,6)  =  relative_power; % create an
array of all the relative power

            end

        end
    end
end

clear data_i
clear duration
clear end_point
clear filename
```

```
clear fn

clear fs

clear i

clear int_all

clear int_selected

clear j

clear labelData

clear relative_power

clear ROI_i

clear SPLn

clear start_point

clear y

clear y_i


%% Algorithm Evaluation

clc

snore = zeros (628,6);

for i = 1:6

    for j = 1:628

        if relative_power_all (j,i) > 0.5

            snore(j,i) = 1;

        else

            snore(j,i) = 0;

        end

    end

end



%% SPL vs. Frequency for averag snoring and non-snore events



clc
```

```
ind = 0;
for i = 1:170

    data_i = labelData.Labels.snoring{i,1};

    % check if it's an empty cell
    % if the cell does not contains
    if isempty(data_i)
        filename = labelData.Source{i, 1}; % extract corresponding file name of
the audio sample
        [y,fs] = audioread(filename); % audio signal: y is sampled data, fs is
sample rate
        % for WFN014A
        %start_duration = [0.1, 1.1, 2.2, 3.3];
        %end_duration = [1.1, 2.2, 3.3, 3.9];

        % for WFN0151A, WFN077A, WFN054A, WFN029
        start_duration = [0.1, 2.2, 4.4, 6.6, 8.8, 11, 13.2, 15.4, 17.6];
        end_duration = [2.2, 4.4, 6.6, 8.8, 11, 13.2, 15.4, 17.6, 20];

        % for WFN054D, WFN054A
        %start_duration = [0.1, 2.2, 4.4, 6.6];
        %end_duration = [2.2, 4.4, 6.6, 8.8];

        for j = 1:4  % cut into 2.2s
            ind = ind + 1;
            start_point = floor(start_duration(j)*fs);  % +1 due to index starts
at 1
            end_point = floor(end_duration(j)*fs);

            % extract selected segment
            y_i = y(start_point:end_point);
```

```
            % power spectrum
            [SPLn,fn] = f_PS(y_i,fs);  % call the f_PS function
            SPL_NS (:,ind) = SPLn;

        end
    end
end

% randomly extract # of non-snore events (2.2s) without repeating numbers
r = randperm(length(SPL_NS),145);  % return (1,x) matrix

% extract corresponding 2.2s event (event of interest)
for k = 1:length(r)
    event_i(:,k) = SPL_NS (:,r(k));
end



% compute mean of non events and plot
ns_average(:,1) = mean(event_i,2);
p1 = plot (fn,ns_average,'Color',"#D95319",LineWidth=3);
hold on

% compute SD
SD_SPL (:,1) = std (event_i,0,2); % compute SD in each row
SD1 = ns_average-SD_SPL;
SD2 = ns_average+SD_SPL;
p2 = plot (fn,SD1,'-c',LineWidth=1);
p3 = plot (fn,SD2,'-c',LineWidth=1);
hold on

% shade SD in between area
shade(fn, SD1, fn, SD2, 'FillType', [1 2;2 1],'FillColor','c')
```

```
% FillType specifies the area between line 1 and line 2 should be filled,
% whether line 1 is above line 2 or vice versa
legend([p1 p2],{'Average SPL' '\pm SD',},'Location','northeast');


xlabel('Frequency (Hz)')
ylabel('SPL (dB), re 20e^{-6} Pa')
title('WFN014A Sound Pressure Level vs. Frequency for Non-Snoring Events')





num_snore = 0;
for i = 1:170
    % check if it's an empty cell
    data_i = labelData.Labels.snoring{i,1};
    % if the cell contains data

    if ~isempty(data_i)
    ROI_i = labelData.Labels.snoring{i,1}.Value; % extract the labelled audios

        for j = 1:length(ROI_i)
            if labelData.Labels.snoring{i,1}.Value(j) == 'Definitely =1.0'
                num_snore = num_snore + 1;
                duration   =   labelData.Labels.snoring{i,1}.ROILimits(j,:);   %
extract duration
                s_filename = labelData.Source{i, 1}; % extract corresponding file
name of the audio sample
                [y,fs] = audioread(s_filename); % audio signal: y is sampled data,
fs is sample rate

                s_start_point = floor(duration(1)*fs)+1;  % +1 due to index starts
```

```
at 1
                s_end_point = floor(duration(2)*fs);
                y_i = y(s_start_point:s_end_point); % extract selected segment

                % power spectrum
                [SPLn,fn] = f_PS(y_i,fs);  % call the f_PS function
                SPLn_all(:,num_snore) = SPLn;

            end

        end
    end
end

% plot the average SPL
sz = size(SPLn_all);
s_mean_SPL(:,1) = mean(SPLn_all,2); % compute mean in each row
p4 = plot (fn,s_mean_SPL,'Color',"#0072BD",LineWidth=3);
hold on

% % compute SD
% s_SD_SPL (:,1) = std (SPLn_all,0,2); % compute SD in each row
% s_SD1 = s_mean_SPL-s_SD_SPL;
% s_SD2 = s_mean_SPL+s_SD_SPL;
% p5 = plot (fn,s_SD1,'-c',LineWidth=1);
% p6 = plot (fn,s_SD2,'-c',LineWidth=1);


legend([p1    p4],{'Average    SPL    for    non-snoring'    'Average    SPL    for
snoring',},'Location','northeast');

xlabel('Frequency (Hz)')
ylabel('SPL (dB), re 20e^{-6} Pa')
```

```
title('WFN054A Sound Pressure Level vs. Frequency')
```