# THE EFFECTIVENESS OF VARIOUS CHATTER DETECTION METHODS UNDER NOISY CONDITIONS

A Thesis
Presented to
The Academic Faculty

By

Lance C. Lu

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
George W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology

August 2020

**THE EFFECTIVENESS OF VARIOUS CHATTER DETECTION METHODS UNDER NOISY CONDITIONS**

Approved by:

Dr. Christopher Saldana, Chair
School of Mechanical Engineering
*Georgia Institute of Technology*

Dr. Thomas Kurfess, Co-Chair
School of Mechanical Engineering
*Georgia Institute of Technology*

Dr. Katherine Fu
School of Mechanical Engineering
*Georgia Institute of Technology*

Date Approved: May 13, 2020

## ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Christopher Saldana, for his guidance throughout the research process. His expertise and discussions have helped me understand this area of research and the research process. Your mentorship through this process has been invaluable. I would also like to thank my co-advisor Dr. Thomas Kurfess for always pushing me to think of new and immerging technology applications. I would like to thank Dr. Katherine Fu for her insights on design thinking and keeping the human in mind. You all have made my time at Georgia Tech incredibly worthwhile.

I would also like to thank the members of the EPICS cohort, whose support and expertise have given me a wide breadth of knowledge in mechanical engineering, digital manufacturing, and machining. You always managed to keep me engaged with your knowledge and humor.

I would also like to thank my mother for supporting my education for all these years. Without her support, I would never have gotten so far in my academic career.

This work was supported in part by DE-EE0008303, NSF CMMI-1646013, NSF CMMI-1825640 and NSF IIP-1631803.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## NOMENCLATURE

| | |
|---|---|
| ADC | Analog To Digital Converter |
| AE | Acoustic Emission |
| ANN | Artificial Neural Networks |
| CF | Crest Factor |
| ClF | Clearance Factor |
| CNC | Computer Numerical Control |
| DOC | Depth of Cut |
| FFT | Fast Fourier Transform |
| FN | False Negative |
| FP | False Positive |
| HBP | Harmonic Bandpower |
| IF | Impulse Factor |
| IoT | Internet of Things |
| kNN | k-Nearest Neighbors |
| Ku | Kurtosis |
| ML | Machine Learning |
| MRR | Material Removal Rate |
| Pk | Peak Value |
| RMS | Root Mean Square |
| RPk | Relative Peak |
| RPM | Revolutions Per Minute |
| SF | Shape factor |
| Sk | Skewness |
| SOF | Statistical Overlap Factor |
| STFT | Short Time Fourier Transform |
| SVM | Support Vector Machines |

| | |
|---|---|
| TN | True Negative |
| TP | True Positive |
| V | Variance |
| WT | Wavelet Transform |

## SUMMARY

Unmitigated chatter can result in poor part quality, accelerated tool wear, and possible damage to the spindle and machine. Several methods have been shown to effectively detect chatter in lab conditions. The implementation of these methods in noisy environments, such as factory floors, has not been well studied, however. In order to achieve reliable performance in a real machining environment, chatter detection methods should be robust to a variety of noises.

This study aims to understand the question of whether machine learning approaches are more robust to high levels of noise in assessing chatter in machining signals. The purpose of this study is to examine the performance of various chatter classification methods under varying background noises. To accomplish this, stable and unstable cuts were made on a milling machine and the audio signal was collected. The audio signal was then superimposed with levels of white Gaussian noise and periodic noise to simulate a noisy data collection environment. A statistical approach, along with several machine learning classifiers were trained and tested on this noisy data.

The performance of these techniques was then compared with respect to the increasing levels of noise. It was found that machine learning approaches achieved satisfactory accuracies of up to 94.1% under noisy conditions. Conventional static threshold techniques, however, failed under most noise conditions. Support vector machines demonstrated an ability to classify noisy data despite limited training. These results indicate that machine learning methods have a significant ability to classify noisy data and may be a promising approach to practical chatter detection.

# CHAPTER 1: INTRODUCTION

In the fourth industrial revolution, manufacturers are increasingly pressured to meet the demands of higher part quality, lower downtime, and faster production. To meet these demands, manufacturers are leveraging the growing fields of Internet of Things (IoT) and machine learning (ML) to automate or assist in their manufacturing processes [1]. One of these processes is milling. Milling is a common material removal operation in manufacturing environments when working with non-axisymmetric parts. The forces generated when the cutting tool engages with the work piece produce significant deflection of the tool-workpiece system. An example tool-workpiece system is shown below in Figure 1.

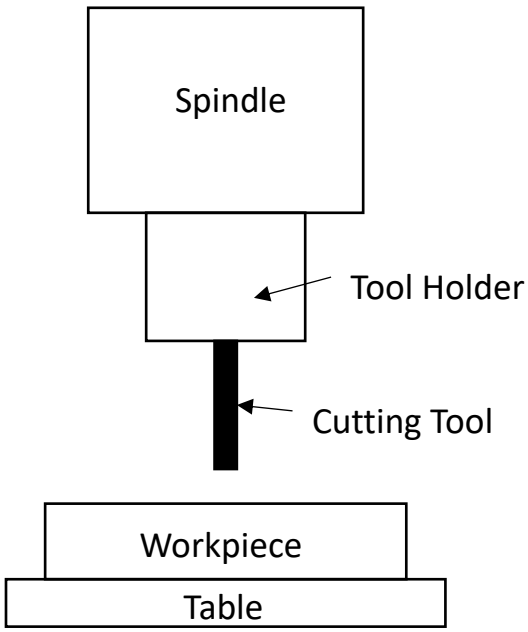


**Figure 1:** Example tool-workpiece system

When this process becomes unstable, regenerative chatter occurs. Chatter is a self-excited unstable vibration that occurs during a metal cutting process, such as milling. Unmitigated chatter can result in poor part quality, accelerated tool wear, and possible damage to the spindle and machine [2]. Figure 2 shows the results of chatter on a milled surface. All these result in increased production time and costs. An article from Deloitte shows that unplanned downtime costs manufacturers an estimated $50 billion each year [3]. In addition to unplanned downtime, chatter can result in machine operators to use less than ideal cutting parameters. Typically, spindle speed and/or feed rate is reduced. This results in a material removal rate (MRR) that is significantly lower than the machines energy limitations. By choosing stable cutting parameters to avoid chatter, MRR can be increased dramatically [12]. This increase can be even more dramatic with newer high-speed milling machines and machining strategies.

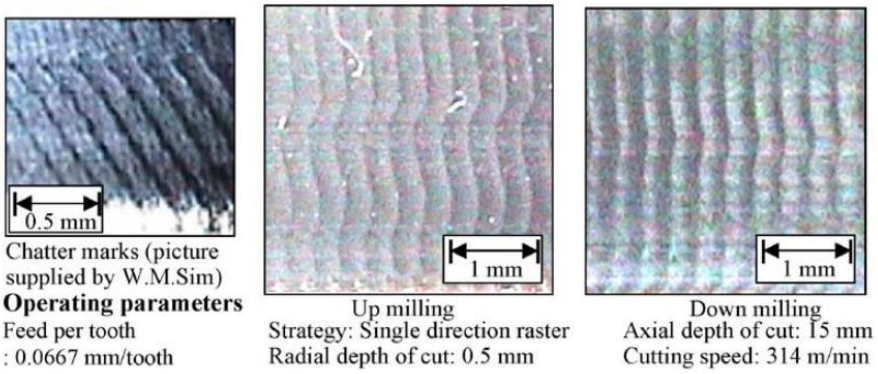


**Figure 2:**Chatter marks [4]

When chatter occurs, the resulting vibration has a frequency that is different from that of the tooth passing frequency. For this reason, frequency-based techniques are

generally effective at detecting chatter. Several threshold-based methods of analyzing an audio signal in the frequency domain for chatter have been developed by researchers over the past few decades [21, 22, 42, 43]. Other methods involve analyzing the time-series and frequency domain, and extracting features to train ML classifiers such as support vector machines (SVM) or artificial neural networks (ANN) [25, 26, 27, 28, 41]. Several commercial products such as the MetalMax Harmonizer [46] and the Okuma Machining Navi [47] are also available to address the problem of chatter.

### 1.1 Motivation

Although audio signals have been proven to be an effective method of detecting chatter in research environments, their effectiveness under noisy conditions has been questioned [22]. The noisy shop floor can cause false alarms in a monitoring system that is not appropriately tuned [27]. In industrial environments, which have been found to have sound pressure levels of up to 90 dB, these false alarms could become a significant issue [50]. These factory sounds have a large range of frequencies they affect. Noise sources include but are not limited to motors, fans, machining centers, talking, air flow, and fluid flow. Very rarely do machining operations happen in complete isolation, so it is important to understand the veracity of such methods in the presence of background noises. Many methods for detecting chatter require setting a static noise threshold that once passed, raises an alarm for chatter [22, 23, 42, 43]. If the threshold is set high so that the system is less effected by noise, the chance for missed detection increases. For practical implementation of these audio signal chatter classification systems in industrial environments, the influence of noise should be understood. Inaccurate classifications due

to noise can lead to non-optimal tooling operations, or, worst-case catastrophic failure with unmitigated chatter. Understanding background noise can increase the reliability of future systems. ML approaches such as SVM classification may have the ability to function accurately in the presence of heavy factory noise, unlike standard thresholding methods. SVMs have shown an ability to classify accurately with limited training and fairly noisy data in other applications such as tool wear [29, 30].

### 1.2 Problem Statement

While the impact of noise levels on threshold-based chatter detection methods has been explored elsewhere [22, 27], the effect of noise on ML approaches is not well understood. ML-based approaches may provide significantly more enhanced performance in the presence of noise when compared to static threshold-based methods [26]. Thus, the present study seeks to address the question of whether ML-based approaches are more robust to high levels of noise in assessing chatter in machining signals. The purpose of this study is to examine the performance of various chatter classification methods under varying background noises. To accomplish this, stable and unstable cuts were made on a milling machine and the audio signal was collected. The audio signal was then superimposed with levels of white Gaussian noise and periodic noise to simulate a noisy data collection environment. The accuracy of these techniques was then compared with respect to the increasing levels of noise. The end goal is to understand the balance between sensitivity, speed of implementation, and performance of various classification systems under noisy conditions.

### 1.3 Structure

In Chapter 2, the background of machine process monitoring will be discussed with a focus on chatter detection. Past work regarding signal processing, feature extraction, feature selection, and classification methods will be overviewed. The experimental set-up and data collection system will be shown in Chapter 3. The signal acquisition process and the sensor specifications will be described. The sampling rates and testing configurations will be detailed. Chapter 4 will present analysis of the experimental data and noise-augmented data, including feature extraction and classification methods. The effectiveness of these systems under varying background noises will be discussed. Chapter 5 will present conclusions and discuss future areas of research stemming from these experimental findings.

# CHAPTER 2: BACKGROUND

Manufacturers are constantly balancing the demands of product quality, product variability, cost, and speed. As a result, manufacturers are increasingly turning to automation to meet these demands and remain competitive, globally. In recent decades, a vast amount of research has been focused on process monitoring to reduce the need for expert operators [1]. Process monitoring also allows machine tool users to step away from planned scheduled maintenance and move towards condition-based maintenance. Condition-based maintenance has advantages over tradition planned maintenance in that it more efficiently minimizes downtime and extends the life of machine tool systems, while reducing life cycle and maintenance costs [2]. Process monitoring in manufacturing is the estimation of process variables such as cutting forces, vibrations, acoustic emission (AE), noise, temperature, surface finish, that are influenced by the cutting tool and cutting parameters inputted [10]. These variables are measured by the appropriate physical sensors. Analog signals detected by these sensors are transformed into digital signals that are then processed. From these digital signals, features are extracted. These features potentially correlate with tool or process conditions of interest to the machine operator. These features are used by decision making systems, such as neural networks, to take an action. Actions may be limited to an alert or suggestion or may go so far as to take automated corrective steps. Figure 3 shows a typical method of monitoring for tool wear.

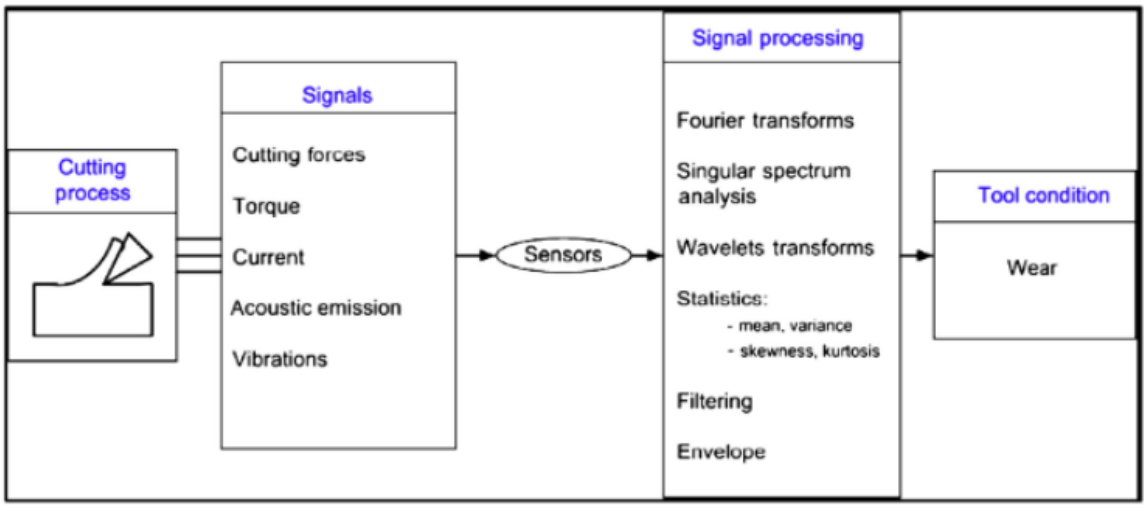


**Figure 3:** Example process monitoring method [11]

### 2.1 Signal Measurement and Processing

As shown in Figure 3, during the cutting process a variety of physical signals are generated. Some of these signals can be measured directly, and others may need to be estimated through measurable phenomena. A wide range of signals has been used for process monitoring including AE, tool temperature, cutting forces (static and dynamic), vibration (acceleration), and surface finish quality, and spindle motor current [2, 6, 10]. Along with measuring the appropriate signal, sensors for process monitoring should also meet the following requirements [13]:

1. Measurement as close to the machining point as possible
2. No reduction in the static and dynamic stiffness of the machine tool
3. No restriction of working space and cutting parameters
4. Wear and maintenance-free, easily changed, low costs
5. Resistant to dirt, chips and mechanical, electromagnetic and thermal influences

6. Function independent of tool or workpiece
7. Adequate metrological characteristics
8. Reliable signal transmission

The processing of a raw signal follows the scheme shown in Figure 4. The sensor measures the physical signal from the process and transforms it into an analog signal. This analog signal is then converted to a digital signal using an Analog to Digital Converter (ADC). Depending on the sensor, pre-processing may be needed before conversion, such as filtering and amplifying [10]. During pre-processing, the continuous data are segmented into portions. The raw signal from a sensor is usually a time-domain signal. This signal can be transformed into the frequency domain, or the time-frequency domain through various means such as the Fast Fourier Transform (FFT), the Short Time Fourier Transform (STFT) or wavelet transform (WT). A wide range of features can be extracted from these domains.

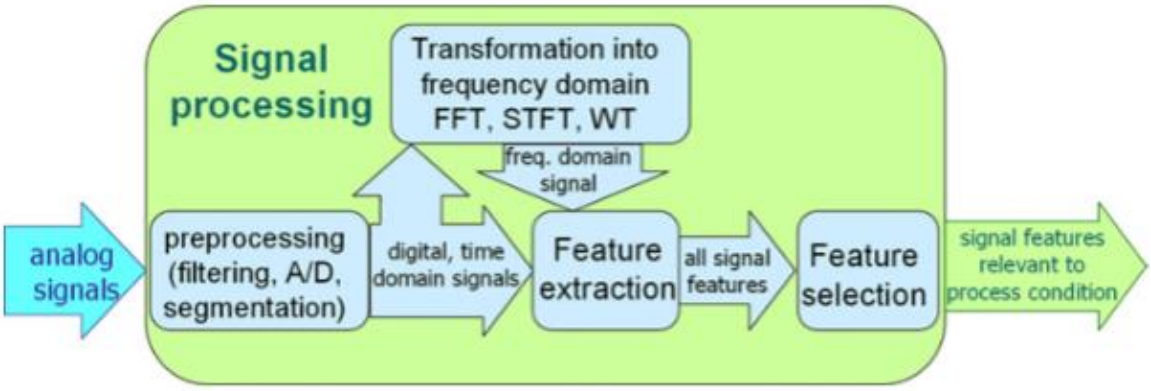


**Figure 4:** Signal processing logical scheme [10]

### 2.1.1 Time Domain

Features from the time domain are extracted that can describe the signal and maintain as much information about the process as possible. Common features to extract include mean, magnitude, root mean square (RMS), variance, standard deviation, skewness, kurtosis, signal power, peak-to-peak value, range, peak-to-valley amplitude, crest factor (CF) and ratios of the signals [10]. In Binsaeid et al, the time domain features of mean, RMS, variance, skewness, kurtosis, signal power, peak-to-peak amplitude, and CF are extracted from force, acoustic emission, vibration, and power signals to detect end milling abnormalities such as flank wear, tool breakage, and tool chipping [14]. Average and peak forces collected from a dynamometer were used in [15] along with spindle speed, feed rate, and depth of cut to predict the surface roughness with a neural network. The standard deviation of the thrust force obtained from a dynamometer was used to detect chatter in a drilling operation in [16] along with chatter suppression by automatic spindle speed selection.

### 2.1.2 Frequency domain

The extraction of features from the frequency domain is done by using a windowed Fourier transform. The Fourier transform breaks down a time-domain signal into a combination of sine waves. The sine waves are then represented in the frequency domain with amplitude, frequency, and phase. One of the most practical and common transforms is the FFT [10]. It is important to note one of the FFTs short comings: spectral leakage.

The FFT is a discrete transformation which will not have perfect whole periods. The waveforms may be truncated, and the endpoints are discontinuous. This leads to discontinuities that are shown in the frequency domain as high frequency components, known as spectral leakage [17]. Applying a window can minimize the effects of spectral leakage by closing the discontinuities and smoothing the signal before the FFT is applied. An example of windowing and its influence on the spectral leakage is shown in Figure 5.

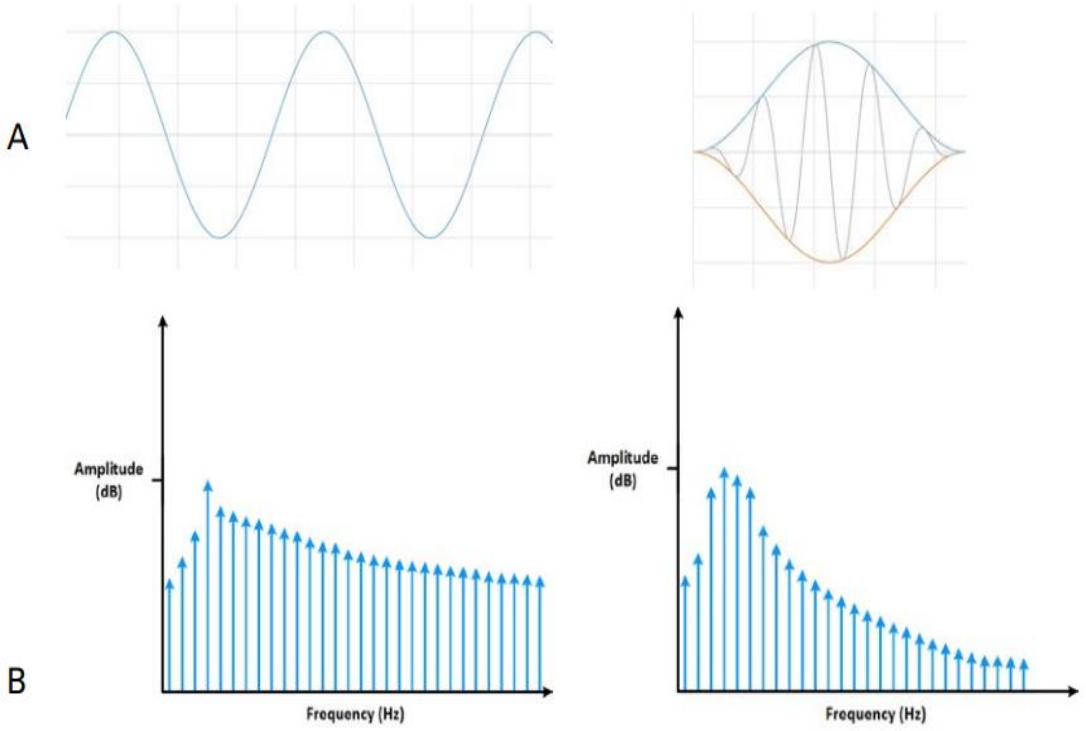


**Figure 5**: Windowing effect on FFT and spectral leakage in the **(a)** time domain **(b)** frequency domain [2]

Numerous windows exist for different applications, but the most appropriate one for audio applications is the Hann/Hanning window [18]. The Hann window, shown in

Figure 6, is a raised cosine function with both endpoints at 0 with a zero slow. The side lobes roll off quickly, making it suitable for audio signal processing.

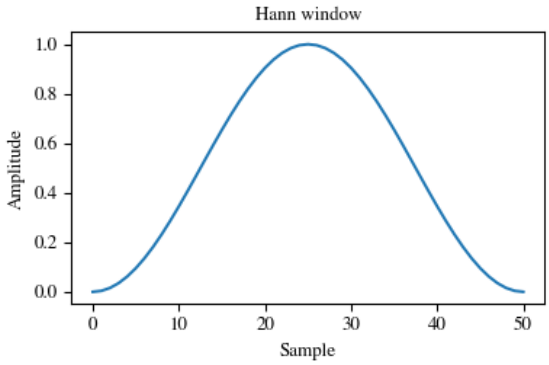


**Figure 6:** Hann window [18]

When the frequency of a signal is of interest, it must be digitized accurately. The signal must be sampled at least two times higher than the highest expected frequency component in the source signal. Half of the sampling frequency is known as the Nyquist frequency. If the sampling frequency is not at least double the highest component frequency, then aliasing will occur as shown in Figure 7.

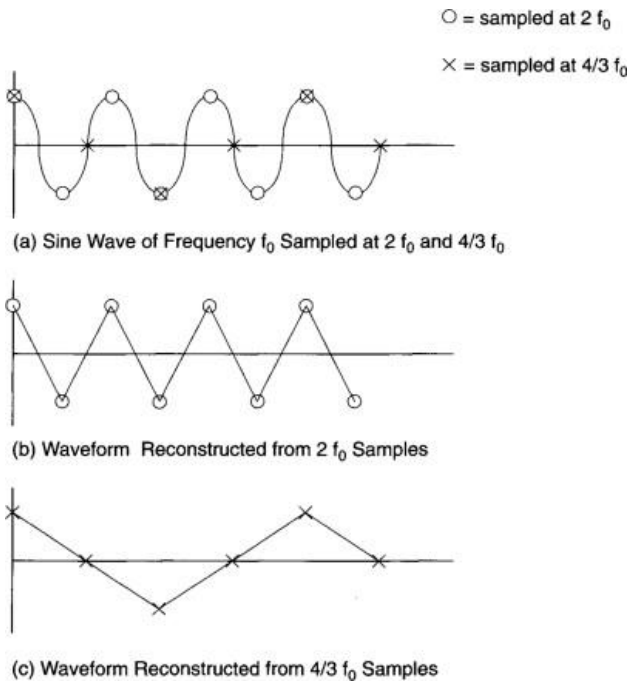


**Figure 7**: Aliasing due to insufficient sampling [19]

In the figure, the signal is a sine wave of frequency $f_0$. If the signal was sampled at double the frequency, then the signal would produce a waveform of the same frequency as shown in (b). If the sampling rate is less than double as in (c), then the frequency recreated would be of frequency (1/3) $f_0$. With a higher sampling rate, the digitized waveform would more closely represent the original signal. If the sampling rate were equal to the signal frequency, the digitized signal would be a constant [19].

Understanding the frequency spectrum of an audio signal generated during a milling operation can give insight into the cutting process. The spectrum contains frequencies from the spindle run-out, the tooth-passing frequency, the possible chatter frequency, as well as any harmonics. The spindle run-out frequency is equal to the spindle speed, and the tooth-passing frequency is equal to the spindle speed multiplied by

the number of teeth. The harmonics are multiples of these frequencies. By taking the FFT, these frequencies can be separated from each other and analyzed.

Features extracted from the FFT may include: amplitude of dominant spectral peaks, signal power in specific frequency ranges, energy in frequency bands, statistic features of the band power spectrum such as mean frequency, variance, skewness, and kurtosis and frequency of the spectrum's highest peak [10]. In analyzing the frequency spectrum from an AE sensor, Guo et al, found that the existence of surface damage increased both the amplitude and frequency of the AE spectrum's second peak [20]. Chip formation and chatter onset during turning were monitored using force from a dynamometer in [21]. They utilized threes ratios of cumulative power spectrum density within specified frequency ranges, obtained from three cutting force components. Chatter, continuous chip formation, and broken chip formation were all successfully classified regardless of cutting conditions.

Delio et al [22] and Smith [23] describe method for chatter detection using frequency domain features. The method relies on the difference between the tooth frequency and the chatter frequency to separate them in the frequency domain. The user inputs several variables including the number of cutting edges, triggering threshold, and the sampling frequency. The triggering threshold level is chosen through experimentation. In their application, there was a pronounced increase in sound level from the machine when chattering so there were a wide range of acceptable thresholds. The system continuously monitors the audio signal during cutting operations. Once the system has been triggered by the sound level exceeding the threshold, the system acquires enough samples to perform an FFT. A peak search routine is performed to find the largest

peak in the spectrum. If the frequency of the peak in the signal does not correspond to the cutter frequencies or harmonics, then the algorithm detects chatter. Their system goes further by commanding the machine to stop and calculate a new spindle speed in order to move towards stability.

The system was shown to detect chatter between 250 to 750 milliseconds and an addition 2-4 seconds to compute the spectrum. The system was shown to be effective in aluminum and cast-iron half immersion and slotting operations. Triggering of the system was noted to be a potential problem, as outside noises could inadvertently trigger the system. They highlight sound isolation and insulation to increase the quality of the signal and prevent false positives.

### 2.1.3 Time-frequency domain

When using FFT, the frequency changes of the signal over time are lost. To combat this, a time-frequency analysis known as the short time Fourier transform (STFT) may be used. This analysis method is especially useful non-stationary signals and processes. The STFT uses a sliding window to calculate the Fourier transform at different times and combining the successive calculations. Thus, the frequency changes over time can be seen across the bandwidth. Marinescu demonstrated that the STFT with AE signals could identify the entrance/exit of each cutting insert into or out of the material [24]. Tilen et al used STFT on sound pressure signals to discriminate between chatter and chatter-free conditions when band sawing. Figure 8 shows an example of their work where chatter occurs distinctly at the beginning and end of the cut [25].

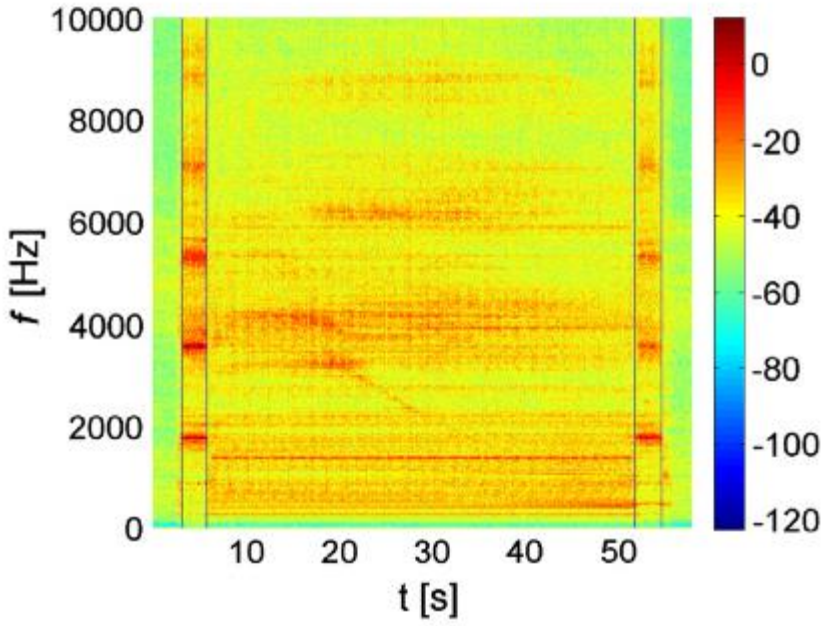


**Figure 8:** STFT with chatter regions on the edges [25]

### 2.1.4 Wavelet transformation

In the Fourier transform, information about time is completely lost. With the STFT, time domain information is recovered, but the resolution is dependent on the size of the window. With a narrow window, the transform will have poor frequency resolution. With a wider window, the time resolution will be poor. The wavelet transformation (WT) provides a compromise on this time-resolution problem by applying windows of different sizes as needed. For low frequency information, a longer interval may be used. For high frequency information, short windows may be used. At high frequencies, good time resolution is achieved at the expense of frequency resolution. The opposite is true at low frequencies.

Time-domain and frequency-domain features mentioned previously can be extracted from the STFT and WT, such as mean, RMS or variance. Yoon and Chin [26] used WT to successfully isolate chatter frequencies for real-time detection in end milling operations.

### 2.2. Feature Selection

Many features can be generated from a signal, but many of them can be poor estimators of the process conditions. Certain features may be more relevant and sensitive to the process being monitored. The goal of feature selection is to remove redundant or irrelevant features while preserving as much information about the signal as possible [10]. Reduction of features allow for quicker training of ML algorithms and faster computation. If the number of features is too reduced, however, the system is more susceptible to disturbance. Several methods of feature selection and reduction have been used in literature.

Statistical measures can be used to determine which features best correlate with the process conditions of interest. Sheffer and Heyns demonstrated two methods of feature selection in [27] by calculating a correlation coefficient and a statistical overlap factor (SOF). A high correlation coefficient *p* indicates that there is a strong correlation between the feature *q* and the process variable *V*. This coefficient can be calculated with the following:

**Equation 1:**

$$p = \left| \frac{\sum_i (q_i - \bar{q})(V_i - \bar{V})}{sum_i (q_i - \bar{q})^2 (V_i - \bar{V})^2} \right| \times 100$$

The SOF can be calculated by using the same feature for two different conditions where $x_1$ is the feature under condition 1and $x_2$ is the feature under condition 2. The SOF determines the degree of separation of a feature between two conditions. A higher degree of separation indicates a more ideal feature. The SOF is calculated with the following:

**Equation 2:**

$$SOF = \left| \frac{\bar{x}_1 - \bar{x}_2}{(\sigma_1 + \sigma_2)/2} \right|$$

Sheffer and Heyns chose features that had high correlation and high SOF. They noted that some human judgement was required as automated feature selected often selected features that were too similar or dependent on one another.

Lamraoui et al used relative entropy to measure the separation "distance" of two distributions of a feature under chatter and non-chatter conditions. The further apart the two distributions, the higher the quality of feature [28]. In general, features that correlate well with the process variable of interest, are distinct from each other, and change more dramatically with the process variable are better suited for classification.

### 2.3 Classification Methods

Once the appropriate features have been extracted from the raw signals, a decision must be made on what condition the machine is operating under. This decision can be made using statistical thresholds, such as the one described in [22, 23], where a threshold was set for the peak frequency amplitude. Statistical classification methods are quick to implement and generally simple to understand. An issue with these methods is that the threshold must be determined experimentally for classification to be meaningful.

Support vector machines (SVM) have been shown to be suitable classifiers for machine process monitoring as well. In [29], SVMs were combined with WT to extract features in order to detect chatter in a boring process. Using the wavelet packet energy and standard deviation, appropriate features were generated. Accuracies of about 95% were achieved, and early onset chatter could be detected. Compared to other ML classifiers, SVMs can overcome problems of multiple local minima and over-fitting, while being able to be trained on minimal data to give a solution. A minimal number of features can be used with SVM as well to get appropriate solutions. In [29], only two features were used to obtain accurate results. Computation times are also relatively fast when compared to other techniques such as artificial neural networks [30].

Artificial neural networks (ANN) are non-parametric machine learning algorithms inspired by the human nervous system [30]. They can accurately model non-linear relationships among features. ANNs were used in [28, 41] to monitor chatter. ANNs however are prone to overfitting and thus require large diversified training data [30].

### 2.4 Chatter Detection

Process monitoring systems have been used to estimate a wide variety of applications including surface roughness, tool wear, tool breakage, dimensional accuracy, temperature, chip conditions, and chatter [1, 10]. The focus of this thesis will be the detection of chatter. Chatter has been researched for more than a century and still poses a major obstacle in optimizing machining processes such as turning, milling, drilling, boring, broaching, and grinding. Unmitigated chatter results in poor surface finish, dimensional inaccuracy, excessive noise, machine tool damage, reduced tool life and

reduced MRR material waste, energy waste, increased machining time, and increased costs [3, 31]. Thus, chatter detection, avoidance, and suppression are all major areas of interest for researchers and machine operators. To avoid chatter, a common practice is to use conservative cutting parameters. This suboptimal MRR may result in decreased productivity compared to a higher, stable MRR. In 2006, Renault S.A.S, an automobile manufacturer, estimated that the cost due to chatter on a cylinder block was around 0.35€ per piece. With 3 million engines per year, the costs of chatter could become very significant [32].

Chatter is classified in two groups: primary chatter, which is caused by the cutting process and secondary chatter which is caused by regeneration of waviness of the workpiece surface. Secondary or regenerative chatter is the most common form of chatter and will be referred to in this thesis as "chatter" [33]. This regeneration of waviness can be seen in Figure 9. The cutter vibrations leave a wavy surface and when the next cutting tooth removes material the waviness results in a varying force on the cutting tool. This varying force builds on itself and becomes regenerative chatter [33, 34, 35]. Chatter continues to build unmitigated until the tool jumps out of the cut or breaks.

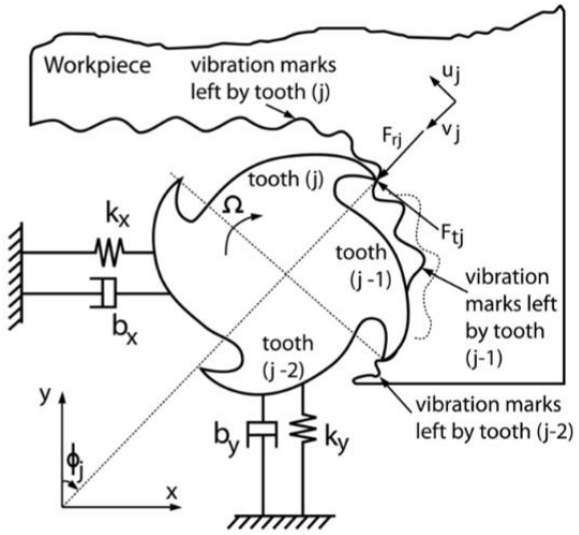


**Figure 9:** Regeneration of waviness in a milling model with two degrees of freedom [33]

Researchers have been able to detect chatter with a variety of instruments such as dynamometers [16, 36, 37], accelerometers [15, 38-41], and microphones [22-25, 42, 43]. Multi-sensor approaches have also been done [14, 44, 45]. Commercial products from Metalmax [46] and Okuma [47] are also available for chatter detection. These products use a microphone to monitor the cutting process. The use of microphones for monitoring the cutting process has been shown to be effective and inexpensive compared to other sensor approaches [22, 38, 43]. Accelerometers are acceptable as well, but the placement of them can cause change the apparent strength of different modes of vibration, leading to sensitivity and noise problems. Tlusty in [4] outlines a procedure to address these placement problems but prior knowledge of the system is needed which may not be achievable for all operations. The sensor must also be capable of detecting vibrations

from multiple sources on the machining center, as chatter can result from the workpiece, tool, spindle and machine structure vibrating [22].

Schmitz et al in [42] sampled the audio signal of a cutting operation once per revolution using a microphone. An infrared detector was used to synch the timing of the sample every revolution. A threshold was set on the variance of the microphone signal. Under unstable conditions, there is a high variance of the measured audio signal. Ismail and Ziaei [48, 49] used acoustic intensity to detect chatter in 5-axis machining. A threshold for acoustic intensity was set to detect stable, moderate chatter, and severe chatter conditions. The detection system was paired with a spindle speed ramping system that would move the operation into a stable cutting zone and suppress chatter. A similar approach is taken by Tsai et al in [43]. Two microphone signals are averaged and converted into an acoustic chatter signal index. Once this index goes above a set threshold, chatter is said to be detected. This index is taken after background noises such as fluid flow and AC power have been filtered out properly.

One of the major disadvantages of using microphones, however, is the prevalence of noise coming from the factory floor in practical implementation. The noisy factory floor can create false alarms in chatter detection systems [27]. False alarms have the potential to greatly slow down production, depending on the actions of the chatter detection system. If the system fails to operate properly in a real factory environment, the system is ineffective. Sound isolation and filtering may be used to decrease the effects of noise on audio signal collection systems [22]. Acoustic intensity can also be used to reduce the effects of background noise [22, 48, 49]

# CHAPTER 3: EXPERIMENTAL METHODOLOGY

Machining experiments were performed at the Georgia Institute of Technology on an EMCOMILL E350 with a Siemens Sinumerik 828d controller shown in Figure 10. The E350 is a three-axis computer numerical control (CNC) milling machine with a top spindle speed of 10,000 revolutions per minute (RPM). This machine was chosen because it has an enclosure and is in a room where other machines are not frequently used. Experiments were conducted when no other machines were operating in the area. With these conditions, the clearest audio signal could be obtained.


**Figure 10:** EMCOMILL E350

A 0.375in x 1.75in (diameter x tool overhang) solid carbide end mill with 2 flutes from Kennametal was used to cut slots in Aluminum 6061. 3 spindles speeds of 3000, 5000, 6000, and 7000 RPM were chosen. From a starting depth of 0.10 inches, the depth

of cut (DOC) was increased until chatter was induced to a max of 0.60 inches. The feed per tooth was kept constant at 0.03 in/tooth. The cutting conditions are summarized in Table 1. Coolant was on for the slotting operations. Workpieces were first face-milled flat to control surface roughness effects.

**Table 1:** Experiment cutting conditions

| | |
|---|---|
| **Workpiece** | Al-6061-T6511 |
| **Endmill** | Diameter = 0.375in<br>Helix angle = 45 degrees<br>2 flutes<br>Solid Carbide |
| **Feed per tooth** | 0.03in/tooth |
| **Radial depth of cut** | 0.375in |
| **Starting axial depth of cut** | 0.1in |
| **Spindle Speed** | 3000, 5000, 6000, 7000 RPM |
| **Tool overhang** | 1.50in |
| **Tool holder** | SK30 |
| **Collet** | ER32 |

A PCB 130F20 microphone was used to collect the audio signal during the experiments as shown in Figure 11. The microphone was placed inside the milling enclosure with 30 inches from the tip of the tool. Microphone specifications are shown in Table 2. The microphone was used with a PCB signal conditioner. Audio signals were

collected in LabVIEW at a sample rate of 48 kHz using a compactRIO-9014 with an NI-9215 module. Processing was done separately in MATLAB. Collected signals were transformed from Volts to Pascals using the sensitivity specification listed in Table 2. 48kHz was chosen to achieve a Nyquist frequency of 24kHz which is the upper range of human hearing and above the specified frequency range of the 130F20 microphone.

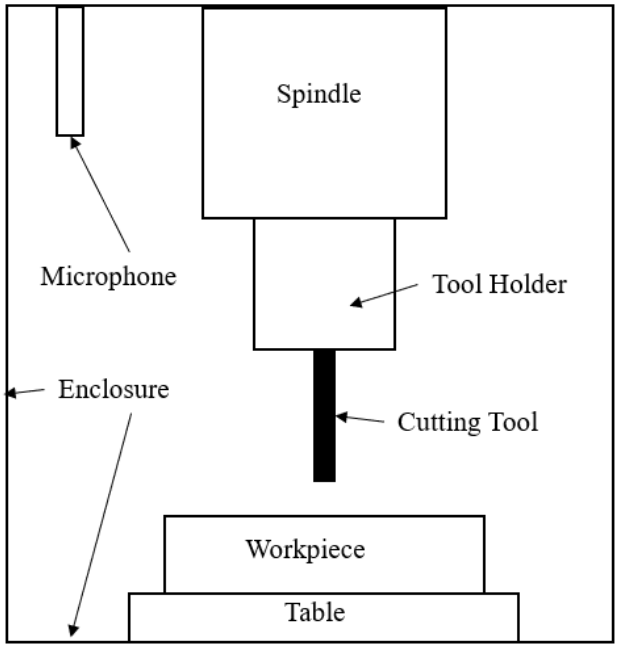


**Figure 11:** Audio collection setup

**Table 2:** Microphone specifications

| | |
|---|---|
| **Model** | PCB 130F20 |
| **Nominal diameter** | 1/4 in |
| **Frequency range** | 10 to 20000 Hz |
| **Sensitivity** | mV/Pa |

# CHAPTER 4: ANALYSIS AND RESULTS

Audio signals were collected at 3000, 5000, 6000, and 7000 RPM. The DOC at these speeds is recorded in Table 3 along with their stability. Chatter was induced at 3000 RPM and 6000 RPM at 0.25 and 0.6 inches, respectively. Past experiments on the EMCO E350 with DOCs above 0.6 inches experienced work holding failures and tool breakage.

**Table 3:** Experimental results

| Spindle Speed (RPM) | DOC (in.) | Stability |
|---|---|---|
| 3000 | 0.15, 0.20, 0.25 | Severely unstable at 0.25 in. |
| 5000 | 0.30, 0.35, 0.40, 0.45 | Stable |
| 6000 | 0.50, 0.60 | Marginally unstable at 0.6 in. |
| 7000 | 0.60 | Stable |

## 4.1 Signal Pre-Processing

A full audio signal can be seen in Figure 12, with an example of a stable cut and an unstable cut. In the unstable cut, the amplitude or volume of the cut grows in strength as the cut progresses due to the regenerative nature of chatter. In a stable cut, the amplitude of the cut remains relatively stable and does not grow. When the tool enters or exiting the workpiece, there is a sharp increase in volume that quickly fades away. These transient behaviors from in the audio signal were truncated so that the signal would be taken at a time when the tool was fully immersed. This time is calculated by equation 3, where tool diameter is in inches and feed rate is in inches per minute. Each tool path

takes approximately 7 seconds from entrance to exit. After truncation, 5 seconds of audio data is gathered.

**Equation 3:**

$$t_{immerse} = \frac{(tool\ diameter\ x\ 60)}{2\ x\ feedrate}$$

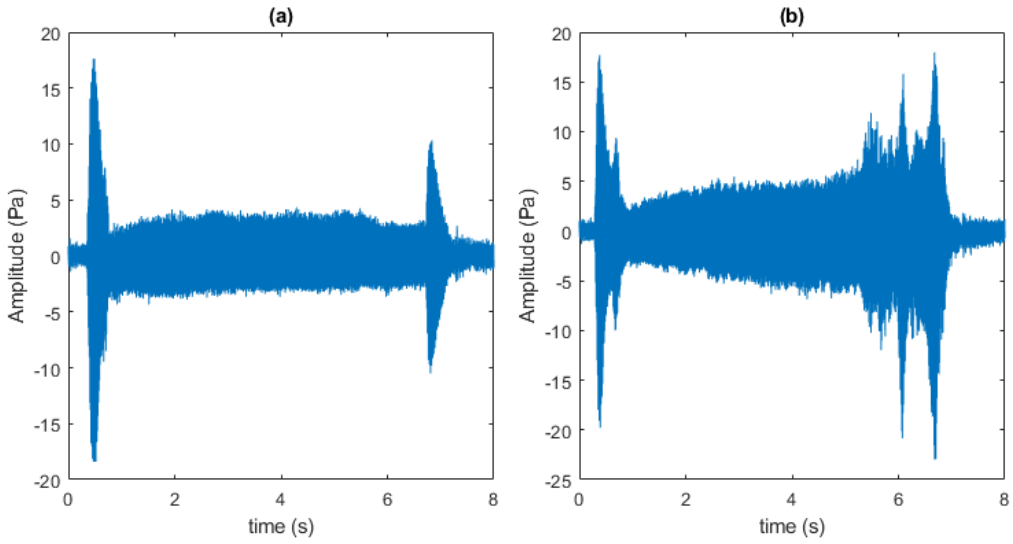


**Figure 12: (a)** A stable cut at 3000 RPM and 0.20in DOC compared with **(b)** an unstable cut done at 3000 RPM and 0.25in DOC

From the full audio signal, the data are segmented into 0.1-second segments with a 50% overlap to increase the dataset size. An example of 0.1s segment of data is shown in Figure 13. The waveform can be more clearly distinguished as well as erratic frequencies resulting from chatter. In Figure 13(b), jagged edges from the chatter

frequencies stand out from the overall behavior of the audio signal. These edges oscillate at a much higher frequency than the main tooth passing frequency of the signal. In Figure 13(a) the jagged edges can also be observed but at much lesser degree. The dominant signal is the tooth passing frequency occurring at 100 Hz. When the segmented signal is transformed to the frequency domain with the FFT and a Hanning window, the tooth-passing and chatter signals become more apparent as seen in Figure 14. The tooth passing frequency is occurring at 100 Hz with its first two harmonics being represented at 200 and 300 Hz. There are also frequencies from the spindle runout and harmonics at 50 Hz and 150 Hz. Chatter and the chatter harmonics are found in Figure 14(b) centered at 1510 Hz. Its harmonics are spaced apart by the tooth passing frequency. The dominant harmonics are shown at 1410 Hz and 1310 Hz. These chatter frequencies and harmonics are not apparent in Figure 14(a) but the tooth passing frequency and harmonics have the same mode of excitation.

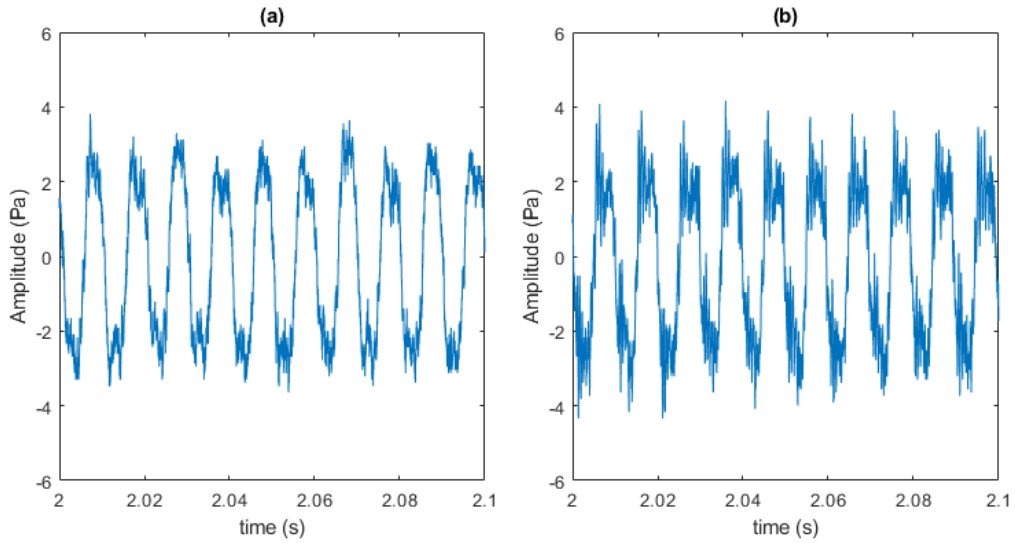


**Figure 13: (a)** A stable cut at 3000 RPM and 0.20in DOC compared with **(b)** an unstable cut done at 3000 RPM and 0.25in DOC

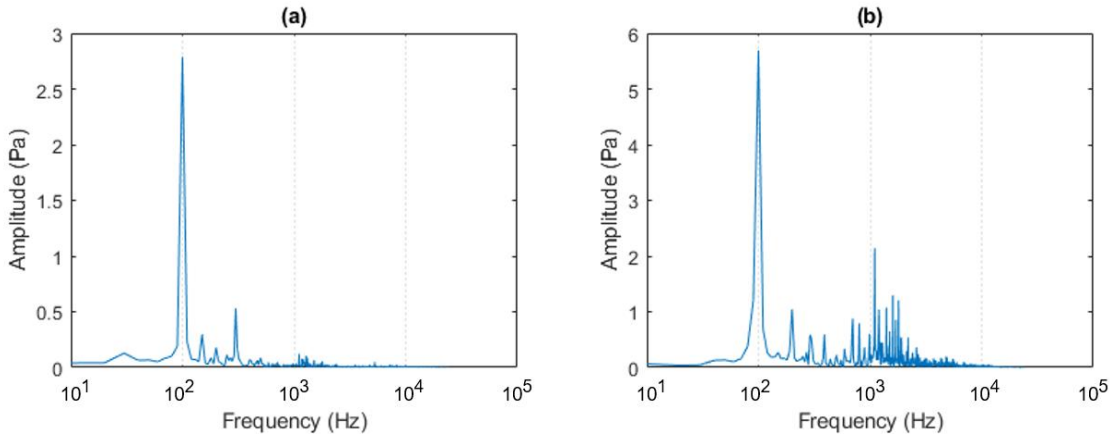


**Figure 14:** (a) FFT of a stable cut at 3000 RPM and 0.20in DOC compared with (b) an unstable cut done at 3000 RPM and 0.25in DOC

### 4.2 Data Augmentation

After segmentation, artificial noise is superimposed on the data by adding the artificial noise signal directly to the original signal. This superposition of noise on the

original audio signal represents the realistic superposition of two audio signals. The artificial noise has two components, a white Gaussian noise component, and a periodic component, which is modeled by a sine wave and its first 2 harmonics. 3 levels of white Gaussian noise were chosen to overlay to create signal to noise ratios of 20, 15, and 10 dBW. The periodic noise has a base frequency at 310 Hz and harmonics at 620 and 930 Hz. The base frequency of 310 Hz was chosen to stand out from the tooth passing frequencies and harmonics of the experimental cutting conditions. The periodic noise frequencies are limited to below 1000 Hz because of the attenuation of sound through air is larger at high frequencies [50]. Doing so allows the noise to pass through the filter with some attenuation. The base frequencies have the highest amplitude 2, 1 and 0.5 Pa. The first and second harmonics have amplitudes that are 30% and 20% of the base frequency amplitude. It is assumed that past the second harmonic, the amplitudes will decrease rapidly and thus have little effect on the original signal. Tables 4 summarizes the periodic and white Gaussian noise signals composition used for data augmentation into 3 noise levels.

**Table 4:** Noise signal composition at 3 levels

| White Gaussian Noise Level | SNR (dBW) |
|---|---|
| **1** | 20 |
| **2** | 15 |
| **3** | 10 |

| Periodic Noise Level | Frequency (Hz) | Amplitude (Pa) |
|---|---|---|
| **1** | 310 | 0.5 |
| | 620 | 0.15 |
| | 930 | 0.1 |
| **2** | 310 | 1 |
| | 620 | 0.3 |
| | 930 | 0.2 |
| **3** | 310 | 2 |
| | 620 | 0.6 |
| | 930 | 0.4 |

The time-series and spectra of these noise signals and the resulting superimposed signals are shown in Figure 15 and Figure 16. The levels for the periodic and white noise components were chosen to demonstrate the strengths and weaknesses of the methods to be compared. 16 different overlays were created using 3 levels of each noise component and the unadulterated signal.

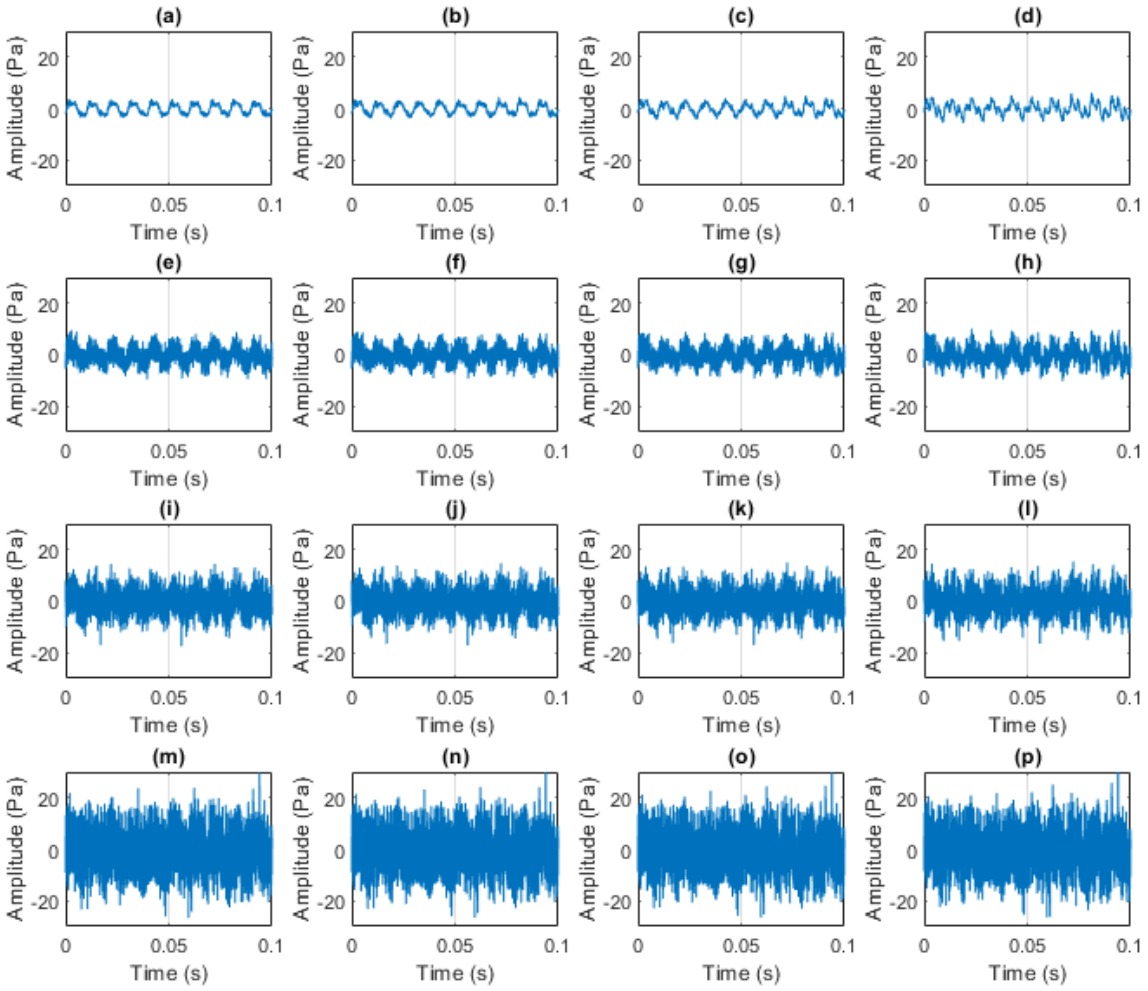


**Figure 15:** Time-series samples of increasing noise levels with **(a)** no noise **(b)** low periodic noise **(c)** medium periodic noise **(d)** high periodic noise **(e)** low Gaussian noise **(f)** low Gaussian and periodic noise **(g)** low Gaussian and medium periodic noise **(h)** low Gaussian and high periodic noise **(i)** medium Gaussian noise **(j)** medium Gaussian and low periodic noise **(k)** medium Gaussian and medium periodic noise **(l)** medium Gaussian and high periodic noise **(m)** high Gaussian noise **(n)** high Gaussian and low periodic noise **(o)** high Gaussian and medium periodic noise **(p)** high Gaussian and high periodic noise

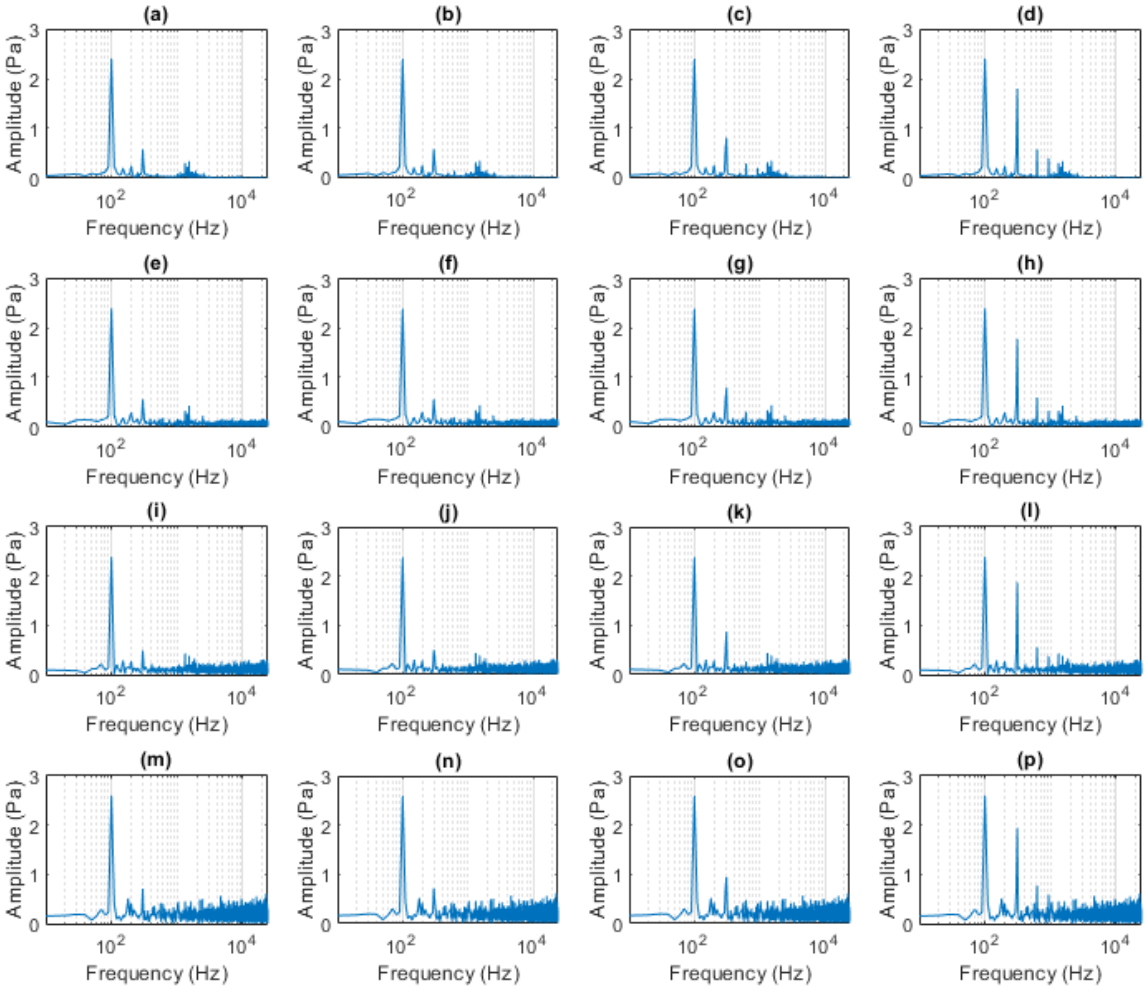


**Figure 16:** FFT samples of increasing noise levels with **(a)** no noise **(b)** low periodic noise **(c)** medium periodic noise **(d)** high periodic noise **(e)** low Gaussian noise **(f)** low Gaussian and periodic noise **(g)** low Gaussian and medium periodic noise **(h)** low Gaussian and high periodic noise **(i)** medium Gaussian noise **(j)** medium Gaussian and low periodic noise **(k)** medium Gaussian and medium periodic noise **(l)** medium Gaussian and high periodic noise **(m)** high Gaussian noise **(n)** high Gaussian and low periodic noise **(o)** high Gaussian and medium periodic noise **(p)** high Gaussian and high periodic noise

### 4.3 Thresholding Method and Results

The threshold method was tested on the original signals and the augmented signals at threshold values of 0.05, 0.1, and 0.15. First, a comb filter is created that filters out the tooth passing frequency, spindle runout frequency, and subsequent harmonics. This comb filter attenuates all chosen frequencies to 0. An example of the comb filter used is shown in Figure 17. This example is generated using a 100 Hz tooth passing frequency. To account for spindle runout, a similar comb filter is used. The FFT is filtered and then any remaining frequencies that have amplitudes above the set threshold are considered chatter. Figure 18 shows an example of the thresholding technique detecting chatter. The main tooth-passing and spindle runout frequency and harmonics were seen at 100, 150 and 200Hz. These are filtered out by the comb. The chatter is occurring in the 1000-1500Hz range, with a max amplitude of 0.6 Pa.

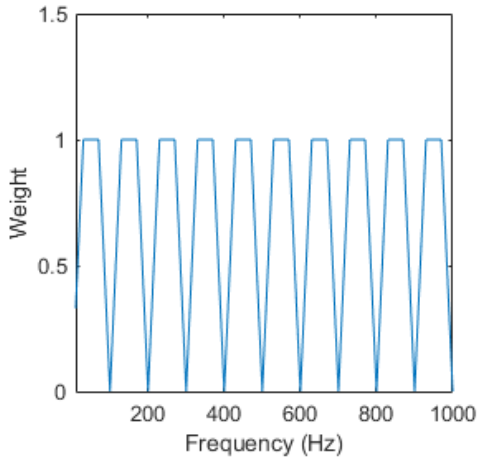


**Figure 17:** Sample comb filter with tooth passing frequency at 100 Hz

The results of the thresholding method on the original data are shown in Table 5. By lowering the threshold, false positives are more likely to occur because noise and frequencies not filtered out by the comb filter have a smaller barrier to overcome. 115 false positives were registered using a 0.05 Pa threshold as opposed to 0 when using a 0.1 Pa threshold. Conversely, choosing a higher threshold can result in missed detections, as shown in the table. Chatter frequencies that may not have fully developed will not be registered because the threshold is higher than their amplitude. A higher threshold makes a system more robust to noise but can lead to many missed detections which can be critical.

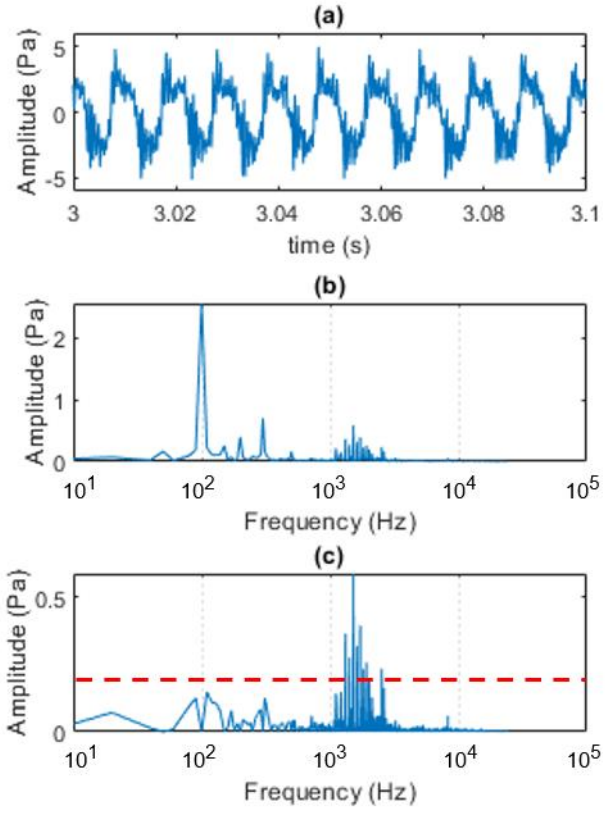


**Figure 18:** Thresholding method detecting chatter while cutting at 3000 RPM with 0.25 DOC. The threshold is set at 0.2 Pa.

**Table 5:** Evaluation of threshold method at 3 thresholds while classifying original audio samples

| Threshold | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy (%) |
|---|---|---|---|---|---|---|---|---|
| 0.05 Pa | 109 | 224 | 115 | 0 | 0.49 | 1.00 | 0.65 | 74.3 |
| 0.10 Pa | 224 | 224 | 0 | 0 | 1.00 | 1.00 | 1.00 | 100 |
| 0.15 Pa | 224 | 153 | 0 | 71 | 1 | 0.76 | 0.86 | 84.1 |

The thresholding method was done again with the augmented data set and its accuracies for each tier of noise from Table 3 is shown in Figure 19. When the threshold is 0.1Pa, an accuracy of 100% is obtained on the original signal set. When the threshold is 0.15Pa, the overall performance is better on the augmented data set. As shown, the higher threshold maintains accuracy at higher noise levels. This is because the threshold has been set above the noise floor. All threshold levels fail to classify chatter meaningfully at the highest noise levels. A 50% accuracy is due to the methods classifying every sample as chatter. Since the data are evenly distributed between chatter and no chatter, the methods have a 50% accuracy and a high false positive count.

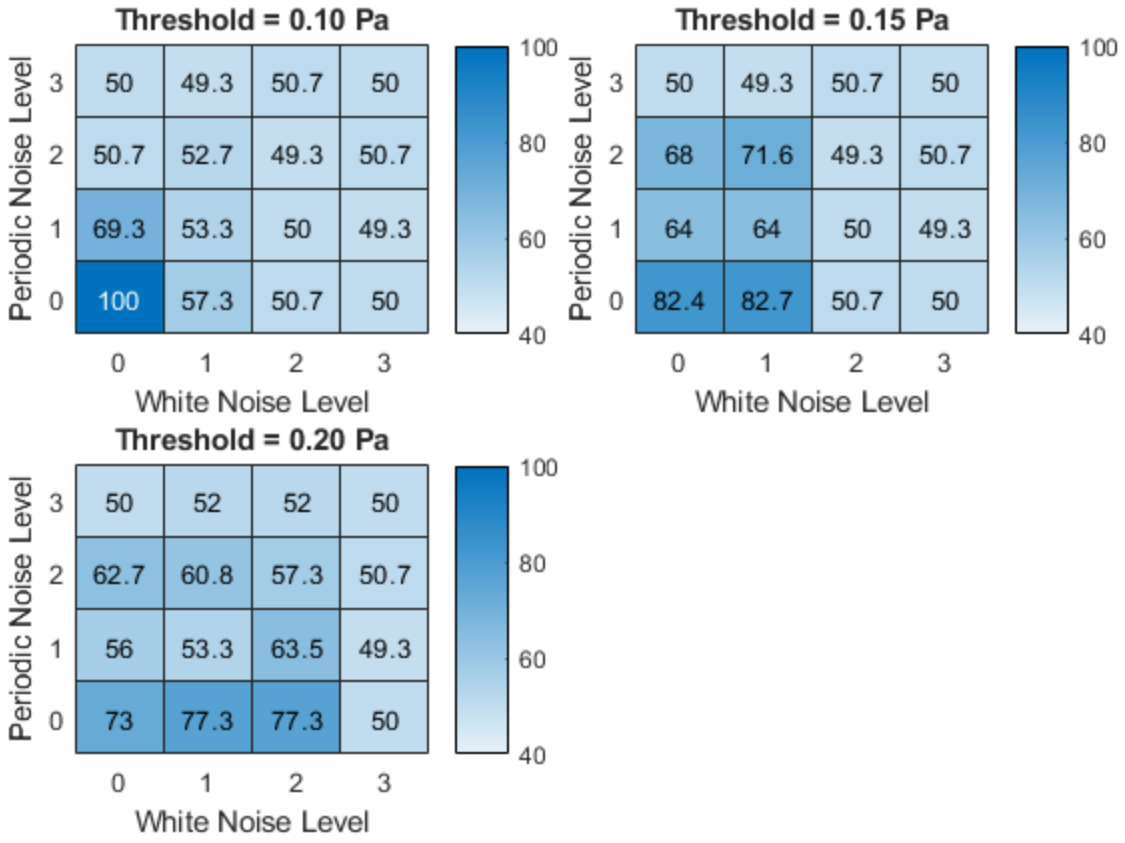


**Figure 19:** Accuracy of threshold method at 3 thresholds while classifying artificially noisy samples

### 4.4 Feature Extraction

Features must be extracted to train the ML classification models. 16 features are extracted from 0.1s segments of audio data from both the time and frequency domains. In the time domain 9 features are extracted: root mean square (RMS), variance (V), skewness (Sk), kurtosis (Ku), peak value (Pk), crest factor (CF), shape factor (SF), impulse factor (IF), and clearance factor (ClF). The time series signals are transformed into the frequency domain with an FFT. A Hanning window is used on the time series data before the FFT is taken to prevent spectral leakage. 7 features are extracted from the

frequency domain: mean amplitude ($M_f$), variance ($V_f$), skewness ($Sk_f$), kurtosis ($Ku_f$), peak amplitude ($Pk_f$), relative peak (RPk), and total harmonic bandpower (HBP). The equations for these features are shown in Table 6 where $x_i$ is the amplitude of a time signal at data point $i$ and $s_i$ is the amplitude of a frequency at $i$. The calculation of the HBP requires the tooth passing frequency, $f_t$.

### 4.5 Training and Validation for ML Approaches

A total of 7,168 samples were used for training and validation of ML classifiers. The number of samples were equally distributed between chatter and stable conditions. The samples were also evenly distributed between the 16 variants of noise, shown in Figure 15-16. Each sample had 16 features to be used for classification. The samples were split into a training set of 5,974 samples and a testing set of the remaining 1,194 samples. Both sets contained even distributions across chatter conditions, and noise variations. When training the models, a cross-validation was done with 5 folds validation to tune the parameters. Decision tree, SVM, kNN, and bagged tree models were trained, each with 4 different training sets. These training sets differed in their size and the noise levels they contained. A breakdown of the 4 training sets is shown in Table 7, with noise levels from Table 3. After the models were fully trained, the common testing set was used to measure their performance. An outline of the training and validation procedure is shown in Figure 20.

**Table 6: Equations for feature extraction**

| Time domain | Frequency Domain |
|---|---|
| $RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$ | $M_f = \frac{1}{n}\sum_{i=1}^{n} s_i$ |
| $V = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}$ | $V_f = \frac{\sum_{i=1}^{n}(s_i - M_f)}{n-1}$ |
| $Sk = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - M)^3}{\sigma^3}$ | $Sk_f = \frac{1}{n}\frac{\sum_{i=1}^{n}(s_i - M_f)^3}{\sigma^3}$ |
| $Ku = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - M)^4}{\sigma^4}$ | $Ku_f = \frac{1}{n}\frac{\sum_{i=1}^{n}(s_i - M_f)^4}{\sigma^4}$ |
| $Pk = \frac{1}{2}[max(x) - min(x)]$ | $Pk_f = max(s)$ |
| $CF = \frac{Pk}{RMS}$ | $RPk = \frac{Pk_f}{M_f}$ |
| $SF = \frac{RMS}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert}$ | $HBP = \sum_{i=f_t}^{Nf_t} s_i, i = f_t, 2f_t, ..., Nf_t$ |
| $IF = \frac{Pk}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert}$ | |
| $ClF = \frac{Pk}{(\frac{1}{n}\sum_{i=1}^{n}\lvert x_i\rvert)^2}$ | |

**Table 7:** Training set noise level composition

| Training Set | White Noise Levels | Periodic Noise Levels |
|---|---|---|
| 1 | 0 | 0 |
| 2 | 0, 1 | 0,1 |
| 3 | 0,1,2 | 0,1,2 |
| 4 | 0,1,2,3 | 0,1,2,3 |

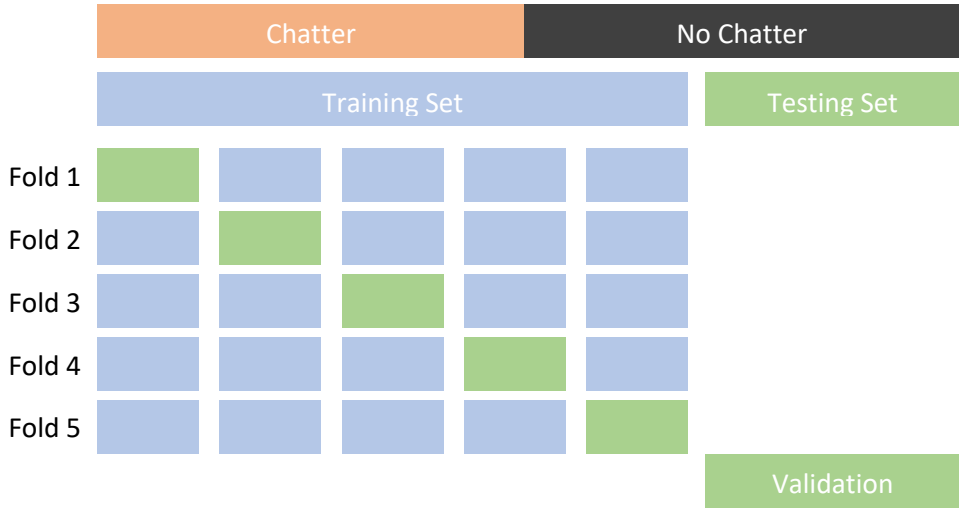


**Figure 20:** Cross-validation with 5 folds and withheld testing set

Models were first trained with 50%, 60%, 70%, 80%, 90%, and 100% of the total training set (training set 4 from Table 7). These training sets had even distributions of noise levels and chatter conditions. The trained models were then validated on the withheld testing set. The results of this testing are shown in Figure 22. The k-Nearest Neighbors (kNN) approach had the highest accuracy of the 4 models regardless of training set size. The performance of the kNN increased as the training set size increased from 91.8% to 94.1%. The performance of the Bagged Trees and the SVM remained

consistent at all training set sizes. kNN, SVM, and Bagged Tree models show capabilities of acceptable accuracy even with limited training. The decision tree model had the worst overall accuracy at all training set sizes. The decision tree's performance declined as it was exposed to more training data. This is an indicator that the decision tree model is overfitting on the training data, resulting in poor generalization.

The performance of the models when trained on 100% data is shown in Table 8. The performance was evaluated on subsets of the testing data where the subsets contained noise at certain levels, as well as the whole testing set. The kNN model once again had the highest accuracy through all the noise level subsets, as well as the lowest occurrence of false positives. The Bagged Tree model has the lowest occurrence of false negatives. The SVM has the highest rate of false positives, and the decision tree had the highest rate of false negatives. As noise increased, all models suffered from decreased precision, but the recall rates remained consistent. Going from noise level 1 to noise level 3, the F1 scores are not changing, indicating a balanced performance between false positives and negatives.

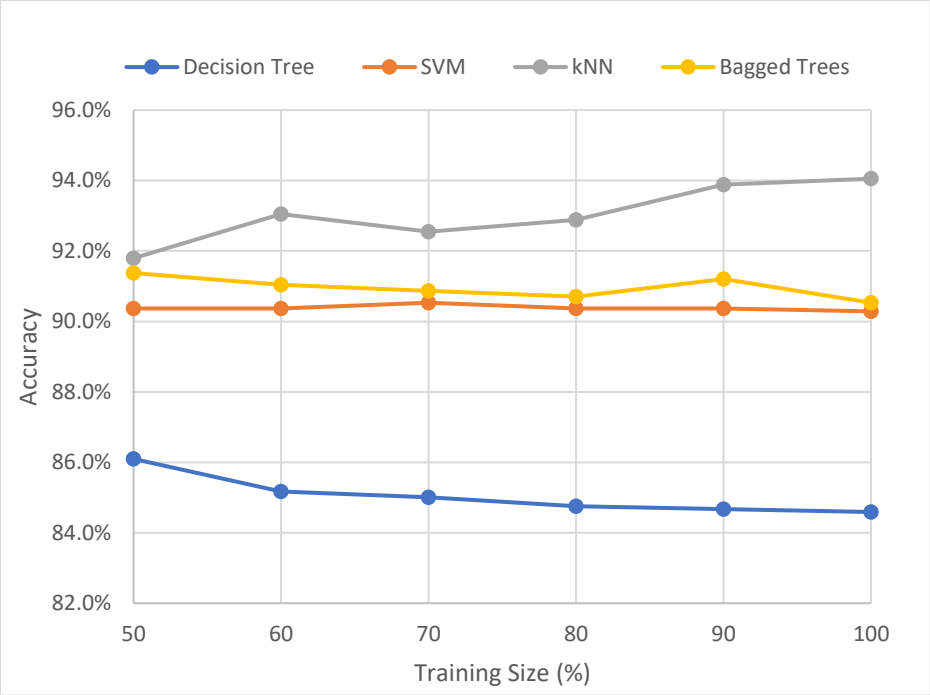


**Figure 21:** Accuracy of classifiers against the testing set while increasing the number of training data samples from 50% to 100% of all training data

The SVM model was trained on withheld subsets of the training data (training sets 1-4). 4 models were trained with data containing only certain noise levels to evaluate its performance in unknown conditions. The results are shown in Figure 22. The accuracy of the various models extends beyond the data they have been trained on. In Figure 22(a), the SVM has improved accuracy at periodic noise levels 1, 2 and 3 despite not having encountered these sets in training. Similar results are shown in Figure 22(b) at levels 2 and 3. The SVM does not handle white Gaussian noise well, without training. In Figure 22(a-c), the levels of white Gaussian noise that the model was not exposed to were not classified accurately. At white noise level 3, all models except for the fully trained model Figure 22(d) exhibit poor performance. Performance generally improved with the

addition of noisier training data, with some drops in performance that can be seen in Figure 22(d). In Figure 22(d) at lower noise levels, the accuracy is lower than models shown in Figure 22(b) and Figure 22(c), but the performance is still acceptable. As the training data becomes noisier and broader, specific accuracy dropped in places, but overall performance improved.

**Table 8:** Evaluation of fully trained ML models compared to threshold technique against subsets of the testing data and the complete testing set. (**Denotes p<0.05 when compared to the threshold method using McNemar's test*)

Original Signal

| Model | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Threshold = 0.15 | 24 | 37 | 0 | 13 | 1.00 | 0.65 | 0.79 | 82.4% |
| Tree | 19 | 37 | 0 | 18 | 1.00 | 0.51 | 0.68 | 75.7% |
| SVM | 37 | 37 | 0 | 0 | 1.00 | 1.00 | 1.00 | 100.0% |
| kNN | 37 | 37 | 0 | 0 | 1.00 | 1.00 | 1.00 | 100.0% |
| Bagged Trees | 36 | 37 | 0 | 1 | 1.00 | 0.97 | 0.99 | 98.6% |

Noise Level 1

| Model | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Threshold = 0.15 | 28 | 20 | 17 | 10 | 0.62 | 0.74 | 0.67 | 64.0% |
| Tree | 27 | 37 | 0 | 11 | 1.00 | 0.71 | 0.83 | 85.3% |
| SVM | 36 | 34 | 3 | 2 | 0.92 | 0.95 | 0.94 | 93.3% |
| kNN | 36 | 35 | 2 | 2 | 0.95 | 0.95 | 0.95 | 94.7% |
| Bagged Trees | 37 | 36 | 1 | 1 | 0.97 | 0.97 | 0.97 | 97.3% |

Noise Level 2

| Model | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Threshold = 0.15 | 37 | 0 | 38 | 0 | 0.49 | 1.00 | 0.66 | 49.3% |
| Tree | 34 | 29 | 9 | 3 | 0.79 | 0.92 | 0.85 | 84.0% |
| SVM | 32 | 30 | 8 | 5 | 0.80 | 0.86 | 0.83 | 82.7% |
| kNN | 32 | 32 | 6 | 5 | 0.84 | 0.86 | 0.85 | 85.3% |
| Bagged Trees | 32 | 30 | 8 | 5 | 0.80 | 0.86 | 0.83 | 82.7% |

Noise Level 3

| Model | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Threshold = 0.15 | 37 | 0 | 37 | 0 | 0.50 | 1.00 | 0.67 | 50.0% |
| Tree | 36 | 28 | 9 | 1 | 0.80 | 0.97 | 0.88 | 86.5% |
| SVM | 36 | 31 | 6 | 1 | 0.86 | 0.97 | 0.91 | 90.5% |
| kNN | 33 | 36 | 4 | 1 | 0.89 | 0.97 | 0.93 | 93.2% |
| Bagged Trees | 36 | 29 | 8 | 1 | 0.82 | 0.97 | 0.89 | 87.8% |

All Data

| Model | TP | TN | FP | FN | Precision | Recall | F1 | Accuracy |
|---|---|---|---|---|---|---|---|---|
| Threshold = 0.15 | 554 | 142 | 455 | 43 | 0.55 | 0.93 | 0.69 | 58.3% |
| Tree* | 495 | 515 | 82 | 102 | 0.86 | 0.83 | 0.84 | 84.6% |
| SVM* | 567 | 511 | 86 | 30 | 0.87 | 0.95 | 0.91 | 90.3% |
| kNN* | 570 | 553 | 44 | 27 | 0.93 | 0.95 | 0.94 | 94.1% |
| Bagged Trees* | 575 | 515 | 82 | 22 | 0.88 | 0.96 | 0.92 | 91.3% |

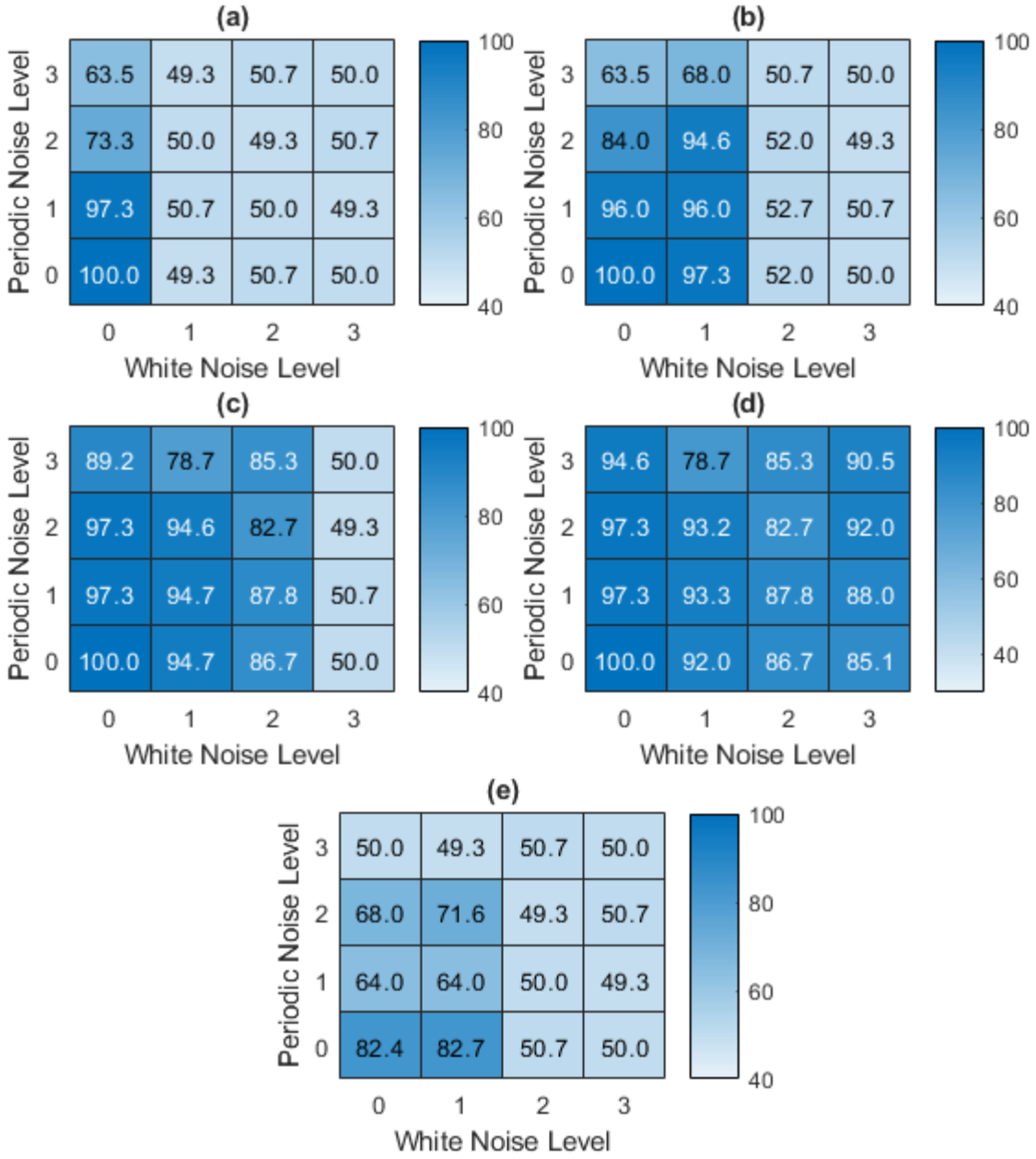


**Figure 22:** Accuracy of the SVM (a-d) compared to threshold technique (e) at the subset level when trained on withheld training sets. (a) was trained with training set 1. (b) was trained with training set 2. (c) was trained with training set 3. (d) was trained with training set 4. (e) is the threshold technique with a threshold of 0.15 Pa.

# CHAPTER 5: DISCUSSION

## 5.1 Comparing Threshold and ML Methods

When compared to the models that have been fully trained or partially trained on noisy data, the threshold method shows considerable weakness at higher periodic and white noise levels. In Figure 22(e), the threshold method begins to deteriorate at periodic noise level 1 and white noise level 2. This drop in accuracy is the result of an increased number of false positives. The SVMs in Figure 22(a) and Figure 22(b) show higher accuracies when classifying signals with periodic noise level 1,2 and 3 despite no exposure to these noise levels in training. The threshold method does show better performance when compared to the SVM in Figure 22(a) with no exposure to noisy data when increasing white noise level. As noise increases past noise level 1, the threshold method starts classifying a much higher level of false positives compared to the ML approaches. The F1 score and the precision of the threshold method decrease consistently with increased noise, unlike the ML approaches. Overall, the threshold method shows acceptable performance at low noise levels, but setting the appropriate threshold based on the expected noise level will improve performance. The ML approaches, when exposed to a varied dataset show robustness to noisy data. All methods show deterioration in performance at the highest noise levels, and an increased occurrence of false positives. The occurrence of false negatives was lower for all models except the decision tree classifier.

Overall, all ML methods showed significant performance improvements compared to the threshold method, shown in Table 8. ML models have the capability to

handle a noisier data set when adequately trained. Calculations of significance are shown in Appendix A.

### 5.2 Misclassifications

When comparing all the approaches, there were several samples that failed to be classified accurately by any method. Two examples of these samples are shown in Figure 24. One case is marginal chatter in Figure 23(a) which is exhibited at 6000 RPM at a DOC of 0.6in. The contributions of the marginal chatter frequency are very weak, and when increasing noise is added, the contributions of the chatter frequency to the extracted features becomes less valuable. This results in a missed detection. The other example is a frequently false positive that occurs at 3000 RPM and DOC of 0.2in. This is likely because of its similarity to the unstable condition of 3000RPM and DOC of 0.25in. When noise begins to increase the amplitude of the higher frequencies where chatter would occur, the spectrum becomes like the chatter condition.

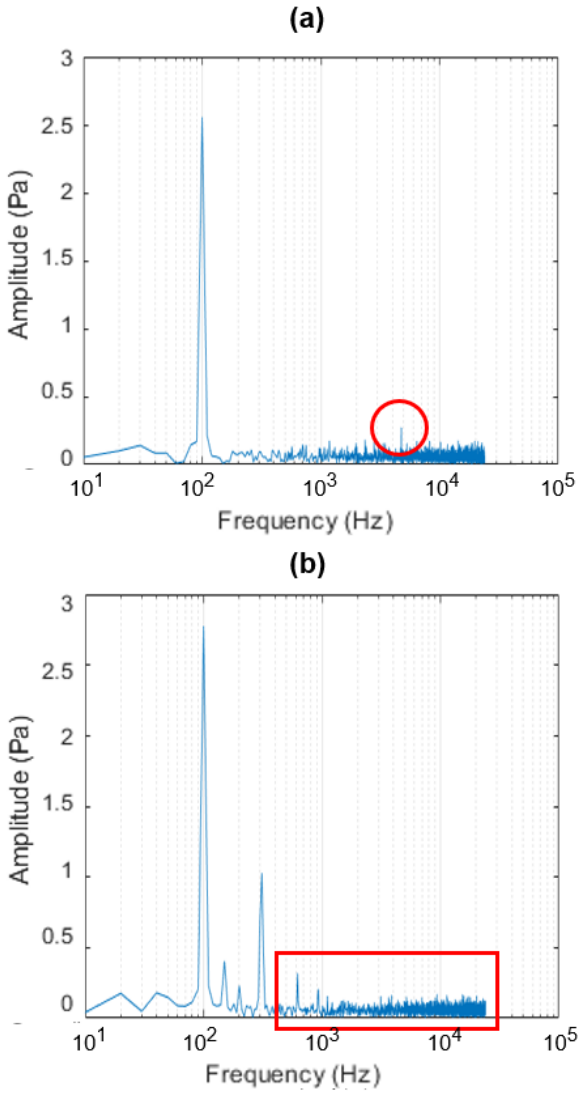


**Figure 23: (a)** Frequently false negative (missed detection). Occurs when spindle speed is 6000 RPM and DOC is 0.6in and signal is sufficiently noisy. **(b)** Frequently false positive. Occurs when spindle speed is 3000 RPM and DOC is 0.2in and signal is sufficiently noisy

## 5.3 Assumptions and Limitations

It was assumed that the spindle RPM commanded was equivalent to the actual spindle speed. This was checked by visual inspection during each cutting operation by observing the

CNC controller's outputted spindle speed, and further verified by taking FFTs with long windows. The tooth cutting frequency aligned with eh commanded spindle RPM. If the real spindle RPM and the assumed spindle RPM were different, this would have a negative effect on the filtering techniques used for the thresholding method, leading to a larger number of false positives.

It was assumed that the cutting tool was in a steady condition throughout the experiment. Changes in the wear and geometry of the cutting tool were assumed to be negligible. These changes would potentially affect the raw audio signal generated by the cutting process, as observed in several tool condition monitoring studies [1, 2, 10, 11] The tool condition was checked by visual inspection before each cutting operation.

The findings of this study are currently limited to this machine and these cutting parameters. It is unknown if the same models could be applied to other machines without new training data. It is also uncertain whether the models trained could detect chatter at frequencies other than the ones it was trained on, although past work indicates this should not be a major issue because of the general consistency of features for detecting chatter [7, 22, 23, 28].

# CHAPTER 6: CONCLUSION

The purpose of this thesis was to understand the balance between sensitivity, speed of implementation, and performance of various classification systems under noisy conditions. 4 different classifiers were trained with 9 different noisy training sets. Their performance was compared to each other and to a common thresholding technique [22, 23]. The threshold technique's performance suffered under heavier noise. With the original clean signals, the threshold method showed accurate results depending on the set threshold. The 4 classifiers chosen performed accurately when trained on the noisy data, with accuracies ranging from 84.6% to 94.1%. This performance is in line with other ML chatter classifiers [28, 29] The SVM demonstrated an ability to classify noise in the testing set that it was not exposed to during the training phase. The SVM was also able to retain accuracy on a limited training set. Both capabilities demonstrate that ML approaches, and specifically SVMs have a robustness to unexpected noise. The performance of the SVM on a noisy factory floor may remain high even when trained on less noisy data.

## 6.1 Contributions

The effects of periodic and white noise on audio signals used for classifying chatter had not been previously studied. An established technique used for classifying chatter [22, 23] with a set threshold was examined under increasingly noisy conditions. The study showed that thresholding techniques had an increased false positive rate with the addition of periodic and white noise. Comparatively, ML approaches showed some

levels of robustness by maintaining their performance even in the presence of excessive noise.

Data augmentation was used to add diversity to training sets. This added diversity led to improved performance under noisy conditions. Artificial noise has the possibility of increasing the size of data sets and improving the robustness of ML systems that may experience noisy data during operation. This could reduce the need for real and diversified data collection and reduce the dataset size needed for implementation of ML models.

### 6.2 Future Work

One major benefit of using threshold techniques is that the system is easily transferable to another machining system. The same cannot always be said of ML models like the ones studied in this thesis. One major area of work is the study of transferable knowledge for these ML models. This would involve gathering data from multiple machines, multiple set ups, and multiple cutting operations. This could be done more easily with a well-curated open-source database with cutting signals from many different set ups. This database could potentially accelerate the development of ML models in this area for many different applications such as tool wear, dimensional conformance, surface roughness, and tool breakage.

Improved feature selection is another area of future work. Features in this thesis were selected based on past literature. However, it would be worthwhile to extract many more features from the audio data as well as other sources and methodically sort out which features have a meaningful effect on the performance of the classifiers. This could

be done either by looking at the end performance of the classifier or by using statistical approaches like the SOF or correlation coefficient.

As manufacturing operations become increasingly digitized, the development of algorithms on edge devices for machine monitoring has gained interest. There is potential to transfer any classifier method to an edge computing device such as the Beaglebone Black, the more powerful Beaglebone AI, or even a cellphone. Newer devices have lowered the barrier to entry to these machine monitoring systems by allowing more computing power. For this to happen, however, these classifiers must be quick and easy to implement with little necessary training. Mobile applications would have to be able to perform well in many non-ideal situations as placement, environment, and operational variables would be different in most cases.

# APPENDIX A

Calculations of significance were done with McNemar's test with an $\alpha = 0.05$.

Example calculation of chi-square value:

**Table 9:** Example contingency table

| | Correct | Incorrect |
|---|---|---|
| **Correct** | A | B |
| **Incorrect** | C | D |

$$X^2 = \frac{(B - C)^2}{(C + B)}$$

**Table 10:** Contingency table of Decision Tree model compared to threshold method

| | | Threshold = 0.15 Pa | |
|---|---|---|---|
| | | **Correct** | **Incorrect** |
| **Tree** | **Correct** | 624 | 386 |
| | **Incorrect** | 72 | 112 |

$X^2 = 215.28$, $p < 0.00001$

**Table 11:** Contingency table of SVM model compared to threshold method

| | | Threshold = 0.15 Pa | |
|---|---|---|---|
| | | **Correct** | **Incorrect** |
| **SVM** | **Correct** | 664 | 414 |
| | **Incorrect** | 32 | 84 |

$X^2 = 327.18$, $p < 0.00001$

**Table 12:** Contingency table of kNN model compared to threshold method

| | | Threshold = 0.15 Pa | |
|---|---|---|---|
| | | Correct | Incorrect |
| kNN | Correct | 668 | 455 |
| | Incorrect | 28 | 43 |

$X^2 = 377.49$, $p < 0.00001$

**Table 13:** Contingency table of Bagged Tree model compared to threshold method

| | | Threshold = 0.15 Pa | |
|---|---|---|---|
| | | Correct | Incorrect |
| Bagged Tree | Correct | 675 | 417 |
| | Incorrect | 21 | 81 |

$X^2 = 358.03$, $p < 0.00001$

## REFERENCES


[1] Liang, S. Y., Hecker, R. L., & Landers, R. G. (2004). Machining process monitoring and control: the state-of-the-art. *J. Manuf. Sci. Eng.*, *126*(2), 297-310.

[2] Quintana, G., & Ciurana, J. (2011). Chatter in machining processes: A review. *International Journal of Machine Tools and Manufacture*, *51*(5), 363-376.

[3] Coleman, C., Damodaran, S., Chandramouli, M., Deuel E. (2017, May). *Making maintenance smarter*. Deloitte. Retrieved March 15, 2020, from https://www2.deloitte.com/us/en/insights/focus/industry-4-0/using-predictive-technologies-for-asset-maintenance.html

[4] Tlusty, J., & Tarng, Y. S. (1988). Sensing cutter breakage in milling. *CIRP Annals*, *37*(1), 45-51.

[5] Toh, C. K. (2004). Vibration analysis in high speed rough and finish milling hardened steel. *Journal of Sound and Vibration*, *278*(1-2), 101-115.

[6] Soliman, E., & Ismail, F. (1997). Chatter detection by monitoring spindle drive current. *The International Journal of Advanced Manufacturing Technology*, *13*(1), 27-34.

[7] Hynynen, K. M., Ratava, J., Lindh, T., Rikkonen, M., Ryynänen, V., Lohtander, M., & Varis, J. (2014). Chatter detection in turning processes using coherence of acceleration and audio signals. *Journal of Manufacturing Science and Engineering*, *136*(4).

[8] Shi, F., Cao, H., Zhang, X., & Chen, X. (2020). A Reinforced k-Nearest Neighbors Method with Application to Chatter Identification in High Speed Milling. *IEEE Transactions on Industrial Electronics*.

[9] Martin, K. F. (1994). A review by discussion of condition monitoring and fault diagnosis in machine tools. *International Journal of Machine Tools and Manufacture*, *34*(4), 527-551.

[10] Teti, R., Jemielniak, K., O'Donnell, G., & Dornfeld, D. (2010). Advanced monitoring of machining operations. *CIRP annals*, *59*(2), 717-739.

[11] Bhuiyan, M. S. H., & Choudhury, I. A. (2014). 13.22—Review of sensor applications in tool condition monitoring in machining. *Comprehensive Materials Processing*, *13*, 539-569.

[12] Snr, D. E. D. (2000). Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. *International Journal of Machine Tools and Manufacture*, *40*(8), 1073-1098.

[13] Byrne, G., Dornfeld, D., Inasaki, I., Ketteler, G., König, W., & Teti, R. (1995). Tool condition monitoring (TCM)—the status of research and industrial application. *CIRP annals*, *44*(2), 541-567.

[14] Binsaeid, S., Asfour, S., Cho, S., & Onar, A. (2009). Machine ensemble approach for simultaneous detection of transient and gradual abnormalities in end milling using multisensor fusion. *Journal of Materials Processing Technology*, *209*(10), 4728-4738.

[15] Chen, J. C., & Huang, B. (2003). An in-process neural network-based surface roughness prediction (INN-SRP) system using a dynamometer in end milling operations. *The International Journal of Advanced Manufacturing Technology*, *21*(5), 339-347.

[16] Tarng, Y. S., and Li, T. C. (1994). "Detection and Suppression of Drilling Chatter." ASME. *J. Dyn. Sys., Meas., Control*. December 1994; 116(4): 729–734.

[17] National Instruments. (2019, March 5). *Understanding FFTs and Windowing.* https://www.ni.com/en-us/innovations/white-papers/06/understanding-ffts-and-windowing.html

[18] The SciPy community. (2018, May 5). *Scipy.signal.windows.hann.* SciPy. Retrieved May 1, 2020 from https://docs.scipy.org/doc/scipy-1.1.0/reference/generated/scipy.signal.windows.hann.html

[19] Austerlitz, H. (2002). *Data acquisition techniques using PCs*. Academic press.

[20] Guo, Y. B., & Ammula, S. C. (2005). Real-time acoustic emission monitoring for surface damage in hard machining. *International Journal of Machine Tools and Manufacture*, *45*(14), 1622-1627.

[21] Tangjitsitcharoen, S. (2009). In-process monitoring and detection of chip formation and chatter for CNC turning. *Journal of Materials Processing Technology*, *209*(10), 4682-4688.

[22] Delio, T., Tlusty, J., and Smith, S. (May 1, 1992). "Use of Audio Signals for Chatter Detection and Control." ASME. *J. Eng. Ind*. May 1992; 114(2): 146–157.

[23] Smith, S., & Delio, T. (1989, December). Sensor-based control for chatter-free milling by spindle speed selection. In *Symposium on Control Issues in Manufacturing* (Vol. 18, pp. 107-114).

[24] Marinescu, I., & Axinte, D. A. (2008). A critical analysis of effectiveness of acoustic emission signals to detect tool and workpiece malfunctions in milling

operations. *International Journal of Machine Tools and Manufacture*, *48*(10), 1148-1160.

[25] Thaler, T., Potočnik, P., Bric, I., & Govekar, E. (2014). Chatter detection in band sawing based on discriminant analysis of sound features. *Applied acoustics*, *77*, 114-121.

[26] Yoon, M. C., & Chin, D. H. (2005). Cutting force monitoring in the endmilling operation for chatter detection. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *219*(6), 455-465.

[27] Scheffer, C., & Heyns, P. S. (2004). An industrial tool wear monitoring system for interrupted turning. *Mechanical Systems and Signal Processing*, *18*(5), 1219-1242.

[28] Lamraoui, M., Barakat, M., Thomas, M., & Badaoui, M. E. (2015). Chatter detection in milling machines by neural network classification and feature selection. *Journal of Vibration and Control*, *21*(7), 1251-1266.

[29] Yao, Z., Mei, D., & Chen, Z. (2010). On-line chatter detection and identification based on wavelet and support vector machine. *Journal of Materials Processing Technology*, *210*(5), 713-719.

[30] Ademujimi, T. T., Brundage, M. P., & Prabhu, V. V. (2017, September). A review of current machine learning techniques used in manufacturing diagnosis. In *IFIP International Conference on Advances in Production Management Systems* (pp. 407-415). Springer, Cham.

[31] Siddhpura, M., & Paurobally, R. (2012). A review of chatter vibration research in turning. *International Journal of Machine tools and manufacture*, *61*, 27-47.

[32] Le Lan, J. V., Marty, A., & Debongnie, J. F. (2006). A stability diagram computation method for milling adapted to automotive industry.

[33] Altintaş, Y., & Budak, E. (1995). Analytical prediction of stability lobes in milling. *CIRP annals*, *44*(1), 357-362.

[34] Altintas, Y., & Ber, A. A. (2001). Manufacturing automation: metal cutting mechanics, machine tool vibrations, and CNC design. *Appl. Mech. Rev.*, *54*(5), B84-B84. Manufacturing automation metal cutting mechanics, Machine too. Vibrations, and CNC Design

[35] Wiercigroch, M., & Budak, E. (2001). Sources of nonlinearities, chatter generation and suppression in metal cutting. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *359*(1781), 663-693.

[36] Liao, Y. S., & Young, Y. C. (1996). A new on-line spindle speed regulation strategy for chatter control. *International Journal of Machine Tools and Manufacture*, *36*(5), 651-660.

[37] Tansel, I. N., Li, M., Demetgul, M., Bickraj, K., Kaya, B., & Ozcelik, B. (2012). Detecting chatter and estimating wear from the torque of end milling signals by using Index Based Reasoner (IBR). *The International Journal of Advanced Manufacturing Technology*, *58*(1-4), 109-118.

[38] Faassen, R. P. H., Doppenberg, E. J. J., van de Wouw, N., Oosterling, J. A. J., & Nijmeijer, H. (2006). Online detection of the onset and occurrence of machine tool chatter in the milling process. In *CIRP 2nd International Conference on High Performance Cutting* (pp. paper-no).

[39] Suprock, C. A., Fussell, B. K., Hassan, R. Z., & Jerard, R. B. (2008, January). A low cost wireless tool tip vibration sensor for milling. In *ASME 2008 International Manufacturing Science and Engineering Conference collocated with the 3rd JSME/ASME International Conference on Materials and Processing* (pp. 465-474). American Society of Mechanical Engineers Digital Collection.

[40] Vela-Martínez, L., Jáuregui-Correa, J. C., & Álvarez-Ramírez, J. (2009). Characterization of machining chattering dynamics: An R/S scaling analysis approach. *International Journal of Machine Tools and Manufacture*, *49*(11), 832-842.

[41] Zhang, C. L., Yue, X., Jiang, Y. T., & Zheng, W. (2010). A hybrid approach of ANN and HMM for cutting chatter monitoring. In *Advanced Materials Research* (Vol. 97, pp. 3225-3232). Trans Tech Publications Ltd.

[42] Schmitz, T. L. (2003). Chatter recognition by a statistical evaluation of the synchronously sampled audio signal. *Journal of Sound and Vibration*, *262*(3), 721-730.

[43] Tsai, N. C., Chen, D. C., & Lee, R. M. (2010). Chatter prevention for milling process by acoustic signal feedback. *The International Journal of Advanced Manufacturing Technology*, *47*(9-12), 1013-1021.

[44] Kuljanic, E., Sortino, M., & Totis, G. (2008). Multisensor approaches for chatter detection in milling. *Journal of Sound and Vibration*, *312*(4-5), 672-693.

[45] Subrahmanya, N., & Shin, Y. C. (2008). Automated sensor selection and fusion for monitoring and diagnostics of plunge grinding. *Journal of manufacturing science and engineering*, *130*(3).

[46] Manufacturing Laboratories Inc. (n.d.). *Harmonizer.* Retrieved May 1, 2020 from https://mfg-labs.weebly.com/store/p40/Harmonizer%C2%AE.html

[47] OKUMA. (n.d.). *Machining Navi | Intelligent Technology*. Retrieved May 1, 2020 from https://www.okuma.co.jp/english/onlyone/process/

[48] Ismail, F., & Ziaei, R. (2002). Chatter suppression in five-axis machining of flexible parts. *International Journal of Machine Tools and Manufacture*, *42*(1), 115-122.

[49] Ismail, F., & Ziaei, R. (2000). Monitoring machining chatter using acoustic intensity. In *World Automation Congress Conference, Maui, Hawaii, USA*.

[50] Guarnaccia, C., Quartieri, J., & Ruggiero, A. (2014). Acoustical noise study of a factory: Indoor and outdoor simulations integration procedure. *International Journal of Mechanics*, *8*(1), 298-306.

[51] Piercy, J. E., Embleton, T. F., & Sutherland, L. C. (1977). Review of noise propagation in the atmosphere. *The Journal of the Acoustical Society of America*, *61*(6), 1403-1418.