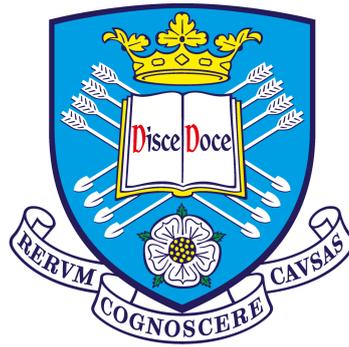


# Automated Testing of Advanced Cutting Tool Materials



A Thesis submitted to the University of Sheffield for the degree of Doctor of  
Engineering in the Faculty of Engineering

by

Chandula T. Wickramarachchi

Department of Mechanical Engineering

University of Sheffield

October 2019



---

## ACKNOWLEDGEMENTS

It has been a pleasure to work on this project for the last few years and none of it would have been possible without the brilliant team I have around me.

Firstly I would like to thank my academic supervisor Dr. Elizabeth Cross for her unwavering support, encouragement and for being an incredible role model. I would also like to thank Professor Keith Worden and the Dynamics Research Group for the inspiration, for teaching me how to be an effective researcher and for introducing me to the world of academia. Special thanks to Dr. Timothy Rogers for his help with the Gaussian process and the Dirichlet process codes used in this work and for the countless meetings to help me understand the maths. I would also like to thank Dr. Ramon Fuentes for his patience when I bothered him with all my questions and also to Dr. Nathan Ray for encouraging me to do this EngD. The DRG is an excellent group to work for due to the high standards of research, but also because of the wonderful group of people. Therefore, I would like to thank Nikos, Tom Gibbons, Paul Gardner, Julian, Kartik, Lawrence, Tim Rooker, Aidan, Marcus, Matt and the rest of the group for helping to create a joyful and supportive place of work. I look forward to the next two years with you all.

I would like to thank Dr. Thomas McLeay for always making sure I was on the right track and being my AMRC supervisor. Dr. Sabino Ayvar-Soberanis, Dr. Jon Stammers and the AMRC has helped immensely with the technical and machining aspects of this work, and for that, I am truly grateful.

This project would not be possible without the sponsorship of Element Six Ltd. and

the support of Dr. Wayne Leahy. It has been an absolute pleasure working with such an established company that value industrial research. I would also like to thank E6 for providing the tools, workpieces and machining time to collect the dataset for this thesis. Special thanks goes to Chris, Luiz, Manuella, John and the technicians team for their support during the machining trials.

The IDC in Machining Science has also been an amazing group to work with, providing support for young researchers to bridge the gap between industry and academia. I would like to thank Clare Clarke, Dr. Francesca Breeden and Dr. Matt Marshall for all the hard work and hours poured into this group. Special thanks to Huseyin, Sam, Máté and Lisa for being there since day one. I would like to thank the EPSRC as this project was part funded by the grant (EP/I01800X/1).

I can honestly say I have had a fantastic four years working and living in Sheffield and it's all thanks to my wonderful friends, Dutty, Sophie, Alex P, Cariad, Simon, Jed, Alex M, Ruth and Dave. A very special thanks to Dr. Julia Carrell for being an incredible friend, for helping me with work and for being my therapist. I would also like to thank Janay, my biggest fan.

My family has always pushed me to do my best and for that, I am truly thankful. Without Amma, Appachchi, Akka and Chooty, none of this would be possible. Special thanks goes to Nanda and the extended family as well as Jan, Dave, Jack and Sophie for their kindness, support and belief in me.

Finally, I would like to thank Sam. This is my proudest achievement and I am so grateful to have you beside me. Thank you for your love, encouragement, and most importantly, for making sure I enjoyed every moment of it.

---

# ABSTRACT

Tool condition monitoring is an area of research where the main focus is predicting time to tool failure. Ultimately, the goal in the manufacturing industry is to establish a production line with a predictive maintenance strategy in order to reduce waste. For the industrial partner and specialised tool materials developer Element Six Ltd., the aim is to predict tool wear in order to increase the speed at which new tool material grades are tested. Owing to the complex nature of machining where numerous parameters affect the characteristics of tool wear, prediction has been a challenging pursuit. For instance, cutting conditions, workpiece and tool material compositions, machining vibrations, forces and temperatures inside the machine are some of the many factors that can influence a tool's behaviour. This makes tool wear predictions extremely difficult. Another crucial challenge in tool condition monitoring is the availability of descriptive labels of tool wear states. It is not possible to directly acquire damage labels whilst the tool is in use. Consequently, collecting machining datasets is expensive and time consuming as the process must be interrupted to obtain damage labels.

It is the aim of this thesis to address these challenges in four steps. Firstly, the experimental procedure used to collect acoustic emission data and tool wear scans from a set of turning tests is detailed. By replicating the procedure already in place at Element Six Ltd., it is possible to ensure that the collected data is representative of the actual machining process, and therefore any prediction algorithms presented later in the thesis are feasible to implement.

The second step focuses on finding novel labels that may provide a better description of tool damage than ones previously available. By adopting a new semi-automated technique of feature extraction from 3D models of worn tools, it is possible to study the wear scar progression across the tool for the first time. Using a pragmatic ranking approach features that best represent the machining process and therefore the generated acoustic emissions can be found.

It is important to extract features that correlate well with damage labels in order to make accurate predictions later on. Consequently, acoustic emissions generated during turning is explored in the time and frequency domain in the following step. Analysis of chips formed during machining is used to aid in further understanding of these signals.

A supervised learning approach using linear response surfaces is then adopted as a prognosis tool to predict damage for previously unseen tools. This method, along with Gaussian process regression, is implemented due to the availability of multiple inputs to predict a single output. For Element Six Ltd., this method can be applied when testing multiple tools of the same material composition following an extensive training period.

Finally, the thesis incorporates the idea of online unsupervised learning method to warn the operator of imminent tool failure. Dirichlet process mixture models adopt the use of incomplete label sets and alleviates the need for pre-labelled training data, allowing the testing of entirely unseen tools. The costs associated with data collection can be reduced significantly and interruptions can be kept to a minimum. The Dirichlet process mixture model allows online clustering as the data is collected without the need to set the number of possible clusters *a-priori*. This will enable operators to visualise when the characteristics of the cutting process change during machining, avoiding the need for exhaustive measures for tracking tool wear, or the early disposal of tools, depending on the context.

---

# TABLE OF CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction to tool condition monitoring . . . . .	1
1.1.1	Machining and Tool Wear . . . . .	3
1.2	The specific challenge for the industrial partner . . . . .	5
1.2.1	PcBN . . . . .	6
1.3	Motivations . . . . .	7
1.3.1	Aims and objectives . . . . .	9
1.4	Outline of the thesis . . . . .	9
<b>2</b>	<b>Literature Review</b>	<b>11</b>
2.1	Online monitoring of tool wear . . . . .	11
2.2	Supervised methods and TCM . . . . .	14
2.3	Unsupervised methods for TCM . . . . .	16
2.4	Tool condition monitoring with acoustic emissions . . . . .	17
2.5	Summary of the literature . . . . .	19
<b>3</b>	<b>Experimental Setup</b>	<b>21</b>
3.1	Current practice for tool wear monitoring . . . . .	21
3.1.1	Background on the machining process . . . . .	22
3.1.2	Limitations and assumptions . . . . .	23
3.2	Feasibility of using AE . . . . .	26
3.2.1	Sensor selection . . . . .	26

---

3.2.2	Additional sensor considerations . . . . .	29
3.3	Main test setup of the EngD . . . . .	31
3.3.1	Tools and tool wear measurements . . . . .	32
3.4	Summary of experimental procedure. . . . .	34
<b>4</b>	<b>Tool Wear features</b>	<b>35</b>
4.1	Wear on PcBN tools . . . . .	36
4.2	Current tool wear measurement techniques . . . . .	37
4.3	Novel method of tool wear feature extraction . . . . .	39
4.3.1	Challenges of the tool feature extraction method . . . . .	42
4.4	Tool wear features . . . . .	42
4.4.1	Features from the novel approach . . . . .	42
4.4.2	Features extracted from standard microscopy . . . . .	49
4.4.3	Extracted features - An example . . . . .	51
4.5	Feature selection for class labels . . . . .	53
4.5.1	Ranking criteria . . . . .	53
4.5.2	Scoring and Weighting . . . . .	56
4.5.3	Results of feature ranking . . . . .	56
4.6	Highest ranking features for damage labels . . . . .	61
4.7	Concluding remarks of tool features . . . . .	63
<b>5</b>	<b>Exploring Acoustic Emission</b>	<b>65</b>
5.1	Acoustic emission in turning . . . . .	66
5.1.1	Causes of acoustic emission in machining . . . . .	66
5.1.2	Acoustic emission from chip formation . . . . .	67
5.1.3	Acoustic emission from tool wear . . . . .	69
5.2	The AE dataset . . . . .	71
5.2.1	Time domain . . . . .	71
5.2.2	In the frequency domain . . . . .	74
5.3	Spectrograms . . . . .	75
5.3.1	Chip formation and AE . . . . .	77
5.3.2	Chip formation harmonics . . . . .	82

5.3.3	Confounding influences and AE . . . . .	83
5.3.4	AE at tool breakage . . . . .	85
5.4	Dealing with confounding influences . . . . .	87
5.5	Candidate features . . . . .	88
5.5.1	Features from spectrograms . . . . .	89
5.5.2	Features from the time series . . . . .	90
5.5.3	Wavelet Transform of the AE signal . . . . .	90
5.6	Conclusions . . . . .	93
<b>6</b>	<b>Supervised Learning</b>	<b>97</b>
6.1	Related work on linear response surface methodology . . . . .	98
6.2	Feature correlation . . . . .	101
6.2.1	Feature selection . . . . .	106
6.3	Linear response surfaces . . . . .	109
6.3.1	Results . . . . .	110
6.4	GP Regression . . . . .	116
6.4.1	Specifying the kernel and learning the hyperparameters . . . . .	118
6.4.2	Feature selection for the GP . . . . .	119
6.4.3	Results using GPR . . . . .	122
6.5	Predicting tool wear on a dissimilar tool . . . . .	128
6.6	Concluding remarks on supervised learning of tool wear . . . . .	131
<b>7</b>	<b>Unsupervised Learning</b>	<b>133</b>
7.1	The Dirichlet Process . . . . .	135
7.2	Feature selection for the DPMM . . . . .	139
7.3	Results and discussions . . . . .	141
7.3.1	DP and confounding influences . . . . .	146
7.3.2	Comparing DP cluster initiations with tool wear curves . . . . .	152
7.3.3	Clustering of dissimilar tools . . . . .	154
7.3.4	Effect of hyperparameters . . . . .	157
7.4	Conclusions . . . . .	159

---

<b>8</b>	<b>Conclusions</b>	<b>163</b>
<b>9</b>	<b>Future Work</b>	<b>171</b>
9.1	Improvements to the experimental setup . . . . .	171
9.2	Exploring tool wear features . . . . .	172
9.3	Understanding acoustic emissions further . . . . .	173
9.4	Extending the supervised learning . . . . .	173
9.5	Improvements to the online clustering . . . . .	174
	<b>Bibliography</b>	<b>177</b>
<b>A</b>	<b>Tool wear features</b>	<b>193</b>
<b>B</b>	<b>Results from the Dirichlet process</b>	<b>201</b>
<b>C</b>	<b>DWT - A brief introduction</b>	<b>207</b>

---

# INTRODUCTION

Automation has the ability to reduce costs and increase productivity across the manufacturing sector and as a result it has been widely accepted as the next step in the industrial revolution. Tool condition monitoring is one of the techniques that can help to address the challenge of automation in machining. With the help of machine learning methods, this thesis aims to predict tool wear in order to automate testing of advanced cutting tools during their design and commissioning phase.

In this chapter, the idea of tool condition monitoring is introduced, followed by detailing the specific challenges faced by the industrial partner that this thesis aims to address. The motivations for this work are then discussed where the aims and objectives are detailed. Finally, a brief outline of the chapters in the thesis is presented.

## **1.1 Introduction to tool condition monitoring**

Tool condition monitoring (TCM) is an area of research where one of the main focuses is on predicting tool wear online. Here, the term online predictions refers to automatic predictions that are made during the machining process (whilst tool

is in contact with the workpiece) without manual interference from the operator. Ultimately, the goal in the manufacturing industry is to establish a production line with a predictive maintenance strategy in order to reduce waste. The fundamental concepts of TCM follows on from condition monitoring, the foundations of which are based on Rytter's hierarchy [1]. This states that for prediction, first damage should be detected, located, and quantified. Here pattern recognition techniques could be used where data is studied in a probabilistic manner [2]. To aid with pattern recognition three main types of learning methods are available for conducting TCM. Firstly, supervised learning can be applied where data from sensor signals are used to make inferences about the damage as long as damage labels are available for modelling. The second method is unsupervised learning where data is clustered according to similarities without labels of damage. Reinforcement learning is the third method where damage labels are discovered by the algorithm through trial and error [3].

In simpler terms, tool condition monitoring to predict wear involves the following steps:

1. Attach sensors to machine and collect data online.
  - Studies undertaken on tooling materials such as carbides and ceramics, for example, have found tool wear to influence the forces, temperatures, acoustic emissions and vibrations generated in machining [4–6]. Therefore, by measuring these outputs, tool wear can be studied without directly examining the tool.
2. Conduct data cleansing followed by feature selection.
3. Make inferences about the tool wear state using the chosen algorithm in real time.

There are numerous challenges associated with TCM owing to the complex nature of machining and availability of descriptive labels of wear states. To expand this discussion, an introduction into machining and tool wear and the need for TCM in industry is presented.

### 1.1.1 Machining and Tool Wear

Metal cutting or machining is the process in which material is removed in the form of swarf, to create a net shape through the relative movement between a tool and a workpiece. There are many types of machining operations in use today, to manufacture a wide range of components for a multitude of industrial sectors. Some of the most common operations are turning, milling, drilling, facing and broaching [7]. Under each of these operations numerous sub-categories exist. In this work the focus is on outer diameter finish turning; the process of removing a fine layer of material from the outer surface of a cylinder in order to achieve a high quality surface finish.

In certain sectors like the aerospace and automotive industries, the final surface finish of machined parts is immensely important as these components can play a safety critical role within the structure. In order to maintain the quality of finished surfaces up to the standards set by industry a combination of steps are undertaken in production facilities. The machining conditions such as speeds, feeds, depth of cut are set according to the workpiece material, machining dynamics, cost and time limitations. Consequently, suitable tool types and shapes are chosen with careful consideration. The type of computer numerically controlled (CNC) machine is also chosen in compliance with the component, materials and cutting conditions. Finally, the set up of the system is thoroughly thought out in order to reduce vibrations, increase durability of equipment and ultimately improve quality of the final component.

One of the most significant yet unavoidable challenges in machining is the continuous wear on tools. Tool wear is the loss of material from the cutting surface due to interactions between the tool and the workpiece [8]. Tool wear is known to cause many issues to machined components, including form and dimension discrepancies, vibrations and chatter as well as insufficient surface finishes [9]. Machining with a worn or damaged tool will lead to surface defects on the component and crack propagation over time, resulting in non acceptance or early retirement of that component. As a result of this, when machining safety critical components, the industry employs a time-based maintenance strategy and discards tools at set times

regardless of wear state. Typically, a tool could be retired with 50-80% of useful life still remaining [10].

Clearly a simple solution here could be to measure the tool wear during machining to increase their time in use. Unfortunately, directly measuring tool wear during machining is almost impossible as the tool is in constant contact with the workpiece. In most cases oils, coolants and swarf can also restrict direct visibility of the tool. Furthermore, the wear scars on tools are too small to measure without microscopes, the majority of which cannot be installed inside the machine due to harsh environments generated during cutting.

The benefits of applying TCM is hopefully now clear. The assumption in this thesis is that by using indirect measurements of tool wear, such as acoustic emissions, force, temperature and vibrations, damage on the tool could be inferred without direct measurement. With the help of machine learning techniques, this entire method may also be conducted in real time.

The question that arises is, why is TCM not widely used in industry? Currently there are numerous aspects impeding the application of fully automated tool wear testing approaches. Due to the costs associated with collecting damage labels, industry can be reluctant to spend money on expensive machining trials to produce data that can be used to predict wear from all operating conditions. The data must be comprehensive and describe tools that wear differently to one another. Usually, data collected during machining does not encompass the entire life of the tool, again due to costs associated with data collection. Confounding influences can affect the datasets, leading to difficulties in the prediction stage. The sensor systems that can be used to collect signals for tool wear prediction can be expensive whilst requiring a lot of space for application. The learning algorithms that use damage labels and sensor features must be versatile in predicting wear in varying environmental and operational conditions. Most importantly, accuracy of predictions is key, as machining with damaged tools must be avoided. Each of these aspects of TCM require thought and care, as well as expensive experimental procedures. As a result, successful implementation of TCM in industry is extremely difficult and rare.

This thesis aims to address the complexities of TCM, by taking the entire process

into consideration, in order to approach the specific challenges faced by the industrial partner.

## 1.2 The specific challenge for the industrial partner

Element Six Ltd. (E6) is a world leading manufacturer of Polycrystalline cubic Boron Nitride (PcBN) tools. At the research facilities of E6, new tool materials are constantly developed to meet the needs of industry. To find the limitations of these new tools they are rigorously tested in time consuming and expensive machining trials.

Ideally, E6 would like to adopt the ‘Green Light Machining’ strategy to increase productivity. In most CNC machines a traffic light system exists where a green light indicates that chips are being produced. Therefore, Green Light Machining refers to the idea of practically eliminating machine downtime<sup>1</sup>. Generally in industry, green lights are only active around 30% of the time [11]. It would be beneficial to E6 if a system was in place that allowed the machine to learn when a tool is about to fail, so that the tool could be changed automatically and continue cutting. Preferably, the ultimate TCM system should either be robust against the new materials continually developed by E6 or entail a simple and fast training period to account for these new properties.

In the absence of such a method, currently at E6 tool wear is measured off-line in a run-to-failure method. This approach interrupts the cutting process intermittently and requires the temporary removal of the tool from the machine to complete wear measurements using optical and scanning microscopy. It is a time consuming and expensive method, causing bottlenecks in the facility. Furthermore, this procedure of measurement adds uncertainty to the process. For example, the removal and return of the tool to the machine leads to positional differences that in turn could affect the cutting conditions.

Another crucial challenge faced by E6 is the duration of test times. Due to the nature of PcBN, the tool wear tests are extremely lengthy. In some cases it has been known

---

<sup>1</sup>The amount of time a machine is not producing chips, including setup times.

to take over an entire working day to complete a standard test on one single tool. A large proportion of this time is due to the offline tool wear measurements. Therefore successful implementation of a TCM procedure would be tremendously beneficial to E6. The next section details the characteristics of PcBN tools that are responsible for the extensive tool tests.

### 1.2.1 PcBN

PcBN is the second hardest material in existence [12]. For this reason, and due to low wear rates, PcBN is utilised as a cutting tool material<sup>2</sup>. Figure 1.1 shows some typical properties of PcBN compared to other cutting tool materials. Wear resistance is a crucial property that allows the tool to operate for longer periods under harsh environments, such as high cutting forces, high temperatures, shock loading and plastic deformation [13]. PcBN also has low chemical wear rates, especially when machining ferrous material, as it does not contain carbon molecules, which tend to burn in high temperatures or diffuse into other carbon base workpieces [14]. Subsequently, PcBN can be used for machining hardened steel. At high cutting temperatures (700°C or higher), the high thermal conductivity of PcBN permits the fast dissipation of heat whilst its toughness and strength aids with resisting fracture [15]. These material properties allow PcBN to be used in finish turning which is a way of securing a desirable surface finish of a machined part. This process is capable of replacing grinding in the final machining stage, reducing the cost and time required for finishing [16] .

Owing to the aforementioned material characteristics, PcBN is expensive compared to carbide tooling [17]. Therefore, the preventative method used in industry where tools are discarded at constant time intervals can be costly.

As PcBN tools are typically used for finish machining, lighter cuts are taken off the workpiece in comparison to rough turning. As a result, the tool wear threshold is generally lower than the average. Subsequently, the wear scar is harder to inspect

---

<sup>2</sup>In the composition, cBN grains are in a binder phase of Titanium Carbo-Nitride (TiCN) ceramics. The grains are between 0.5-5 $\mu$ m in size.

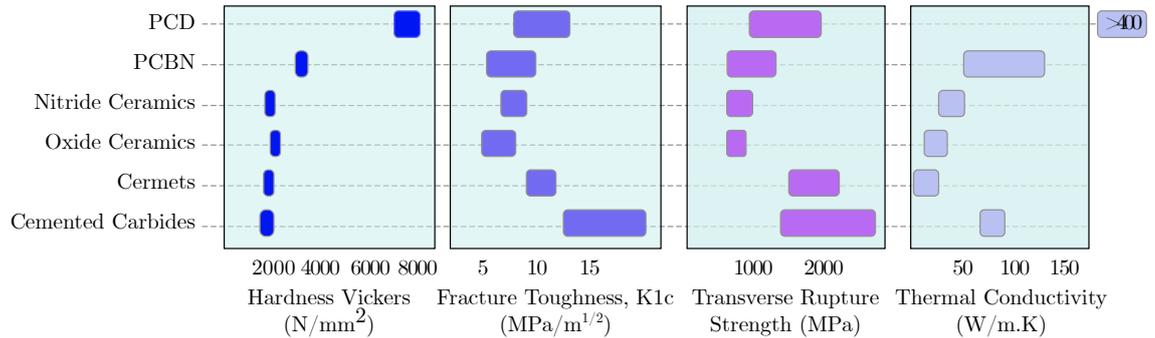


Figure 1.1: Typical values of material properties for main classes of cutting tools (after [15]).

while the corresponding signals may not capture the wear progression in great detail. Both of these factors can lead to TCM of PcBN tools relatively more challenging than most other tool materials.

### 1.3 Motivations

Much of the research conducted on tool condition monitoring focuses on investigating whether features from a given sensor signal could be used to predict tool wear. Few authors have attempted to use the information and the available machine learning techniques to actually predict tool wear on an entirely unseen tool<sup>3</sup>. With the help of acoustic emissions generated during hard turning, this thesis aims to capture signal features that are capable of predicting wear and failure on completely unseen tools.

In previous attempts made by researchers to predict tool wear on unseen tools, often the environmental and operational condition remain constant. Due to the nature of the test setup at E6, confounding influences affect the machining process and therefore any collected features. This is also often the case in realistic machining set-ups in industry. It is the aim of this work to build a TCM system capable of dealing with confounding influences whilst retaining prediction accuracy for previously unseen

<sup>3</sup>Here, an unseen tool is a tool that has not been used in a training set. It is usually either a new tool that has not been used for machining or a tool that has not been directly measured to obtain measurements of tool wear.

tools. The resulting system could therefore be more versatile than previous attempts which could allow for easier implementation in industry.

Thanks to advances in measurement technologies it is now possible to capture the entire wear scar of a tool in three dimensions without damaging or altering its surface. By doing so, the wear scar can be studied in detail, a prospect that was not achievable previously. This thesis aims to take advantage of these 3D wear scars and discover tool wear features that may be better suited for prediction than their predecessors.

In a review of around a decade's worth of research, Dornfield et. al. [18] found that acoustic emission generation is more sensitive to precision machining as the signals are not hindered by low frequency disturbances. Though acoustic emissions have been widely accepted as an output of machining that is correlated with tool wear, its full potential is rarely realised in research, especially in the area of hard turning with PcBN. Firstly, most studies collect acoustic emissions during parts of the process usually due to the significant time needed for data analysis and substantial data storage requirements. Owing to the transient nature of the acoustic emissions, this method risks the disregard of important information that could encapsulate the wear state of the tool. To address this shortcoming in current research, acoustic emissions are collected continuously throughout the entire life of the tool in this EngD. Secondly, only a few researchers have attempted to explore the frequency domain of the acoustic emission signal in order to make predictions about the tool's wear state. In this work, spectrograms are utilised to investigate the frequency changes of acoustic emissions during tool life. Furthermore, there is little literature linking acoustic emissions to the physical process of machining. The authors that have previously studied this field have concentrated all efforts on orthogonal machining [19]. As many of the turning operations in use today are conducted under oblique conditions, this thesis aims to further the knowledge of chip formation and acoustic emission generation in realistic oblique turning processes.

For a futuristic 'Green Light Machining' process where automatic tool changes are desired, the sensing system should be capable of measuring outputs from each tool within the machine to make predictions. For this reason, dynamometers are not suitable due to their size and cost. A network of acoustic emissions sensors on the

other hand would be easy to incorporate without disrupting the machining process. This thesis investigates whether it is possible to monitor all tools in the machine with a single sensor or whether a network is needed for Green Light Machining.

Researchers studying tool wear and acoustic emission often haven't conducted many repeats, presumably as a result of storage and computing power required for analysis. In this work it is the aim to collect a dataset at E6, under realistic test conditions, that encompass information on the effect of tool material and sensor location on acoustic emission and its prediction capability. These variations will be studied across multiple repeats for successful training and testing of prediction algorithms, making the dataset collected during this work one of the most comprehensive and representative of real life applications to date.

### **1.3.1 Aims and objectives**

The aims and objectives of this thesis are the following:

- To further the understanding of tool wear mechanisms in turning.
- To build a detailed knowledge of chip formation in turning.
- To explore the feasibility of unmanned tool testing using a lathe, where automatic tool wear inspection is integrated within the operation.

## **1.4 Outline of the thesis**

Figure 1.2 visually indicates the chapters of this thesis and how they relate to one another.

- The related work conducted in the field of tool condition monitoring is reviewed in Chapter 2. This chapter investigates the use of acoustic emission in tool condition monitoring in turning and expands into the type of learning methods used by past researchers.
- In Chapter 3, the experimental setup used in this thesis is described.

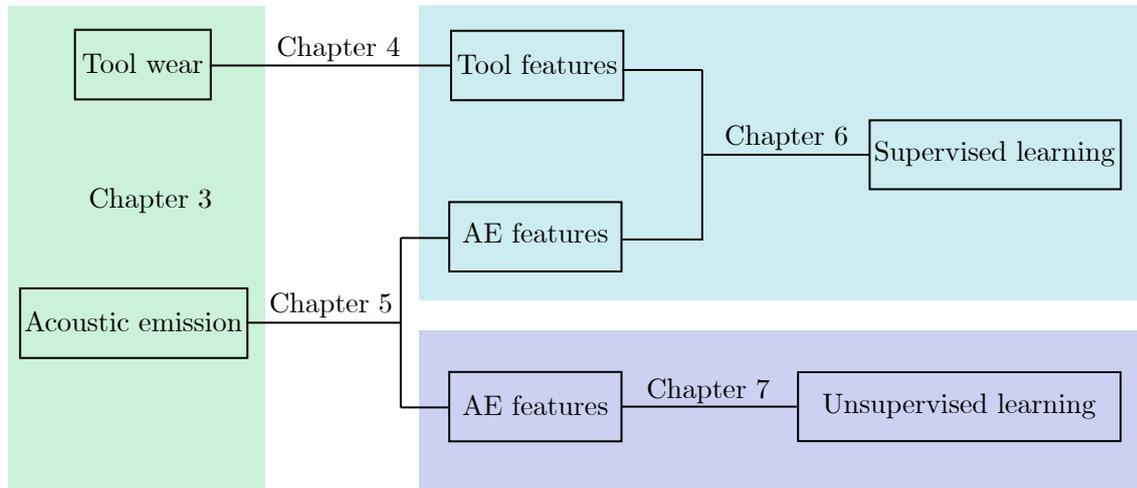


Figure 1.2: The connections between the chapters in this thesis.

- Chapter 4 focuses on the wear scars measured in Chapter 3 in order to find novel features that may be more descriptive of tool wear than ones in use today. This chapter addresses the first aim in section 1.3.1.
- A detailed knowledge of the chip formation is built in Chapter 5 where acoustic emissions collected using the setup in Chapter 3 are analysed in detail. Features from the acoustic emission signal are then selected as targets for prediction in the following two chapters.
- Using the features from Chapter 4 and 5, supervised learning to predict tool wear is conducted in Chapter 6. Here, the linear response surface methodology is applied to the dataset and the results are compared with Gaussian process regression. This chapter begins to undertake the third aim in section 1.3.1.
- To address the shortcomings of the work in Chapter 6, Chapter 7 proposes a unsupervised learning method that is capable of clustering acoustic emission features online in order to predict imminent failure of the tool.
- Chapter 8 details the conclusions from the previous chapters.
- Finally, in Chapter 9 the future work planned for this project is presented.

---

# LITERATURE REVIEW

This chapter provides an overview of the literature on the topic of tool condition monitoring, spanning the types of learning methods used in this thesis. In Chapters 4 to 7, more relevant literature reviews explore the specificities covered within that chapter.

In Section 2.1 related work on tool condition monitoring is presented. Past examples of work conducted on supervised learning is presented in Section 2.2 followed by a review of the literature on the subject of unsupervised learning in Section 2.3. Finally literature with a focus on using acoustic emission as the observed signal for TCM is explored in Section 2.4.

## **2.1 Online monitoring of tool wear**

Numerous literature reviews on different aspects of tool condition monitoring have been conducted in the past. As discussed in Chapter 1, a number of machine learning and pattern recognition techniques can be used for TCM, i.e. supervised and unsupervised learning. Within each of these techniques researchers have employed specific algorithms for diagnosis and prognosis, using a multitude of different sensors and features.

In one of the first reviews on TCM in 1990 the authors argued that the sensing techniques reviewed, such as vibration and cutting force signals, were too specific for industrial application [20]. The authors argued that the sensing system should be flexible for a variety of machining variables. In 1995, another review stated a number of demands of industry for a practical TCM system [21]; fast responses, reliability, robustness, low cost, simplicity, ease of use, small space requirements, minimum modifications and maintenance were some of the most prevailing needs. A review of neural networks for TCM in 1997 found that most methods reviewed were capable of online application although no proof was given to substantiate this claim [22]. This can often be the case in TCM literature. Following a review conducted in 2000, it was suggested that TCM research is not successful in industry as the sensor information is often inadequate at reflecting the complexity of the process [23].

In 2002, a review summarised that 138 publications focused on using artificial neural networks for tool wear monitoring during the turning process [24]. The variations of the networks for supervised learning were the multilayer perceptron network, recurrent neural network and supervised neuro fuzzy system, to name a few. The author concluded that although there were promising results, there was still much more work to be done on the software aspect of TCM such as preprocessing and data analysis in order to design a system that can be applied in industry.

In 2005, a review of the state-of-the-art methods of TCM suggested that there exists a lot of untapped potential in the feature selection methods where wavelet analysis and short-time Fourier transform are underutilised [25]. The authors also state that artificial neural networks are only viable in large scale production environments where a lot of training data is available. Another drawback of neural networks is stated as the sensitivity to cutting conditions.

A review on state of the art methods applied to TCM in 2008 summarised that all currently available methods are not computationally fast enough to detect sudden or imminent failure [26]. As reliability is key for industry, the authors recommend the application of simple TCM systems over more complex ones.

In 2010, a general review on machining monitoring systems criticised the lack of details and clear guidelines as key issues in publications of methodologies within the

topic [27]. The authors also suggest that the following drawbacks of each learning approach is responsible for the shortcomings in global industrial application; the lack of clarity in the design, absence of physical meanings, low extrapolation capability and the time consuming method of obtaining parameters. For neuro-fuzzy methods, the disadvantages come from the large number of parameters that must be learned or defined and contradictory rules as well as limited inputs sizes. The authors indicate that the low learning ability and generalisation capability as well as the inability to input knowledge of the physics into the algorithm are the limitations of fuzzy logic methods. Finally, in Bayesian networks, the issues are suggested as the need for large datasets, high computational costs and the need for variable discretisation.

A recent review of TCM literature found that there is a demand for advanced signal acquisition and preprocessing steps for extraction of features that are insensitive to the process and workpieces, in order to apply TCM techniques in industry [28].

A review on TCM in milling conducted in 2018 concluded that a number of aspects should be improved for accurate predictions in an industrial setting [29]. Firstly, the requirement for extensive training periods incur a large cost, leading to TCM systems that are not capable of tool prediction. The authors specify a need for learning methods that do not require large amounts of training samples. It is suggested that more work needs to be done on the configuration of sensors and the feature extraction method as well. Most publications focus on diagnosis which the authors consider to be less important than prognosis. Due to the challenges of prognosis, however, most researchers do not attempt to predict remaining useful life (RUL) as simple models cannot handle this task. The review calls for additional research on this topic. Fuzzy clustering has been merited for its ability to accommodate uncertainties but the authors acknowledge the need for a technique that does not require the presetting of number of clusters as this can affect the performance of the algorithm. Transfer learning has been suggested as a solution to prediction across different tool and workpiece types in future work.

In summary of the literature reviews of TCM, it is clear that there is a need for a simple, robust and non-invasive system that can predict wear on unseen tools that can be applied in industry.

A closer look at supervised learning methods used for TCM in the past is explored in the next section to estimate the capability of these techniques in an industrial setting.

## 2.2 Supervised methods and TCM

Many authors have attempted to learn and predict tool wear using supervised learning techniques in the past. Much of the literature on prediction and monitoring only evaluates whether a chosen feature set and algorithm used to train the model is able to predict a given response. A meaningful prediction should instead be conducted by training models and testing with previously unseen data. Ideally these predictions should be robust against the changes in environmental and operating conditions that affect the behaviour of tools.

A review by Siddhpura et al. [30] summarised the literature between the years 1951-2012 focussing on flank wear prediction according to direct and indirect measurement methods, the domain in which the features are extracted, and the type of classifier used. At the time, neural networks were favoured above other classifiers such as fuzzy logic, neurofuzzy & hidden Markov models and support vector machines (SVM) due to the high fault tolerance, high adaptability and the capability to suppress noise.

Das et al. [31] used neural networks to predict wear on carbide inserts. The model struggled to predict the measured wear in high speed and feed conditions due to unstable wear mechanisms such as built-up-edge formation. In another attempt, the authors [32] suggest the error of prediction is small though quantitative results are not presented. The model sometimes under predicts the measured values significantly which could lead to machining with a damaged tool. Dimla et al. [33] used multi-layer perceptron neural networks where classification of a number of wear states were possible when the model was trained and tested on the same tool. However, when testing on a different tool to training, the model was incapable of tool state classification for sharp and worn tools. Sun et al. [34] details the drawbacks of applying neural networks for the tool condition monitoring problem; using the average performance as an index for classification accuracy, the unequal size of data

distributions in each tool state, using maximum classification rate as a loss function and assuming an even cost function are attributed as the problems of using neural networks.

Binsaeid et al. [35] used a combination of 135 signal features to predict gradual flank wear as well as breakage and chipping using radial basis function neural networks, SVM and multilayer perceptron neural networks. The authors found that 25 signal features correlated well with tool wear, reducing the computational time needed for classification. SVM outperformed the other methods. By using a 10 fold cross-validation, over 90% classification accuracy was achieved when using SVMs with sensor fusion where data from acoustic emissions (AE), force, vibration and power were used. However, this method is not proven to work on unseen tools. The RUL of a milling cutter was predicted by Benkedjouh et al. [36] using nonlinear feature reduction and support vector regression. The model was trained and tested using cross validation. A nonlinear power model must be calculated for each cutter in order to calculate the RUL. Wang et al. [37] uses a GPR model to predict tool wear. The authors train and validate the model using data from only one tool, therefore, it is unclear if this method would generalize and perform well on an entirely unseen tool. Dutta et al. [38] used support vector regression to predict flank wear on a turning operation using machined surface images. As the training and testing sets were chosen through cross validation, the method has not been shown to work on unseen tools. As the images of the surface are collected at the end of each operation, this method would not be able to predict tool wear at any given time, therefore would not be useful in the context of this thesis.

A naïve Bayes classifier was used by Karandikar et al. [39] to monitor tool wear in an end-milling operation. Both continuous and discrete tool wear cases were investigated. The authors trained the model using data from three tests at varying spindle speeds and tested the model using data from tests under different cutting conditions. The authors state that a good fit was found, however, a quantifiable measure has not been presented. The posterior distribution does not seem to capture the actual value of flank wear across the entire life of the tool. The authors suggest that this may be due to the uncertainty in the feature values due to measurement

challenges.

### 2.3 Unsupervised methods for TCM

In order to avoid an extensive training procedure, unsupervised learning algorithms can be applied to sensor data where collecting labels of wear can also be avoided. A decade's worth of research on applying artificial neural networks (ANN) in TCM was reviewed by Sick [24]. The author discourages using unsupervised network paradigms where tool condition is used as an output as a training phase is required. Alongside this, the clusters produced during this phase cannot be interpreted without knowledge of the output which is normally gained by collecting wear labels. Previous work on unsupervised learning of the machining process conducted by McLeay [40] focuses on using Gaussian mixture models (GMM) to cluster different damage and operational classes occurring on the tool. The author sets number of clusters *a-priori* for each tool by using the cluster information from the previous tool investigated. Due to variance in tool behaviour, it is challenging to classify wear states of unseen tools using this technique.

Auto-Regressive (AR) time series modelling of the  $AE_{RMS}$  signal was carried out by Liang et. al [41] to predict flank wear. The authors identified separable clusters associated with flank wear widths when plotting sets of AR parameters against one another. No attempt was made to cluster previously unseen data.

Wang et al. [42] uses integrated particle filters with SVM and AR models to predict RUL of a tool using force measurements. By measuring force for the first 200 cuts, tool wear is predicted for the next 200, 150, 100 and 50 steps by using the SVM or AR model to predict the force measurement first, then using that prediction to predict tool wear using the particle filter. The results are promising, although training of the AR model and SVM is conducted offline and it is unclear whether this method was successful on a previously unseen tool.

As suggested by much of the literature, AE is a suitable sensing method that can be utilised for a system such as this. A further exploration of AE in TCM is therefore presented next.

## 2.4 Tool condition monitoring with acoustic emissions

AE signals can be used as an indirect measure of the tool wear process as AE is sensitive to damage. However, due to the costs associated with collecting, storing and analysing AE, authors choose not to measure AE during the entire life of the tool.

AE is the release of elastic energy in the form of stress waves during the deformation of a material. These stress waves can be captured by piezoelectric transducers. As a non-destructive evaluation method, AE is often used in preventative maintenance of structures and machinery [43]. Although AE is not usually adapted into continuous monitoring systems, it has been employed in attempts to predict tool wear in metal cutting in the past [44–46]. AE signals obtained during machining are made up of both continuous and transient components. The continuous signal is produced by plastic deformation and frictional effects occurring in the cutting zone [47] and transient signals are an indication of collisions and breakages of chips and the tool [41, 48].

Numerous studies have been conducted on examining the AE signal throughout tool life and many have found that flank wear has a correlation with changing features of AE such as root mean square (RMS), skewness, kurtosis, AE count rate, amplitude and frequency. It has been found that low values of AE amplitude are caused by worn tools, whereas the high values are present owing to machining with sharp tools [49]. In 2003, Dornfeld et al. [18] summarised a decade's worth of research on monitoring ultra precision machining processes with a focus on the use of AE sensors. The sensitivity of acoustic emissions to small changes in the machining process, as well as its capability to detect subsurface damage, qualifies AE to be a forerunner in the choice for the sensing technology in tool condition monitoring.

Li [47] conducted a review on tool condition monitoring with AE in turning covering three main areas; generation of AE, signal processing and tool wear prediction. According to the review, sensitivity of AE to tool wear is responsible for the vast quantity of research in the TCM field, and signal processing is paramount to ensuring process parameters do not affect the TCM algorithms.

Kamarthi et al. [50] found AE to be a good choice for TCM as the signal is usually uncontaminated by background noise and vibrations. AE was also preferred to other sensor types as it reflects the change in tool condition with speed. Moreover, as mounting an AE sensor does not alter the machining process, it is preferred over large sensors such as dynamometers in industrial applications.

Jemielniak [51] performed a wavelet packet transform (WPT) of AE signals in rough turning and found that kurtosis and the energy of decomposed WPT coefficients correlated well with the used up portion of tool life. However, AE was only measured during the middle of each pass of the workpiece and the signals were significantly down sampled as there were far more observations than predictors for supervised learning. By limiting original data collection in this way, important features of the process could be missed, especially at the start and end of each pass where the force acting on the tool is non-stationary. When predicting tool wear of milling cutters, Ray et al. [52] found time-domain AE features can be used to classify wear states, using SVM. The authors also employ the Gaussian process nonlinear Auto Regressive model with eXogenous inputs (GP-NARX) to treat the tool wear as a continuous process instead of classifying wear into discrete states. Promising results show that wear on the cutting edge radius can be predicted continuously, though training data is required from the full tool operational range.

The use of acoustic emission measurements to understand tool wear in hard turning using PcBN has only been conducted once previously. The study in question was based on basic time domain analysis where the skewness and kurtosis of the signal's RMS values were compared against flank wear. The author found the shape of the bursts coupled with its amplitude could give an idea of the severity of flank wear but concluded that the results were not reliable or robust enough for practical implementation [53]. Subsequently, work in this thesis is carried out in order to understand whether AE signals taken from PcBN tooling could provide information regarding tool wear states when analysed in the frequency domain.

## 2.5 Summary of the literature

It is clear from the literature that there is still a need for an accurate TCM system that is robust, simple and inexpensive to implement in industry. Due to numerous complexities, it has been challenging to design a TCM method that is capable of accurate predictions in the past.

This thesis aims to use data collected from AE sensors alongside supervised and unsupervised methods in attempt to design a TCM method that may be capable of diagnosis and prognosis. This will be particularly tailored to the outer diameter turning process with oblique cutting for PcBN tools. In each following chapter, any particular issues with sensors, feature extraction and prediction that relate to this work highlighted in the literature will be discussed.



---

## EXPERIMENTAL SETUP

In order to address the challenges of automated tool testing presented in Chapter 1, an experimental series was designed to capture behaviour of interest, associated with the process of tool wear. As discussed in Chapter 2, acoustic emissions of the cutting process may be used as correlates to tool wear. Subsequently, AE sensors were installed to the existing test procedure at E6 where data was collected across multiple tools spanning different material compositions, making this the largest and most comprehensive AE and tool wear dataset collected for tool condition monitoring.

This chapter introduces the current practice at E6 in Section 3.1. Section 3.2 expands on the feasibility of incorporating AE measurements into the current practice, followed by Section 3.3, which describes the main test setup of the EngD.

### 3.1 Current practice for tool wear monitoring

At the facilities of E6, PcBN tools are used to machine case hardened<sup>1</sup> workpieces until catastrophic failure. These are referred to as ‘end of life’ tests. As it is not

---

<sup>1</sup>Case hardening is the process of rapidly heating and cooling a material to harden its outer surface. Case hardened workpieces are used in this setup as it is a common application for PcBN, i.e. the machining of case hardened steels used for automotive transmission components like gears or pinions.

possible to measure tool wear during machining, these tests are sectioned so that four continuous cuts, referred to as passes of the workpiece, are carried out at a time, followed by inspection of the tool. The inspection consists of removing the tool from the machine to measure flank and crater wear using a 2D optical microscope and a 3D scanning microscope respectively. Subsequently, a decision is made on whether to machine a further four passes or to stop the process due to tool failure. A schematic showing these steps can be found in figure 3.2.

The current test procedure also includes measurements of additional parameters:

- The surface roughness of the workpiece after machining to study the tool's performance.
- Hardness of the workpiece before and after machining to ensure tool interaction is within material tolerance.
- Diameter of the workpiece to confirm depth of cut is as expected.
- Force generated in the machining process used as an indication of tool failure.

### 3.1.1 Background on the machining process

A brief background of outer diameter (OD) machining is presented in this section. Figures 3.1a and 3.1b display the process in detail.

In figure 3.1a, the workpiece is rotated about the z axis using the spindle. The tail-stock is pressured against the workpiece to reduce unwanted vibrations. The tool insert is secured to the tool holder which progresses from right to left in the z axis starting from point S. The speed at which the tool moves in the z axis is the feed. The distance between the outer edge of the workpiece in the x direction to point S is the radial depth of cut (DOC). This is highlighted in figure 3.1b. Due to the rotating workpiece and the sharp cutting edge of the tool, the relative motion produces a chip (or swarf), which reduces the dimensions of the workpiece. The speed at which chips are removed is the cutting speed. Reaching the other end of workpiece, point E (the end point) completes what is referred to as a pass. The tool moves back to the right of the workpiece and this process is continued three more

times. The machining conditions mentioned here are set by E6 according to the workpiece material, type of tool and the required material removal rate.

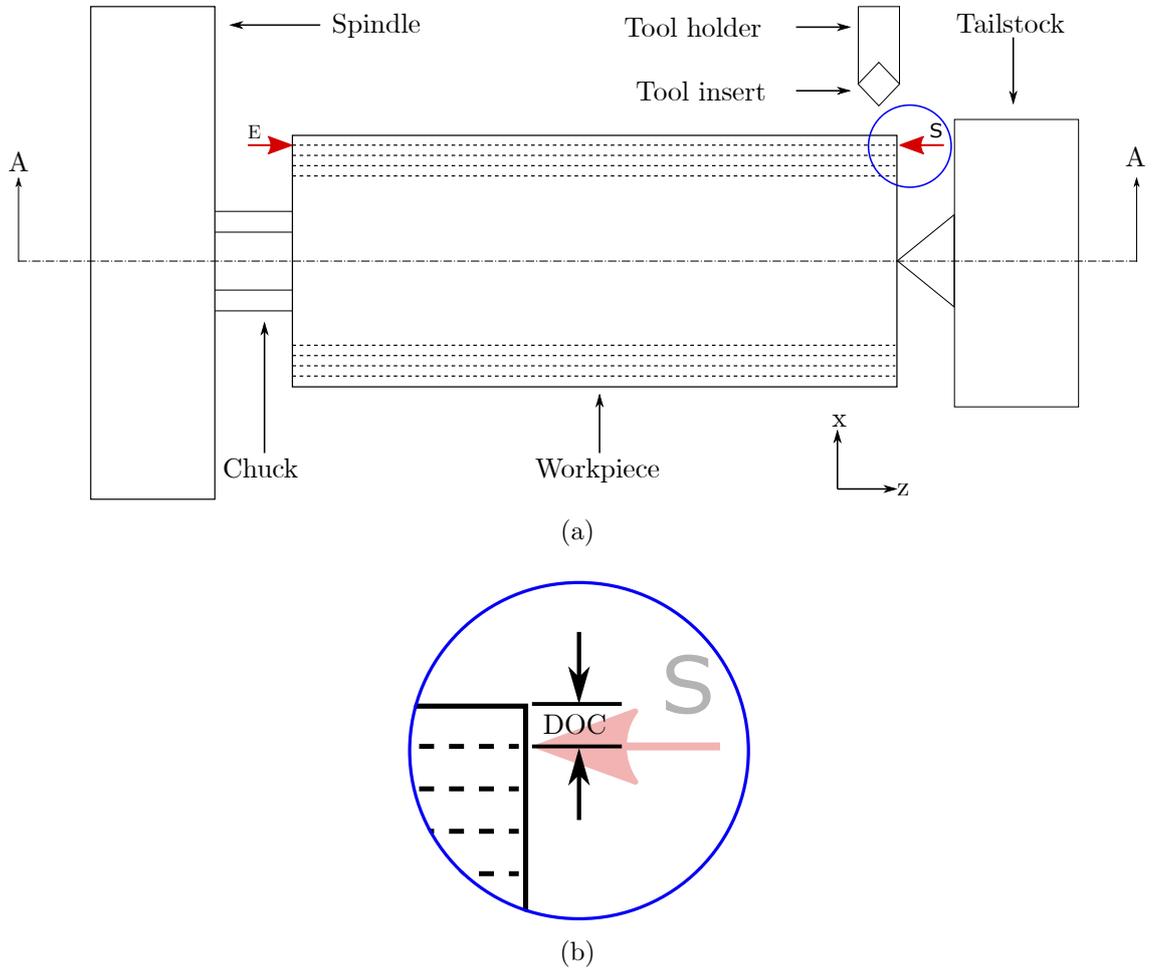


Figure 3.1: (a) Drawing showing the set up of the workpiece and tool including the start (S) and end (E) of a single pass of the workpiece. (b) Figure showing the start of cut with depth of cut indicated.

### 3.1.2 Limitations and assumptions

The current practice at E6 has some limitations; it is therefore important to make assumptions about the collected data by carefully considering these limitations:

- Removing the tool for measurement is unavoidable as it is the only way to learn information about wear. Without these measurements it is not possible to obtain damage labels or make decisions about whether to continue machining or to stop. Unfortunately, when a tool is removed, it is not possible to secure it back to its exact position prior to measurement. As a result, the value of DOC is uncertain for the next pass. DOC is pivotal to the mechanics of machining; varying the depth of cut can affect the forces, stresses, and dynamics of the process. Therefore, it is not good practice to conduct machining when the exact DOC is unknown. With this in mind, it is important to assume any data generated during the first pass of the four pass interval is unreliable.
- At the aforementioned four pass intervals, the workpiece is replaced by a new one. This is because the case hardening only penetrates a small layer of the workpiece's outer diameter, and is likely to be machined off after four passes. In order to maintain relatively consistent conditions for the tool, only workpieces within a specified hardness tolerance are selected for testing. However, it is assumed that the workpiece hardnesses, even within tolerance, can have an influence on the tool and collected data. It is therefore important to consider the hardness of workpieces during data analysis and machine learning.
- On average, it takes over 8 minutes to complete four passes and associated tool wear measurements shown in part two of figure 3.2. Removing the tool from the machine, walking to and from the microscopes, and setting up at the microscopes are some of the hidden time consuming steps that impact this process. Over 40% of the time spent completing four passes are spent on actions associated with these measurements. Furthermore, figure 3.2 does not show the steps required to analyse the tool wear measurements. Studying and understanding the features of 3D scans and images from the optical microscope require time and advanced technical skills. Detailed breakdowns of tool measurement inaccuracies are presented in Sections 4.2 and 4.4.2. Therefore, it is clear that tool wear prediction without directly measuring the tool can save valuable time for the test operators and analysts.

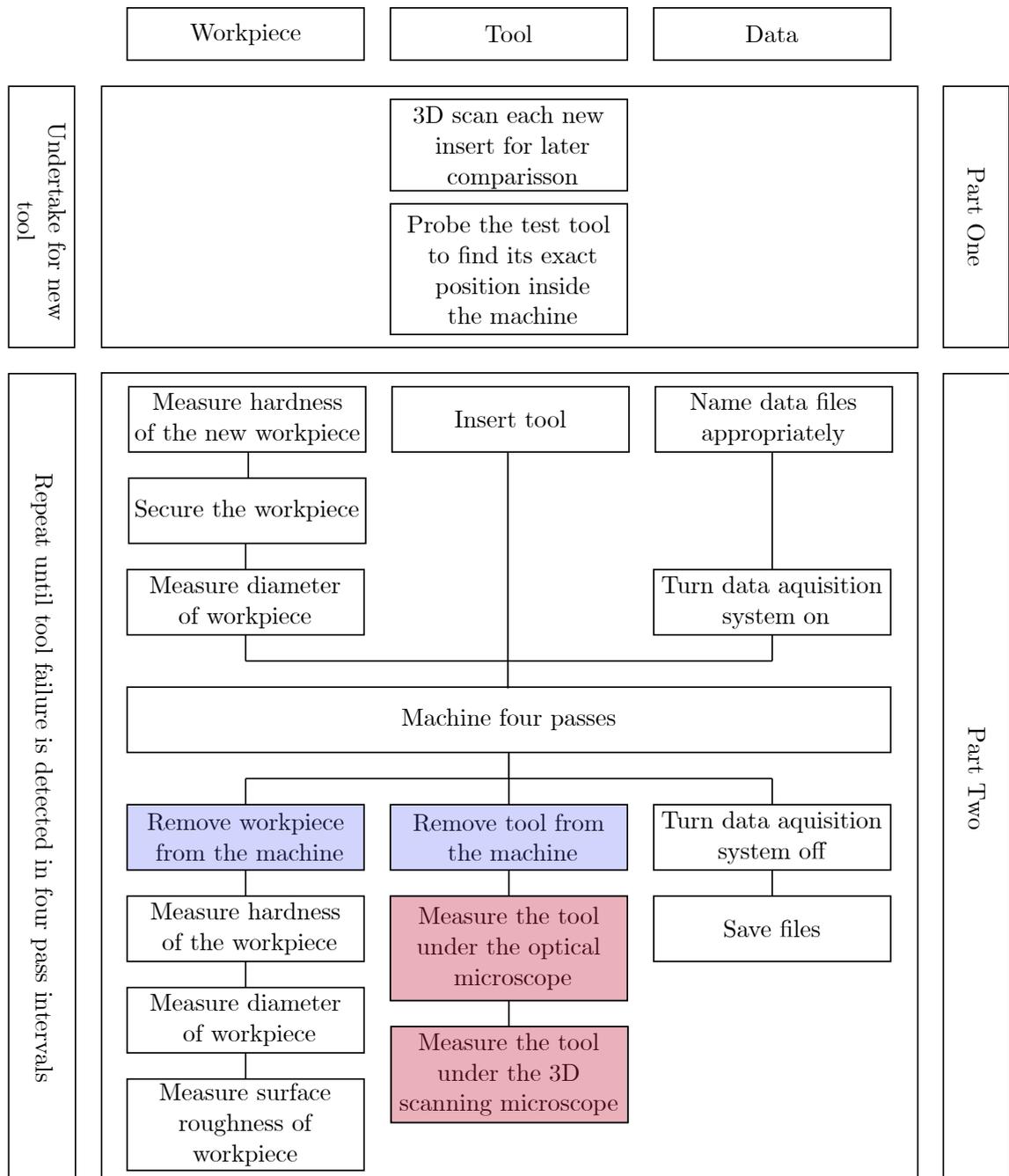


Figure 3.2: These are the steps taken during tool wear tests. The steps highlighted in blue are limitations of the experimental process that affect the collected data. The steps in red are time consuming and can be avoided by tool condition monitoring.

## 3.2 Feasibility of using AE

A study was conducted to understand the feasibility of incorporating AE sensors to the current test rig. The main aims of this work was to investigate the selection and positioning of sensors. Preliminary analysis of data acquired from this study revealed a bank of information regarding the type, placement, and sampling frequencies that are required to obtain useful signals for feature extraction. These findings were used to finalise the sensor selection and placement for the main data collection setup detailed in section 3.3.

### 3.2.1 Sensor selection

Material composition of tools stipulate the environment in which they can be used for machining. Due to the susceptibility towards thermal shock, PcBN cannot be used with coolant. Subsequently extremely harsh environments are generated in the machine, especially at the cutting zone (figure 3.3). Therefore it was clear from the feasibility study that the sensors need to be placed strategically to reduce the risk of entanglement, lacerations and fire damage, whilst retaining a high signal to noise ratio. Subsequently, the machining environment limited the type of sensors that could be used in tool wear tests.



Figure 3.3: This image demonstrates the harsh environment inside the machine with burning swarf emitting from the cutting zone.

The feasibility study indicated that the frequency response range of Mistras Micro - 30D differential sensor (referred to as Sensor 1) with a stainless steel casing was suitable for this work. The stainless steel casing was useful in the extreme machining environment and the sensitivity was suitable to place the sensor near the tool. The positioning of this sensor can be seen in figure 3.7c. The position that the sensor could be attached to was restricted due to the movement of the turret and tool. This sensor is differential in nature and so it is electrically isolated from the background electrical noise of the sensor network and wires within the system. Data collected between the machined passes during the feasibility study confirmed that the signals are not affected by the insignificant background noise.

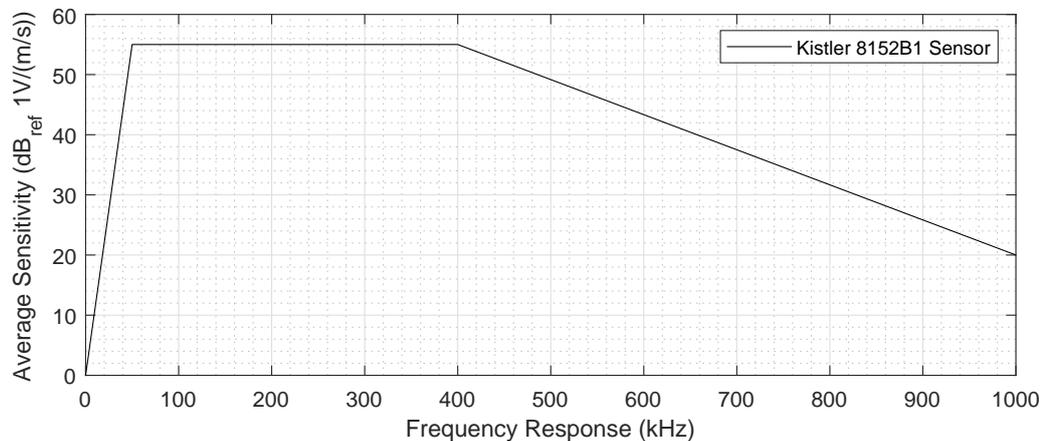


Figure 3.4: The ideal frequency response of the Kistler 8152B1 sensor.

A Kistler 8152B1 sensor (referred to as Sensor 2) performed well during feasibility work when placed away from the cutting zone due to its high sensitivity (figure 3.4). When placed close to the cutting zone, the signal was clipped due to large amplitudes of generated AE. Thus, Sensor 2 can be placed on the machine turret<sup>2</sup> away from the tool and central to all other tool holders. A futuristic, fully automated tool testing procedure could benefit from this approach; one central sensor capable of detecting acoustic emissions from all tools in the turret alleviates the need to attach sensors on all tools, reducing costs and complexity of the entire system.

<sup>2</sup>The turret inside a lathe is capable of holding numerous tools at one time.

As seen from figure 3.4, the sensitivity of the sensor starts to decrease beyond 400kHz. Consequently, 1MHz sampling rate was chosen. This results in a frequency range of 0 - 500kHz when taking the Nyquist criteria into account.

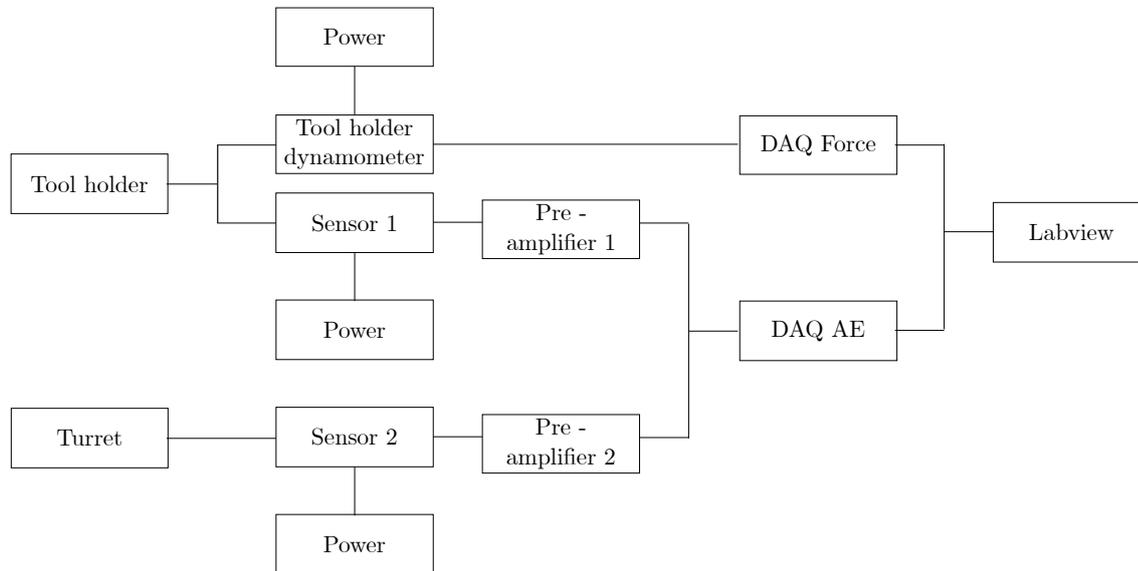


Figure 3.5: The set up with sensors and the accompanying hardware connections.

A Kistler 9129AE1 tool holder dynamometer was used in the current set up and was retained for this experiment to aid with sensor fusion later on. The force signal may be useful when studying the unknown behaviours of the acoustic emissions. In a fully automated rig, it is not financially or spatially viable to setup these dynamometers for every tool on the turret and so replacing force measurements with AE would be beneficial in the future.

Figure 3.5 shows the setup of sensors and the associated hardware. A pre-amplifier is a device used when the required gain or frequency bandwidth is unknown. Inside, a bandpass filter is used to filter out the low and high frequencies of the signal. Mistras 2/4/6 and Kistler type 5125 pre-amplifiers were used alongside Sensor 1 and 2 respectively. A National Instruments (NI) compact data acquisition (DAQ) chassis (NI cDAQ 9174) was used to synchronise data from the two AE sensors. A NI type 5070 multichannel charge amplifier was used to synchronise the three dimensional<sup>3</sup>

<sup>3</sup>Three dimensions are the directions of cutting, feed, and thrust forces.

signals from the tool holder dynamometer.

It was clear from the feasibility study that the AE signals received from these sensors are rich in information. However, the physical cause of the variation in these signals were challenging to decipher. Consequently the feasibility of introducing a high speed camera, to visualise the process, was assessed. Section 3.2.2 introduces the camera and expands on its capability.

### 3.2.2 Additional sensor considerations

Number of other technologies were considered for this work in order to find meaningful and informative features for TCM later. Due to numerous shortcomings, these were not considered in the final setup.

#### High speed videos

To understand the chip formation and AE in more detail, a FASTCAM Mini UX100 high speed camera capable of recording up to 800,000 frames per second was used alongside the test rig in a feasibility study (figure 3.6a). The camera was synchronised

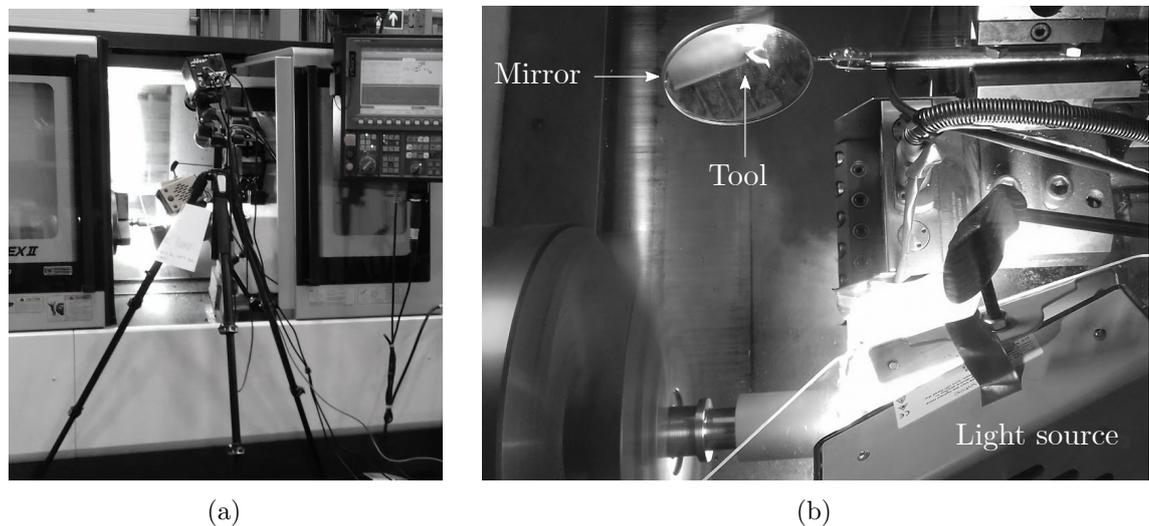


Figure 3.6: (a) The camera set up. (b) The set up of the mirror and light source inside the machine.

to the AE signal using a trigger, specifying the exact section of the data corresponding to the recorded video. Given the frame rate and aliasing issues, up to 400kHz of the frequency component of the signal could be studied using the camera.

Vast amounts of light is needed to capture these videos due to the short exposure times of the camera. To compensate, a bright light source needs to be fitted inside the machine directed at the cutting zone (figure 3.6b). It is not possible to record the entire tool pass as video length and file size is governed by the frame rate; a higher frame rate results in a very large data file that takes a long time to process. Due to swarf bunching and focusing issues, only the start of the pass for some of the tools were recorded using the high speed camera. A mirror was used to focus the camera on the tool due to the cutting edge orientation.

Due to these limitations, the videos could only be captured up to around 50,000 frames per second. Faster frame rates require a significantly higher amount of light at much lower resolutions. As a result, the recorded videos were not detailed enough at the current frame rates for analysing AE signals at 1MHz sampling frequency.

### **Temperature measurements**

In a previous feasibility study, thermocouples were attached to the cutting tools to investigate the temperature generation in PcBN during hard turning. When positioned on the surface of the tool, thermocouples measured the local heat generation from the bunching of swarf, reducing the usefulness of the data for analysis. The restricted view and the localised temperature increase due to swarf bunching also eliminated the use of thermal cameras for thermography. When thermocouples were attached near the cutting edge through an electrical discharge machined hole, the positioning of the thermocouple within the hole was a significant variable that affected the temperature readings. It was not possible to attach the sensor to the correct position without the guidance of computed tomography scans that added significant time to the setup procedure. As a result, the use of thermocouples were disregarded in the main setup of this EngD.

### 3.3 Main test setup of the EngD

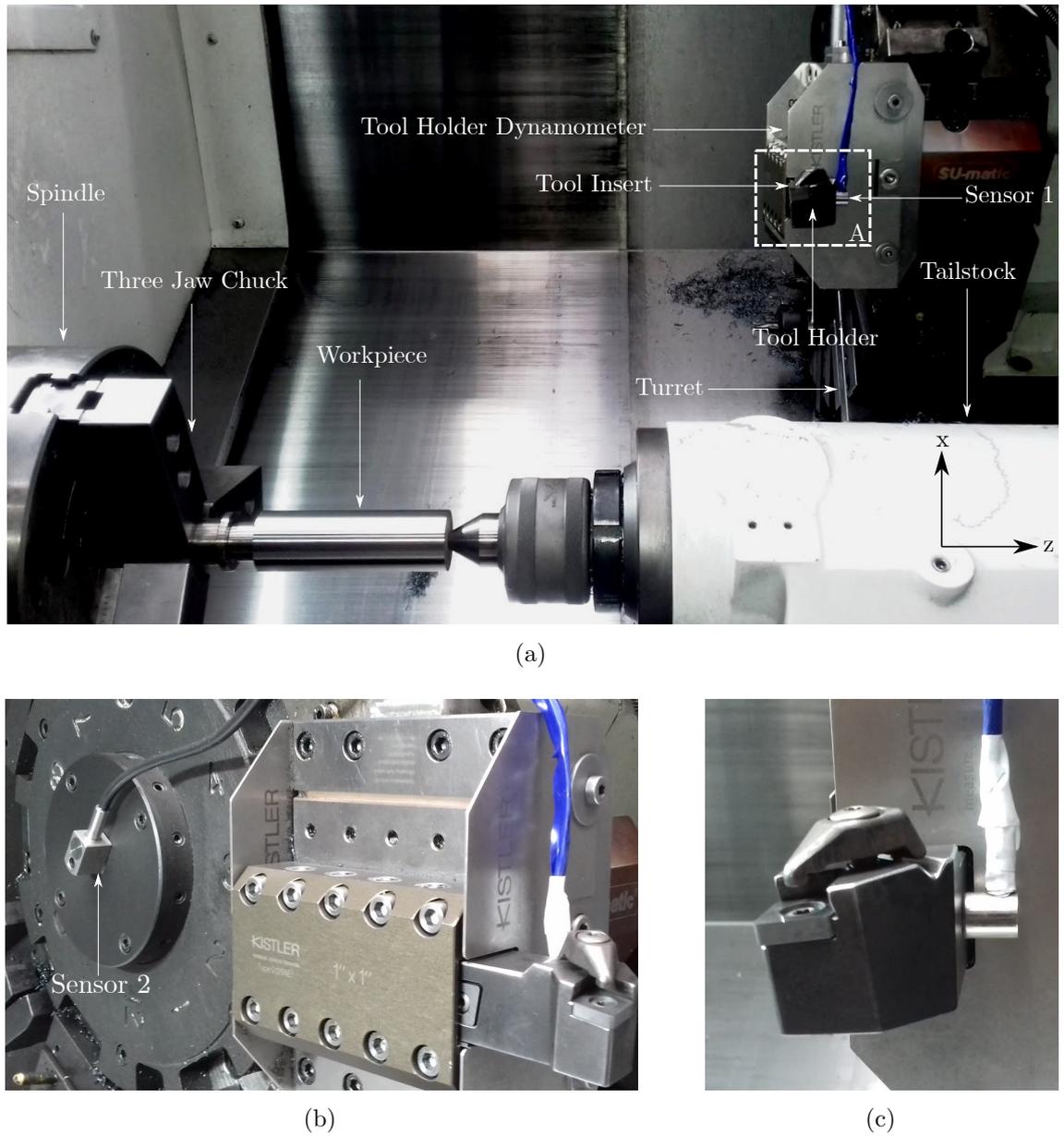


Figure 3.7: The new rig showing (a) The entire set with the tool holder highlighted in box A. (b) The set up (looking from the z direction) showing Sensor 2. (c) An enlarged view of the tool holder (box A) with Sensor 1.

The main findings and any machine learning algorithms suggested by this EngD will be implemented at the facilities of E6. Therefore, it is beneficial to collect the test data for the algorithms from the same facilities. This is because many factors of the factory floor could influence the data. For example, the lathes used in these tests may behave dynamically differently to others in other factories. The type and layout of machinery in the factory could also affect the data as some vibrate and behave differently to others. The setup inside the machine, stipulating the movement of tools and workpieces, is often varied for different machines. The main test set up is combined with the current practices at E6 in order to reduce the effects caused by these variations. Additionally, updating the current procedures, as opposed to entire redesign, assures the rig is realistic in an industrial setting.

The main test setup is shown in figures 3.7a to 3.7c. The decisions made regarding equipment used here are explained in the previous sections in this chapter. An Okuma Space Turn lathe LB3000 EXII was used for this experiment. Addition of AE to the current rig introduces previously unavailable frequency information about the system. To assess and understand these AE signals, swarf was collected.

The type of tools used for data collection and the tool wear measurements are presented in greater detail in section 3.3.1.

### 3.3.1 Tools and tool wear measurements

Two different types of tools were used for collecting AE; the only variation being the material composition, or ‘grade’ of material. The grade of tool specifies the percentage of cubic boron nitride (cBN) particles within its material composition. The grade has a direct effect on the behaviour of the tool, and defines its suitability to a type of machining operation. For a continuous turning operation<sup>4</sup> such as this one, low concentrations of cBN particles are required in the material composition. By controlling the variability of tools in terms of shape, size and orientation, the dissimilarities in AE can be limited. Henceforth, the two grades used will be names Grade A and Grade B.

---

<sup>4</sup>In a continuous turning operation, the tool does not break contact with the workpiece during a pass. This is a ‘low’ severity operation as there is no cyclic loading on the tool.

Material Composition	Tool Label	Pass Intervals
Grade A	A3 to A9	4
	A10	1
Grade B	B2 to B8	4
	B9 & B10	1

Table 3.1: Specified in this table are the tool grades, names, and pass intervals.

In order to train or test any future algorithms, multiple tools from each grade were tested under identical cutting conditions. To track the trends in tool wear with a higher resolution, during some of the repeat tests, the tool was measured after every pass. Table 3.1 details the grade, and label given to each tool, and the rate at which it was measured.

Tool wear measurements are conducted using an optical microscope and a 3D scanning microscope to measure flank and crater wear respectively; the wear scars are too small to be seen by the human eye. The 3D microscope works by taking multiple 2D scans of the object in question from top to bottom and then layering these scans consecutively to make a 3D model. The accompanying software of the microscope is capable of conducting analysis and producing quantitative evaluations about the scanned object. An in depth discussion on this is found in Chapter 4.

When using the optical microscope, lighting and human errors are major factors of measurement inaccuracy. Lighting can often trick the human eye into observing an incorrect wear land. Also, length of the average wear land is a subjective measurement leading to variation amongst operators. The 3D microscope is also sensitive to light and unwanted foreign objects or material deposition on the measured surface. Care must be taken when using these microscopes, leading to longer measurement times.

### **3.4 Summary of experimental procedure.**

An experiment setup was designed around the current procedure in place at E6 to reduce costs, save time and maintain reliability following feasibility work. The designed focused mainly on the addition of AE sensors to collect data. Ultimately, the data collected using the new test setup will be used in a machine learning algorithm, that should alleviate the need for time consuming, and therefore expensive measurements of tool wear.

To find correlations between tool wear and the measured signals, the images and scans taken from the microscopes are analysed in Chapter 4. Data from the AE sensors are processed and analysed to obtain features for machine learning algorithms in Chapter 5.

---

## TOOL WEAR FEATURES

Tool wear is complex. It is not possible to categorically quantify tool wear using just one value. Therefore, descriptive labels of tool wear are often established using a combination of different wear types and characteristics that vary over the tool life.

The focus of this chapter is to investigate tool wear in more detail than previous work conducted in this research field. With the help of three dimensional tool scans throughout the life of tools, the aim of this chapter is to find characteristics of tool wear that may be more informative and useful than the ones currently available. By doing so, it may be possible to further understand the mechanics of tool wear and the cutting process, and consequently, generated acoustic emissions.

Wear scars seen on PcBN tools obtained from the experimental setup in Chapter 3 are presented in Section 4.1. Current practices involved with measuring these wear scars and their drawbacks are discussed in Section 4.2, followed by a novel method of extracting tool features in Section 4.3. The extracted tool features are detailed in Section 4.4. In order to find the most representative damage labels, feature selection is carried out in Section 4.5, and the selected features are presented in Section 4.6.

## 4.1 Wear on PcBN tools

In the last few decades, the application of PcBN tools for machining hardened steel has been a significant advancement. Prior to the use of PcBN, it was not possible to machine hardened steel with the required tolerances. As a consequence, research into the field increased greatly in order to understand the effect of machining process parameters on tool wear [54, 55].

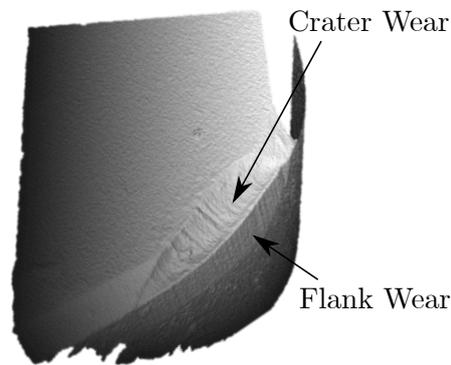


Figure 4.1: 3D model of a PcBN tool with flank and crater wear indicated.

There are two main types of wear present on PcBN tools used for turning case hardened steel according to ISO 3865:1993<sup>1</sup> [56]. As seen in figure 4.1, flank wear<sup>2</sup> acts on the flank face of the tool and crater wear<sup>3</sup> on the rake face of the tool. Additionally, notch and built-up-edge (BUE) wear can also be found on PcBN tools, though these are not substantial compared to flank and crater wear. These wear types are caused by: abrasion, adhesion, diffusion and chemical wear [57, 58]. An extensive literature review on PcBN and tool wear on finish hard turning can be found in [59]. In general, flank wear is caused by the contact between abrasive cBN particles and hard particles in the workpiece, and has an adverse effect on the surface finish of the machined workpiece [60, 61]. Diffusion and adhesion of material due to the increased temperature in the cutting zone, is the main cause of crater wear,

<sup>1</sup>The international standard for tool-life testing with single-point turning tools.

<sup>2</sup>Quantified as the average width of the wear land,  $V_B$ .

<sup>3</sup>Quantified as crater ratio,  $K$ ; the ratio between distance to crater depth and crater depth.

leading to reduction in process reliability [59]. Moreover, increase in crater depth can cause the cutting edge to weaken, leading to tool failure [14].

## 4.2 Current tool wear measurement techniques

The majority of research conducted on measuring, modelling and predicting tool wear relies heavily on visually studying damage using an optical technique [23,62–64]. Compared to unworn surfaces, wear scars are reflective, hence, tool wear can be captured on camera with relative ease, though problems can arise in the presence of BUE or material deposition [20]. Some authors opt for using scanning electron micrography (SEM), given the higher resolution and precision compared to optical microscopes [65–67]. These techniques are non-destructive, allowing the study of wear scars throughout the life of the tool. However, by their nature, these measurements interrupt the machining process, and especially in the case of SEM, increase measurement time. These methods also introduce depth of cut (DOC) uncertainty due to the repositioning of the tool post measurement. Given the DOC has an influence on the tool's performance, the intermittent inspection invariably has an effect on tool wear.

In an attempt to automate the tool wear measurement process, numerous researchers have explored the use of computer vision [68–74]. This technique uses images and videos of tools to analyse wear by attaching cameras to the experimental setup. In this case, the tool is not removed from the holder, thus reducing the variation in DOC, post measurement. Kerr et al. [70] applied non-linear operators in the pre-processing of tool images taken using a charged coupled device (CCD) TV camera to determine whether images can be used to obtain information about tool wear. Additionally, higher order statistical moments derived from 2D histograms and spatial frequency content were also studied. The authors suggest that though tool wear values can be extracted from some image properties, illumination, viewing conditions and contamination of the tool has an influence on the results, and therefore this technique is not robust for TCM currently. Pedersen [71] discussed challenges in identifying the differences in notch and flank wear when using CCD camera images.

Jurkovic et al. [72] uses a CCD camera alongside a laser diode with a linear projector to capture the surface deflection of the wear land by projecting laser lines on to the tool. It is stated that binary grey scale values can be used to obtain information about crater and flank wear. This technique is not able to provide a location of the damage using pixel values, and the projected laser lines does not have a sufficient resolution to capture the crater profile adequately. Pfeifer et al. [73] used a series of tool wear images from numerous positions and illuminations to infer wear values by studying the differences in image features. Satisfactory results of flank wear can be extracted, though it is not stated if notch wear can be distinguished, or whether crater wear can be detected. Sortino [74] applied statistical filtering to obtain flank wear values from turning and milling tools. The method is not adequate for wear lands smaller than 0.1mm, due to dependence on pixel width. This will not be suitable for the turning operation at E6 as the tools are often discarded at 0.1mm flank wear threshold.

As mentioned previously in section 4.1, edge breakage is closely related to crater wear. Therefore, it is important to measure the crater in detail. Most computer vision techniques are not focused on this issue. Lahiff et al. [14] performed three methods of measuring the crater. Firstly, a profilometer was used along the crater depth but the accuracy was unsuitable for this process. Secondly, white light interferometry was undertaken, but the resolution was found to be low at steep sections of the crater. Also, the tools needed to be gold-coated for analysis. Finally, the author partitioned the tools in half to study the cross section or 'profile' of the tool. Though the results were promising, this is a destructive technique that cannot be undertaken in an industrial setting.

Consequently, it is clear that a technique of measuring flank and crater wear through a non-destructive and repeatable method would be useful for tool feature extraction. By using optical interference fringe projection microscopy, Faraz et al. [75] studied the effect of edge rounding (or sharpness of cutting edge) of drilling tool profiles on composite materials. This work was one of the first to study tool wear features other than crater and flank wear. Zhao et al. [76] attempted this using tool profiles obtained from a 3D scanning microscope, similar to the one used in this EngD. The

edge radius of the tool was extracted using a Gaussian fitted circle method. The forces, tool life, and surface quality were found to be closely related to the edge radius. However, the fitted circle on the tool profile could be improved for higher accuracy. Ray et al. [52] (again, using a similar 3D scanning microscope to this EngD) also studied cutting edge radius and observed the changes in wedge angle throughout tool life, as well as many other features provided by the 3D scanning microscope software. Some tool features were found to be highly linearly correlated with indirect measurements of tool wear.

Exploring tool profiles provided by advanced technology of 3D scanning microscopes has established a new method of examining unconventional features that may be more informative than their predecessors, i.e flank land width ( $V_B$ ) and crater ratio ( $K$ ). Therefore, section 4.3 focuses on extracting new features from 3D tool scans that may be useful for tool life prognosis and well correlated to acoustic emissions of the machining process.

### 4.3 Novel method of tool wear feature extraction

The 3D scanning microscope software has the capability to split the 3D scan into a number of 2D cross sections, which are rich in information. An example of where these cross sections are applied to a scan by the software is shown in figure 4.2. The green lines represent the cross sections and the grey image is the tool model. For the tools studied here, 40 lines are applied to the model and around 15 to 17 of these cover the worn area of the tool. Henceforth, these cross sections will be referred to as tool profiles.

The microscope exports 2D profiles as Cartesian coordinates. An example of a profile from a new and a worn tool are shown in figures 4.3 and 4.4 respectively. Once plotted, the differences between these profiles can be used as 2D features in addition to flank land width ( $V_B$ ) and crater ratio ( $K$ ). Unlike previous work conducted on tool profiles discussed in section 4.2, the tool profiles are not averaged over the entire wear scar. Therefore, this method provides a higher resolution across the wear land. Moreover, studying the progression of wear features across the length of the scar

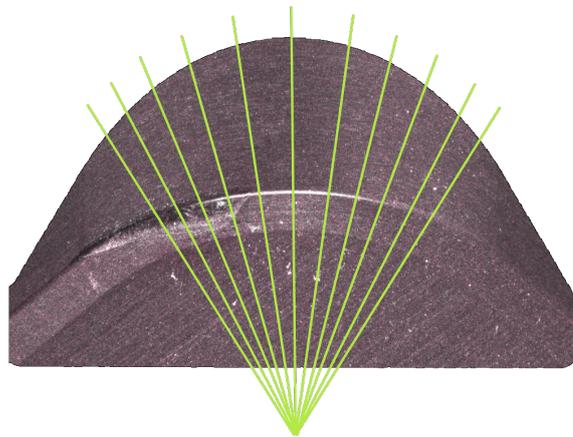


Figure 4.2: A reconstructed image demonstrating the way in which the 3D scanning microscope allocates the 2D cross sections or tool profiles to a tool scan.

has not been undertaken in the past. By doing so, it may be possible to locate the section of tool wear scar that has the most significant impact on acoustic emissions.

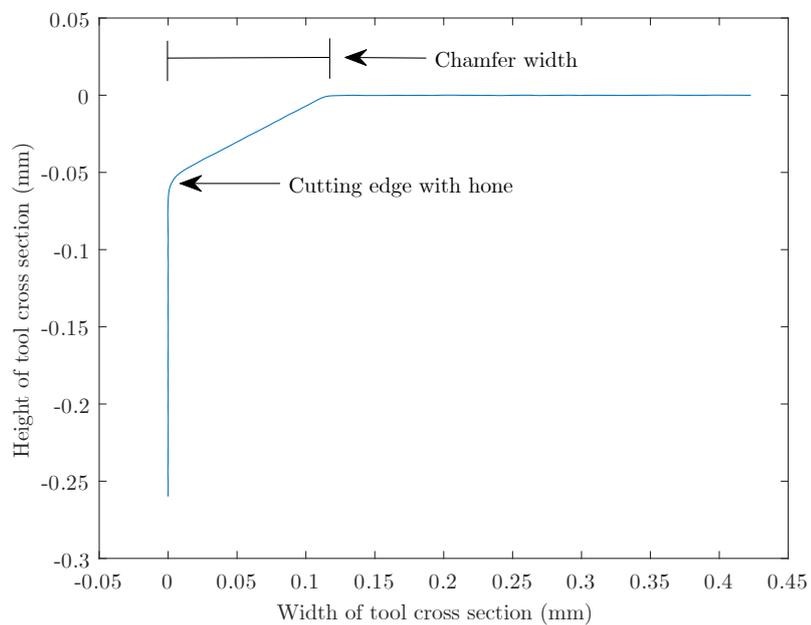


Figure 4.3: An example profile (or cross section) of a new tool with the cutting edge and chamfer highlighted.

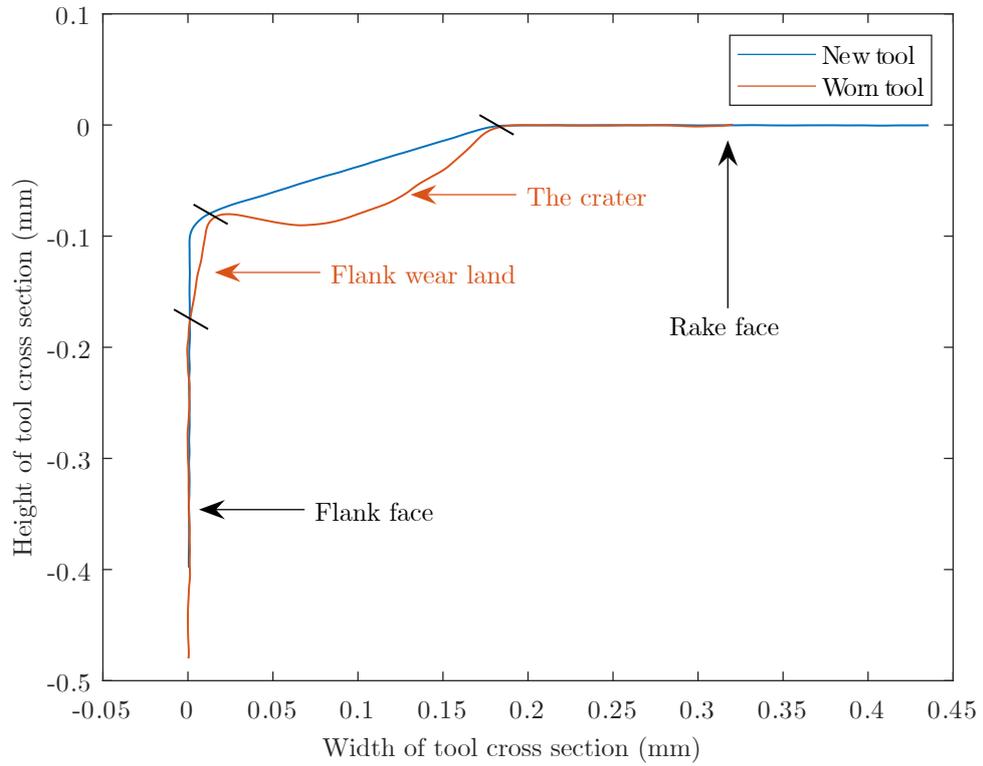


Figure 4.4: Difference between a new and a worn tool profile. The figure highlights the rake and flank face as well as typical flank and crater wear on the tools used in this work.

In this work, the profiles from three sections of the wear scar (leading edge (LE), centre (C) and trailing edge (TE)) are examined and compared. These areas are shown in figure 4.5. The profiles within each section (eg. centre) is averaged here.

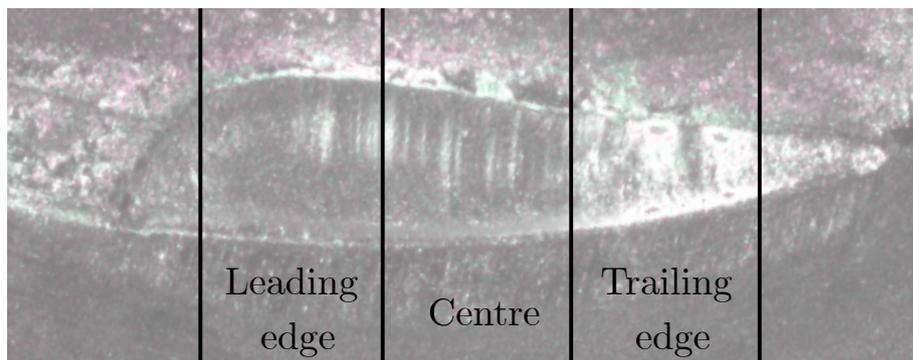


Figure 4.5: An example of a tool wear image divided into three sections of the wear scar.

### 4.3.1 Challenges of the tool feature extraction method

As the 3D scans are of an extremely small area of the tool, complications can arise when analysing the models. The positioning and orientation of each tool at the point of measurement is important for later comparison. If the tool scan in a respective pass is captured at a considerably different position to the previous scan, the profiles will not be aligned. If the orientation of each scan is markedly different to one another, the resulting feature values will be affected. In order to reduce the effect of these issues, a jig was used to secure the tools in place during measurement.

Resolution of tool scans can suffer when measuring deep craters. This is due to the relative angle between the microscopic lens and the tool wear scar. Unfortunately, a compromise must be made between imaging the flank face and the resolution of the rake face as it is not possible to achieve a detailed model of both surfaces.

## 4.4 Tool wear features

Numerous features from the flank and crater wear scars can be extracted using the profiles discussed in section 4.3. These will be presented in Section 4.4.1. Additionally, standard features currently extracted using the optical and 3D scanning microscopes will be presented in Section 4.4.2 for comparison.

In this work, the extracted features are aligned to four categories describing the feature type. Features which are associated with the flank and crater wear are allocated to ‘flank’ and ‘crater’ categories respectively. All features that quantify a change in area between profiles are designated into an ‘area’ category. The cutting edge shape that is usually indicative of tool sharpness is appointed to a ‘sharpness’ category. Finally, global volumetric difference between the worn and unworn tool surfaces are assigned to a ‘volume’ category.

### 4.4.1 Features from the novel approach

Ideally in an industrial setting, an automated tool feature extraction process is favourable when compared to the current procedure of manual measurement, owing

to higher accuracies and speed; automatic extraction is not affected by human error and also bypasses the need to manually align tool scans and highlight wear zones. With this in mind, the extraction process in this work is designed to reduce the level of human intervention when possible. The algorithm used to obtain tool features in this work generally use a combination of algebra and trigonometry to calculate the features from Cartesian coordinates automatically. However, the calculations require reference points that can be challenging to capture in an automated way. Five of these reference points (x and y coordinate pairs) are manually chosen for each profile at the start of the extraction process. These will be explained in detail when discussing the relevant extracted features. As the manually picked reference points can be affected by human error, the feature extraction process is repeated three times and averages are saved for analysis.

The cutting edge (illustrated in figure 4.3) is the first point of contact between the tool and the workpiece. Naturally, the area surrounding the cutting edge is therefore first to wear. Subsequently, it is important to identify the cutting edge of the new tool profile first, to use as a reference for making comparisons in the consecutive worn profiles. For the profiles used in this work, the cutting edge is automatically detected by identifying the point at which the direction of the x coordinates undergo the largest change. Here the standard deviation of x-coordinates is used to capture this directional change (using y coordinates is also sufficient).

Figure 4.6 graphically represents the features that can now be extracted by comparing new and worn profiles. Previously unexplored features obtained by the 2D profiles are marked (N) for new.

### **Flank**

Due to the orientation in which the tool is mounted for measurement, the microscope captures the tool rake face with higher resolution than the flank face (the rake and flank face has been annotated in figure 4.4). As a result, the average length of flank wear land ( $VB_B$ ) as specified by the ISO standard 3865:1993 [56] is challenging to detect automatically when calculating the difference between the worn and new profiles. Therefore the end of the flank wear land has to be chosen manually for all profiles.

1. **Average flank wear length,  $VB_B$**  - The length of the wear land on the flank face, calculated here as the vertical distance between the unworn cutting edge (or worn cutting edge 'A', the point perpendicular to the unworn cutting edge) and the end of flank wear land (the manually picked reference point). These points have been presented in figure 4.6a.

Flank wear is the most common type of damage present in cBN inserts [59]. The wear land affects the surface finish of the machined component and therefore has a significant influence on the machining process. Usually, the average length of the wear land across the wear scar is used to quantify flank wear ISO standard 3865:1993 [56]. In this work, the average wear land is calculated three times within the sections detailed in section 4.3. This way, a local value of average flank wear can be achieved.

Wear scars shine under lighting, enabling easy inspection. The lighting can, however, trick the eye into measuring the wear scar from worn cutting edge 'B' to flank wear land in figure 4.6a when using optical microscope for manual measurement. The resulting behaviour of flank wear therefore can seem non-monotonic. Indeed, the measurement should be conducted between worn cutting edge 'A', the point perpendicular to the unworn cutting edge, and the flank wear land. Therefore measuring flank wear land with tool profiles has the potential to reduce manual measurement errors.

2. **Flank wear angle (N)** - Angle between the new tool profile and the flank wear land, calculated here by finding the angle between the original cutting edge,  $VB_B$  and the worn cutting edge 'A'. Figure 4.6a displays this angle. A reference point is chosen manually from the flank face to calculate this angle. It is challenging to pick this reference point automatically by the algorithm as the flank face is not always smooth between the flank wear land and unworn cutting edge 'A' due to the nature of the wearing process.

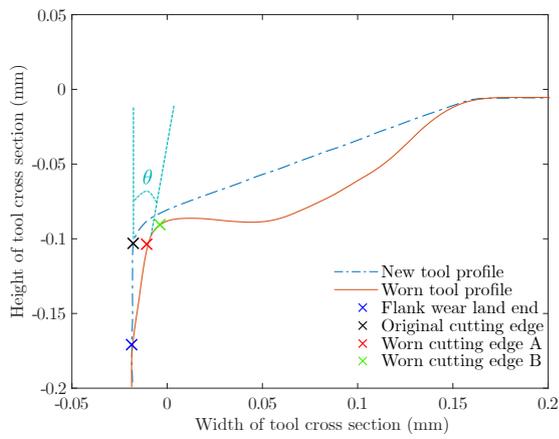
Research into tool wear often does not investigate the effect of flank angle on the machining process as it is not possible to measure the angle without acquiring tool profile information. This angle is influenced by flank wear and reduces the uncut chip thickness, affecting the final dimensions of the workpiece.

3. **Flank ratio (N)** - This is the ratio of flank angle over  $VB_B$ . With increasing sliding distance, it seems that at times, the wear is progressing further into the tool, increasing the flank angle, whilst the length of the wear land does not increase significantly. At other occasions, the wear land is appearing to increase but the depth of the flank wear seems to remain relatively unchanged. The flank ratio is expected to take these two mechanisms into account, and present a single feature that could be more indicative of the tool wear progression.

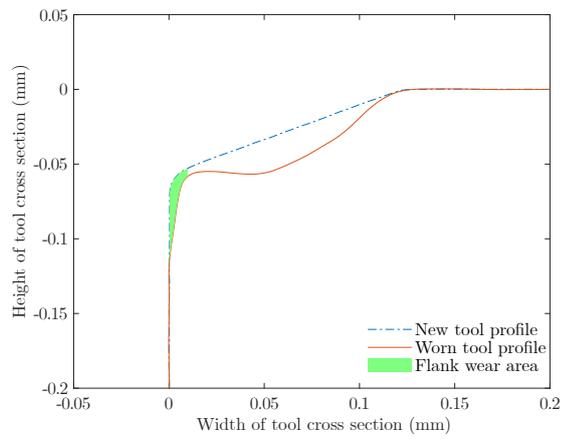
### Crater

As crater wear progresses, its shape and size changes drastically. As a result, the algorithm struggles to pinpoint the start and end of the crater across all profiles. Therefore, for ease, these points are picked manually in this work.

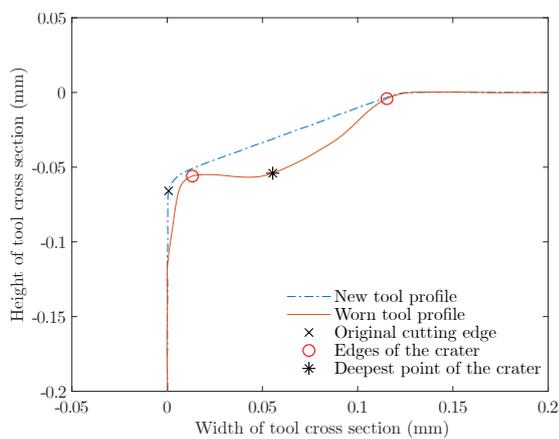
4. **Maximum crater depth** - The perpendicular distance from the new tool chamfer to the deepest point of the crater is the maximum crater depth. An example has been shown in figure 4.6c. The black star represents the deepest point. This is a well researched characteristic of crater wear as it can influence cutting edge failure. However, this is a challenging feature to measure without having access to cross sections of the tool or a metrology system capable of scanning the crater. Therefore this feature is not studied often.
5. **Crater centre distance** - The length between the original cutting edge and the deepest point of the crater. This feature can be challenging to measure without making comparisons to the new tool, therefore it is difficult to find researchers that have attempted to measure this feature.
6. **K ratio** - The ratio of the crater depth over the crater centre distance. This is the standard feature usually used to quantify crater wear. Though it is a standardised feature, it is difficult to quantify as the relevant features are challenging to measure.
7. **Radius of the crater (N)** - The radius of the circle in 4.6d that approximately fit the crater as the crater is not a perfect segment with a smooth arc. The shape of the crater could influence chip evacuation as the chamfer surface has changed. Therefore, it is useful to obtain a feature that can quantify this behaviour.



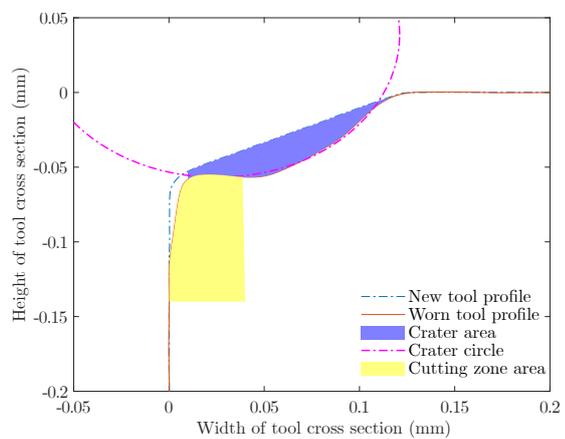
(a) Flank features



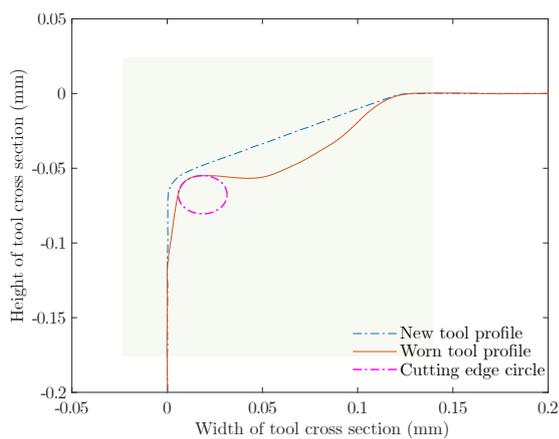
(b) Area feature



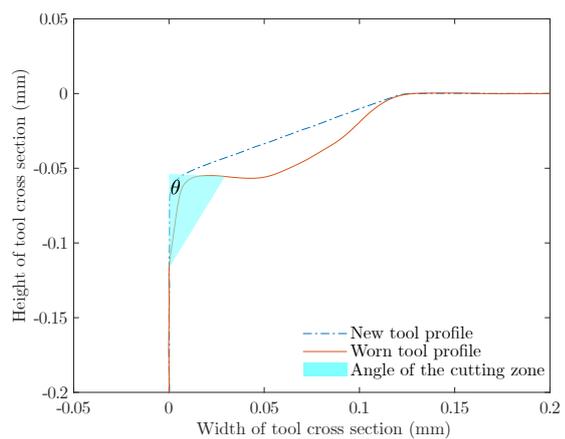
(c) Crater features



(d) Crater and area features



(e) Sharpness and area features



(f) Sharpness features

Figure 4.6: Figures (a) to (f) present the features extracted from tool profiles. The points in figures 4.6a and 4.6c have been used to plot the areas in figures 4.6b and 4.6d.

## Area

8. **Flank wear area (N)** - The approximate area of material removed from the flank face, as highlighted in green in figure 4.6b. This is the area under the new tool profile, and flank wear land up to the worn cutting edge 'B'; points annotated in figure 4.6a.

Area of flank wear may be more informative than  $VB_B$  as this feature takes into account the material loss from the rake face removed due to flank wear as well.

9. **Crater area (N)** - As shown in figure 4.6d, crater area is the material removed from the rake edge to form the crater. In this work, this feature is calculated by finding the area under the new and worn tool profiles between the start and end of the crater reference points that were picked manually.

Given this value is calculated across the wear scar, it could be more informative than the common crater measurement, maximum crater depth. It is not possible to calculate this feature without obtaining and overlaying cross sections of the new and worn tools.

10. **Area of the cutting zone (N)** - In figure 4.6d, the cutting zone area can be seen in yellow. Here, the area under the worn tool profile has been calculated within the limits of x smaller than 0.04 and y bigger than -0.14 to obtain areas that can be compared across all profiles and all tools. This area becomes weak throughout tool life and is a common failure mode of the worn tool.
11. **Overall difference in area (N)** - Difference in area between the worn and new profiles within the grey shaded area in figure 4.6e specified for all profiles between  $x = -0.02, 0.14$  and  $y = -0.17, 0.02$  to aid comparison. The area was picked by studying all profiles. This parameter gives a localised value for the amount of wear on a tool.

## Sharpness

12. **Radius of the cutting edge or edge roundness (N)** - Figure 4.6e shows the circle that fit worn cutting edge. The coordinates that are used to fit this

circle is found by taking the value of worn cutting edge 'B' and isolating the points that are up to 10% away from the value in the y direction.

The new cutting edge is drastically different from the worn edge. This new shape is usually the main prerequisite for tool failure. The value of edge roundness is usually extracted directly from the microscope software. Though sometimes accurate, the results can be unreliable as the software is prone to incorrectly isolating the cutting edge and leading to false edge radius values.

13. **Angle of the cutting zone (N)** - This is the new effective rake angle within the chamfer shown in figure 4.6f, calculated as the angle between the unworn tool flank face, and onset of the crater. The angle of the cutting edge is a feature that represents the 'sharpness' of the worn tool. This is similar to the radius of the cutting edge. However, it is sometimes difficult to fit a circle to lightly worn tools. Therefore, the angle of the cutting zone may provide a better insight into the sharpness of the tool at early stages of tool wear.

There are incidents where the sharp cutting edge smooths instead of failing, leading to longer tool life. At these instances, this angle may provide insights into the cutting edge behaviour.

The features that have been extracted up to now can be identified physically using tool profiles. As the profiles can be plotted using Cartesian coordinates, they can also be treated as data that can be used for statistical analysis. Therefore the statistical moments of each profile's coordinates from leading, centre and trailing edge of the wear scar is also studied in this work. Each moment is an average between the x and y coordinate.

14. **Mean (N)** - The average of the worn tool profile.
15. **Standard deviation (N)** - Calculated variation found on the tool profile coordinates.
16. **Skewness (N)** - A parameter describing the symmetry of the distribution of worn profile coordinates.
17. **Kurtosis (N)** - Quantifies the tail shape of worn profile coordinate distributions.

#### 4.4.2 Features extracted from standard microscopy

Section 3.3.1 presents a brief introduction to the type of tool wear measurement techniques that were undertaken during the main tool wear test of this EngD. A detailed list of the standard wear features from the optical microscope followed by the 3D scanning microscope are presented here. As before, the features are separated into the wear categories mentioned previously.

The optical microscope is used to collect an image of the tool's flank face during testing due to its speed and ease of use. The images are used to visually inspect the tool between passes and manually derive the flank wear values. Currently, this technique is the only method of examining the tool during testing at E6.

#### Flank

18. **Flank wear measured using the optical microscope** - Extracting this features requires skill and prior knowledge of flank wear. For example, notch wear may be present in the edges of the wear scar shown in figure 4.7a. It is important not to include notch wear in the measurement, but an inexperienced operator may not be able to distinguish between notch and flank wear. This issue is avoided in the novel feature extraction method stated above as profiles from the extreme edges of the wear scar are not used for extraction. Another source of inaccuracy is caused by incorrectly positioning the datum line<sup>4</sup>. This is the line that should sit on the cutting edge of the tool. The results from a badly chosen reference line could lead to flank wear appearing to reduce in size as tool life increases. In reality, this is not possible. Figure 4.7b presents the reference lines that should be used to capture average flank wear width from an image.

The 3D scanning microscope software can be used to calculate global three dimensional features such as depth and volume of defects. This is done by operators manually overlaying scans from the unworn and worn tools, and undertaking time consuming difference calculations once the tool test is complete. Therefore, this method of

---

<sup>4</sup>Effectively measuring flank wear from worn cutting edge 'B' instead of worn cutting edge 'A' in figure 4.6a.

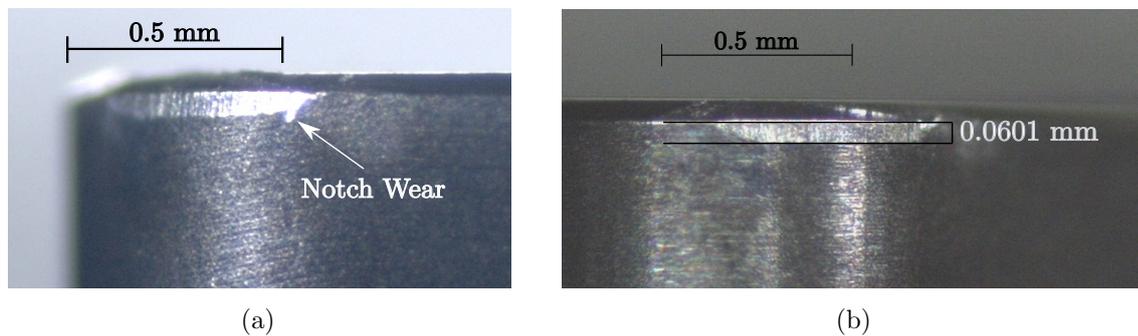


Figure 4.7: (a) Image showing notch wear. (b) Flank wear land measured on an image obtained using the optical microscope.

feature extraction is not conducted in-process. The technique used to select the area of interest for comparison can be a source of inaccuracy in the results. If the chosen area is not focused on the tool wear, the results will not be indicative of actual wear. For example, in figure 4.8a, the deviation area shown in blue and red, which represents the defects below and above the reference surface respectively, are largely highlighting the tool wear and material deposits. Whereas in figure 4.8b, it is clear that the areas of highest deviation is not entirely due to tool wear. This type of inaccuracy usually stems from misalignment of the scans. To compensate, it is possible to isolate the wear scar during analysis. Though this increases accuracy of results, it also increases the time taken to conduct detailed analysis.

The 3D features selected from the microscope are:

### Crater

19. **Maximum deviation below reference surface** - The distance to the deepest point of the crater is picked by the 3D microscope when scans are overlaid. This feature is not affected by the misalignment of the scans as the crater depth is often larger than the deviation caused by an inaccurate overlay. This feature is similar to maximum crater depth from the novel extraction method. The difference here is that this feature is a global measure whereas the feature in the novel method can be calculated locally to sections of the wear scar. In this chapter, these features are named as such for clarity.

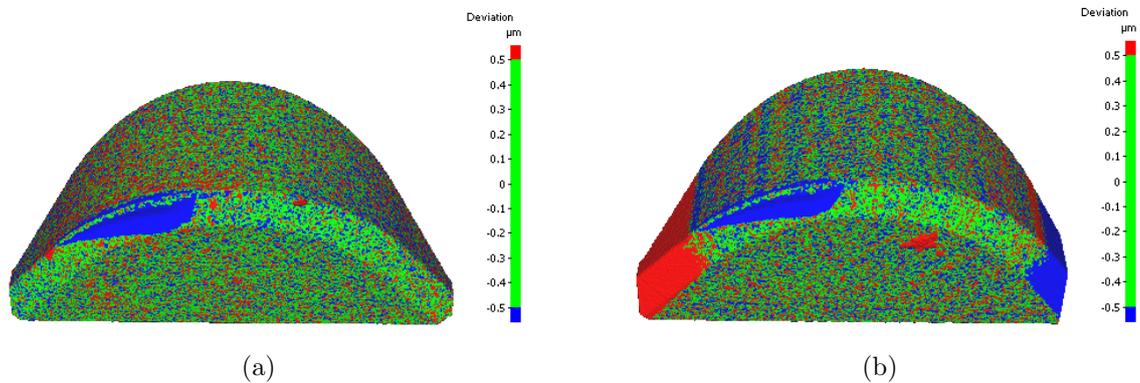


Figure 4.8: (a) Correctly aligned 3D scans highlighting tool wear area. (b) Misaligned 3D scans that misrepresent the defects of the worn tool.

## Volume

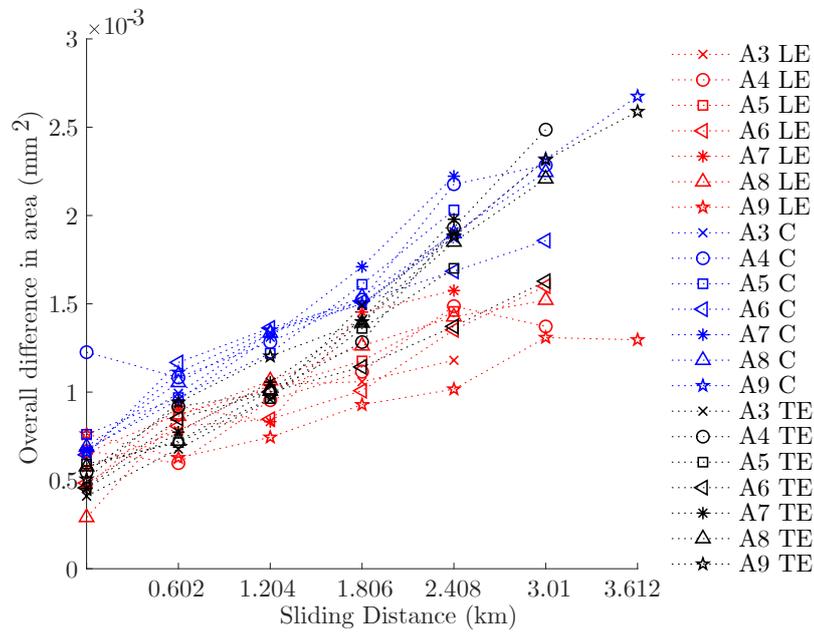
20. **Maximum volume below the reference surface** - A global representation of the volume loss from the rake face (crater wear) and the flank face (flank wear). This feature can be affected by the scan overlay issues if the wear scar is not isolated.

## Area

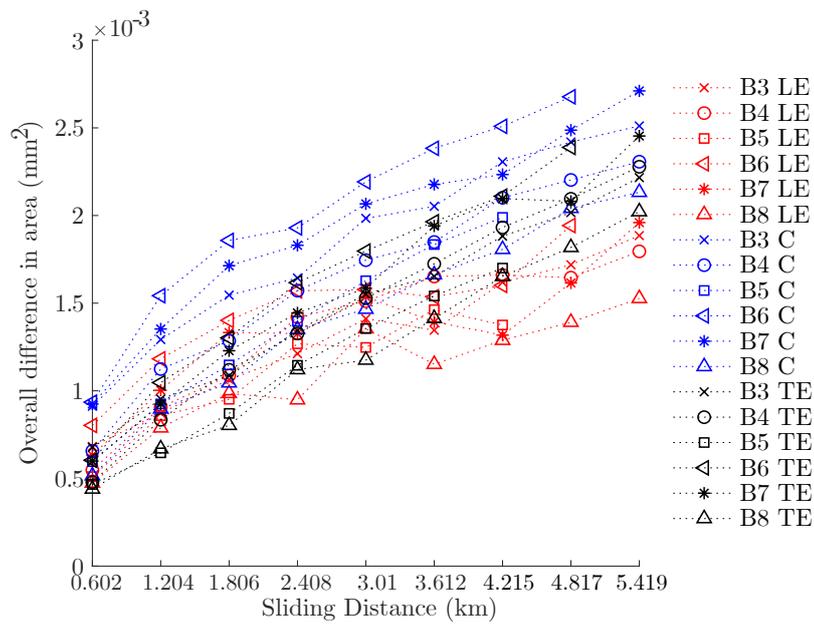
21. **Whole area of defects (ISO 8785)** - Sum of the areas of deviation above and below the reference surface. This feature in particular is prone to inaccuracies due to previously mentioned scan misalignment issues.

### 4.4.3 Extracted features - An example

According to the methods described in Sections 4.4.1 and 4.4.2, all 21 features were quantified for all tools in each grade introduced in Section 3.3.1. To visualise the typical trends that may be present, the overall difference in area between the worn and unworn profiles (feature 11) for grades A and B have been shown in figures 4.9a and 4.9b respectively as examples. This feature was chosen as an example as it was expected to increase monotonically with sliding distance due to material loss of the tool caused by wear. It should be noted that sliding distance here is the distance



(a) Grade A tools



(b) Grade B tools

Figure 4.9: Overall difference in area between worn and new profiles throughout tool life showing data from all sections (LE - leading edge, C - centre and TE - trailing edge).

that the tool has travelled during machining. The entire list of extracted features have been plotted and presented in Appendix A.

Immediately, it is obvious that wear scar section has an influence on the behaviour of this particular feature. Many other features also exhibit this characteristic. Another important observation is that all tools of the same grade appear to behave relatively similarly within each wear scar section for this particular feature, suggesting that a supervised tool wear prediction algorithm that works on one tool could perform well across all tools.

## 4.5 Feature selection for class labels

Out of the total set of features listed in Section 4.4, some will represent the tool wear scar and its progression better than others. The features that are most representative of tool wear can be used as descriptive class labels of damage that can be used in prediction algorithms as responses later. Therefore, it is useful at this stage to discriminate all the features against an intuitive criteria set in order to find the ones that outperform the others. A filter approach is undertaken here; this is the process of selecting tool features or damage labels independent of the algorithm used for inference and other outputs of the process, specifically in this case, the AE features. In other words, redundant tool features are removed from consideration prior to studying the indirect measurements of tool wear.

In the next section a series of suitable ranking criteria are discussed to assess which features may be the most useful for tool condition diagnosis and prognosis. Each of the features discussed above will be assigned a ranking score for each of the criteria introduced. This ranking score will be relative rather than absolute. The importance of each of the criteria will be weighted before feature selection.

### 4.5.1 Ranking criteria

- **Monotonicity** - Tool wear increases over the course of tool life. Therefore the features should be predominantly monotonic as tool wear is irreversible.

A test for monotonicity using the sign method is presented in equation 4.1 where  $j$  is the tool number,  $x$  is the feature vector,  $M$  is the number of tools monitored, and  $k = 1 \dots N$  where  $N$  is the number of measurements in each tool feature [77].

$$\text{Monotonicity} = \frac{1}{M} \sum_{j=1}^M \left| \sum_{k=1}^{N_j-1} \frac{\text{sgn}(x_j(k+1) - x_j(k))}{N_j - 1} \right| \quad (4.1)$$

- **Difference** - The features that undergo a large change during the life of the tool may be more descriptive of the damage as they may be responsive to the process. Therefore the largest difference between the start and end of tool life was calculated for the normalised features using equation 4.2.

$$\text{Difference} = \frac{1}{M} \sum_{j=1}^M \left| (x_j(1)) - (x_j(N_j)) \right| \quad (4.2)$$

- **Stability** - Features with low variance can be considered as more stable with increasing time across all tools. The standard deviation of the features for all tools within the same grade were calculated, and features with lowest variation were ranked higher than others.
- **Entropy** - Features with non uniform distributions are low in entropy, suggesting the tools behave dissimilarly to one another. Marginal entropy of the features were calculated using equation 4.3 [78] where  $P(x)$  is the probability of  $x$ . The features with highest entropy ranked higher than others.

$$H(X) \equiv \sum_{x \in A_X} P(x) \log \frac{1}{P(x)} \quad (4.3)$$

- **Smoothness** - A smooth feature does not have a large local curvature. The leave-one-out (LOO) method using ordinary least squares was applied to find the best regression fit for each feature up to and including second order polynomials. The order with the lowest error was picked to be the best polynomial fit (equation 4.4). The normalised mean squared error (NMSE)

was used to rank the features from the best to worst fit to its best polynomial using equation 4.5.

$$\text{LOO}_{\text{error}} = \mathbf{X}^T \mathbf{y} - (\mathbf{X}^T \mathbf{X})^{-1} \hat{\boldsymbol{\beta}} \quad (4.4)$$

where  $\hat{\boldsymbol{\beta}}$  is the coefficient vector of the least-squares hyperplane when considering vectors.

$$\text{NMSE} = \frac{\sqrt{\frac{\sum_k^{N_j} (|\hat{y}_k| - |y_k|)^2}{N_j}}}{\max(y_k) - \min(y_k)} \quad (4.5)$$

- **Correlation** - If a feature is highly correlated with other variables, it is considered to be high in density. Mutual Information between features is calculated using equation 4.6 (where X and Y are two vectors written here in information theory notation) to quantify this characteristic. Mutual information quantifies how two variables move with one another. Two variables with a high mutual information are highly correlated, though the type of correlation is unknown. At this stage, features that are highly correlated to one another are ranked higher than others. For example, features that are not monotonic will not be correlated well with features that are monotonic. This means that using correlation will enable the discarding of features that have unusual trends.

$$I(X; Y) \equiv H(X) - H(X|Y) \quad (4.6)$$

Where the conditional entropy is,

$$H(X|Y) \equiv \sum_{x \in A_X A_Y} P(x, y) \log \frac{1}{P(x|y)} \quad (4.7)$$

Once ranked, it is beneficial to score and weight the features according to the importance and relevance of the ranking criteria.

### 4.5.2 Scoring and Weighting

The 21 features listed in section 4.4 were split into three groups for scoring:

1. Highest ranked 7 features are given a score of 3.
2. Features between 8 and 14 receive a score of 2.
3. A score of 1 is given to the lowest ranked features.

By scoring in batches, it is possible to identify the features that perform similarly.

Of the six ranking criteria, some are more significant than others, and therefore should influence the ranking to a higher degree. For example, due to the nature of tool wear, a monotonic feature with a large variation between its value at the start and end of a tool's life is more useful than one which has very low variation that does not increase over time. Additionally, a feature that is highly correlated with others does not necessarily suggest it is representative of the system. Consequently, it is important to weigh the criteria in order to highlight the best features.

The most representative features should be changing with time to reflect the physical process of tool wear. They should also be highly responsive, where the change in behaviour is substantial. The features should also be robust and stable across all tools of the same grade. Due to these reasons, monotonicity, largest difference and stability are weighted two times higher than the other ranking criteria.

Of course, this is a subjective process and the decisions on weighting will affect which features are deemed the most suitable. This pragmatic approach is robust if the weighting criteria is changed as the scoring based on the physical behaviour of the tool wear features affect the score more than the weighting.

### 4.5.3 Results of feature ranking

The results of the feature scoring and weighting for both tool grades from the centre of the wear scar have been presented in table 4.1 as an example. The original scores were doubled for the monotonicity, largest difference and stability criteria. In order

to rank these features, the weighted average of the scores were calculated. The results are presented in table 4.2.

Feature number	Tool feature	Monotonicity	Difference	Stability	Entropy	Smoothness	Correlation
8	Flank area	6	6	6	3	2	3
11	Overall difference area	6	4	6	3	3	3
20	Max. volume below ref. surface	6	6	6	2	3	2
4	Maximum crater depth	6	4	6	3	2	3
9	Crater area	6	4	6	2	3	3
10	Area of the cutting zone	6	4	6	3	1	2
19	Max. deviation below ref. surface	6	4	4	2	3	3
6	K ratio	4	6	2	2	3	3
1	$VB_B$	4	2	4	3	3	3
13	Angle of the cutting zone	4	6	2	3	2	2
18	Flank wear from optical microscope	4	2	4	3	3	2
3	Flank ratio	4	6	2	1	2	1
12	Radius of the cutting edge	2	4	4	2	1	2
14	Mean of profiles	2	6	4	1	1	1
2	Flank wear angle	2	6	2	1	2	1
21	Whole Area of defects	4	2	2	2	2	2
5	Crater centre distance	2	4	4	1	1	1
15	Standard deviation of profiles	2	2	6	1	1	1
17	Kurtosis of profiles	4	2	2	2	1	2
7	Radius of the crater	2	2	4	1	2	1
16	Skewness of profiles	2	2	2	1	1	1

Table 4.1: The weighted score given to wear features of tool grades A and B from the centre of the wear scar. Monotonicity, largest difference and stability are weighted higher.

Flank wear area, a novel feature, has the highest weighted average score for all tools from the centre of the wear scar. In fact, six of the seven highest ranked features were extracted using the novel approach and four out of the seven top features have not been studied previously in literature. Though not a novel feature, maximum crater depth calculated using the profiles ranked better than maximum deviation below the reference surface, calculated by the microscope software.

Feature number	Tool feature	Weighted average score
8	Flank area	2.89
11	Overall difference area	2.78
20	Max. volume below ref. surface	2.78
4	Maximum crater depth	2.67
9	Crater area	2.67
10	Area of the cutting zone	2.44
19	Max. deviation below ref. surface	2.44
6	K ratio	2.22
1	$VB_B$	2.11
13	Angle of the cutting zone	2.11
18	Flank wear from optical microscope	2.00
3	Flank ratio	1.78
12	Radius of the cutting edge	1.67
14	Mean of profiles	1.67
2	Flank wear angle	1.56
21	Whole Area of defects	1.56
5	Crater centre distance	1.44
15	Standard deviation of profiles	1.44
17	Kurtosis of profiles	1.44
7	Radius of the crater	1.33
16	Skewness of profiles	1.00

Table 4.2: The tool wear features for grades A and B from the centre of wear scar ranked according to the weighted average.

In general, the highest scoring features are monotonic whilst undergoing a significant overall change during the tool test, they are stable through time, whilst simultaneously presenting low variation between tools. These features are also correlated strongly with others.

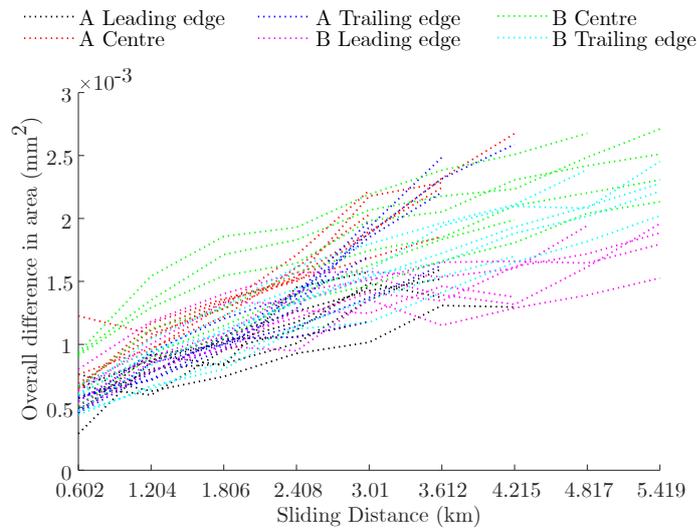
In order to visualise how the features have been ranked, examples of highly ranked to low scoring features are shown in figures 4.10a to 4.10c.

The highest scoring features are all monotonic, where most tools from both grades performed similarly. Figure 4.10a presents an example of a high scoring feature. It is possible that area and volume features rank higher in general than others as they are multidimensional in nature. 2D and 3D features capture the differences between the profiles better than 1D features as they are less affected by inaccuracies of the extraction process and model orientation, increasing stability.

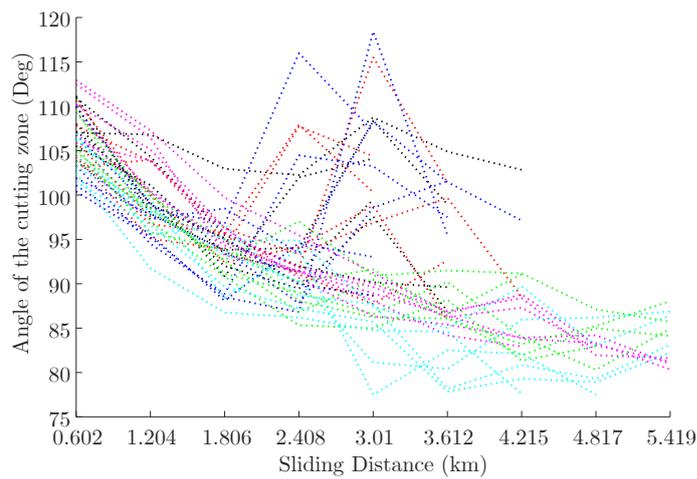
The two grades behave very differently in cutting zone angle as the edge wear is different for tool A compared to tool B (figure 4.10b). The large and rapid increase in the angle of the cutting zone for grade A tools is due to the smoothing of the cutting edge. Although this causes a reduction in the weighted average score, this feature could be useful in understanding the AE generated by the machining process as this change in edge shape could affect the chip formation and machined surface of the workpiece. A similar trend can also be seen in the cutting zone radius. Due to the sudden decrease in sharpness of grade A, the features are not monotonic, and the largest difference and mutual information are reduced whilst stability and entropy are increased. These features do not have a good polynomial fit due to low smoothness, resulting in high NMSE values.

The radius of the crater is an unreliable feature; it is challenging to extract owing to low resolution of the crater surface measurement. As the crater is not shaped as a perfect segment, it is difficult to capture its circumference and therefore radius by fitting a circle to it. The sharpness change of grade A and B near the end of tool life also has an effect on the crater radius and subsequently, this feature does not score highly.

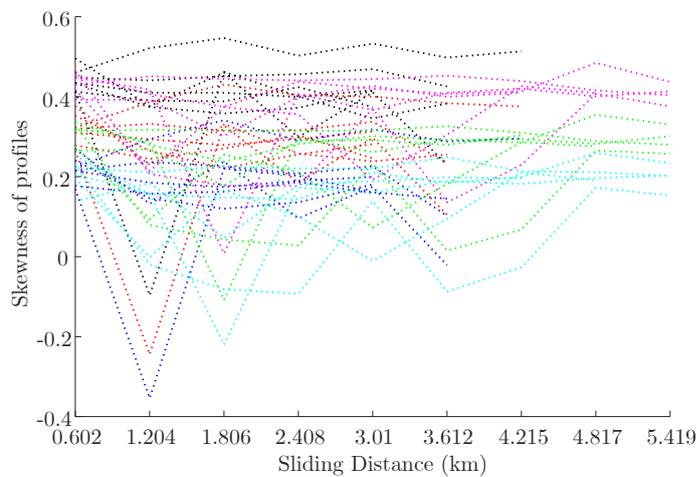
Flank ratio plateaus once the break in stage of the tool is complete as flank angle reaches the relief angle due to the mechanics of flank wear. Consequently, the



(a) Overall difference in area



(b) Angle of the cutting zone



(c) Skewness of profiles

Figure 4.10: Three examples of the ranked features where the ranking is high to low from (a) to (c).

resulting flank angle and flank ratio are not monotonic, have high entropy and low correlation with others.

The distance to the crater centre can vary in consecutive passes as the adhesion process responsible for the crater wear is inconsistent. Material from any part of the crater can potentially adhere to the chip and be removed, forming the new deepest point. This feature, by its nature, is therefore not monotonic, and does not have a large difference. It is also highly unstable. Due to this behavioural unpredictability, two tools are unlikely to behave similarly, leading to a low average score.

The profile features concerning the first four statical moments are not monotonic, and do not have a large difference. All tools behave differently, even within grade. It is not a surprise that the statistical features (mean, variance, skewness and kurtosis) of the profiles ranked low as it is challenging to represent the whole profile with just one single value. Unlike the others which have been extracted guided by physical changes, and only focus on part of the wear scar, these features attempt to generalise the whole profile. Therefore, they do not progress through time in an informative way. The progression of skewness values throughout tool life can be seen in figure 4.10c.

## **4.6 Highest ranking features for damage labels**

It is of interest to find the best features from each wear category in order to understand how tool grades and section of the wear scar influence the features. Therefore, the highest scoring features from each category for both tool grades across the wear scar have been presented in table 4.3.

From the results presented in table 4.3, it can be seen that features from tool A and B behave similarly; overall difference in area perform as the best feature for grade A, whereas flank area scores highest for grade B and angle of the cutting zone out-performs radius of the cutting edge for grade A. For grade A, except for sharpness, all three wear scar sections produced similar feature scores, while for grade B, the exception was in flank wear features. Within tool grade, variance across

Category	Highest ranking tool feature	A Leading edge	A Centre	A Trailing edge	B Leading edge	B Centre	B Trailing edge
Flank	$VB_B$ (1)	2.11	2.00	2.11	2.33		2.11
	Flank wear - optical (18)					2.00	
Crater	Maximum crater depth (4)	2.56	2.56	2.56	2.44	2.44	2.44
	Max. dev. below ref. surface (19)					2.44	
Area	Crater area (9)	2.78					2.78
	Overall difference area (11)	2.78	2.78	2.78			2.78
	Area of the cutting zone (10)			2.78			
	Flank area (8)				2.89	3.00	2.78
Sharpness	Radius of the cutting edge (12)	2.33			2.67	2.33	2.22
	Angle of the cutting zone (13)		2.00	2.00			
Volume	Max. vol. below ref. surface (20)	2.78	2.56	2.89	2.78	3.00	3.00

Table 4.3: The highest scoring tool wear features from each category shown for each tool and section of the scar.

the wear scar was lowest in the crater wear category; maximum crater depth was 2.56 for all sections on grade A and 2.44 on grade B.

These results show that tool wear can be different throughout grades and along the wear scar. However, it is also clear that few features performed relatively similarly across these differences. In order to find correlations between these labels of damage and the acoustic emissions of the machining process, this tool feature set will be reduced down to the most consistent in performance. For grade A, overall difference in area, maximum crater depth and maximum volume below the reference surface, are chosen. For grade B, flank area, maximum crater depth and maximum volume below the reference surface are selected.

## 4.7 Concluding remarks of tool features

It is well known that the edge preparation of a new tool has an effect on the machining process [79]. Similarly, the worn cutting edge of a tool has an influence on the chip formation and evacuation, surface finish of the workpiece and most importantly in this case, the length of tool life. By extracting features from 3D models of worn tools, it may be possible to better understand tool wear progression.

From attaining cross sectional profiles from the 3D models, 21 wear features were extracted from three sections, allowing the study of tool wear across the wear scar as well as throughout tool life. The novel technique of tool feature extraction allowed the examination of 13 previously unexplored features. When scrutinized against the feature ranking, scoring and weighting criteria, 4 out of the 13 novel features were amongst the best performing.

It is important to note here the subjectivity of the ranking criteria. Though the features were ranked according to practical measures, it is possible that different criteria could lead to a change in the order of the results. The criteria could be extended to including variables such as time taken for measurement, and repeatability of measurements in order to rank features that are consistent and easy to collect. In the future it would be intriguing to find the effect of this particular criteria when judged against another when considering the final feature rankings.

Four features from three categories have been chosen to represent the wear progression; these are, flank area, overall difference in area, maximum crater depth and maximum volume below the reference surface. The chosen features will be used as damage labels for supervised learning in Chapter 6.



---

## EXPLORING ACOUSTIC EMISSION

Some tool condition monitoring (TCM) algorithms require training on input data in order to predict responses such as flank wear. In Chapter 4, the extracted target features of the tool wear test was discussed in detail. Subsequently, this chapter focuses on the inputs, acoustic emissions.

Though acoustic emissions generated during the machining process have been studied in the past (as reviewed in Chapter 2), it is fair to say that insight into the generation of AE during cutting, in terms of what can be learnt from the structure and content of these signals, is still lacking. In addition to this, there exists gaps in the literature where oblique turning is concerned. It seems that much of the past research has been conducted on orthogonal turning, possibly due to the simplicity in the cutting mechanics. However, as oblique turning is used abundantly in industry and research, it is important to study acoustic emissions in this setting. Furthermore, little work has been dedicated to investigating the effect of PcBN tooling on the frequency content of generated acoustic emissions in oblique turning. It is the aim of this chapter to expand the understanding of the acoustic emissions signals using a combination of chip analysis and signal processing. A greater comprehension of the generated time series may allow the extraction of more informative input features, consequently improving the predictive performance of any learning algorithms that may be used later in this work.

Building on the review in Chapter 2, related work undertaken on acoustic emissions and tool wear is reviewed in Section 5.1. The acoustic emissions dataset used in this work is introduced in Section 5.2, followed by an investigation of the data in the frequency domain in Section 5.3. A solution to the challenges associated with handling acoustic emissions data for this particular experimental procedure is suggested in Section 5.4. The type of features that may be suitable to represent the tool wear process as input features in tool condition monitoring algorithms are presented in Section 5.5.

## 5.1 Acoustic emission in turning

In general, the release of acoustic emissions (AE) is due to fatigue fracture and friction, according to a review by Hensman [80]. In fatigue fracture, stress is released by crack extension, extension of the plastic region and faces of the crack rubbing against one another. Frictional effects such as adhesion, impact wear, erosion, abrasion and oxidation are also detailed as sources of AE.

The most common use of AE measurements is in the field of non-destructive testing; the release of elastic stress waves during crack formation is detected by sensors in the form of AE bursts. When measuring AE of the turning process, however, the received signal appears considerably different. It is no longer in the form of separable bursts; instead, it is a continuous, stochastic time series. Furthermore, due to the structure of the machine and tool, the AE signal generated at the tool tip is reflected, refracted and scattered, prior to reaching the sensor [81]; meaning that understanding the exact nature of the AE source is extremely challenging.

### 5.1.1 Causes of acoustic emission in machining

Many researchers have investigated the causes of AE generation in turning. Uehara et al. [44] hypothesised that the sources of AE originated from the shear plane, as well as the tool-chip and tool-workpiece interfaces. These sources produce continuous AE energy. The tool motion increases stress on the workpiece, and as a consequence, (up

to the yield point) the AE count rate increase due to the rise in average dislocations velocity. At yield, a rapid surge in AE occurs as the dormant dislocations start to move [82]. Plastic deformation of the workpiece causes dislocation of the substructure, releasing elastic strain energy as a vibrational wave, increasing thermal energy and culminating in AE generation [83]. In contrast, tool breakage and chip collisions are known to produce transient bursts of energy [47].

Summarizing the above, acoustic emissions from metal cutting can be generated from five sources. These are, deformation in the shear zone, chip interaction on the rake face of the tool, breakage of chips and tools, tool and workpiece entanglements with chips and rubbing of the tool and workpiece [84].

AE generation in machining of metals is mainly controlled by the strain rate, applied stress and volume of material loss. Therefore, the rake and clearance angle as well as the cutting conditions have an effect on AE [84].

### 5.1.2 Acoustic emission from chip formation

In metal cutting, material is removed from a workpiece in the form of chips or swarf, to achieve a final net shape. Chips are formed as a result of plastic deformation when tools interact with the workpieces. At the specified cutting conditions of the turning operation for this EngD, sawtooth chips are produced.

Few authors have attempted to relate AE to chip formation. In [44], by analysing the sawtooth chips and the cutting force signals alongside the AE, it was suggested that the variation in speed of the chips sliding over the rake face of the tool caused the force signal to become periodic and the amplitude of the AE signal to vary similarly. The authors suggest that this variation in speed is due to the change in shear angle during machining.

With the help of quick-stop optical micrography however, it was found that adiabatic shear bands formed during plastic deformation of the workpiece creates sawtooth chips [85]. Adiabatic shear bands grow due to the periodic imbalance of shear stress in the upper and lower regions of the primary shear zone, releasing elastic strain

energy [86]. Thermal softening from the high temperatures produced by plastic deformation conducts away from the primary shear zone and reduces workpiece material yield strength in the contact zone. Then the tool periodically begins to indent the colder material ahead of this zone. The shear zone therefore moves away from the tool until this indentation process repeats itself and produces sufficient global shear to initiate a new shear zone [87]. Though [86] related  $AE_{RMS}$  to chip formation and found that the transition from continuous to saw tooth chips could be captured by the increase in  $AE_{RMS}$ , no attempt was made to link the frequency of chip formation to frequencies observed in the AE signal.

When studying AE and chip formation, many other authors also opted for studying the energy of the AE signal ( $AE_{RMS}$ ). In early work, Dornfeld et al. [84] found that the  $AE_{RMS}$  is proportional to the AE signal energy. However,  $AE_{RMS}$  was not found to be heavily influenced by the breakage of chips or the tool workpiece interactions because the main source of energy originated from deformation and tool-chip interactions.

It should be noted that much of the literature is focused on the orthogonal cutting process because it is a simpler mechanism for analysis. However, the machining process of this work is an oblique cutting process which will inevitably affect the chip formation and the generated AE; unlike in orthogonal cutting, a uniform stress distribution cannot be assumed across sections of the chip for oblique cutting as the tool approaches the workpiece at an angle [88]. Thus, the velocity of flow varies across the chip. Consequently, chips tend to be curved with varying thickness and deformation rates along its width.

In an attempt to link sawtooth chip formation and AE generation, Neslušan et al. [89] attempted to calculate segmentation frequency from measuring chips collected during test with varying cutting speeds. The author chose the highest peak of the FFT as the segmentation frequency, though the results may be distorted as the signal processing steps undertaken are not clearly explained. Therefore, in this thesis, one aim is to study the frequencies associated with periodic saw tooth chip formation when tool wear progress from start to end of life, because cutting conditions such as the rake angle and depth of cut can alter in the presence of tool wear. Furthermore,

by identifying the chip formation frequency and by process of elimination, it may be possible to discover the cause of other generated AE frequencies and their evolution with tool wear. More on tool wear and AE generation is discussed in the following section.

### 5.1.3 Acoustic emission from tool wear

The tool-chip and tool workpiece interactions lead to tool wear, generating AE. Researchers have made many attempts to identify the specific aspects of the AE signal that are affected by tool wear. For example, with increasing flank wear, an increase in energy of the higher frequencies (500-750kHz) have been observed [6]. As the effective rake angle increases due to crater wear, the  $AE_{RMS}$  has been seen to decrease [46]. As the tool fails by cracking, chipping or fracture, AE with large amplitudes are generated [48]. Extracting features that show a correlation with tool wear can then be used as inputs into learning algorithms to predict tool wear.

Kannatey-Asibu et al. [45] compared the second, third and fourth order statistical moments of the  $AE_{RMS}$  signal against tool wear and found skewness and kurtosis to be sensitive to wear progression. With the aid of reverse cutting tests, Blum et al. [90] were able to identify the effect of flank wear on the mode of the AE signal. The test bypassed the chip formation process by reversing the direction of the workpiece motion, thereby isolating only the tool workpiece interactions.

When analysing AE signals during chip formation and tool breakage, Kannatey-Asibu et al. [91] found that the changes in the frequency domain were informative features for tool condition monitoring. Taking this idea one step further, Diniz et al. [46] filtered the AE data collected from a finish turning process to identify the frequency bands associated with tool wear and surface roughness.  $AE_{RMS}$  of these frequency bands were studied and it was found that the standard deviation of the  $AE_{RMS}$  has a high dependency on surface roughness and tool wear.

Kamarthi et al. [50] stated that due to assumption of periodicity and time insensitivity in FFT analysis, it may be useful to apply a discrete wavelet transform on the sometimes non-stationary AE signal instead. Xiaoli et al. [92] performed a wavelet

transform on AE data from a machining centre in order to decompose the signal into different frequency bands. The RMS of the wavelet coefficients were then used as features into a fuzzy neural network as the features were found to be insensitive to change in cutting conditions.

A short time Fourier transform (STFT) is a way of studying the AE signal in the frequency domain without compromising time resolution. Marinescu et al. [93] used spectrograms to study one tool rotation of a milling operation and to identify the precise moment the tool contacts the workpiece as well as making observations about tool wear. The data from middle of each tool pass were used for analysing the progression of AE power. Though this is computationally less expensive than looking at data from the whole process, transient events could be missed by selecting sections of passes. Also, as a tool pass in milling lasts for a fraction of the time when compared with turning, a significant amount of information would be discarded if only a middle section of a turning pass is investigated. Olufayo et al. [94] also used spectrograms to identify interesting frequency ranges from AE signals generated during a milling operation. Frequency bands 125-250 kHz displayed high energy, which the authors suggest is due to rubbing of the tool. Wavelets were used to decompose the signal into five frequency ranges and the detail coefficients were used as inputs in a neural network. Results presented are incomplete as only error values of training the network were presented. Mian et al. [95] opted for the use of spectrograms to study AE generated during sections of a micro-milling operation. According to the authors, the frequency bands associated with inter-granular micro-fracture, inclusions bursting and cleavage micro fracture were unexpectedly discovered owing to spectrogram analysis. Though it is unclear how these discoveries were verified, it is possible that Mian et al. visually inspected the spectrograms and assumed that the microstructure of the workpiece is responsible for the features observed in the AE frequency.

It is clear that many features of the AE signal are indicative of tool wear. These features can be extracted from the time and frequency domain by finding the energy of the signal, statistical moments and by conducting Fourier or wavelet transforms. In this chapter, the AE signal will be studied in detail taking all of these into consideration. The use of spectrograms will be extended to visualise the progression

of energy of frequency bins across the entire life of the tool. By doing so, it may be possible to uncover features that are more sensitive to tool wear than before.

## 5.2 The AE dataset

The size and complexity of the acoustic emission dataset collected during this work appears to be unmatched by previous literature on the topic, possibly owing to the typical sampling rates associated with AE, and the duration of end of life turning tests; as a result, data handling, storing and analysis can be extremely challenging. Acoustic emission data was collected throughout the entire life of each tool introduced in Chapter 3, encompassing the complete tool wear progression. The AE released by each of the 17 tools were captured by two sensors, giving an insight to the effect of sensor location. By testing two differing tool grades, it is possible to understand the effect of material composition on tool wear, and generated AE.

To aid with speed and ease of storage, the original time series was separated into smaller segments of  $1 \times 10^6$  data points representing one second of data.

### 5.2.1 Time domain

In turning, the tool is in constant contact with the workpiece, hence the generated AE signal is a time series. Though there may be separable bursts within the time series, it is challenging to isolate and analyse these events due to the large numbers that are present. Even if separation was possible, it is infeasible to judge the source of each burst due to the many likely causes. Therefore, AE signals must be averaged in some way in order to analyse or visualise trends in the data.

Figure 5.1 and figure 5.2 present the AE signals from Sensor 1 and 2 respectively for tools B6<sup>1</sup> and B7 across the entire life of the tool. A general trend of decreasing amplitude can be seen in figure 5.1 as time and tool wear increase. Moreover, a sudden increase in amplitude can be seen after data points number  $1.5 \times 10^6$ . This

---

<sup>1</sup>All tools discussed from this point on in the thesis were introduced in Chapter 3, Section 3.3.1.

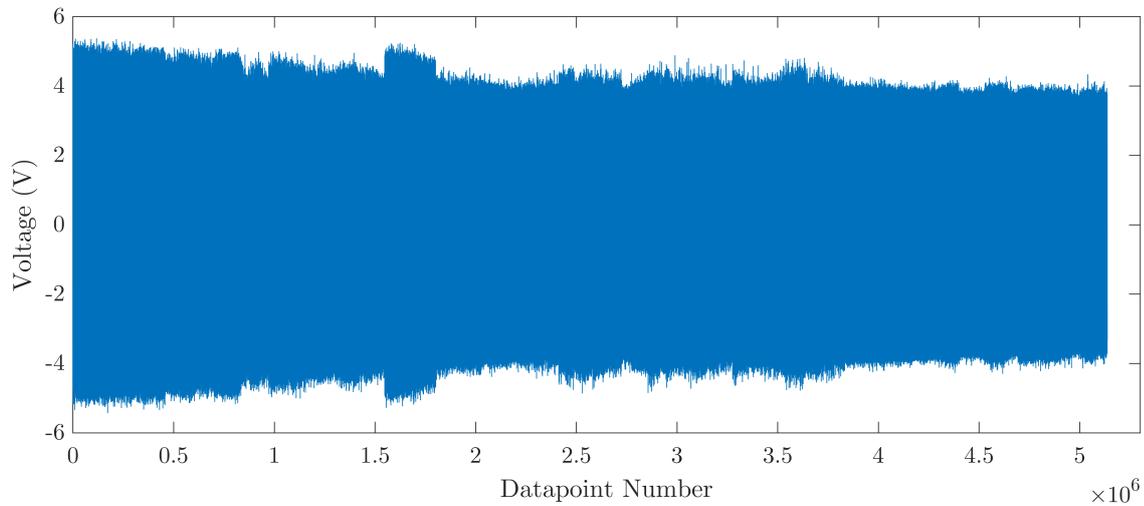


Figure 5.1: A down-sampled time series from Sensor 1 of tool B6 to demonstrate the behaviour of AE in the time domain.

could be a result of inconsistency in depth of cut (DOC) post a tool inspection as detailed in section 3.1.2. Numerous transient events can be seen in figure 5.2. The assumption here is that these bursts are mainly caused by the flying chips colliding with the machine interior.

From figures 5.1 and 5.2, it is also evident that the location of sensors has an effect

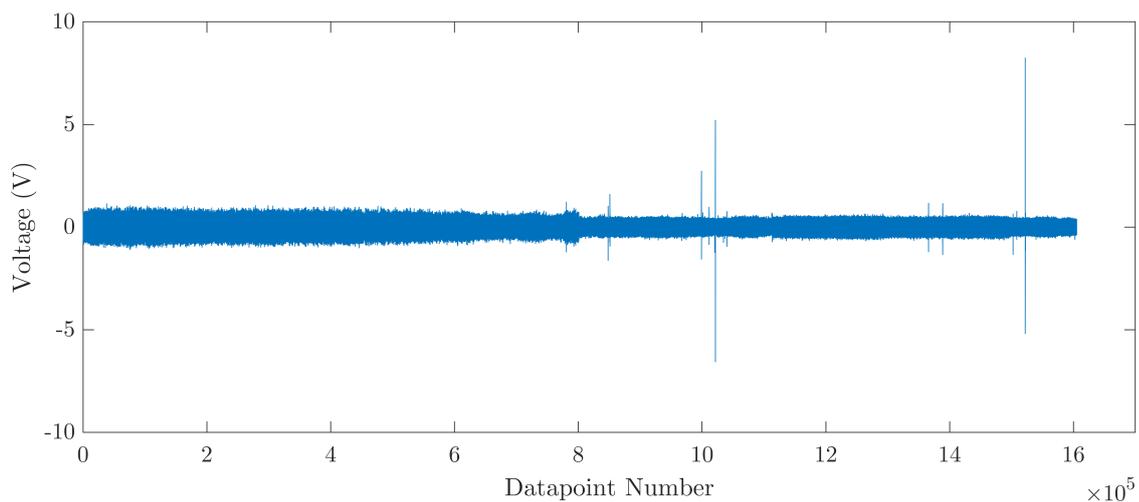


Figure 5.2: Time series of tool B7 from Sensor 2 down-sampled to aid with visualisation.

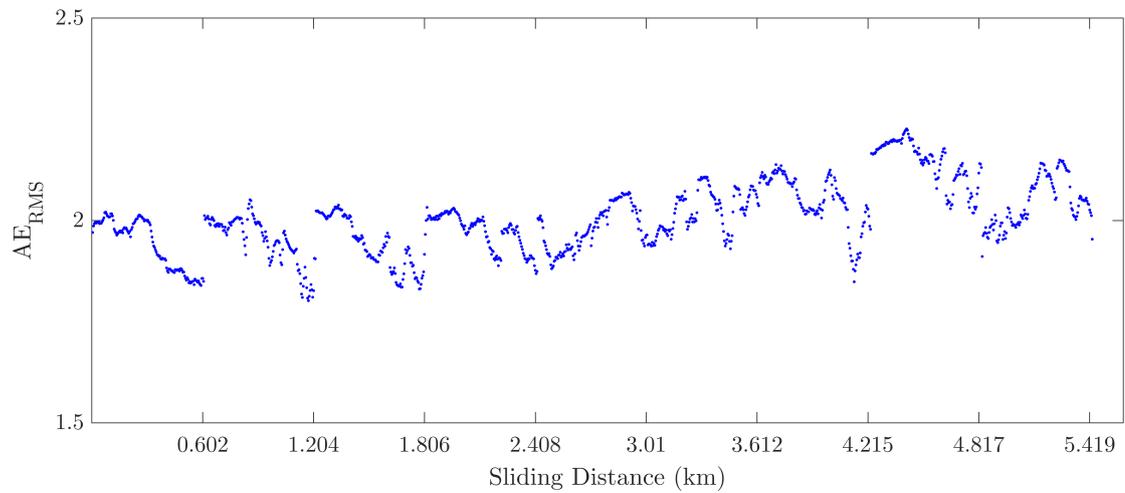


Figure 5.3:  $AE_{RMS}$  of Sensor 1 signal from tool B6. The first pass of each four pass intervals has been removed. Each data point represents a second of the AE signal ( $1 \times 10^6$  original data points).

on the amplitude of the collected signal; reflection, refraction and attenuation due to the AE passing through parts of the machine is the most likely cause.

The corresponding root mean square of the signal,  $AE_{RMS}$ , of the raw data for the entire life of tool B6 from Sensor 1 and tool B7 from Sensor 2 are shown in figure 5.3 and 5.4 respectively. Each data point in these plots represent one second of data

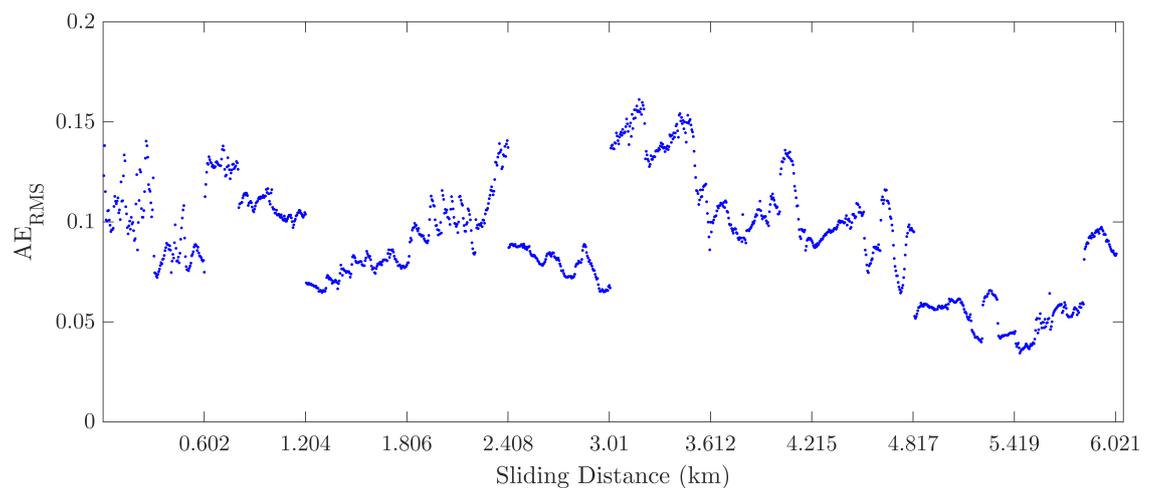


Figure 5.4: Tool B7  $AE_{RMS}$  from Sensor 2.

which is an average of  $1 \times 10^6$  raw data points.

When comparing figures 5.1 and 5.3 or figures 5.2 and 5.4, it is clear that the  $AE_{RMS}$  does not simply represent the amplitude of the signal.  $AE_{RMS}$  is a measure of energy of the events captured by the signal and so a seemingly decrease in amplitude of the overall signal could in fact correspond to an increase in  $AE_{RMS}$  if more events have occurred [96]. However, the raw time domain signal is much too large to observe these events. Consequently, in order to identify the nature of these events, further analysis is carried out in the frequency domain.

### 5.2.2 In the frequency domain

The frequency content of AE signals may be rich with information. Observing the signal in the frequency domain may also help to reduce the dataset size. The signal can be converted into the frequency domain using methods such as the Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT). A short time Fourier transform (STFT) is used in this work as it was found to have adequate time and frequency resolution for initial analysis. Due to the length of AE signals generated in turning, time resolution is desired to capture the many continuous and transient events present within the signal. A STFT enables the visualisation of AE data from the entire tool in one plot, giving an idea of the overall behaviour. This may aid the extraction of features that are changing as tool wear progress.

As mentioned previously, for ease of analysis, the original AE signal was split into segments comprising  $1 \times 10^6$  data points (corresponding to one second of data). In order to transform these segments into the frequency domain using STFT, windowing was used here. Particularly, 50% overlapping Hamming windows with a length of 256 data points were used in this work. For each second of data, this resulted in 7800 windows. The average of these windows were calculated for each segment (one second of data) and stored as a spectrogram. As a result, the spectrograms consists of around  $129 \times 1500$  data points on average, where there are 129 frequency bins and around 1500 seconds of AE data across the life of the tool.

It is assumed when performing the Fourier transform that the signal is periodic, which is not the case here as the continuous AE signal is a collection of bursts.

Therefore, the window type and length as well as the sampling rate can have a significant impact on the observed frequency content. As a result, spectrograms should be studied with caution. To investigate the effect of the window type on the results, the Hanning window and the flat top weighted window were also tested but increased leakage (spread of energy) was observed.

### 5.3 Spectrograms

The results of STFT can be visualised in the form of spectrograms that show the energy of the frequency content with the colour intensity. An example of the frequency domain transformation of AE data measured from Sensor 1 and 2 for tool B6 are displayed in figures 5.5 and 5.6 respectively. Due to the depth of cut inconsistencies, where relevant, the first of the four passes of the workpiece are eliminated from the plots in the following chapters of this thesis. The data from the tool during full contact of the workpiece has been presented. In all cases, the results have been presented as a function of sliding distance, the distance travelled by the tool.

As the sliding distance increases, some frequency bands exhibit change in intensity. These changes could be due to multitude of reasons that will be explored further in

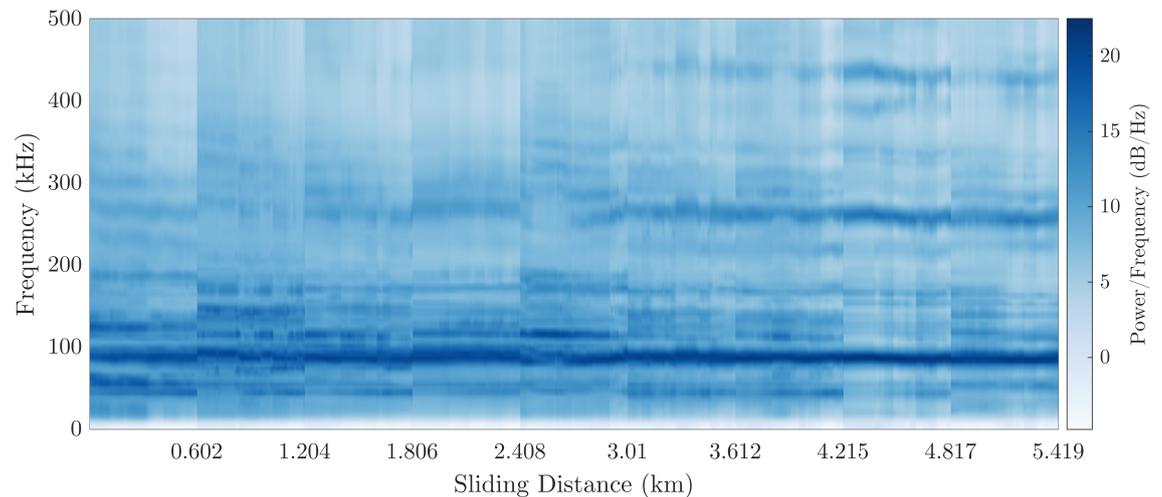


Figure 5.5: The results of STFT of tool B6 AE data measured using Sensor 1. The power of the signal is presented as the average power per unit bandwidth.

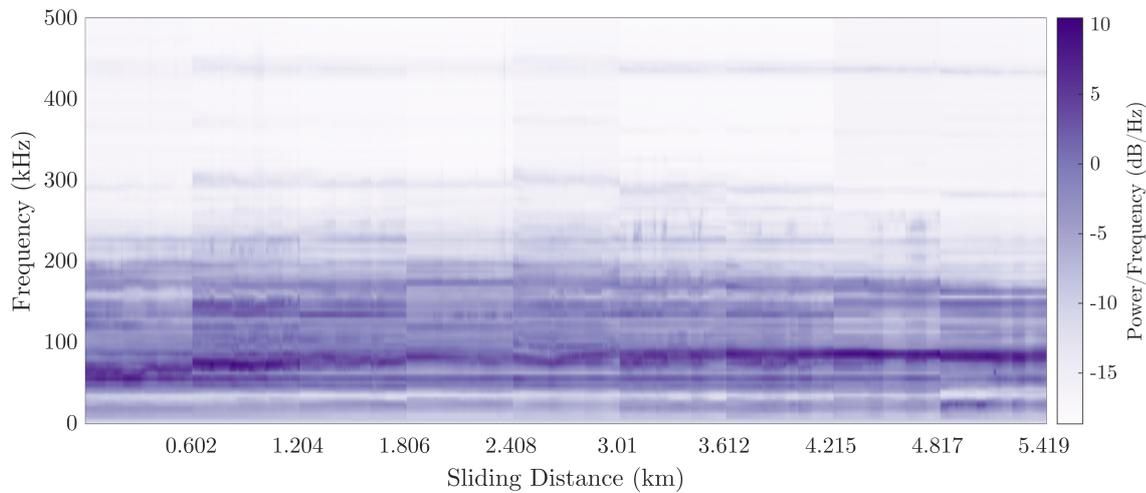


Figure 5.6: Tool B6 spectrogram of data from Sensor 2.

this section. By observing the trends in frequency in this format, it is possible to identify abnormal or non-stationary behaviours.

Interestingly, the spectrograms of data from the two sensors are very similar when considering the prominent frequency bins, but are marginally different in terms of the overall energy observed; a lower amount of energy is measured from Sensor 2. The structure of the sensor is assumed to have an influence on the results. It is also likely that the position of the sensor affects the collected data, as Sensor 2 is located much further away compared to Sensor 1.

For both sensors, most of the signal's energy is spread across the frequency range of around 40-250kHz. Beyond which, some energy is observed in the high frequency part of the spectrograms. The behaviour of the frequency band with the highest energy slightly differs for the two sensors at the start of tool life (up to 1.2km) where Sensor 1 has multiple distinctive frequency bands and Sensor 2 has just one high energy frequency band. It is possible that these small variations are due to the low signal to noise ratio at Sensor 2. The cause of these behaviours are explored further in Sections 5.3.1 to 5.3.4 to find features that correlate well with the tool wear features obtained in Chapter 4.

### 5.3.1 Chip formation and AE

The area with highest energy of the spectrogram is highlighted in figure 5.7, where data from Sensor 1, tool B4 is presented as an example. It can be seen that in early tool life, the energy is highly dispersed. As tool wear progress and sliding distance increase, the energy is concentrated into a sharp peak or a lower spread of energy. This behaviour is seen across all tools and may be explained by analysing the chip formation during cutting.

Unlike orthogonal turning, in oblique cutting, the tool approaches the workpiece at an angle. As a result, the chips formed in oblique cutting move away from the tool faster at the leading edge<sup>2</sup> than the trailing edge. The varying velocity across the chip affect the mechanisms involved in the adiabatic shear band growth, changing the chip shape. Subsequently, the hypothesis made here is that the energy of acoustic emissions generated during the formation of chips in oblique cutting are spread across a wide range of frequencies.

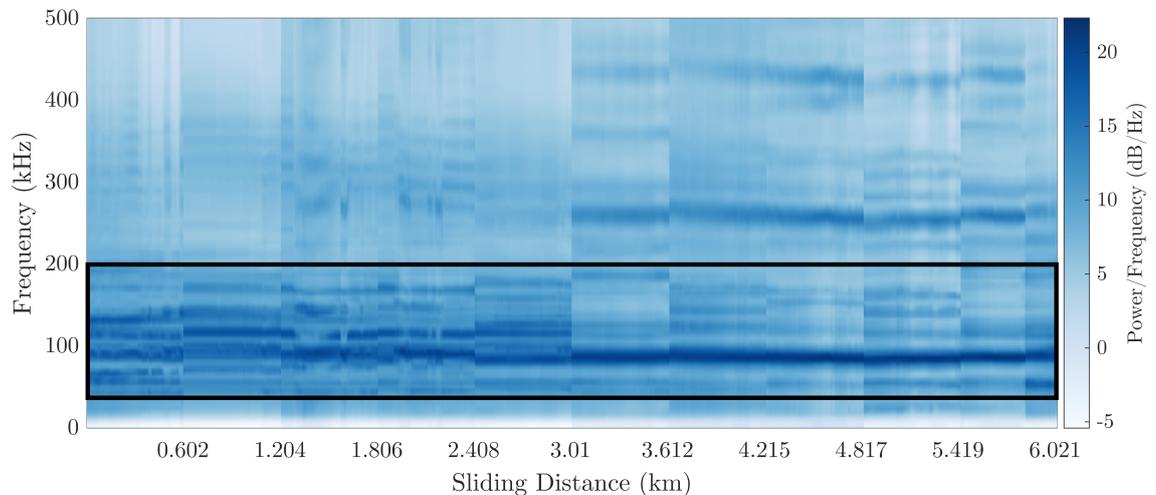


Figure 5.7: Spectrogram of AE data collected from Sensor 1, tool B4. Frequencies with the largest concentration of energy has been highlighted.

Figure 5.8 presents an image of a typical chip. Usually, the leading edge of the chip has the highest depth of cut due to the angle at which the tool approaches

<sup>2</sup>Figure 4.5 illustrates the leading and trailing edge of a wear scar.

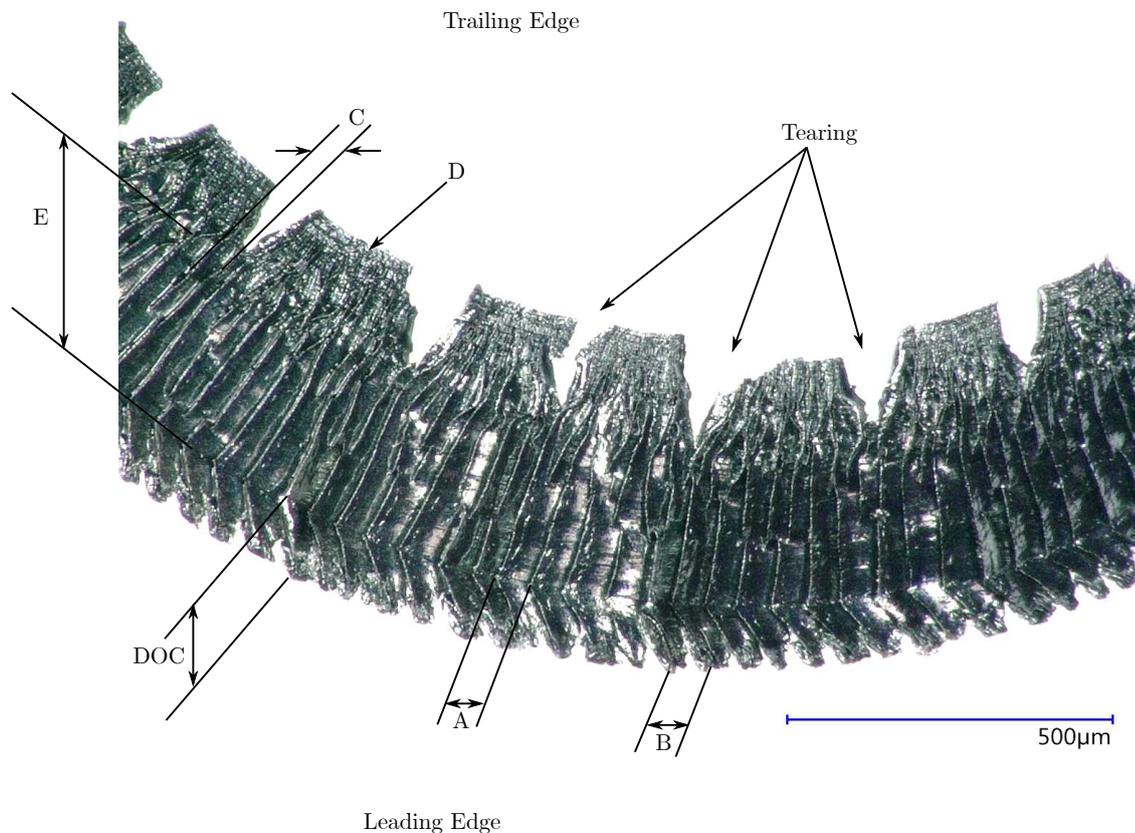


Figure 5.8: A typical chip with the leading and training edge annotated. A = distance between two saw tooth, B = distance between tearing of the shear bands, C = distance between a pair of smaller peaks, D = area where the chip is slowest and has very thin shear bands, E = main section of the saw tooth formation.

the workpiece. In the trailing edge, the chip tears in order maintain the velocity of the flow at the leading edge. Due to the changes in velocity, in some instances the adiabatic shear band growth can be non-uniform across the tool, leading to varying peak distances, A, B, C and D.

Scanning electron microscope (SEM) imaging can provide a closer look at chip formation. The saw tooth effect and the inconsistent shear bands across the chip is clear in figures 5.9a to 5.9d.

The chips formed during early stages of tool life tend to have non-uniform adiabatic shear bands. Some shear bands can appear across the whole length of the tool

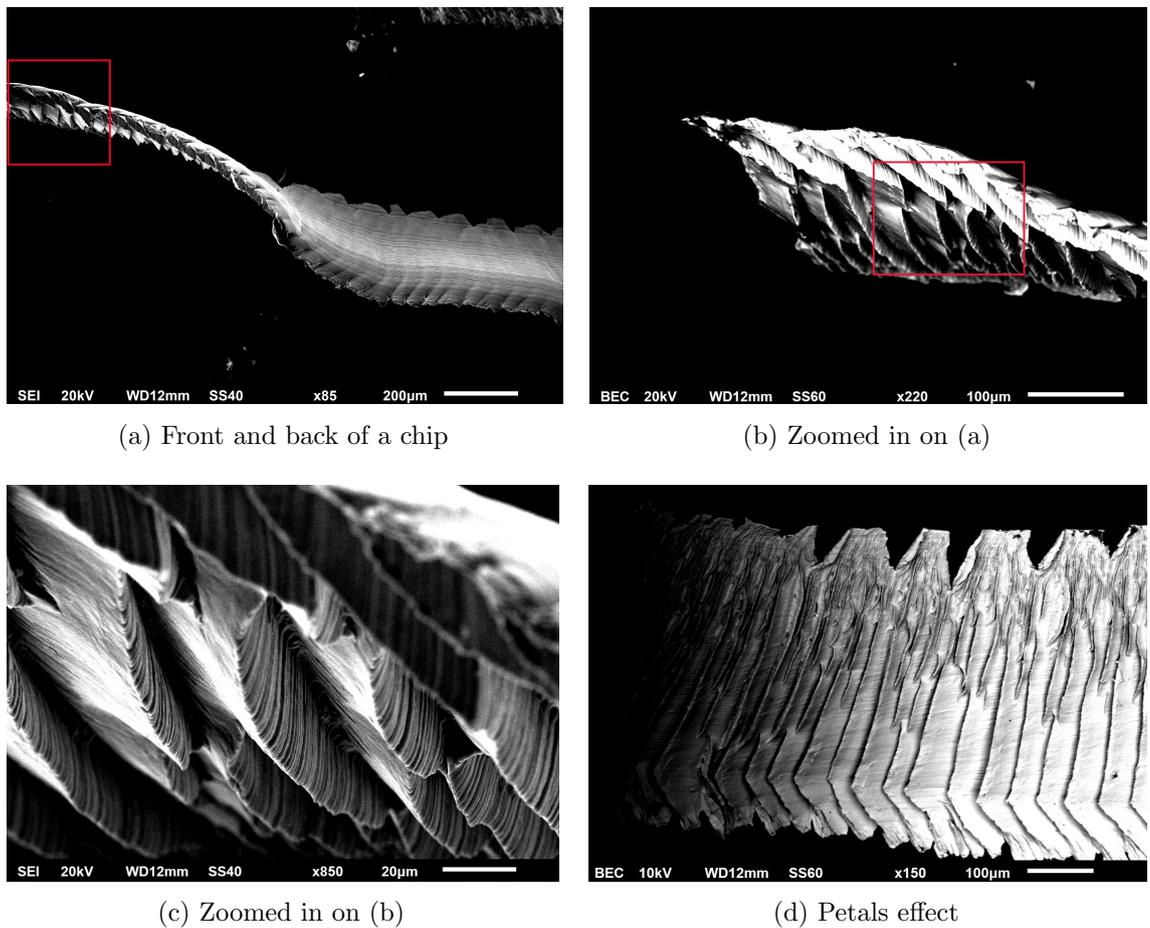


Figure 5.9: SEM images taken from a chip after 8 passes, 1.204 km.

whereas other, shorter bands can appear between the long ones (figure 5.9d). A chip collected after 4 passes of the workpiece, or a sliding distance of around 600m displays this behaviour in figure 5.10a. In this figure, uniform tearing can be seen, the frequency of which will be similar to chip formation frequency and may be observed in the spectrogram. The back of the chip can also be informative (figure 5.10b). It is observed that during early stages of cutting, the back of the chip has a profile matching the front, named folds [85]. The frequency at which the folds are being created will also be visible in the spectrogram.

As the sliding distance increases, the cross section of the chips become more uniform and the tearing occurs fewer times with higher severity (figure 5.10c). The back of

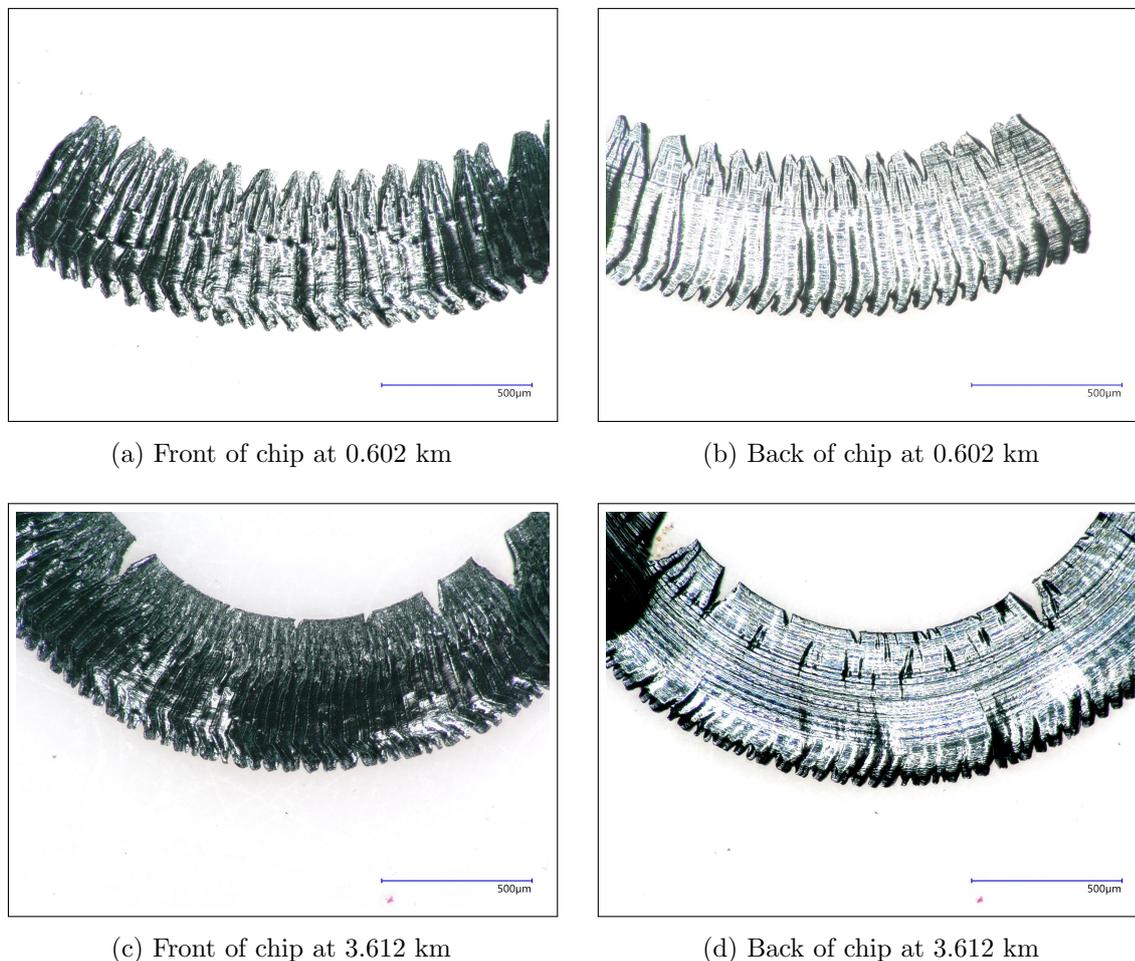


Figure 5.10: A sample of chips from tool B4 used for measurements.

the chips are also smoother, with grooves now travelling along the chip instead of across the chip (figure 5.10d). The change is most likely due to the change in rake angle and profile of the tool due to wear. Consequently, this is the most likely cause of the energy of the chip formation frequencies focusing to a narrower band.

A sample of the images were studied in an attempt to verify that the high energy bands seen in the spectrograms between 20-400kHz are indeed from chip formation. The distances presented in figure 5.8 were measured using image processing software and the lengths were converted to frequencies using the cutting conditions values. Table 5.1 displays the assumed chip formation frequencies calculated from SEM

Distances	Sliding Distance (km)									
	0.60	1.20	1.81	2.41	3.01	3.61	4.22	4.82	5.42	6.02
	Frequencies (kHz)									
Distance A	48.3	68.4	57.4	72.0	56.5	58.1	62.1	72.5	74.2	75.9
Distance B	43.1	65.6	58.3	89.8	148.2	73.8	78.4	86.3	115.9	86.7
Distance C	84.1	111.9	109.9	135.3	115.7	118.1	115.1	150.1	110.6	98.8
Distance D	411.2	349.4	331.6	448.9	395.1	397.4	220.4	597.1	259.7	216.5

Table 5.1: The calculated chip formation frequencies from tool B4 throughout tool life.

images for tool B4. The results are taken from the three sections of the chip, left, middle and right and an average is shown here. The calculated frequencies in Table 5.1 align with the high energy frequency bands between 40-200kHz seen in the spectrograms. It should be noted, however, that the sample of the chips analysed is extremely small when compared to the volume of chips that are produced by the tool. Therefore, this analysis is only able to provide a glimpse of the process and does not represent all the behaviours that occur during cutting. Furthermore, the chips are curly and therefore can be difficult to measure under the microscope. Often it is not possible to lay the chip flat without causing damage to the fragile structure. Therefore, it can be challenging to measure the very small distances accurately. For this reason, large variances within measurements can occur. Having said that, the important point to note here is that most of the measurements fall within the chip formation frequency range of 40-200kHz.

Interestingly, this work may be the first significant and comprehensive attempt at linking mechanics of chip formation in oblique cutting to the generated acoustic emissions signals in literature. The relationship between AE and mechanics of the process can lead to selection of more informative features for machine learning later. Consequently, the results of any condition monitoring technique can be better related to the actual mechanics of the process. For this reason, validating these findings is key and methods of validation are discussed in Chapter 9.

From this point in the thesis, this frequency range will now be referred to as the chip formation frequency.

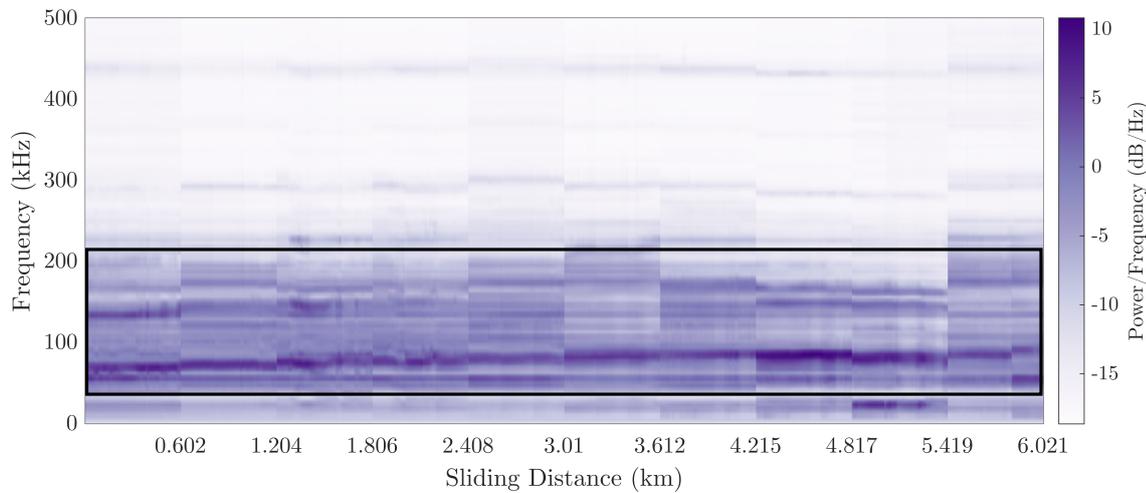


Figure 5.11: Spectrogram of AE frequency content measured using Sensor 2, tool B4.

When studying Sensor 2 results, a similar behaviour is observed. The highest energy band in the spectrogram shifts as the sliding distance increases. More specifically, the frequency of the high energy content is lower than results from Sensor 1 at the start of tool life; around 65kHz as opposed to 85kHz. As the sliding distance increases, so does the frequency at which the high energy behaviour occurs. At latter sliding distances, the high energy frequency band matches the results from Sensor 1. An example of this behaviour is highlighted in figure 5.11.

As Sensor 2 is located much further away compared to Sensor 1, it is possible that only the highest energy content is visible in the spectrograms, therefore the spread of energy of chip formation is much lower for Sensor 2 than Sensor 1 at the start of tool life.

### 5.3.2 Chip formation harmonics

The annotated boxes in figure 5.12 highlight the higher frequency bands with high energy. The energy of these high frequency bands markedly increase as the sliding distance increases. The reader should note that the spectrogram has been presented in the log scale to highlight the behaviour of these bands in this chapter. Prior to scaling, the energy of these bands are significantly low compared to the energy of

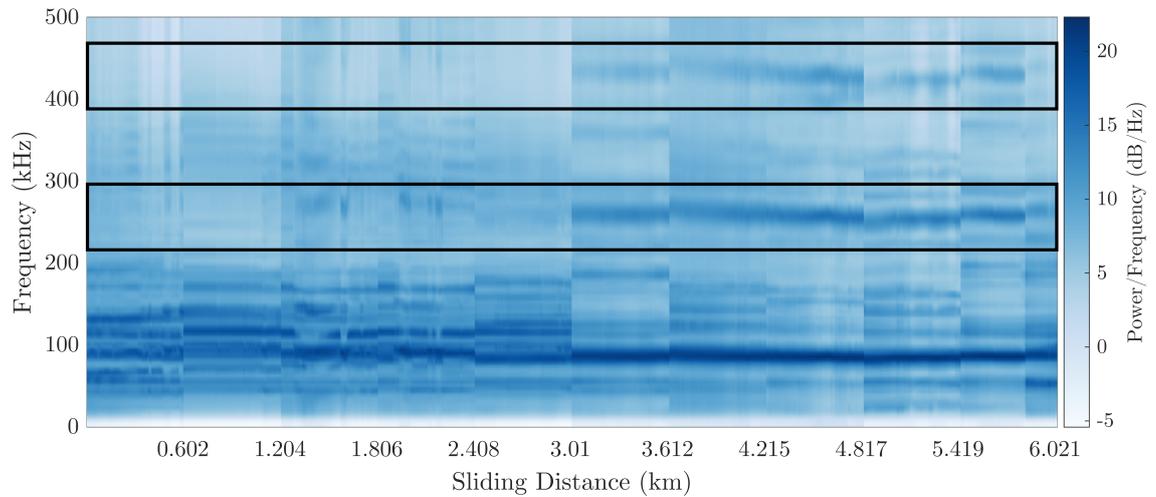


Figure 5.12: The frequency content of tool B4 obtained from Sensor 1.

the chip formation frequencies.

It is possible that the high frequency content seen in these figures are harmonics of chip formation frequency. The cause of these harmonics may be due to various reasons. Harmonics of the fundamental frequency can be expected in the case of a nonlinear systems due to discontinuous dynamics within the system.

Interestingly, increasing sliding distance, and therefore tool wear, causes the harmonics to intensify. This is possibly due to the chip formation mechanisms discussed in section 5.3.1. As the energy of chip formation intensify due to tool wear, and the dispersion of energy decreases across frequency bands, the intensity of the harmonics could increase. On the other hand, this behaviour could be due to the bands at the trailing edge of the chip occurring at a higher frequency near the end of tool life (figure 5.10c) compared to start of tool life (figure 5.10a).

### 5.3.3 Confounding influences and AE

In some instances, the energy content throughout the entire frequency range is noticeably different between two sliding distances, as highlighted in figure 5.13, box A. At these intervals, the general trends can alter and high energy bands can shift across frequency bins.

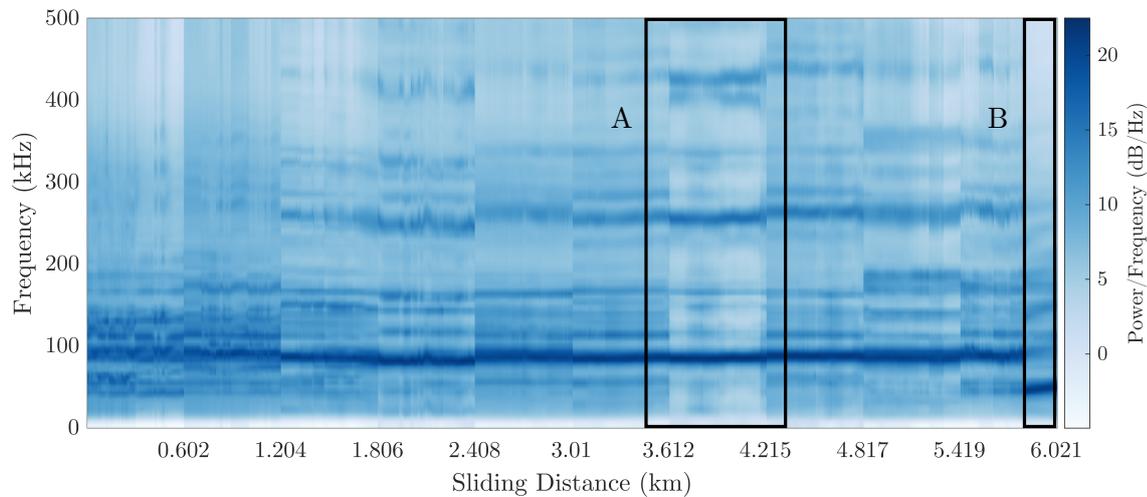


Figure 5.13: Sensor 1 STFT results for tool B7. Box A highlights the confounding influences and box B highlights data after tool breakage and is discussed in Section 5.3.4.

It is likely that a global change in energy within four passes of the workpiece is due to the hardness of the workpiece; recall that the workpieces are replaced every four passes in this process<sup>3</sup>. Figure 5.14 displays the measured hardness of the workpiece before and after each four passes for tool B7. It is clear that the AE from box A section of figure 5.13 was collected when machining a workpiece with a high hardness at the beginning of the pass. Work hardening has also taken place throughout the four passes as the workpiece is harder at the end than at the start.

Increase in hardness increases the energy required to remove material from the workpiece [85]. This in turn affects the generated AE. In this case, it is assumed that energy concentrated on the frequency bands associated with chip formation are captured by the sensor more readily, masking the power across the other bins.

At these specified intervals, the tool is removed for analysis. Upon retuning to the holder, it is possible that the tool is secured in a different position. Although it is assumed that after one pass of the workpiece the tool moves back to its natural position due to the large forces acting on it during machining, it is possible that the tool remains at the new position for the entirety of the four passes of the workpiece.

<sup>3</sup>In section 3.1.2, the reason for changing the workpiece is explained.

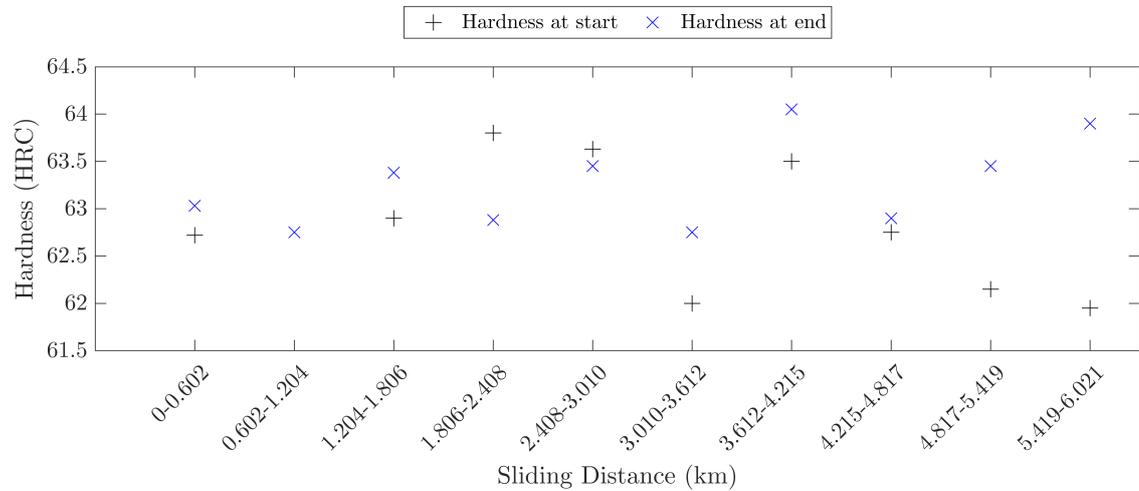


Figure 5.14: Hardness values at the start and end of each 4 pass intervals of the workpiece when machining with tool B7.

In this case, the depth of cut at which the machining has taken place would be different to the previous passes of the workpiece which could also be the cause of the abnormalities seen in the spectrogram; an increased depth of cut can increase  $AE_{RMS}$  [97] leading to higher energies associated with chip formation.

As the workpiece changes and tool removals are unavoidable in this experimental procedure, the only option of dealing with these confounding influences is to use data pre-processing to reduce the effects of hardness and DOC on the AE data, or chose an algorithm that is capable of dealing with the affected data. A further discussion on this can be found in Section 5.4.

### 5.3.4 AE at tool breakage

Box B in figure 5.13 highlights an example of AE behaviour at tool failure. There is a major shift in the highest energy frequency band as well as suppression of energy in the harmonics.

The catastrophic failure on tool B7 can be seen in figure 5.15. The entire flank face around the wear scar has been removed during failure. Consequently, the cutting edge and the nose radius has been severely damaged. As a result, the depth of cut, effective

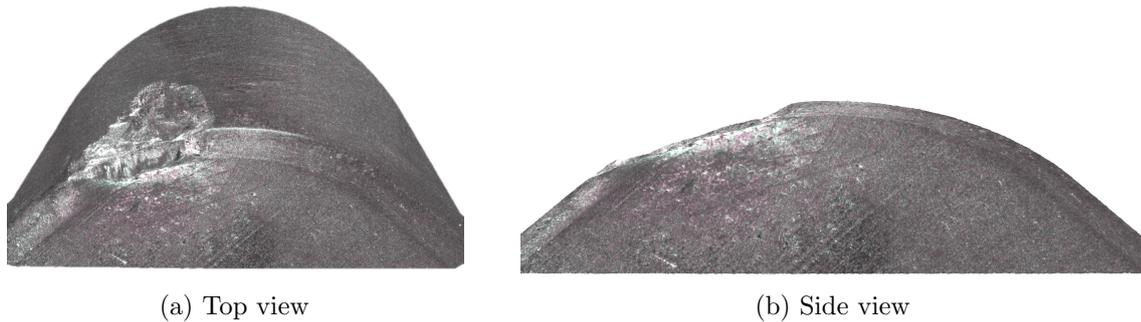


Figure 5.15: Broken tool B7

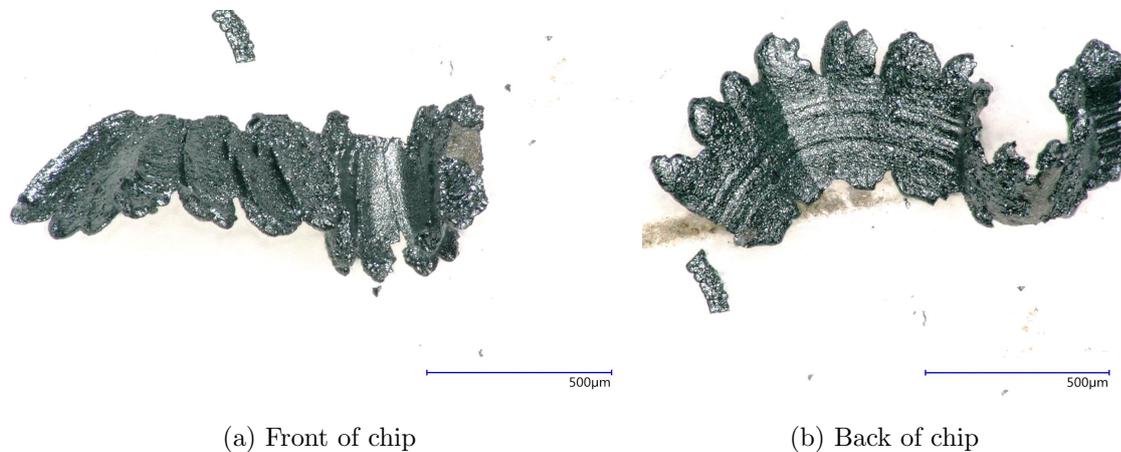


Figure 5.16: Chip produced at tool breakage.

rake angle, effective clearance angle and contact area will be altered, increasing the forces and temperatures as well as varying the chip formation characteristics.

In this case, the depth of cut seems to have reduced the thickness of the chips formed, and the increased temperatures have burned the chips (figures 5.16a and 5.16b). The distance between each sawtooth band has increased, and the formation of the sawtooth is constant across the chip from the leading to trailing edge. As a result, chips are formed at a lower frequency where most of the energy from the process is concentrated into forming the uniform chips. Furthermore, a tool break is similar to an impulse where the resulting energy content will be spread across a large range of frequencies, leading to the masking of harmonics [91].

## 5.4 Dealing with confounding influences

The impact of workpiece hardness on the collected AE data can be categorised as a confounding influence. In the field of structural health monitoring, seasonal trends can often influence the data collected from structures. For example, temperature variation during summer and winter affects the behaviour of structures that in turn affect the acoustic emissions. As a result, any damage detection and prediction algorithm must be capable of distinguishing the difference between these environmental trends and actual damage to the structure. Cross et al. [98] approached this challenge by introducing the idea of cointegration, first used in econometrics. When non-stationary features are cointegrated, a stationary residual that is void of the environmental trends may be created, that can be used as a damage sensitive feature. Amongst others, Principal Component Analysis (PCA) can also be used as a tool when dealing with confounding influences. PCA takes linear combinations of features in order to form new variables that represent the data in a more informative projection. When ordered according to variance, principal components with low variance can be used as features that are damage sensitive, yet unimpeded by prominent environmental trends [99]. Though these techniques are successful in suppressing seasonal trends, they require a set of representative training data *a-priori*. For this reason, it is beneficial to obtain features that are unaffected by hardness of the workpiece in the feature selection stage where possible.

A simple test can be performed here where the variation in power values at workpiece changes are calculated in order to understand which frequencies are least affected by change in hardness. The results for tool grades A and B are presented in figures 5.17 and 5.18 respectively.

As chip formation and tool wear are affected by the hardness of the workpiece, the corresponding acoustic emissions generated by chip formation are significantly influenced. Beyond 200kHz, the impact is lessened. As a result, using frequencies above 200kHz can be a way of avoiding the confounding influences on the dataset.

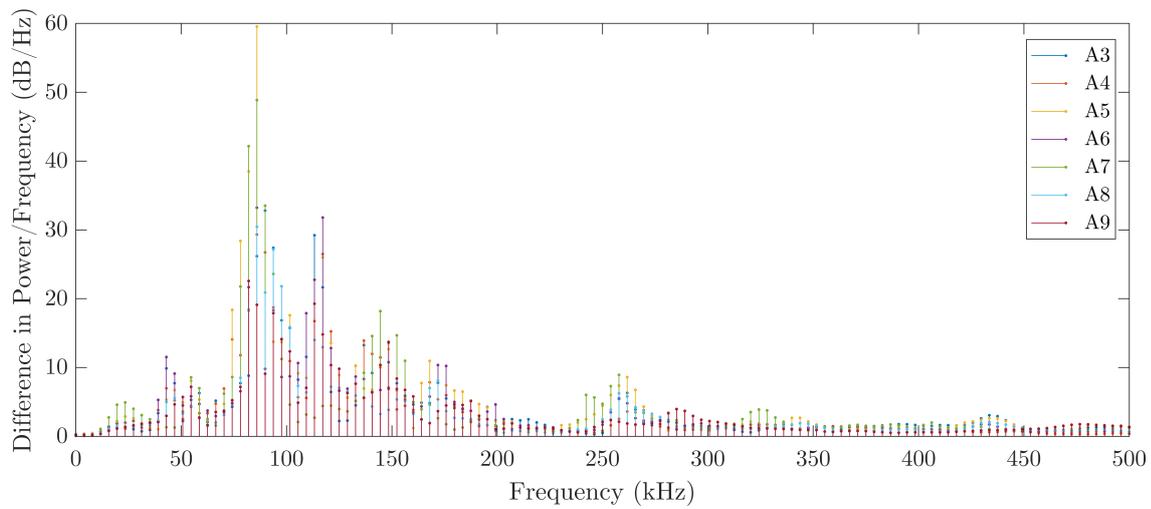


Figure 5.17: The average difference in power of the frequency bins at workpiece changes. Results for Sensor 1, tool grade A.

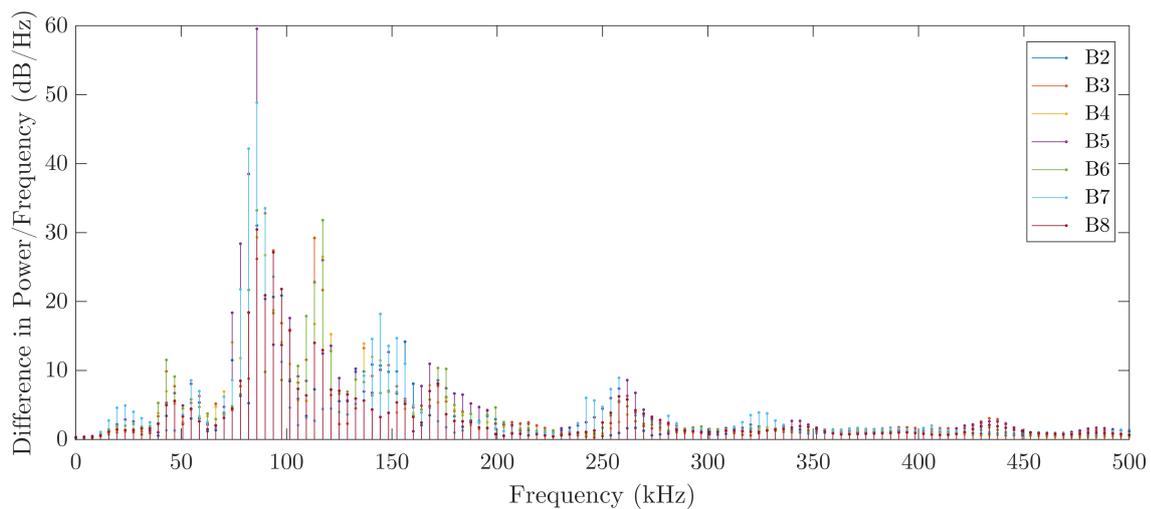


Figure 5.18: Average difference in power of the AE signal when workpieces are replaced. Sensor 1 results of tool grade B.

## 5.5 Candidate features

A better understanding of the physical mechanisms that cause the acoustic emissions in turning allows for meaningful features from the signal to be extracted. AE features from the time and frequency domain that may perform well in prediction algorithms

due to their correlation to tool wear are presented in this section.

Two types of features are selected as candidates in this section. Features that increase progressively with tool wear may be useful for predicting the state of the tool at any given time. However, features that are dormant until imminent tool failure may also be of interest.

### 5.5.1 Features from spectrograms

When selecting features, past research has opted for studying the FFT of AE signals from sections of passes, giving a significantly generalized indication of the frequency content. However, tracking informative frequency bands across the life of the tool has only been conducted in a handful of instances [100–102]. The frequency bins that are related to chip formation and its harmonics may be especially useful for tracking tool wear as plastic deformation of materials can be determined non-destructively by studying the harmonics the acoustics generated [103] and because acoustic harmonics are able to detect microcracks in materials [104]. Figure 5.19 presents frequency bins around the harmonics of chip formation, that may be useful for predicting imminent failure of the tool.

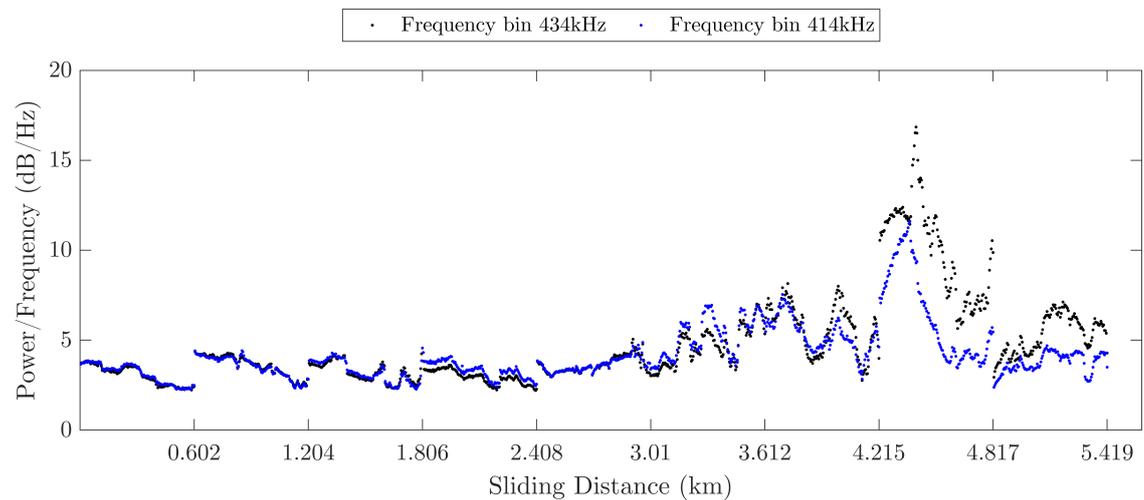


Figure 5.19: Frequency bins that increase in intensity near the end of tool life. Results from tool B6, Sensor 1.

### 5.5.2 Features from the time series

Due to the large size of the AE signal, the data must be heavily averaged in order to practicably extract time sensitive features. However, by studying the distribution characteristics of the time series in small segments, it is possible to capture information that may represent the mechanisms that are producing the workpiece surface, such as tool wear [45]. Out of the first four statistical moments, the standard deviation, skewness and kurtosis of the time series can be used to track the behaviour of acoustic emission data.

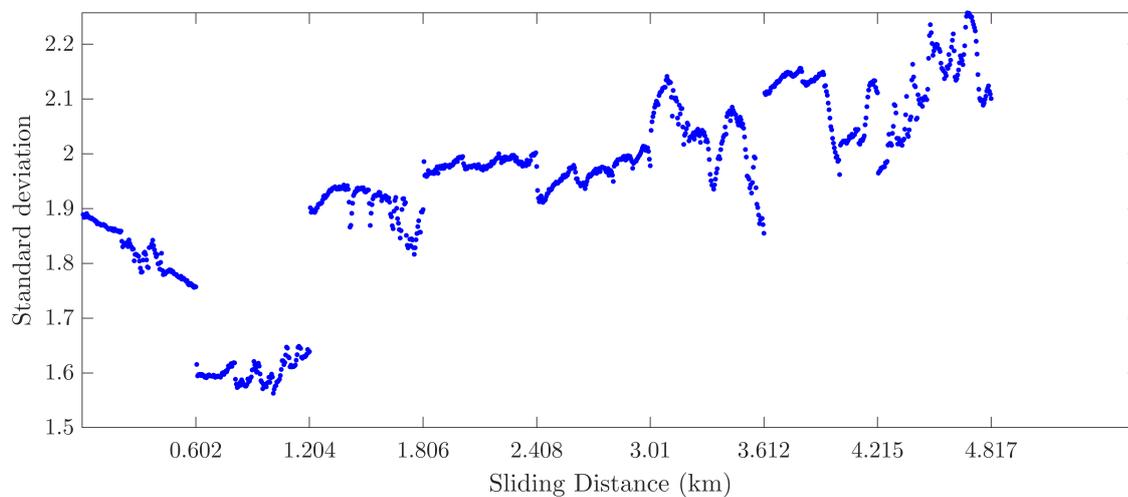


Figure 5.20: Standard deviation of the AE time series collected from tool B5, Sensor 1. Each data point represents one second of AE data.

Figure 5.20 displays the standard deviation taken per second of data throughout the life of tool B5 from Sensor 1. In general, it can be seen that standard deviation gradually increases with sliding distance from around 1.2km. Monotonic tool wear features may correlate well with AE features of this type.

### 5.5.3 Wavelet Transform of the AE signal

By applying a discrete wavelet transform (DWT)<sup>4</sup>, it may be possible to reconstruct the AE signal comprising only of frequencies from a turning mechanism of interest.

<sup>4</sup>A brief introduction to DWT can be found in Appendix C.

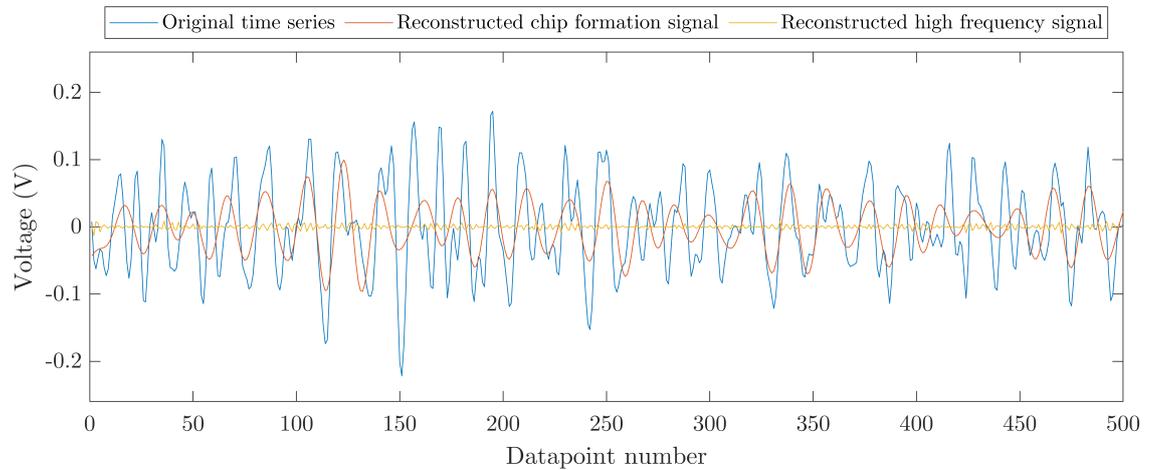


Figure 5.21: A visual example of the reconstructed time series where 0.0005s of data is displayed for clarity.

For example, a three level DWT can be applied to isolate the frequencies 62.5-125kHz, capturing the entire frequency range associated with chip formation. As tool wear is intrinsically linked to the mechanisms of chip formation, this range was isolated using a Daubechies 20 wavelet; orthogonal wavelets are better suited for DWT and this wavelet, from observations, is similar in behaviour to the AE signals generated during turning. The reconstructed signals corresponding to chip formation frequencies (62.5-125kHz) and the high frequencies associated with harmonics (250-500kHz) have been shown in a comparison plot with the original time series in figure 5.21.

The reconstructed time series can be treated similarly to the original signal and statistical features can be extracted as before. Figure 5.22 shows the standard deviation of AE data once DWT has been used to reconstruct the signal with frequencies associated with chip formation. As expected, the standard deviation is high at the start of tool life due to the large spread of energy as a result of chip formation mechanisms. This feature may be inversely correlated with increasing tool wear and therefore can be a good visual indicator of tool degradation.

Similarly, the high frequency content of the signal shows interesting trends, such as the harmonics of chip formation. Therefore, the  $AE_{RMS}$  and other statistical components of the reconstructed signal from the high frequency range can also be

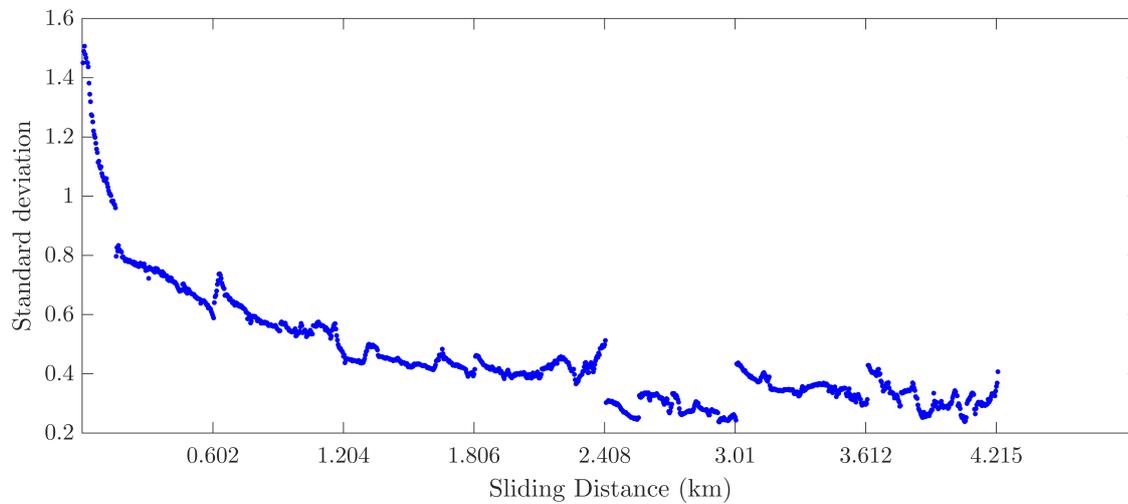


Figure 5.22: Standard deviation of the AE time series collected from tool A8, Sensor 1 after DWT of chip formation frequencies. Each data point represents one second of AE data.

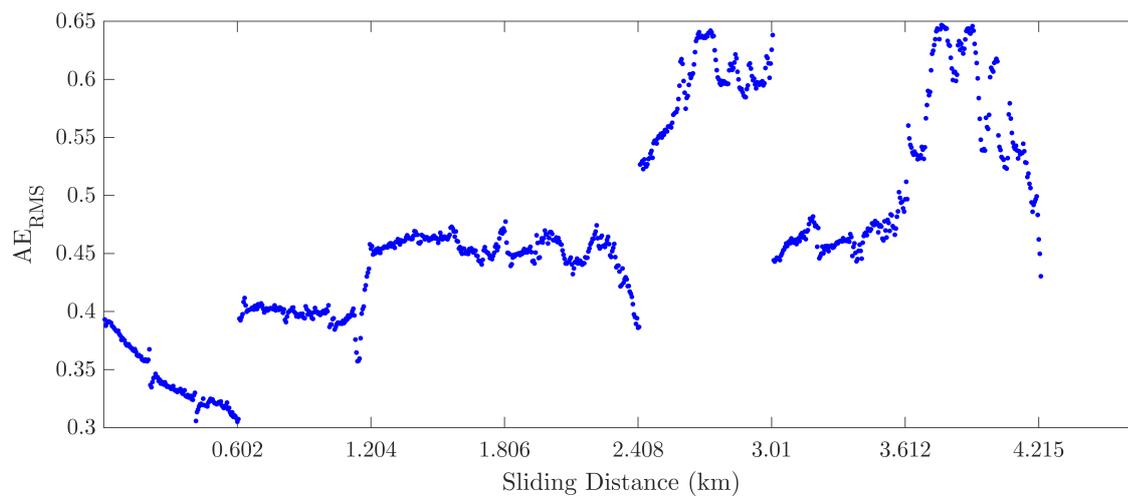


Figure 5.23:  $AE_{RMS}$  from tool A8, Sensor 1 after DWT of frequency range 250-500kHz. Each data point represents one second of AE data.

studied as a feature; see figure 5.23.

It is also possible to use the detailed coefficients of the DWT as input features into learning algorithms [105]. However, due to the size of the AE dataset in this work, the number of detailed coefficients is too large to use as inputs. Most, if not all,

algorithms would struggle to handle these datasets and computational time would be significant.

## 5.6 Conclusions

The process of hard turning generates complex AE signals due to the numerous physical mechanisms involved. For example, the workpiece, cutting conditions, tool wear, tool material, type of turning and the test set up are some of the many aspects of the process that have an influence on the acoustic emissions generated. Nevertheless, its high sensitivity to the chip formation process and tool wear establishes AE as an effective indirect measurement of the machining process.

This chapter has shown that significant amounts of information from the AE signal can be derived from the frequency domain. With the help of spectrograms and detailed chip analysis, an attempt was made to infer the physical mechanisms involved in generating AE from start to end of tool life for oblique cutting with PcBN. It was found that chip formation can be closely linked to frequencies between 40-200kHz likely due to the flow of the sawtooth chips. Given past research has mainly focused on orthogonal turning where the chip formation process is much less complex, this work may be one of the first to investigate the effect of oblique cutting on the formation of chips and make links between the chip formation process and the frequencies observed in the generated acoustic emissions signal. A better understood relationship between the signal inputs and the mechanics of the process can be valuable when deciphering the results of TCM algorithms.

It is important to validate any assumptions made in order to be certain that the frequencies associated with chip formation are indeed those. As previously mentioned, the current method of chip analysis has many drawbacks that make it challenging to measure chips. The author suggests using a high speed camera at frame rates of around 500k frames per second during machining in order to observe chip formation in real life. The reader is referred to Section 3.2.2 for an example setup. With advanced image processing, it may be possible to validate the frequency of chip formation against acoustic emissions generation. Another more simple solution may

be using a 3D scanning microscope to measure the chips, allowing for analysis of curved chips that are produced in oblique turning. However, the issue of analysing a very small fraction of chips from a large sample would still be problematic in this case as it is difficult to state, without any doubt, that the analysed sample is representative of that particular pass of the workpiece as a whole.

Harmonics of the highly intense frequency bands associated with chip formation were observed near the end of tool life, possibly caused by a nonlinearities of the system and further amplified due to tool wear. This behaviour has not been noted in past literature and as a result, features based on these harmonics have not been used as inputs to any machine learning algorithm previously. In this thesis, these harmonics are exploited to obtain features that are indicative of tool failure.

The necessary workpiece changes that cause the confounding influences within the dataset were studied and a simple solution of feature selection within the high frequency content of the spectrogram was suggested. Dealing with confounding influences has not been a concern in past literature and therefore obtaining features that are robust against workpiece changes was discussed in this thesis possibly for the first time in the tool condition monitoring community.

In order to validate the origin of the confounding influences seen here, a controlled test could be considered. Ideally, two scenarios would be tested, where, the tool is not removed from the holder during the entire test to investigate exactly how the depth of cut is causing the changes seen in the AE. A dedicated microscope that permits the inspection of the tool at the machine would be sufficient to achieve this. The second scenario of testing the effect of workpiece hardness is, however, much harder to implement without having a thorough understanding of the hardness profile of each workpiece used in the experiment. In order to understand the affect of the workpiece, the simplest solution would be to test the tools on a true hardened<sup>5</sup> bar in a face turning<sup>6</sup> operation. In this scenario, the workpiece hardness would also be controlled.

---

<sup>5</sup>True hardening refers to consistent hardness across the workpiece. It is not possible currently to undertake outer diameter tuning on a true hardened bar as the hardness is inconsistent across the length of the workpiece.

<sup>6</sup>Face turning is where the tool traverses from the outer diameter to the centre of the workpiece.

Finally, 21 candidate features have been suggested as inputs to learning algorithms in Chapters 6 and 7. Features that progressively change across the life of the tool have been identified as potentially useful when comparing to monotonic tool wear features from Chapter 4. Chapter 6 will attempt to find relationships between these feature sets, and hence, it may be possible to predict the wear state of a tool at a given time. Furthermore, by studying the whole frequency domain across the entire tool life, it has been possible, maybe for the first time, to find features with increased activity just prior to tool failure. These features may be considerably helpful when predicting imminent tool failure in Chapter 7.



---

## SUPERVISED LEARNING

The ability to observe tool wear states at all times would be extremely helpful to machine operators in making decisions regarding the remaining life of the tool. However, due to physical restrictions, direct observation is not possible. This chapter therefore explores the feasibility of accurately predicting the behaviour of tools represented by features (also referred to as labels) selected in Chapter 4 using the corresponding acoustic emission features from Chapter 5, negating the need for direct tool wear measurements.

In this chapter, supervised learning methods are applied to the aforementioned feature sets. Specifically, the linear response surface methodology is employed that fits a plane through the input features in order to make predictions about the output. The results are then compared with Gaussian process regression to evaluate whether a nonlinear model is required for this data.

Section 6.1 reviews the related work on supervised learning in this context. Next, the two sets of features from previous chapters are examined simultaneously in Section 6.2 to discover suitable AE features for supervised learning. Subsequently, Sections 6.3 and 6.4 explore tool wear prediction through linear response surfaces followed by Gaussian Process regression. The predictive capability of these models when applied to tools that do not follow the expected wear progression is then studied in Section 6.5.

## 6.1 Related work on linear response surface methodology

As discussed in Section 2.2, researchers have attempted to use supervised learning in the field of tool condition monitoring to predict tool wear. Here, linear response surface (RSL) methodologies as well as Gaussian process regression (GPR) are explored as a model that may be used to predict tool wear using the AE dataset.

Response surfaces models were originally developed to find optimum conditions when conducting chemical experiments [106]. In the field of machining research, response surfaces are usually used in conjunction with design of experiments to quantify the effect of machining conditions and parameters on the outputs of the process such as tool wear or surface roughness [107–110]. Gupta [111] compared the goodness of fit of response surface methodology for predicting surface roughness, tool wear and power consumption, against artificial neural networks and support vector regression. However, none of these researchers have attempted to predict responses in real time for unseen tools (i.e. those not used to formulate the training set for a learning algorithm) or surfaces using input data collected online. Although the response surface models can be used to predict tool wear on unseen tools, this has rarely been undertaken using the RSL methodology. In structural health monitoring research, however, RSL methodology obtained high accuracies when predicting variables such as the lowest modal frequency of a bridge when using operational and environmental conditions as inputs [112]. When considering linear features in this thesis, the simplicity of the linear model may allow for fast predictions without the need for complicated algorithms.

When there exists many challenges related to feature collection, preprocessing and selection, it is difficult to know for certain the exact form of the underlying trend in the data. In the tool wear feature set collected in Chapter 4 for instance, it is not possible to state, without any doubt, that a given feature is linear with time because the sampling frequency may be insufficient; it is unlikely that one wear value representing four passes of the workpiece can accurately describe the entire change in the tool's cutting surface. In occasions such as these, GPR may be a suitable candidate to implement because the functional or model form does not have to be

specified for prediction. For a comprehensive discussion of Gaussian process models and their uses (especially in structural health monitoring), the reader is directed to [113], where the author has expanded on kernel selection for engineering, choice of cost function and hyperparameter optimisation.

GPR has been used in many different engineering applications to successfully predict responses. Using recorded flight parameters, Fuentes et al. [114] predicted aircraft structural loads by applying GPR. When the GPR model was tested on data collected from the same aircraft used to obtain the training samples, an average error of under 10% was achieved. In the offshore wind sector, GPR has been successfully utilised to predict the health state of one blade from training samples collected from another given structural properties of blades are correlated to one another [115]. At the presence of damage however, the weakening of the correlation between the properties of the damaged and undamaged blades lead to high residual errors in the model that was used as an indicator of damage.

In the field of medicine, to correct nonlinear trajectories of cable-driven surgical robots, Mahler et al. [116] used GPR with velocity of the robot as an input. A significant reduction in positional errors and angular errors were achieved along with an increased speed of 3.8 times over past results. In natural language processing, GPR has been used to automatically identify periodicity in texts because it plays an important role when analysing temporal dimension of text. The authors used this information to successfully identify tweet hashtags depending on the rest of the text in Twitter, a social media platform. To improve synthetic speech, Koriyama et al. [117] implemented GPR with a kernel function based on temporal acoustic events that is available for syllables and accent phrases. The GPR model outperformed Hidden Markov models, though the improvement was marginal.

In TCM, GPR is used rarely to predict tool life. A few examples have been detailed here. Kong [118] uses GPR to predict tool wear of cemented carbide tools in a turning operation. The authors merit this technique for its ability to model Gaussian noise quantitatively, resulting in better prediction accuracy compared to artificial neural networks and support vector machines. The authors were able to predict tool wear using integrated radial basis function based kernel principal component

analysis (KPCA-IRBF) and cross-validation methods. It is unclear which features were used for the model and whether the test set is a replication of the training set. Therefore it is uncertain whether prediction was performed on a previously unseen tool. Wang et al. [119] attempted to diagnose tool wear states of milling tools by implementing Continuous hidden Markov models and GPR. Through cross validation of the training set comprising of cutting force data, the authors found the number of tool wear states that should be present in the feature set. Then, by using data from these states as inputs to the model, where the number of inputs increased with each new added state, the authors predicted the remaining useful life of the tool close to the actual life. Hong et al. [120] used GPR to predict tool wear on a deep hole drilling gun where the results were compared with linear regression, support vector regression, and an ensemble method. GPR outperformed the other methods when comparing root mean squared error and the R2Score (assumed to be the  $R^2$  value from mathematical expression presented). The authors are unclear about the number of observations in the feature set and the nature of the inputs and outputs; there is no information on the conditions in which the training and testing set were collected under. As a result, it is difficult to evaluate the success of this work. Zhang et al. [121] used GPR to investigate whether surface roughness could be predicted using a combination of cutting conditions where prediction accuracy of around 80% was achieved. This method was not used for tool condition monitoring. Instead, the purpose was to find relationships between the variables, not to predict surface roughness in real time for previously unseen workpieces using inputs that are collected online.

It seems that the potential of implementing response surfaces and Gaussian process regression to predict wear has not been realised in the TCM community. In this chapter, it is hoped that these models will be able to predict tool wear on unseen turning inserts using a combination of AE and force features, following a rigorous feature selection and training procedure.

## 6.2 Feature correlation

The performance of machine learning techniques rely heavily on the feature selection process. The selected AE features in this work must be sensitive to changes in the cutting tool due to wear. By examining the linear correlations between the two feature sets, it is possible to extract the most descriptive features whilst furthering the understanding of the relationships between AE and tool wear. The aim of this section is to find signal features that not only describe the tool damage well, but can also be collected with ease, and ideally, online.

One of the unfortunate drawbacks of the feature selection process is the requirement to compress, summarise or down sample the AE dataset in some way, in order to match the size of the tool feature set; the number of tool wear feature observations are limited by the nature of the machining process, as discussed in Chapter 3. Subsequently, the AE data collected over four passes of the workpiece corresponding to around 130 seconds of data was averaged<sup>1</sup> and is represented by a single data point in this work. Inevitably, a significant amount of information is compromised during this process.

The small number of available damage label observations also limit the number of input variables that can be used in the RSL methodology [106]. Consequently, only a small number of signal features need to be correlated well with tool wear.

Due to the considerable impact on the surface integrity of components when machined with worn tools, it is of no use to predict the value of tool wear once breakage has taken place. Prediction of imminent failure is far more useful to an operator. With this in mind, only the data up to the last tool measurement prior to failure are used for prediction in this chapter.

In general, it is easier to understand and predict features that correlate linearly to one another. Therefore existence of such relationships should be explored first as they are preferred over complex behaviours. Here, equation 6.1 [122] is used to calculate the strength of linear correlations (also referred to as Pearson correlation),

---

<sup>1</sup>Mean values of the data from each four passes of the workpiece is used as the feature here.

where  $n$  is the number of observations,  $x$  is the AE feature,  $y$  is the tool feature,  $\bar{x}$  is the mean of  $x$  values,  $\bar{y}$  is the mean of  $y$  values and  $\sigma_x$  and  $\sigma_y$  are the variance of  $x$  and  $y$  respectively.

$$r = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sigma_x} \right) \left( \frac{y_i - \bar{y}}{\sigma_y} \right) \quad (6.1)$$

The entire feature set comprises 40 and 51 observations for grade A (A3-A9) and B (B3-B8) tools respectively, where each data point represents four machined passes of the workpiece, or a sliding distance of around 600m.

The strength of the linear correlation between tool features extracted from the middle of the wear scar, and AE features from Sensors 1 and 2 for tool grades A and B are presented in figures 6.1 and 6.2 respectively. The strengths of the linear relationships are shown here as a value between 0 and 1, where 1 represents a perfectly linear relationship and 0 suggests that the features do not share a linear relationship. It should be noted here that all AE features represent the energy of the signal. Also, principal component analysis (PCA) was carried out to reduce the dimensionality of the dataset in order to find features that represent the highest variance in the data; PC1 and PC2 refers to the first and second principal components respectively. Section 5.4 gives a brief introduction to PCA.

These results can either be affected by the tool grade, the progression of tool wear or by the type and position of the sensor. It is immediately clear from figure 6.2 that most AE features from Sensor 1, grade B are generally well correlated with the corresponding tool features, with around 9 out of 21 features scoring higher than 0.7. In stark contrast, for Sensor 2, RMS of the chip formation is the only feature to score higher than 0.7 for grade B. The reality is even bleaker for grade A results, where the highest strengths are in the vicinity of 0.45 for both sensors. Consequently, it can be assumed that the sensors, and the tool grade, influence the feature correlation results significantly. It is assumed that the low signal to noise ratio of the data collected from Sensor 2 as discussed in Section 5.3 is the reason for these notably low correlation strengths.

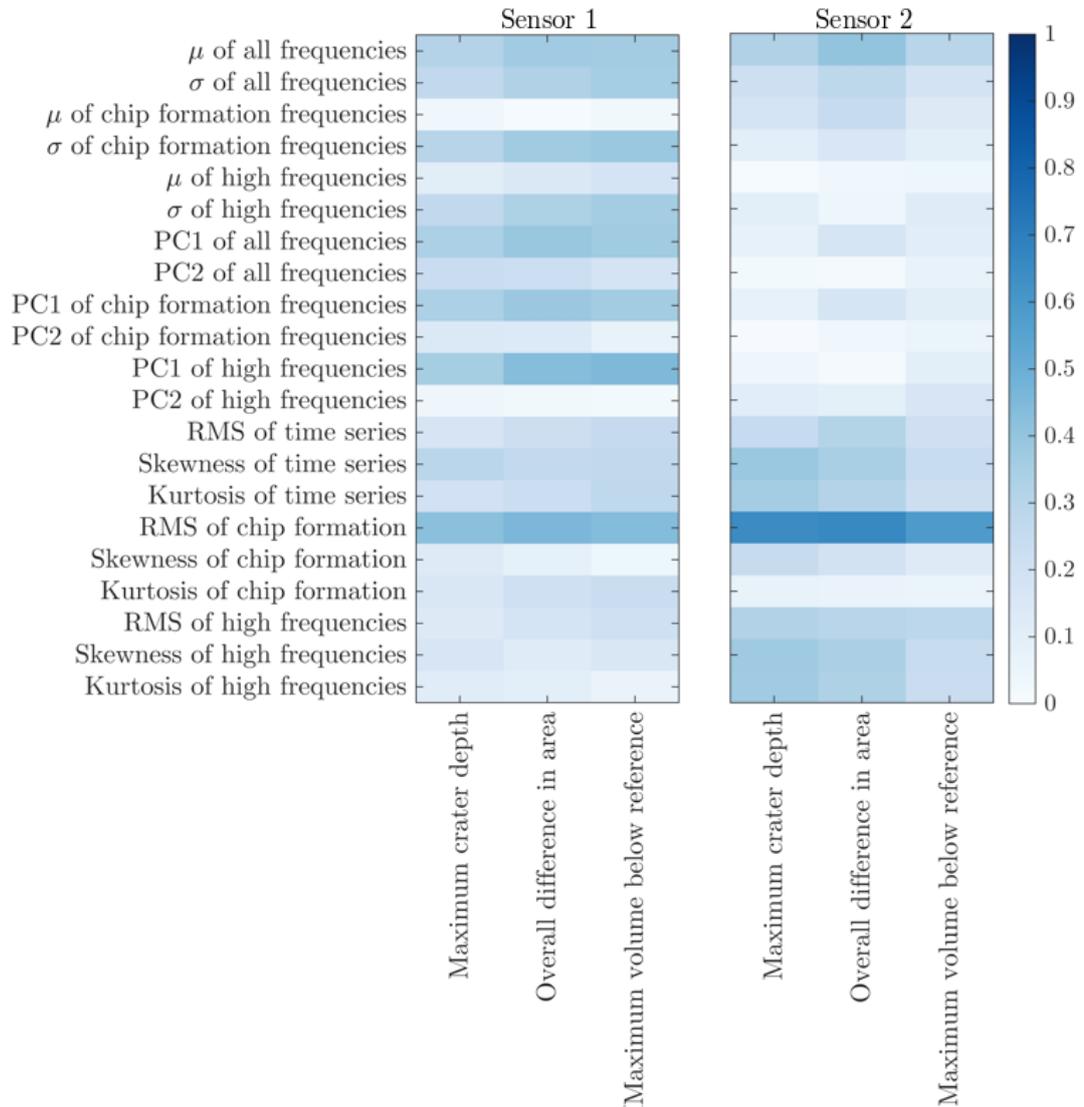


Figure 6.1: Linear correlation between the chosen tool features (Chapter 4) and AE features (Chapter 5) from grade A tools. Results from Sensor 1 is on the left and Sensor 2 on the right. The strength of the correlations is shown without taking the direction into account.

At this stage, it is useful to examine the features that present a strong linear correlation with tool wear from figure 6.2. Maximum crater depth presents lower correlation strengths across all AE features as it scores lower than flank area and maximum volume below reference surface (in section 4.5.3). For tool B, Sensor 1, the

strongest relationships are between the features that capture the standard deviation of the AE signal in the frequency domain. An example of this has been presented in figure 6.3. This suggests that the amplitude of the frequencies are nonstationary (in the variance) and this nonstationarity is linear. This result is not surprising as tool

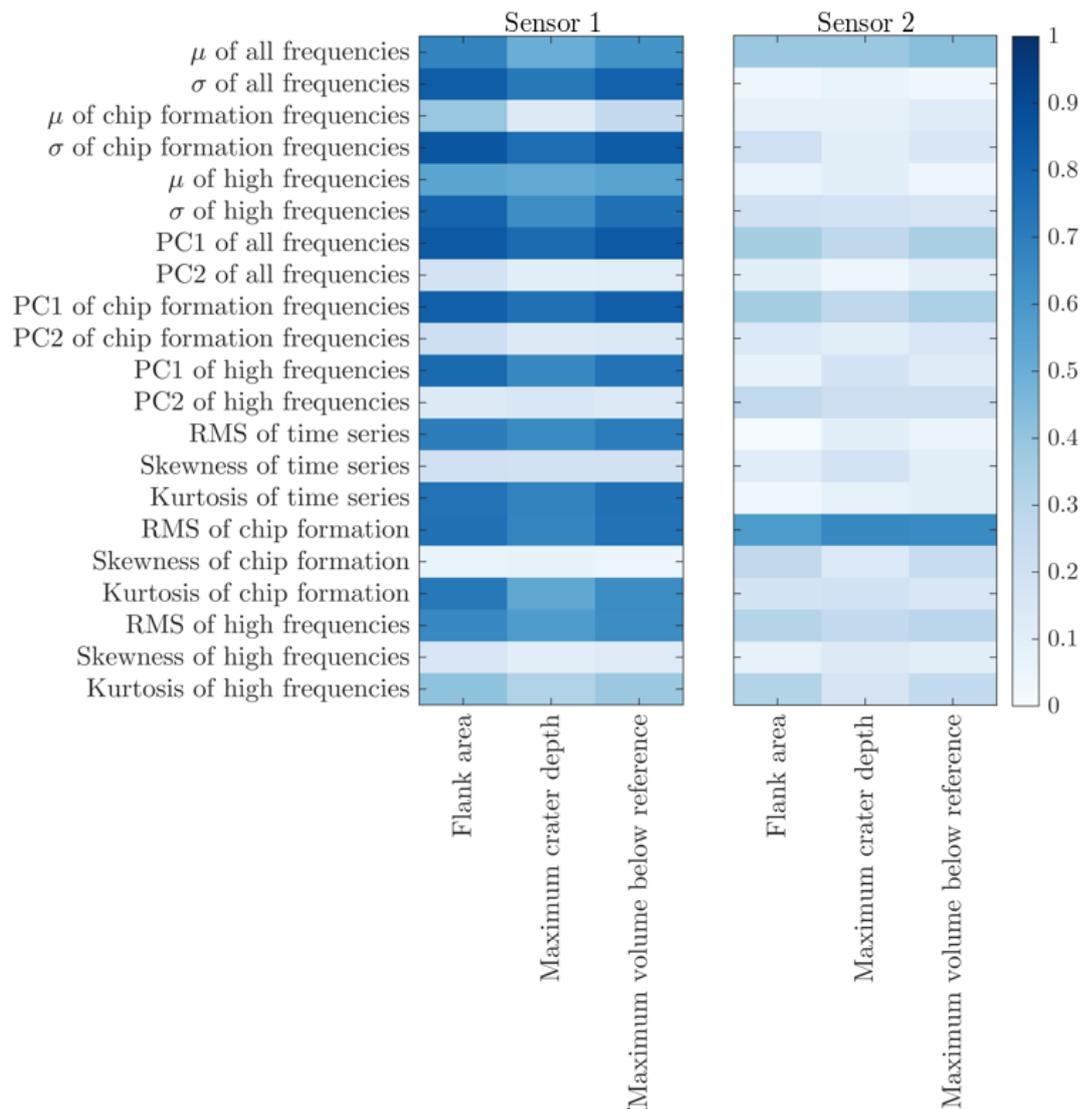


Figure 6.2: Linear correlation between tool features and AE features from grade B tools. Results from Sensor 1 is on the left and Sensor 2 on the right. The strength of the correlations is shown without taking the direction into account.

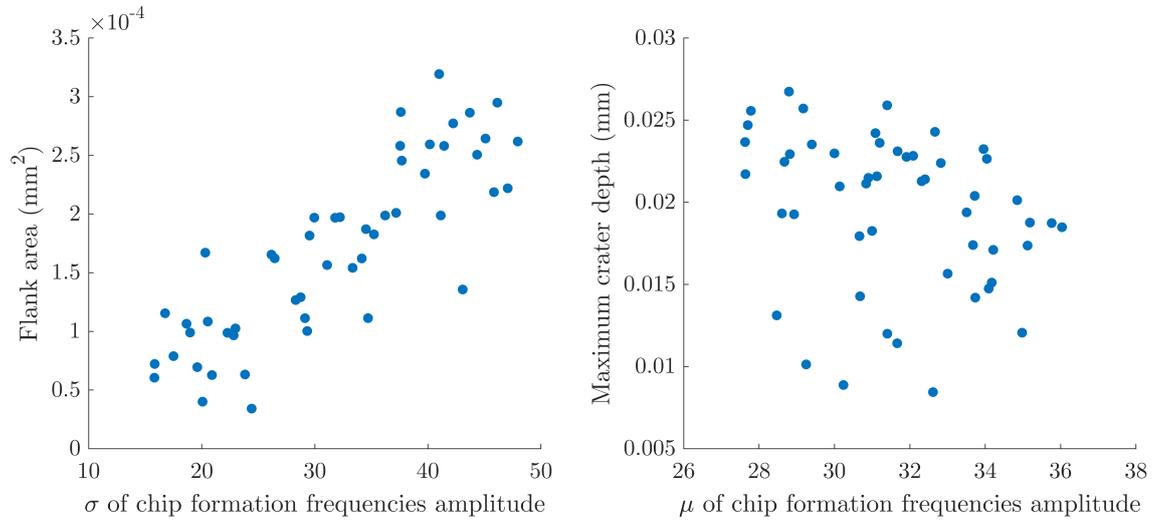


Figure 6.3: An example of a highly correlated feature pair on the left and a weakly correlated pair of features on the right. Both signal features were collected from tool grade B, Sensor 1.

wear leads to an increase in energy in the AE frequencies as seen in Chapter 5. It is assumed that the location of Sensor 2 is the cause of the low correlation values

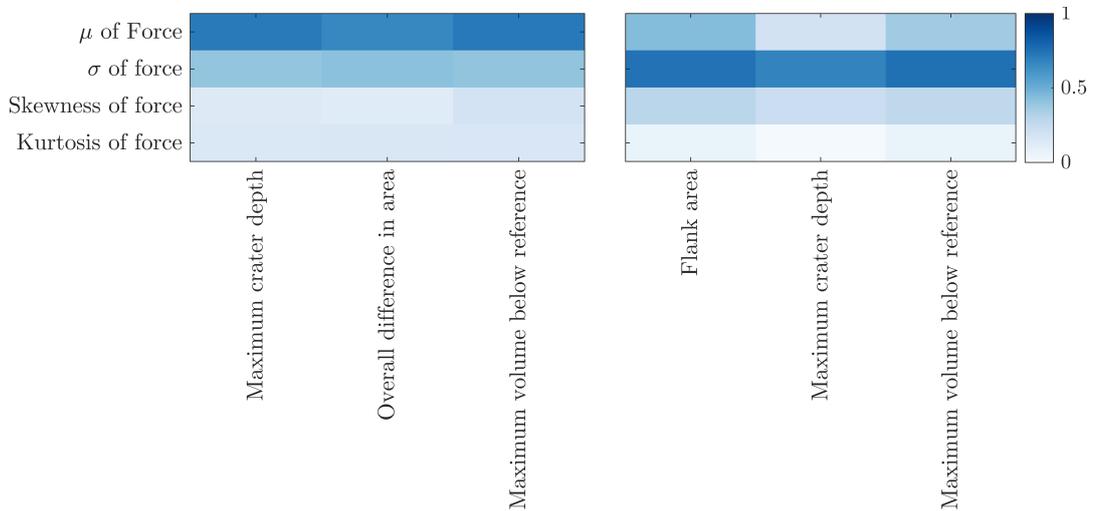


Figure 6.4: The strength of linear correlation between statistical moments of the force and tool features for grade A tools (left) and grade B tools (right).

seen in figure 6.2. As the sensor is situated on the turret, its signal is attenuated significantly as it must travel through numerous mediums and structures, leading to a low signal to noise ratio; the extracted features are affected as a result. However, correlation strengths above 0.7 exist between RMS of the chip formation frequencies and the tool wear features for Sensor 2. This is assumed to be due to chip formation frequencies containing the most energy of the signal.

The first four statistical moments of the resultant force signal collected from the tool holder dynamometer are also calculated here for comparison. From figure 6.4, it can be seen that the mean of the force results are highly correlated with tool wear for grade A tools with strengths between 0.66-0.72 and standard deviation for grade B with strengths around 0.67-0.75.

### 6.2.1 Feature selection

With all the correlations reviewed, the selection process for input features can now commence. As the results from grade A tools did not present many significant linear relationships, the strongest two features,  $\mu$  of force and RMS of chip formation from Sensor 2 are chosen as the inputs. For tool type B however, number of feature sets could be tested for predictive performance according to how well they correlate with tool wear and how easy they are to collect online. The signal features from grade B are evaluated against the criteria stated below. The top 10 signal features for grade B Sensor 1 are presented in table 6.1, where the features used for prediction are colour coded against the feature set for clarity.

For further analysis the features are collected into three sets:

#### Feature set 1 (FS1)

The prediction performance of the input features that are correlated strongest with tool wear. These are the top three signal features in each wear category in table 6.1.

**Feature set 2 (FS2)**

Though principal component scores are correlated well (blue features in table 6.1), these features cannot be calculated online. For this reason, the second scenario evaluates how well the model performs with inputs of the strongest online AE features and force measurements (the online features and the force features are in purple and green respectively in table 6.1).

**Feature set 3 (FS3)**

Collecting the force data for sensor fusion requires the installation of a large and expensive tool holder dynamometer on the turret. A futuristic set up (that includes force measurements) capable of automatic tool changes as a result of tool condition monitoring would require a dynamometer on every tool of the turret. In contrast, attaching AE sensors on the tool holders is an inexpensive and compact solution. Therefore, a third scenario is explored where only the best AE features (that can be collected online) are used to execute the model. Out of the features that can be collected online, however, some features such as the kurtosis of the chip formation and high frequencies require the reconstruction of the time series following a wavelet transformation. Although this is possible to conduct online, the computational power leads to longer time consumptions. As a result, features that demand the shortest amount of time to be calculated are better suited for online applications.

Normally when considering inputs for linear regression models, multicollinearity between the features must be considered. Multicollinearity exists if there are strong correlations between the input features themselves. Usually, it is advised to avoid using inputs that are well correlated to one another as it complicates the model interpretation; multicollinearity makes it difficult to make inferences about the role of each input variable on the response, as numerous response functions with varying regression coefficients are able to prove a good fit. In this work, however, the goal is to predict tool wear as accurately as possible. Therefore, the focus here is the predictive error, although the model structure remains an important consideration. Consequently, any model that fits the data with the lowest error values will be

Rank	Flank area	Maximum crater depth	Maximum volume below reference surface
1	$\sigma$ of chip formation frequencies 0.85	PC1 of all frequencies 0.79	PC1 of all frequencies 0.83
2	PC1 of all frequencies 0.83	PC1 of chip formation frequencies 0.77	$\sigma$ of chip formation frequencies 0.82
3	$\sigma$ of all frequencies 0.82	$\sigma$ of chip formation frequencies 0.76	PC1 of chip formation frequencies 0.81
4	$\sigma$ of high frequencies 0.80	$\sigma$ of all frequencies 0.72	$\sigma$ of all frequencies 0.79
5	PC1 of chip formation frequencies 0.80	PC1 of high frequencies 0.70	PC1 of high frequencies 0.76
6	PC1 of high frequencies 0.78	Kurtosis of time series 0.68	$\sigma$ of force 0.75
7	RMS of chip formation frequencies 0.75	$\sigma$ of force 0.68	$\sigma$ of high frequencies 0.75
8	$\sigma$ of force 0.74	RMS of chip formation frequencies 0.67	Kurtosis of time series 0.75
9	Kurtosis of time series 0.74	$\sigma$ of high frequencies 0.65	RMS of chip formation frequencies 0.74
10	Kurtosis of chip formation 0.72	RMS of time series 0.65	RMS of time series 0.70

Table 6.1: The strongest 10 signal feature correlations with tool wear features. The features used for prediction are colour coded as: purple - AE features that can be captured online, blue - AE features that cannot be captured online, green - statistics of the force features.

serviceable as multicollinearity does not affect the predictions of new observations or the mean of a response [123].

With the finalising of the feature selection process, it is now possible to undertake supervised learning using RSLs. The next section introduces the concept and presents the findings.

### 6.3 Linear response surfaces

The response surface model allows the use of multiple input variables to establish a regression using data that best fit the response. Just as in linear regression with a single input where a line is fit through the observed data, the idea in these models is to fit a plane through the multiple inputs that best describe the response. The generic model is stated in equation 6.2. In this model,  $y$  is the response,  $x$  is the input variable,  $n$  is the number of input variables and  $\beta$  are the parameters to be found.

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n \quad (6.2)$$

Higher order effects such as polynomial terms are more challenging to estimate simply because a large number of observations (experiments) are required to model increased number of variables. Given the size of the dataset available for this work, only the simple, linear case is considered here.

The normalised mean square error (nMSE or MSE) will be used as a measure to evaluate model performance in this work (shown in equation 6.3, where  $\sigma$  is the variance.). If instead of a prediction model, only the mean of the data was used to make predictions, the resulting MSE would be 100%. Therefore, MSE values below 100% signify the existence of correlation and above 100% suggests that the performance is worse than if the mean of the data was used instead of the model [112].

$$\text{MSE} = \frac{100 \sum (\text{model errors})^2}{n (\sigma [\text{predictions}])^2} \quad (6.3)$$

In this work, each response (tool wear features chosen in Chapter 4) is modelled separately, where tools A3-A6 (21 observations for each input) are chosen to train the grade A tools (this was a random selection). For grade B, the training set is made up of tools B3-B5, containing 25 observations of each predictor. Tools A7-A9 (18 observations) and B6-B8 (26 observations) are used for testing grades A and B respectively.

### 6.3.1 Results

Using the best possible AE features from the linear correlations, the response surface method is used to predict the tool wear features for previously unseen tools. Due to the weakly correlated AE features for tool A, the two best performing features, RMS of chip formation frequencies from Sensor 2 and  $\mu$  of the force signal are used. The results with the corresponding test MSE values can be seen in figures 6.5a to 6.5c. It should be noted that the model was performed independently for each tool wear feature.

At first glance, the fit looks good with the prediction with AE features mostly following the tool wear trend. In all cases, the model generalises well from training to testing. The testing MSE values of the models vary between 28-46%, depending on the response feature. For flank area, the trend has been captured with errors of around 28%. The marginal drop in test MSE values for overall difference in area may be due to the weaker signal correlation strengths. Given the same input signal features were used for estimating all three tool wear features, it is interesting to note what may be the reason for the decrease in model accuracy for maximum volume below reference surface with test MSE value of around 45%; what is causing the overestimate at point 26 in figure 6.5c?

Figures 6.6a and 6.6b show the wear scars of the tool as a difference between the new and the worn tool scans for tool A4 after pass 24 and tool A9 after pass 28 respectively. For both tools, this was the penultimate wear scan before catastrophic failure. Tool A4 makes up the training data points 5 through to 10, and data from tool A9 are in the test set represented by points 33 to 39. It is clear from these

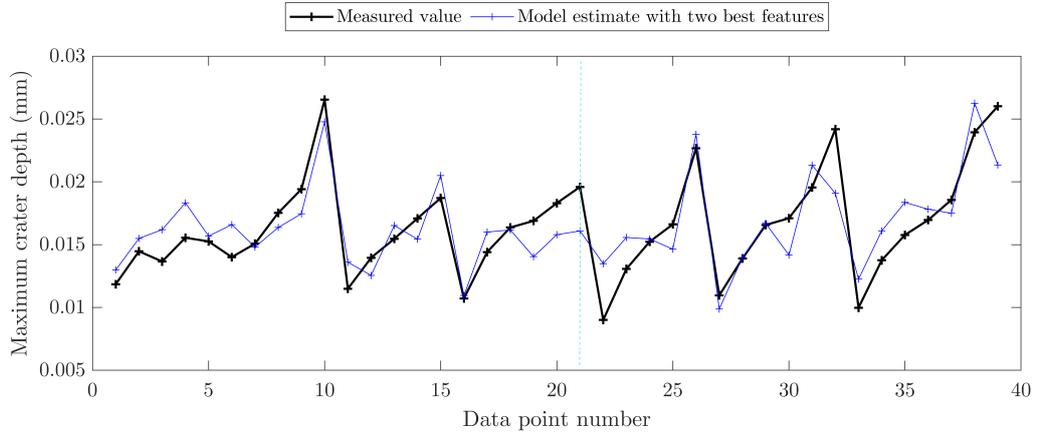
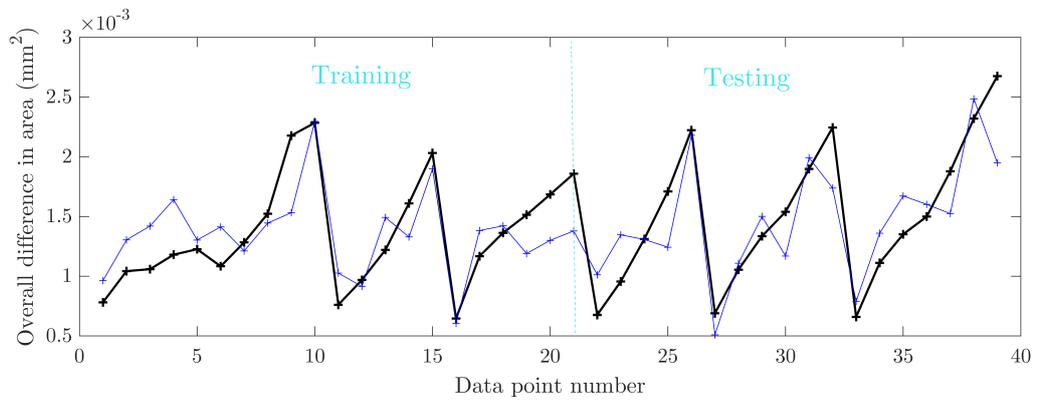
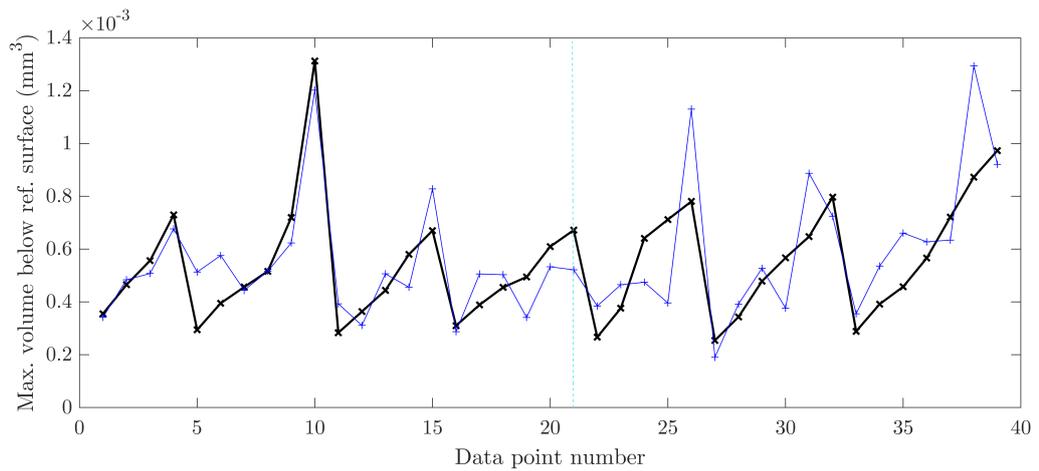
(a)  $MSE_{\text{test}} = 28.35$ (b)  $MSE_{\text{test}} = 32.85$ (c)  $MSE_{\text{test}} = 45.92$ 

Figure 6.5: Prediction results using the RSL for grade A tools. In all cases, the model was trained on data up to data point number 21 representing the tools A3-A6. The model was tested on tools A7-A9. Figure 6.5b display the testing and training split.

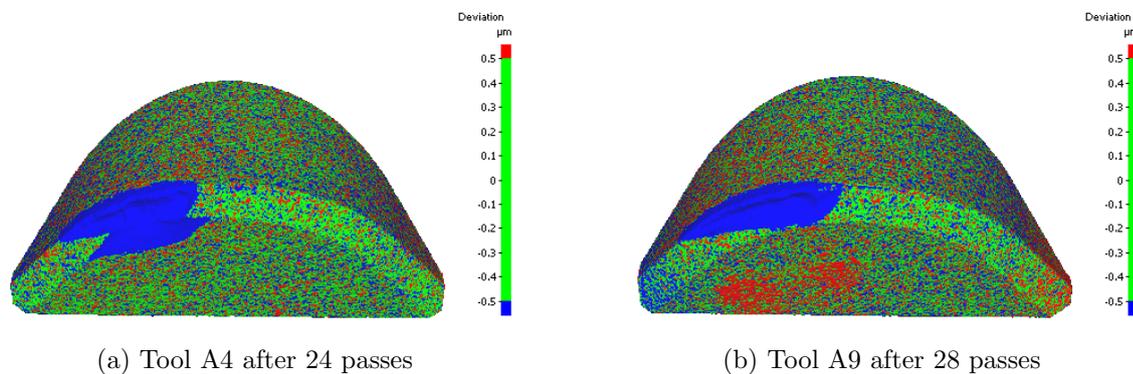


Figure 6.6: Tool scans highlighting the difference between a new tool and a worn tool for (a) worn tool with material missing from rake face and (b) worn tool without material missing from the rake face.

figures that the wear scar is quite different in figure 6.6a compared to figure 6.6b. This difference is the loss of material from the rake face for tool A4. A wear scar of this nature, in rare occasions, may not meet the threshold for failure as cutting edge may still be intact. In this case, the tool wear features from the flank face and some crater wear features may not be affected as the damage is localised to the rake face of the tool outside of the wear scar. However, the maximum volume loss between the two features is sensitive to this type of wear. At this point, a number of questions should be asked. Firstly, should a tool that has damage outside the wear scar be used for machining? The measured values for surface roughness, hardness or depth of cut at pass 24 does not suggest that this tool should be discarded. In that case, should the chosen method of feature extraction be robust against damages outside the wear scar? Finally, should a tool that wears slightly differently to others, or have damage that does not affect the cutting process (and, therefore, the signal features), be removed from the training and testing datasets? The author believes that for this particular example, feature selection should be robust against damage that does not affect the machining process as the aim of using this model is to only predict tool wear. Consequently, when using the RSL methodology, care must be taken when using features that do not concentrate on just the wear scar, such as maximum volume below the reference surface. For a detailed investigation on tools that wear

dissimilarly to others in the dataset, and how this may affect the prediction of RSL models, the reader is referred to Section 6.5.

At this point, it would be interesting to investigate whether the RSL model would achieve a good accuracy for grade B tools as well. Figures 6.7a to 6.7c display the results for grade B tools, where models using FS1, FS2 and FS3 are shown in blue, red and green respectively.

Across all feature sets, the models have achieved a good fit where the lowest test MSE scores are achieved when modelling flank area with an average around 26%, whereas highest average of 41.31% is found for maximum crater depth. It is also beneficial to obtain information regarding the feature set that modelled the data most accurately in order to understand if an entirely online process could be achieved. With test MSE scores of 28%, FS3 performed just as accurately as FS1 when modelling flank area. Again, the models have generalised well from training to testing.

For all grade B wear features, there is a notable overestimation when using FS2 at point 33 in figures 6.7a to 6.7c; this is the penultimate wear value before failure for tool B6. It seems that this overestimation may be due to force data as FS2 is the only feature set to contain data collected from the dynamometer. Upon inspection it seems that the resultant force has increased significantly during the second half of pass 31; none of the other tools from the training or testing sets display this behaviour. The physical driving mechanism behind this behaviour is currently unclear. As the acoustic emissions are not affected by the same mechanism, it is possible that the increase in force did not affect the chip formation mechanisms and hence could be a result of swarf bunching around the dynamometer. It is not possible to verify this theory without a visual evidence from the process such as a video recording. If this increase in force is a result of some outside influence that does not affect the tool wear process, then care should be taken when using force measurements; the features used in the prediction algorithms should be robust against environmental conditions. Another noteworthy overestimation can be found in figure 6.7a at points 36 and 37 from all feature sets. This data belongs to tool B7. Between passes 8 and 20, the input features increase much more than for other tools in the training and testing sets. Upon studying the depth of cut, hardness, and surface roughness measurements,

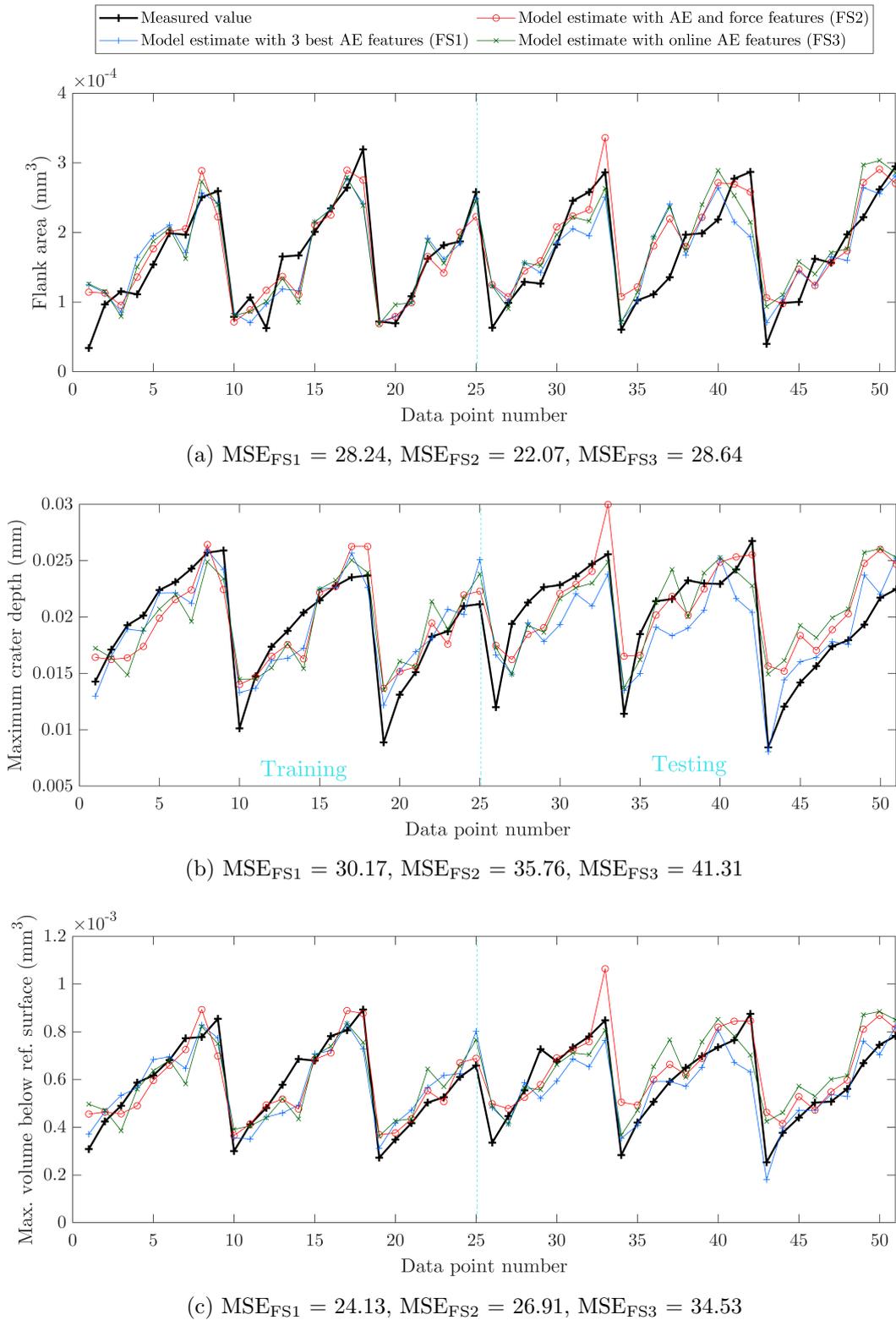


Figure 6.7: Prediction results using the RSLs. In all cases, the model was trained on data up to data point number 25 representing the tools B3-B5. The model was tested on tools B6-B8 (data points 26-51). Figure 6.7b display the testing and training split.

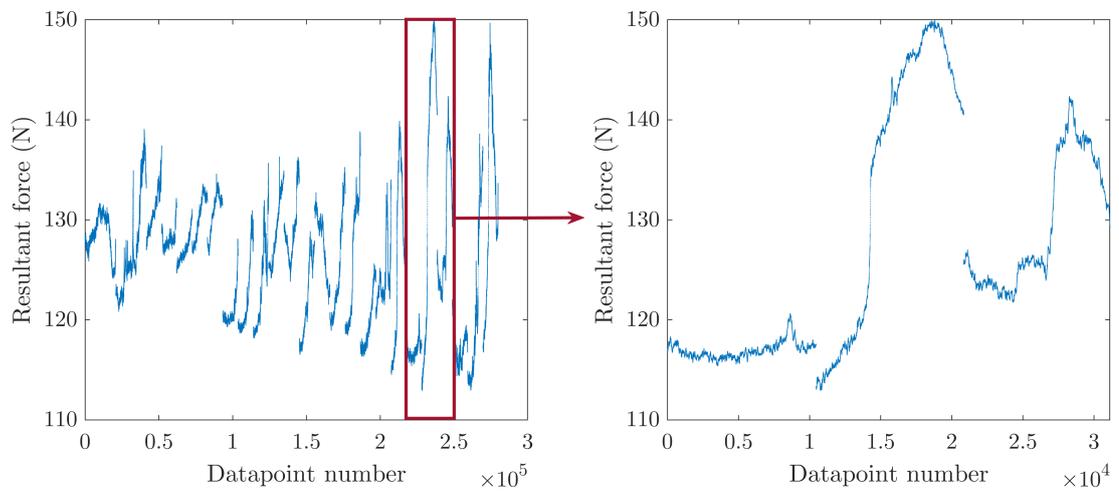


Figure 6.8: Resultant force data collected from tool B6 where the figure on the left shows data from the entire life of the tool, and figure on the right highlights passes 30-33.

the reason for the increase remains inconclusive. A combination of confounding influences and wear feature extraction measurement errors may be the reason for disparity between the measured and estimated values at this point.

Table 6.2 displays a summary of the training and testing MSE values for both tool grades and across all feature sets.

Even though the linear response models are able to track the progression of tool wear, MSE values of over 20% suggest that some variance between the real target and the model estimate is present. It is important to understand if this variance is large enough to prompt a situation where an unacceptably damaged tool is used for machining due to the performance of the model. Visually, the results are inconsistent across the test tools as the models sometimes predict larger wear values and sometimes significantly smaller values as can be seen in figure 6.7b between data point number 26-43. One way to improve these MSE values may be to increase the size of the training set or collect more wear data during the machining process. Furthermore, it may be beneficial to investigate whether a model that is not bound by restrictions of linearity would perform well when considering the linear and non-linear signal features that describe tool wear. Polynomial models could be explored here. However,

due to the limited number of observations available for modelling, polynomial models are not suitable for this dataset. For this reason, Gaussian process regression is explored in the next section.

## 6.4 GP Regression

The Gaussian process (GP) is capable of performing inference over functions that pass through its input data. For small datasets (around 100 data points) such as this one, GPs are well suited as the computation power and storage required are low;

			Flank area	Maximum crater depth	Overall difference in area	Max. volume below reference surface
A	RMS of chip formation & $\mu$ of force	Train	-	29.75	40.26	21.91
		Test	-	28.34	32.85	45.93
B	Highly correlated AE features (FS1)	Train	21.17	17.50	-	21.01
		Test	28.24	30.17	-	24.13
	Sensor fusion (FS2)	Train	17.08	24.43	-	18.98
		Test	22.07	35.76	-	26.91
	Online AE features (FS3)	Train	22.19	31.36	-	30.22
		Test	28.64	41.31	-	34.53

Table 6.2: The percentage error (MSE) using RSLs.

GPs must perform matrix inversions leading to a complexity of  $\mathcal{O}(n^3)$ . Also due to the non-parametric nature of GPs, an in depth and complicated architecture of the model is not required. GPs also automatically define the uncertainty of the model, making them favourable over techniques such as artificial neural networks.

A brief introduction into GPR and prediction is presented here; the explanations closely follow Rasmussen [124] and Barber [125]. Just as how the mean and the variance can describe a multivariate distribution, functions of the mean,  $m$  and covariance,  $K$  are able to describe a Gaussian process (equation 6.4a).

$$\mathbf{f} = \mathcal{GP}(m, K) \quad (6.4a)$$

The mean function in equation 6.4b encodes the *a-priori* expectation of the unknown function. Here,  $\mathbf{X}$  is the matrix of all the inputs and has size  $D \times n$ , where  $D$  is the number of dimensions (variables), and  $n$  is the number of observations. It is common to set the mean function to zero when prior knowledge of the process is unknown. In this case, usually, the outputs are normalized.

$$m(\mathbf{X}) = \mathbb{E}[f(\mathbf{x})] \quad (6.4b)$$

The covariance function presented in equation 6.4c gives a measure of similarity between two vectors, or functions of vectors in this case. This assumption makes inference possible as it is assumed that similar data points have similar function values. The covariance function must be symmetric and positive semidefinite. The covariance function plays an important role in specifying the prior in GPs; adjusting the hyperparameters here can change the behaviour of the sample functions. Here,  $\mathbf{x}$  and  $\mathbf{x}'$  are input variables.

$$K(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \quad (6.4c)$$

The prior distribution of functions is shown in equation 6.5. The prior specifies the type of functions that are expected.

$$\mathbf{f}_{prior}(\mathbf{X}^*) \sim \mathcal{GP}(\mathbf{m}(\mathbf{X}^*), \mathbf{K}(\mathbf{X}^*, \mathbf{X}^*)) \quad (6.5)$$

As it is time consuming and computationally expensive to draw multiple samples from a prior distribution of functions and eliminate the undesired ones that do not fit the data, it is more efficient to condition the prior on the training data. With training data  $\mathbf{X}$  and testing data  $\mathbf{X}^*$ , and adding Gaussian noise into the model  $y = f(\mathbf{x}) + \varepsilon$ , the joint distribution under the prior is specified in equation 6.6. Here  $\sigma_n^2$  is the variance from the noise.

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{y}^* \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I & K(\mathbf{X}, \mathbf{X}^*) \\ K(\mathbf{X}^*, \mathbf{X}) & K(\mathbf{X}^*, \mathbf{X}^*) \end{bmatrix}\right) \quad (6.6)$$

Once conditioned using standard results for the conditional properties of a Gaussian, the targets are distributed as described in equation 6.7. Here, the new mean function gives an indication of the prediction and the covariance function permits the calculation of the model confidence.

$$\begin{aligned} \mathbf{y}^* | \mathbf{X}^*, \mathbf{X}, \mathbf{y} &\sim \mathcal{N}(K(\mathbf{X}^*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I]^{-1} \mathbf{y}, \\ &K(\mathbf{X}^*, \mathbf{X}^*) - K(\mathbf{X}^*, \mathbf{X})[K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I]^{-1} K(\mathbf{X}, \mathbf{X}^*)) \end{aligned} \quad (6.7)$$

#### 6.4.1 Specifying the kernel and learning the hyperparameters

The choice of the kernel and the hyperparameters should depend on the dataset. Without much prior knowledge however, it is advantageous to use a kernel that has the ability to describe the underlying trends in the data. The Gaussian kernel or the squared exponential (SE) kernel is a popular choice as it is infinitely smooth<sup>2</sup>. The dataset in this work however, is not smooth and is assumed to have linear trends. Therefore it is beneficial to apply a kernel that reflects this behaviour. On the other hand, using a linear kernel is not advised here as it will not be able to capture any underlying nonlinearities (should they exist) in the data [113]. Therefore, a Matérn $\frac{3}{2}$  kernel<sup>3</sup> is used in this work. The form of the kernel is specified in equation 6.8 [126], where  $K_\nu$  is a modified Bessel function and  $\nu$  is a constant that specify the number of differentiations.

<sup>2</sup>The SE kernel can be differentiated an infinite number of times.

<sup>3</sup> $\frac{3}{2}$  specifies that this kernel can be differentiated once.

$$k_{Matern}(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{\sqrt{2\nu} |\mathbf{x} - \mathbf{x}'|}{l} \right)^\nu K_\nu \left( \frac{\sqrt{2\nu} |\mathbf{x} - \mathbf{x}'|}{l} \right) \quad (6.8)$$

The hyperparameters (length scales and noise parameters) in this work are learned using a quantum particle swarm optimiser [113].

#### 6.4.2 Feature selection for the GP

As the GP is capable of modelling non-linearity within the data, it is important to explore the existence of feature sets that may exhibit these relationships. Mutual information (MI) gives an indication of how variables move with one another without giving any information regarding the type of correlation; MI is used to find the average amount of information that one feature conveys about another in this context [78]. Though these complex relationships may be harder to understand, they may give a better insight into the feature sets. Equation 6.9 is used to calculate MI:

$$I(X; Y) \equiv H(X) - H(X|Y) \quad (6.9)$$

Where the conditional entropy is,

$$H(X|Y) \equiv \sum_{x \in A_X A_Y} P(x, y) \log \frac{1}{P(x|y)} \quad (6.10)$$

The calculation of MI is affected by the parameters used in the formulation; when evaluating the joint distributions, the number of bins used will have an effect on the results. Furthermore, the number of observations in the feature set can also influence the value of the calculated MI. Values of MI for example distributions are displayed in figure 6.9 as a guide to the reader. The numbers were calculated using a bin size of 10 and with 100 observations. Features that have no association with one another have low MI values where the value increases with features that are correlated well with one another, whether the correlation is linear or a higher order.

Figures 6.10 and 6.11 present the mutual information for grade A and B tools respectively. The MI of grade A tools indicates that correlations exist between the

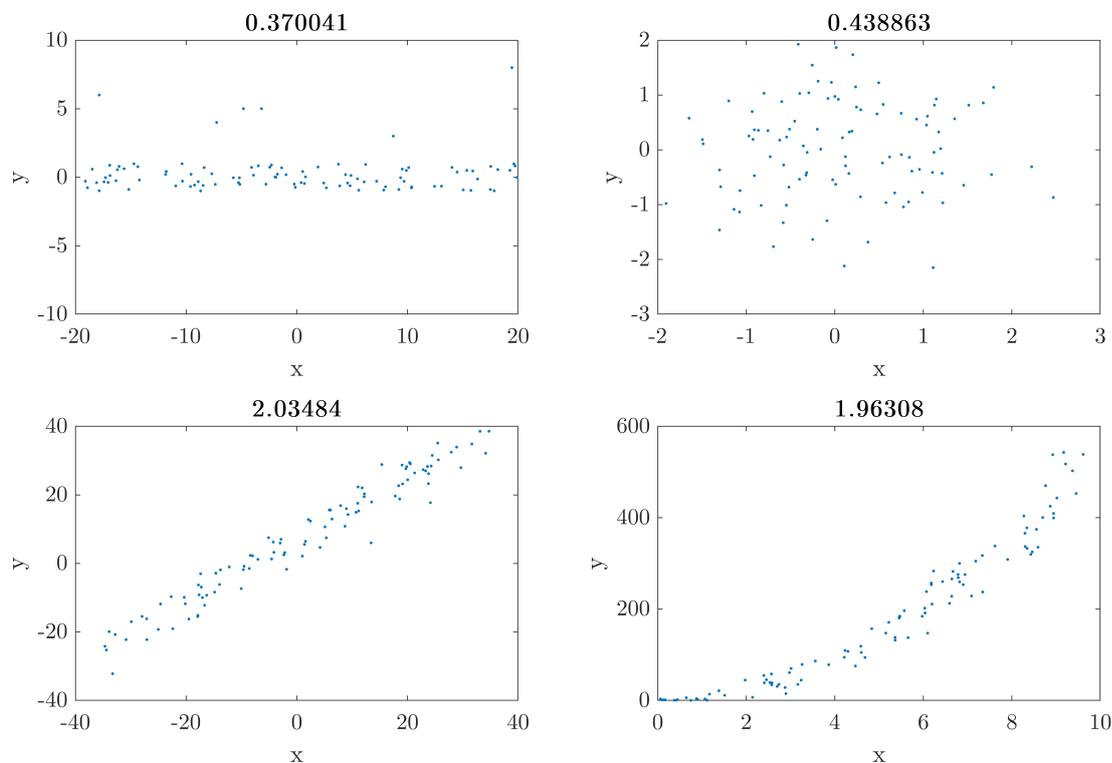


Figure 6.9: Examples of cluster shapes generated using random two dimensional data and the corresponding mutual information. The following cluster types are displayed: top left - flat cluster, top right - circular cluster, bottom left - sloped cluster with low variance, bottom right - quadratic type cluster with low variance.

tool and AE features, where some of the stronger relationships are from the signal collected from Sensor 2. Due to the analysis conducted in the previous section, it can be assumed that these relationships are not linear. The results for tool B are somewhat different with the highly correlated features emerging from Sensor 1.

Due to the nature of the GPR there are no restrictions on how many input variables that can be included in the model, unlike the RSL methodology. However, the curse of dimensionality states that as the number of dimensions increase, more data is needed to define the projected space. As more features are added, the metrics used to calculate the distance between the data breaks down in the absence of large number of observations. Due to the restrictions set by the number of target values in this

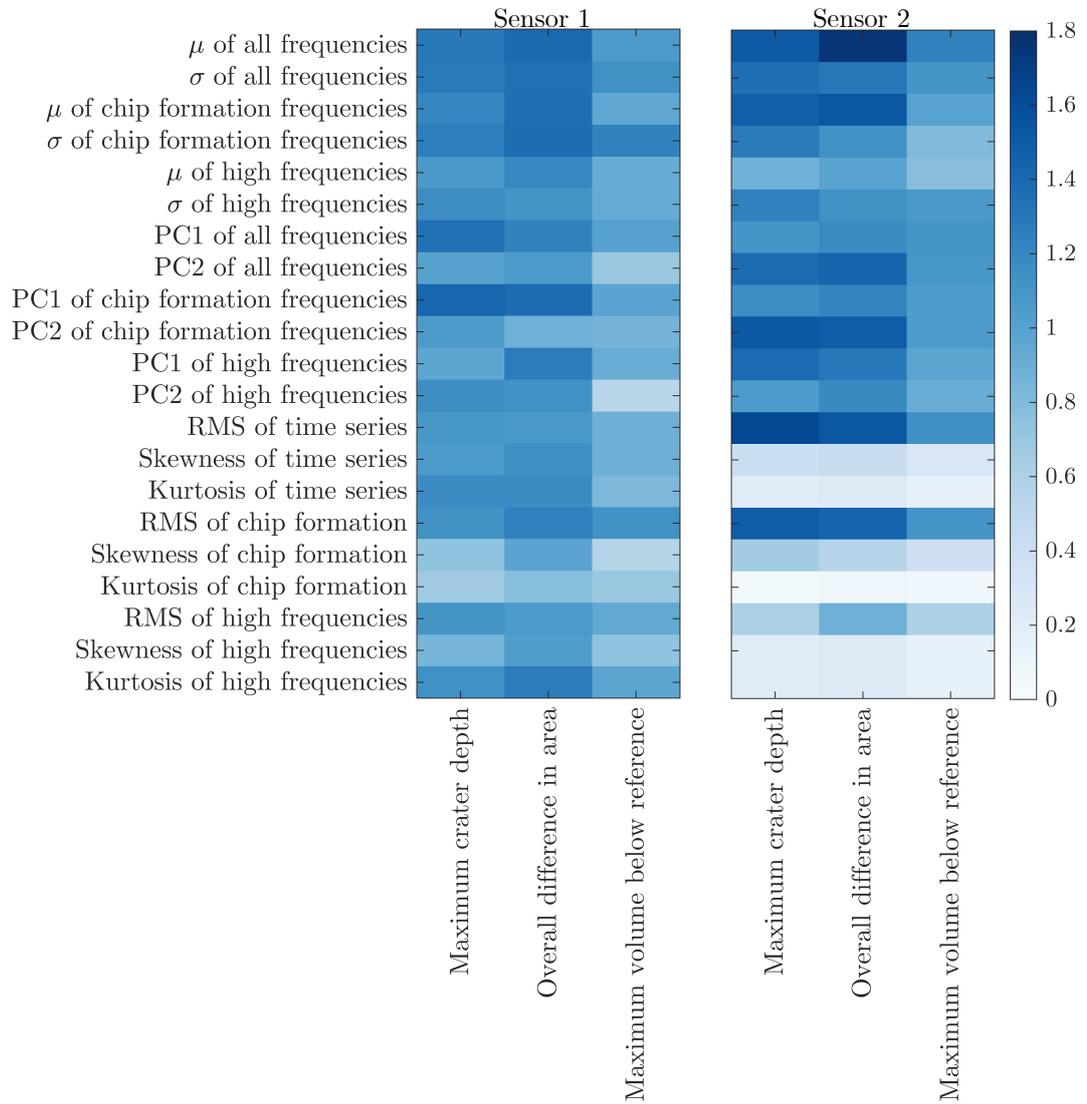


Figure 6.10: Mutual information of tool features against AE features for grade A tools. Results from Sensor 1 is on the left and Sensor 2 on the right.

work, the number of observations that can be included in the algorithm are limited here. Consequently, the number of input variables that are used in the GPR is also limited. Therefore, five features that present the highest MI from tool grades A and B will be used in the GPR as inputs. Here, only the features that can be collected online are considered. The chosen features are presented in Table 6.3.

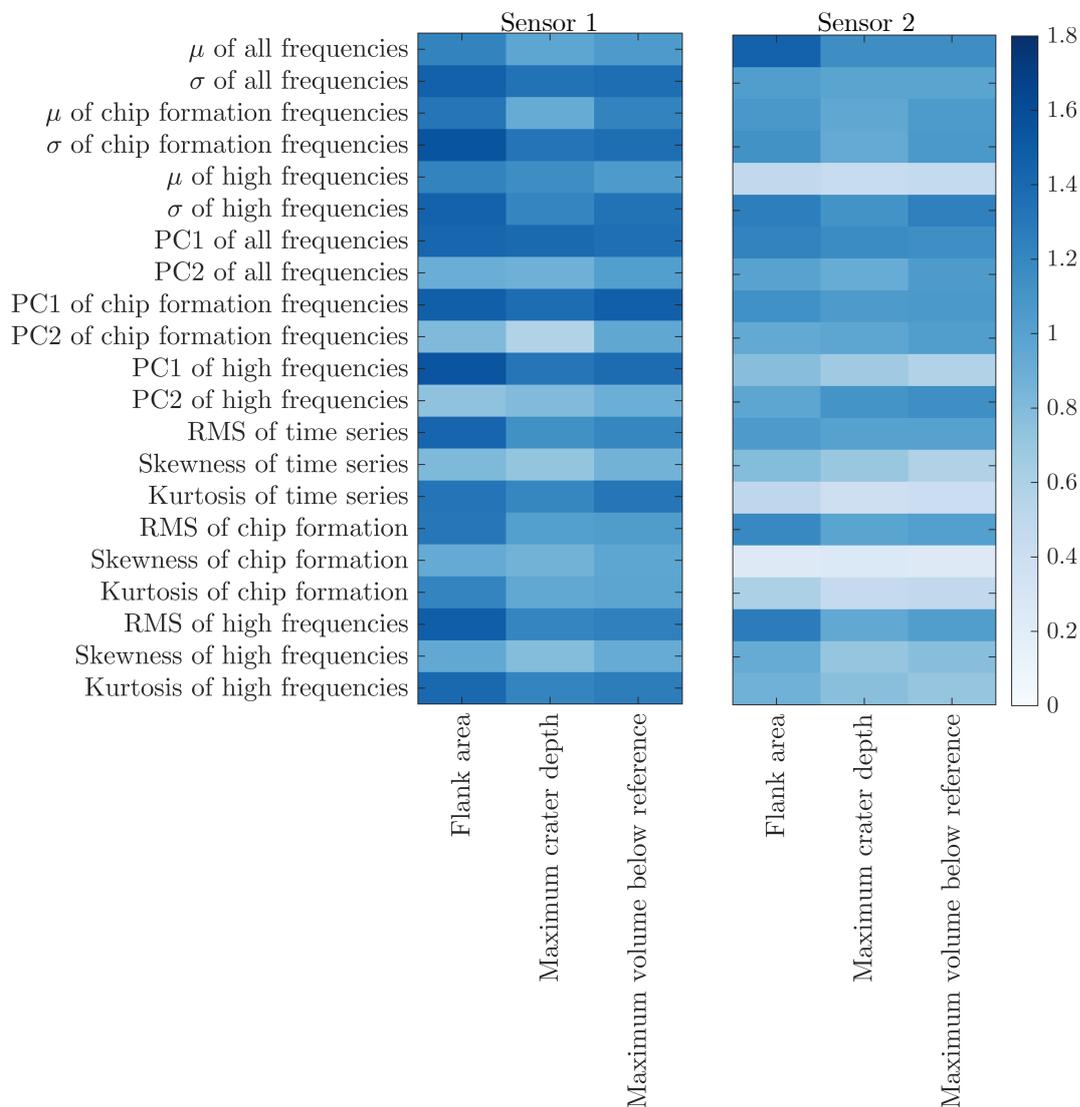


Figure 6.11: Mutual information of tool features against AE features for grade B tools. Results from Sensor 1 is on the left and Sensor 2 on the right.

### 6.4.3 Results using GPR

The feature set in table 6.3 are fed through the Gaussian process described earlier. As an example, figure 6.12 displays the GPR results of flank area for grade B tools, with the results from RSL methodology using FS3 for comparison. The true data and the mean of the estimate functions are presented. The confidence interval of

	<b>Grade A tools Data from Sensor 2</b>	<b>Grade B tools Data from Sensor 1</b>
A	RMS of time series	$\sigma$ of chip formation frequencies
B	$\mu$ of all frequencies	$\sigma$ of all frequencies
C	RMS of chip formation	$\sigma$ of high frequencies
D	$\mu$ of chip formation frequencies	RMS of high frequencies
E	$\sigma$ of all frequencies	kurtosis of high frequencies

Table 6.3: Features with highest MI chosen as inputs to the GPR model.

$\pm$  three standard deviations is also highlighted. The model was trained with the same tools as the ones used in Section 6.3.1 for easy comparison of model predictive capability.

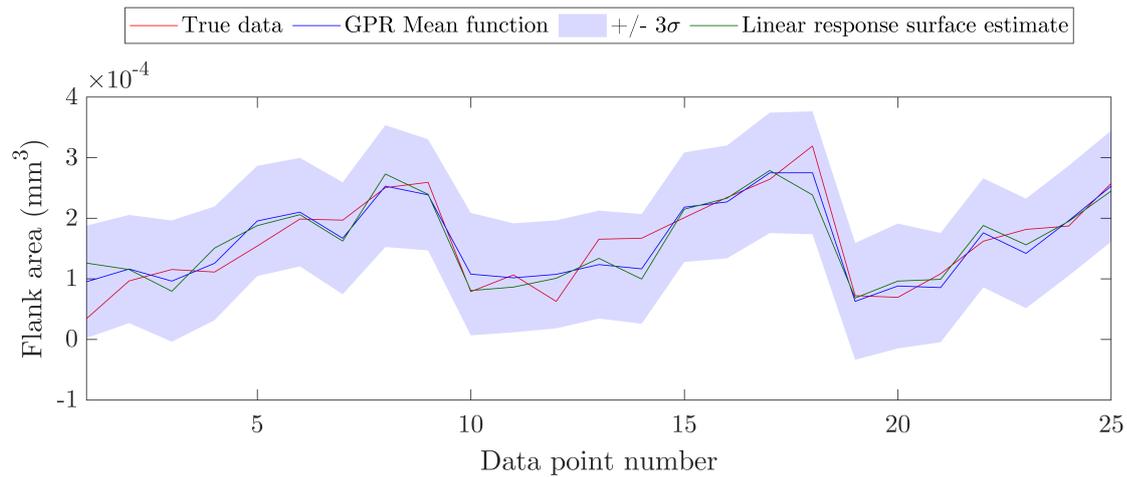
From figure 6.12, it can be that in general, the model is able to follow the trend of the true data during both the training and testing stage. The confidence intervals do not diverge significantly throughout the testing stage, suggesting that the testing data are similar in behaviour to the training data.

The MSE scores achieved from using GPR for both tool grades are presented in Table 6.4. From this table, it seems that lower training and testing MSE values are achieved by the GPR for grade B tools compared to RSL methodology with FS3 (online) features (table 6.2). As the GPR is not bound to modelling linear trends or the number of input variables, it is assumed that the increase in the number of inputs and the inclusion of nonlinear features help to improve the model prediction.

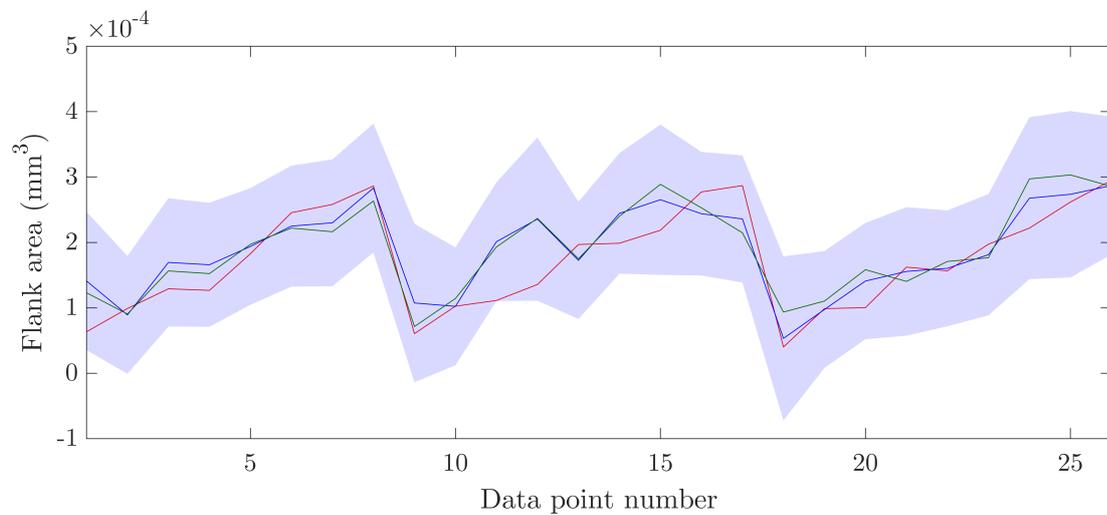
Another notable observation from Table 6.4 is that the GPR has achieved much higher testing MSE values for grade A tools. It seems that the model is not generalising well from training to testing when considering the MSE scores. The low MSE scores could be due to several reasons. Firstly, as the model trains well on the training data and does not perform well on testing, an assumption can be made that the datasets are dissimilar. Figure 6.13 presents the input AE features against the tool wear target features for grade A and B tools. There seems to be a larger discrepancy between the training and testing data for grade A tools compared to grade B. As the

GPR is conditioning on data it sees in the training phase, if the testing data looks different, the model cannot predict that trend well.

Another thing to note is that the dataset collected from tool grade A is smaller than



(a) Training stage.  $\text{MSE}_{\text{test}} = 14.20$



(b) Testing stage.  $\text{MSE}_{\text{test}} = 27.70$

Figure 6.12: An example GP regression result predicting flank area using the online AE features (FS3) for grade B tools. (a) Training stage where tools B3-B5 were used to train the algorithm. (b) Testing on tools B6-B8.

grade B. This highlights the issue of noise; due to the size of the dataset (i.e. low number of observations), it is challenging to know for certain if the data is noise or if the data describes the underlying trend. As a result, the GPR is not able to make the distinction, leading to a mean function that may not be similar to the target values. Consequently, it is assumed that the averaging of input features may affect the model's predictive capability and is a significant limiting factor when applying supervised learning methods. One solution here would be to interrupt the process more often to collect wear labels, although this would be costly and introduce depth of cut uncertainty to the measurements more often. This is one of the main drawbacks of using supervised learning in a process with high costs associated with data collection; without a large number of input data that can clearly define the underlying function, it is difficult to use an algorithm that can confidently suggest a tool wear value at any given time. The question then arises, why does the RSL methodology perform better for grade A tools? It is possibly because, unlike GPR,

		Flank area	Maximum crater depth	Overall difference in area	Max. volume below reference surface
Tool grade A features	Train	-	6.15	6.51	0.0022
	Test	-	67.77	76.89	59.11
Tool grade B features	Train	14.30	11.58	-	16.65
	Test	27.70	36.40	-	26.40

Table 6.4: The percentage error (MSE) using GPR.

the response surface model is restricted to linear trends and is fed only linear inputs. This means that the model form is already embedded in the algorithm, resulting in lower MSE scores.

It is also important to evaluate whether MSE values are a suitable metric for checking if the model prediction is good or not. To help with this discussion, the reader is referred to figures 6.14a and 6.14b where the GP results for grade A tools predicting

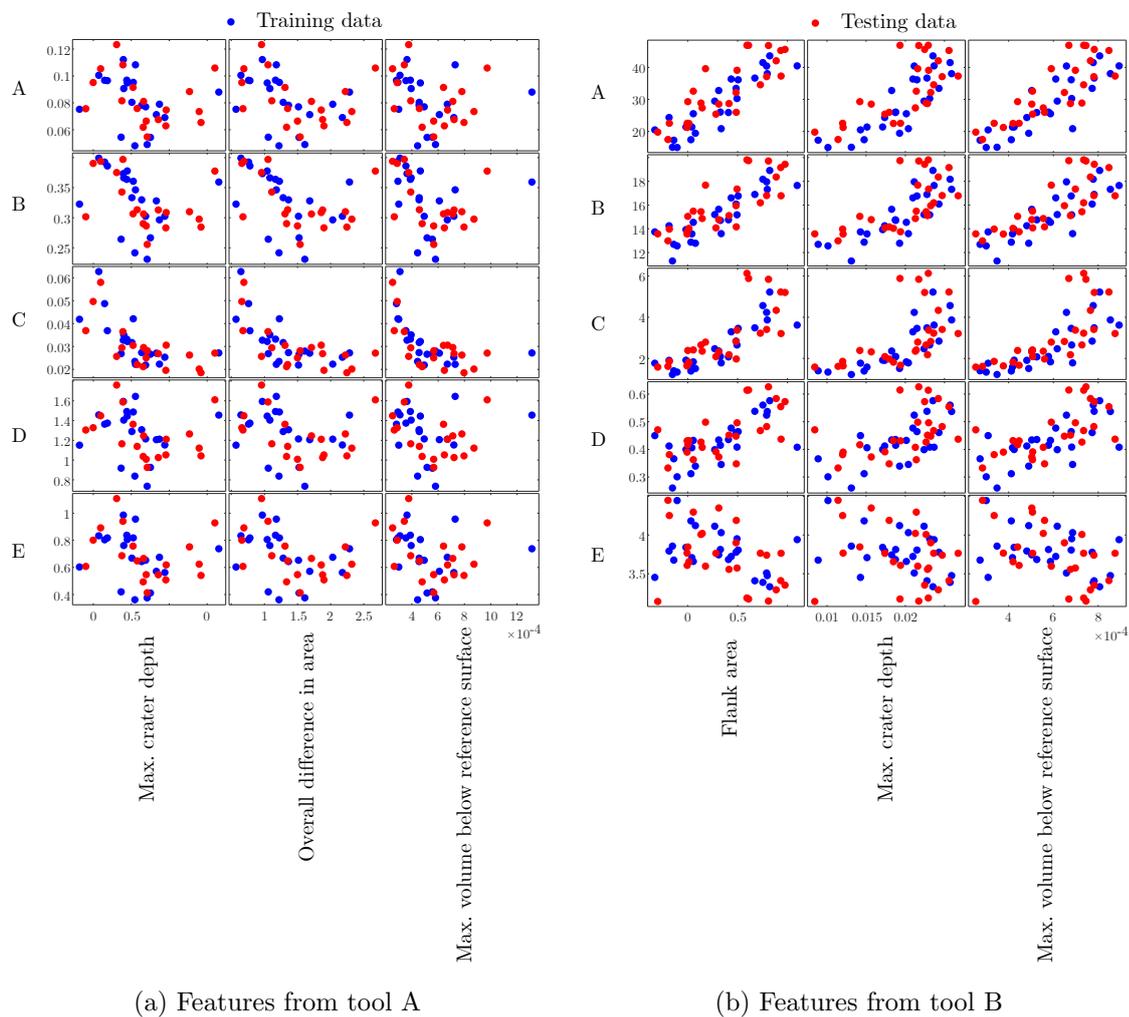


Figure 6.13: Input features (refer to Table 6.3 for feature names A to E) against the target wear features used in the GPR with training and testing data highlighted for (a) Grade A tools and (b) Grade B tools.

overall difference in area between the worn and the new tool profiles are presented. From figure 6.14b it can be seen that even though the mean function of the model does not follow the true data well, the true data is within the  $\pm 3\sigma$  confidence bounds for almost all the values. This suggests that the model would be able to confidently find the probability of the prediction being within the range of the target, suggesting MSE scores are not a good indicator of accuracy here. However, due to the large variances, the confidence bounds are wide relative to the progression of the tool wear. It is therefore clear that the width of the confidence bounds are an important parameter in this problem; a small incremental increase in the tool wear value within the confidence bounds could lead to the retirement of that tool for example, if the confidence bounds are wider than the entire life of the tool.

When using RSL methodology on the tool grade B dataset, it is suggested that online AE features are used to predict the novel tool wear feature, flank area. This is due to a number of reasons. Firstly, flank area has low normalised MSE values during the testing stage for both RSL and GPR models, making it the target that can be modelled most accurately and consistently across all scenarios. Secondly, there is no difference in performance between the online AE features and the highest ranked AE features for flank area as the MSE values are both 28% for the test case, suggesting that the RSL methodology can be implemented entirely online.

It is not possible to suggest a suitable solution for predicting wear on the grade A tools using a supervised learning method as the dataset size is not adequate for this work. This signifies a larger problem where due to the costs associated with data collection, the industry is reluctant to collect data for a supervised method that may not provide results. However, from a research perspective, it is challenging to provide a proof of concept without collecting data for analysis. It is therefore beneficial to possibly discover avenues where wear labels are not required for prediction. This idea is explored further in Chapter 7.

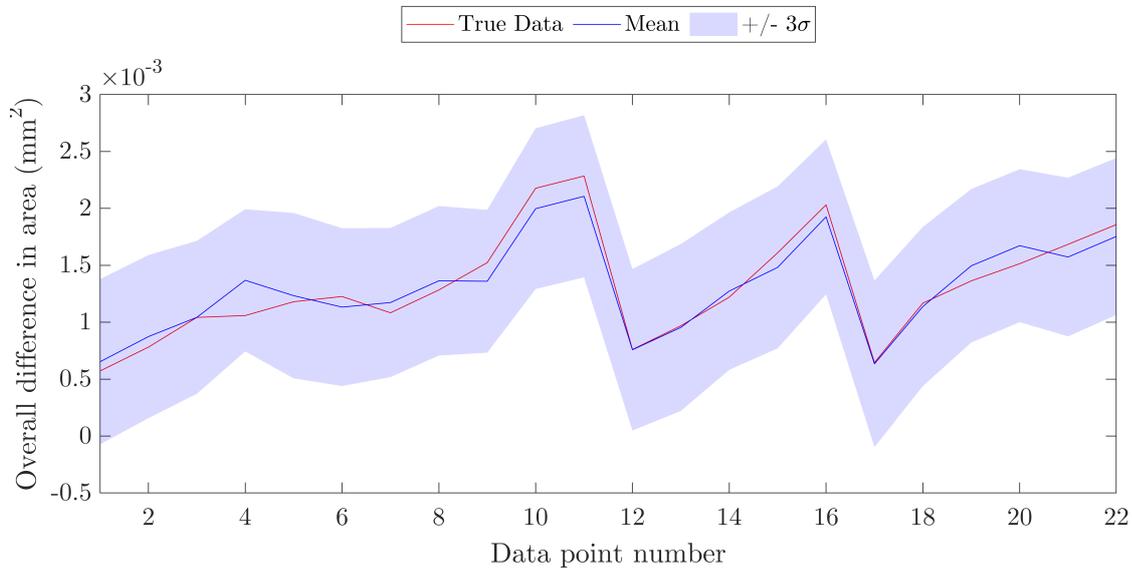
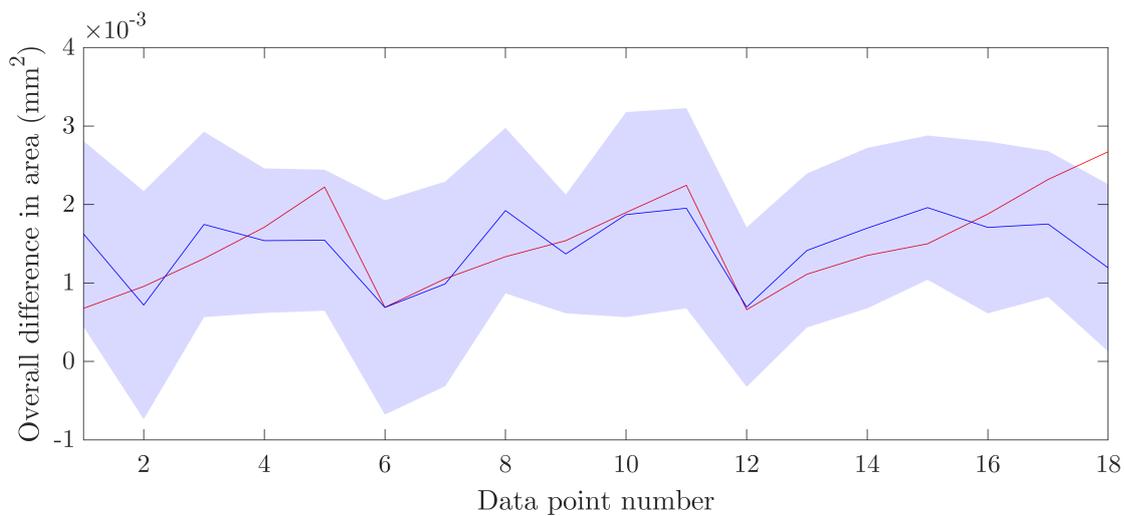
(a) Training stage.  $\text{MSE}_{\text{test}} = 6.51$ (b) Testing stage.  $\text{MSE}_{\text{test}} = 76.89$ 

Figure 6.14: An example GP regression result predicting overall difference in area using online features with high MI. (a) Training stage where tools A3-A6 were used to train the algorithm. (b) Testing on tools A7-A9.

## 6.5 Predicting tool wear on a dissimilar tool

At this point, it may be interesting to investigate whether the RSL or the GPR model is robust for predicting wear of dissimilarly behaving tools; these are tools made with

the same material composition but have worn differently to others under the same cutting conditions. This is not an uncommon trait during tool wear as occasionally, instead of failing under load, a tool's sharp edge formed at the intersection of crater and flank wear tend to smooth over and retain its strength. In this event, the tool's life extends beyond others. Within the dataset collected for this thesis, tool B2 exhibit this behaviour.

The results from applying the RSL methodology for prediction is presented in figures 6.15a to 6.15c. As before, three sets of input features were examined. To aid with comparison with previous results, the same data set was used for training the current model (tools B3-B5). From visual inspection of the plots, it is immediately clear that for tool B2 (points 51-61), the model estimates do not follow the general trend of the measured targets in any of the three cases; the estimated values are significantly below the target values in almost all cases. In reality, this prediction would lead to machining with a broken tool as in an automated system relying on this model, the tool would not be retired in time.

A normalised MSE score of above 100% means that the model is performing worse than if just the mean of the test data were used as the estimate. All of the GPR test results are approaching this score, as seen in table 6.5.

For GPR, as the distributions from which the data points have been sampled are probably not similar, it is possible that the covariance function is not representative of the difference, reducing the model's prediction capability. The model is unable to capture the underlying trend as the tool behaves so differently to others, that GPR with the hyperparameters and kernels chosen for this work cannot be used to predict dissimilar tools.

The coefficients learnt from the training data do not produce useful results when used for prediction in the RSL methodology. The significant difference in behaviour between the tools in the training and testing set is assumed to be the cause of this.

It seems that training under all operational conditions is paramount to performance when predicting tool wear with the models discussed in this work. Due to the material compositions of tools and the harsh environments they are used in, tools

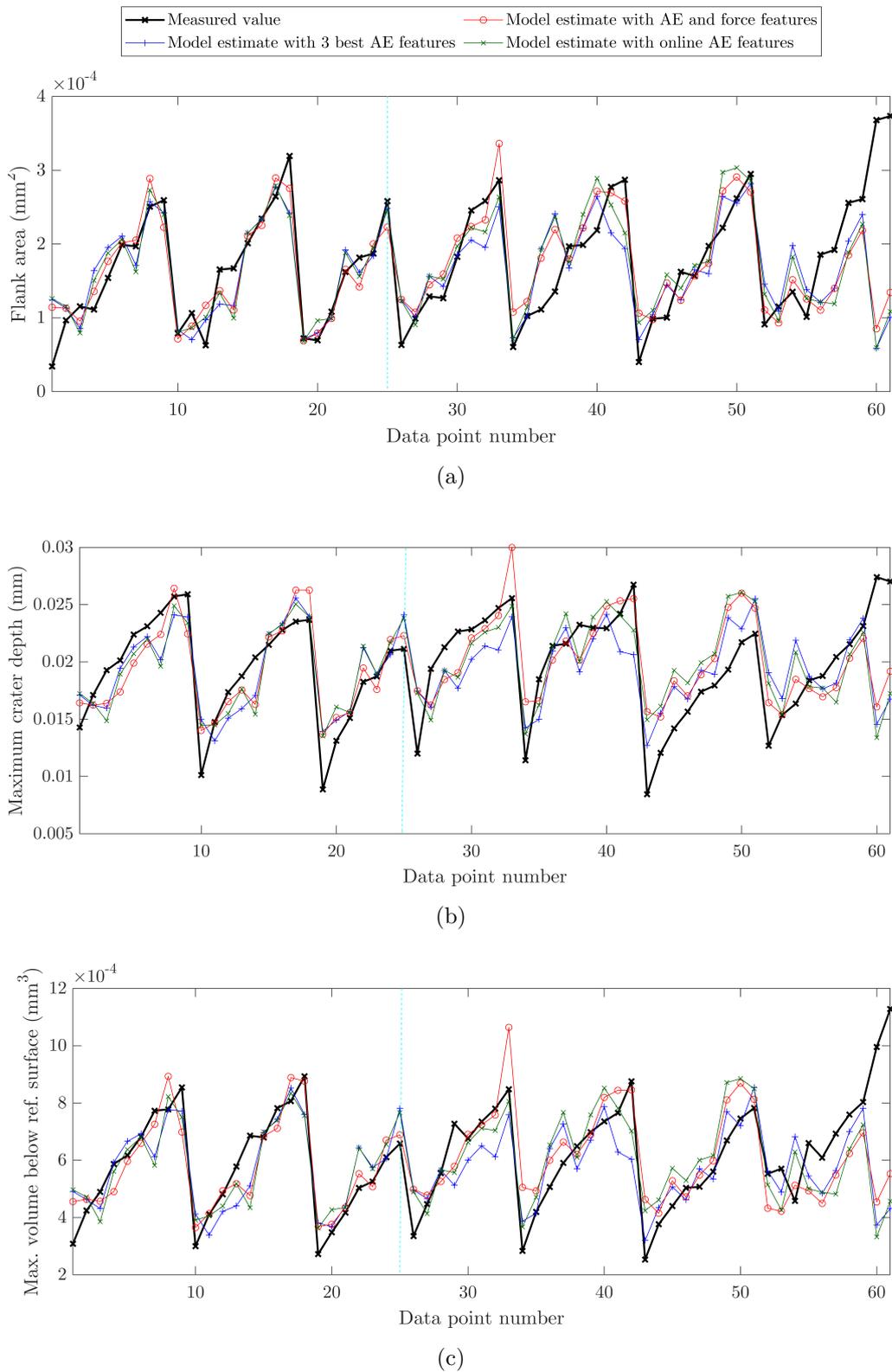


Figure 6.15: Predicting the tool wear features for grade B tools using RSL methodology where tools B3-B5 were used to train the model (points 1-25), and tested on tools B6 to B8 and B2 (points 26 - 61).

		Flank area	Maximum crater depth	Max. volume below reference surface
Training	RSL	22.19	31.32	30.22
	GPR	14.29	12.14	16.27
Testing	RSL	66.24	61.85	74.53
	GPR	86.90	82.57	100.62

Table 6.5: Comparing the percentage error (MSE) for RSL and GPR models predicting tool wear of dissimilar tools.

can fail prematurely. For the purposes of modelling dissimilar tools, a technique that does not require a bank of training data may be more suitable.

## 6.6 Concluding remarks on supervised learning of tool wear

The RSL methodology was adopted in this chapter to predict the chosen tool wear features from Chapter 4 using AE features extracted from Chapter 5. The aim was to find a model that is capable of predicting the state of the tool at any given time, so that a machine operator can make judgements and decisions on whether to continue machining with the current tool or to discard it.

Though in the past numerous researchers have attempted this task, few predicted damage on previously unseen tools. With the help of RSLs, it was possible in this

chapter to model tool wear features that capture the area and volume loss of the tool with MSE values of around 20-30%. The results were improved by GPR, possibly owing to the flexibility of the model. However, this chapter also identified many shortcomings associated with supervised learning, the most prominent being the costs associated with data collection in order to train and test models for accurate predictions.

The feasibility of modelling an identical tool that has worn differently to others in the dataset was also tested. Ideally, any model used in a critical process such as this one should be able to handle datasets that are not completely similar to one another as random effects can influence the progression of tool wear. In this case, both RSL and GPR struggled to model the tool with most MSE values above 60% for RSL and 80% for GPR. The assumption here is that the dissimilarities between the training and testing data are causing these inaccurate predictions.

There are numerous other avenues to be explored here. For example, the tool features used in this chapter are entirely from the middle of the wear scar. There may exist stronger relationships between the tool wear and signal features from the leading and trailing edge of the tool that is yet to be explored. Furthermore, numerous other features were found in Chapter 4 that were not used as the targets here. These may be better suited for supervised learning and could be investigated in future work. Manipulating the signal features to obtain better correlated ones is also an option. Furthermore, different kernels and kernel combinations could be tested.

One of the main drawbacks of this method of supervised learning is the significant reduction in the size of the signal feature set. Given the challenges associated with damage label measurement and extraction, it would be ideal if a method exist that does not require the averaging of the input data or interpolating the targets. However, to know the wear state at all times, the targets must be included in the model. On the other hand, knowledge of the tool wear state is arguably only important near the end of tool life. Therefore, a method capable of predicting imminent tool failure may be more useful to an operator than a method that provides the state of the tool. The following chapter explores this idea by applying a unsupervised learning technique to the machining data.

---

## UNSUPERVISED LEARNING

In the previous chapter, the feasibility of using supervised learning to predict tool wear from AE data was examined. Though promising results were found, it is clear that the feature and model selection process and the wear progression of each tool affect the model's prediction capability quite significantly. For this reason, it would be preferable to apply a model that does not depend on making a comparison with the behaviour of other similar tools; an assumption is made when supervised learning is conducted where all the tools are expected to wear in a similar manner to those in the training set. The wear progression of tools can also vary depending on inclusions and microstructure; properties that cannot be easily measured or predicted accurately with supervised methods due to their random nature. In addition, the collection of damage labels is time consuming and expensive, and in supervised learning in this situation, the input dataset must be reduced in size significantly to match the damage labels. This results in potentially a large loss of information.

Given the aforementioned challenges of collecting wear labels, and the difficulty associated with applying supervised learning methods, an important question should be asked. Is it desirable to understand the exact state of the tool at a given time, or would it be more beneficial to have an indication of when a tool is starting to behave differently as time progresses? Knowing the state of the tool is useful when the study

in question is focused on preventing tool failure by changing machining parameters, usually in research environments. Alternatively, if the only concern of the operator is knowing when a tool may start to fail, as it is in established machining processes in industry, it may not be worthwhile collecting extensive wear labels that interrupt the machining process and increase test times.

An online unsupervised learning method is proposed in this chapter to tackle the latter scenario. The model allows the use of incomplete label sets and alleviates the need for pre-labelled training data, therefore reducing the costs associated with data collection. Here, Dirichlet process mixture models (DPMM) are applied to the dataset collected in Chapter 3. The DPMMs allow data clustering as they are collected, without the need to set the number of possible clusters *a-priori*. The Dirichlet process (DP) permits the covariance function to change with the input data [127], resulting in a model that can handle dissimilar datasets compared to RSL and GPR. Also, as the DP treats the data as a mixture and does not require the input data to be correlated with any outputs, larger datasets can be handled by the model, meaning the AE dataset does not need to be averaged heavily as with the previous chapter. This will enable operators to visualise when the characteristics of the cutting process change during machining, avoiding the need for exhaustive measures for tracking tool wear, or the early disposal of tools, depending on the context.

DPMMs can detect changes in data and create new clusters as new data arrive. Unlike outlier analysis, this method can lead to observations of behavioural changes within the data. Therefore this method is possibly more suitable for the dataset used in this work as confounding influences must be taken into account. DPMMs cluster data online and does not require a training stage. Moreover, DPMMs do not need damage labels for the clustering process. It is possible to add damage labels to the clustering process making the entire approach semi-supervised. By doing so, false positives may be reduced and multiple normal states may be learned by the algorithm. The changes in data collected during the machining process (i.e. acoustic emissions) detected by the DPMMs could alert the operator of when the tool condition is deteriorating, prompting an investigation.

DPMMs are used across numerous areas of research for clustering data. In the field of natural languages processing, the study of translation and text generation has led Vlachos et al. [128] to use DPMMs for clustering verbs. Dreyer et al. [129] applies the technique to organise and predict words. In medical research, DPMMs have been used to classify brain tissue from magnetic resonance imaging (MRI) scans [130]. Though GMMs perform effectively for well-defined tissue types, by not having to set number of clusters *a-priori* in DPMMs, the authors were able to classify abnormal brain data. In structural health monitoring (SHM), Rogers et al. [131] applied DPMMs to a three-storey building structure to successfully identify simulated damage. Furthermore, on the Z24 bridge dataset, DPMMs were able to detect damage whilst the structure was undergoing fluctuating environmental conditions.

The only previous use of DPMMs in the field of TCM was conducted by Liang et al. [132]. The authors computed the remaining useful life (RUL) of milling tools by supervised learning, using the DP. During the offline training stage, a wear model was built with force data generated by the tool alongside actual wear values, and the results were used to weight the clusters of the DPMM. The learned model performed well and estimated conservative RUL for testing data. However, this work is not fully online and assumes that the tools used for testing performs similarly to the tools used for training, which is not always the case.

In this Chapter, an introduction into DPMMs is given in Section 7.1 followed by the feature selection process in Section 7.2. Promising results are presented in Section 7.3 where clustering in the presence of confounding influences and clustering dissimilar tools within the subset are explored. Moreover, the effect of hyperparameters on the clustering process is also examined.

## 7.1 The Dirichlet Process

An introduction to Dirichlet process Gaussian mixture models (DPGMM) is given in this section where a Gaussian is used as the base distribution. A comprehensive explanation can be found in [131] and [133].

DPGMMs can be used to cluster Gaussian and non-Gaussian data. The DPGMM can be seen as an Infinite Gaussian Mixture Model (IGMM), where a mixture of Gaussian distributions are used to cluster the data. Here, the number of Gaussian distributions can tend to infinity, allowing the modelling of any non-Gaussian datasets; a cloud of data that does not look like a Gaussian can be split into an infinite number of small clusters that are, themselves, Gaussian. IGMMs can also be used to learn information about the probability of each data point belonging to each cluster. The idea here is to find clusters within the data whilst also finding the parameters (size, shape) and the labels for those clusters (cluster 1,..., cluster n). It is not possible to find the labels and the parameters of the clusters in one step, and so, Gibbs sampling is used in this work to infer the joint distribution. It is not a requirement for the DPGMM to have knowledge of the number of clusters that may exist within the data before the algorithm is applied; the Gibbs sampler is able to initiate new clusters if data already evaluated by the algorithm is sufficiently different to the current data point.

The generative model of the DPGMM is shown in equations (7.1a) to (7.1e). In Gaussian mixture models, an assumption is made where each data point  $\mathbf{x}_i$  ( $i = 1, \dots, N$ ,  $N$  is the number of data points) is sampled from a Gaussian distribution (equation 7.1a), a cluster of data with label  $c_i$ . For each Gaussian distribution, the conjugate prior is used (Normal Inverse-Wishart ( $\mathcal{NIW}$ ) distribution over the mean (equation 7.1b) and the covariance (equation 7.1c) with hyperparameters,  $\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \boldsymbol{\kappa}_0, \nu_0$ ) to achieve a closed form solution for the posterior. The data is normalised and the prior cluster has a zero mean and unit variance Gaussian in this work.

$$\mathbf{x}_i | c_i \sim \mathcal{N}(\boldsymbol{\mu}_{c_i}, \boldsymbol{\Sigma}_{c_i}) \quad (7.1a)$$

$$\boldsymbol{\mu}_{c_i} | \boldsymbol{\Sigma}_{c_i}, c_i \sim \mathcal{N}(\boldsymbol{\mu}_{c_i} | \boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}_{c_i}}{\boldsymbol{\kappa}_0}) \quad (7.1b)$$

$$\boldsymbol{\Sigma}_{c_i} | c_i \sim \mathcal{IW}(\boldsymbol{\Sigma}_{c_i} | \boldsymbol{\Sigma}_0, \nu_0) \quad (7.1c)$$

$$c_i | \boldsymbol{\pi} \sim \text{Mult}(\boldsymbol{\pi}) \quad (7.1d)$$

$$\boldsymbol{\pi} \sim \text{Dir}(\boldsymbol{\alpha}) \quad (7.1e)$$

The cluster labels are sampled from a multinomial distribution (equation 7.1d). The mixing proportion,  $\boldsymbol{\pi}$ , is the probability of data belonging to each cluster. In order to calculate these probabilities, the Dirichlet distribution is used as it is the conjugate prior to the multinomial distribution (equation 7.1e).  $\boldsymbol{\pi}$  is controlled by the strength parameter  $\boldsymbol{\alpha}$ , the DP prior.

Ultimately, the aim of this process is to find the posterior distribution over the cluster labels, from which the most likely label could be chosen. It is very difficult to find the probability of all cluster labels given all data (equation 7.2) as it is not possible to simultaneously sample each cluster to find the cluster parameters ( $\boldsymbol{\mu}_{c_i}, \Sigma_{c_i}$ ) whilst also finding the mixing proportion of all clusters.

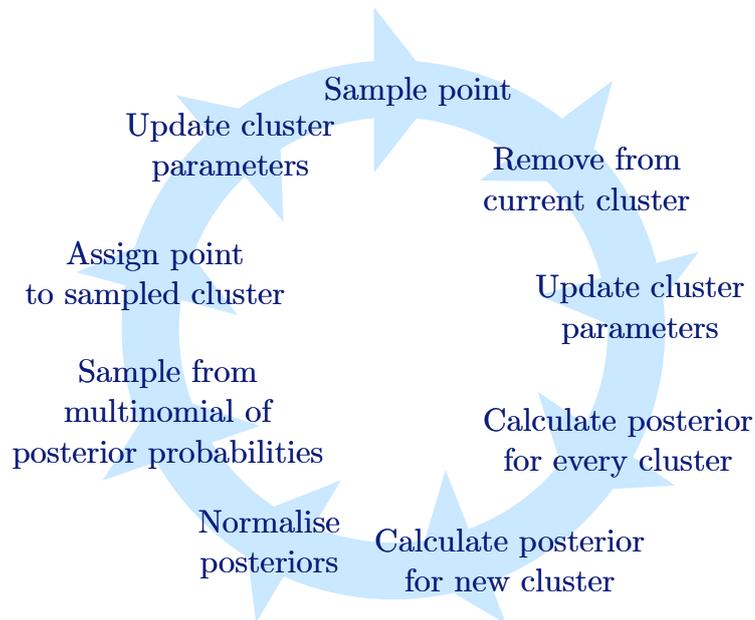


Figure 7.1: The process of Gibbs sampling. Image reconstructed from [134].

$$p(\mathbf{c} | \mathbf{X}) \quad (7.2)$$

The collapsed Gibbs sampler can be implemented to solve this. The collapsed Gibbs sampler sequentially samples new sets of cluster parameters based on samples of labels and new sets of labels based on parameters (equation 7.3). Figure 7.1 shows the steps followed by the Gibbs sampler. Each of these sampling steps can be done in closed form because of the selection of conjugate priors. In other words, it is a valid Markov chain Monte Carlo method that finds the distribution over the cluster labels and the distribution over the cluster parameters sequentially. The collapsed Gibbs sampler is performed over a window of data, the length of which is specified according to available computation power.

$$p(c_i | \mathbf{x}_i, X_{-i}, \mathbf{c}_{-i}) \quad (7.3)$$

Equation 7.3 is the first probability distribution of interest. It is the probability that point  $\mathbf{x}_i$  has the cluster label  $c_i$  given the model has seen all data  $X_{-i}$ , all other clusters  $\mathbf{c}_{-i}$  and the new data point. This posterior probability has a multinomial distribution. Here,  $-i$  means all the data points except for the data point  $i$ .

To compute this, the Gibbs sampler assigns the data randomly to clusters, and removed data points sequentially to update the cluster parameters and find the cluster best suited for that data point. The prior likelihood of drawing data from an existing cluster ( $k = 1, \dots, K$ ) can be found in equation 7.4, where  $N_{-i,k}$  is the number of data points in the current class and  $N$  is the total number of data points once the one being considered has been removed.

$$p(c_i = k | \mathbf{c}_{-i}, \alpha) = \frac{N_{-i,k}}{N + \alpha - 1} \quad (7.4)$$

The prior for a new cluster ( $k^*$ ) is calculated using equation 7.5.

$$p(c_i = k^* | \mathbf{c}_{-i}, \alpha) = \frac{\alpha}{N + \alpha - 1} \quad (7.5)$$

The posterior predictive likelihood of the point belonging to each cluster should now be calculated, which is the probability of assigning the current data point to cluster

$k$ , given the value of the data, the current cluster the data is in, all the data already in that cluster and some hyperparameters.

To calculate the likelihood of the DPGMM, equation 7.6 is used, referred to as the posterior predictive distribution. This states the likelihood that the data point was generated by the cluster defined by the posterior distribution over the parameters, where  $\mathcal{D}_k$  is all the data the algorithm has seen in a given cluster. For the DPGMM model, this is represented by a multivariate- $t$  distribution which has heavier tails than a Gaussian, facilitating smaller clusters to accept new data points.

$$p(x | \mathcal{D}_k) \tag{7.6}$$

The posterior is normalised by the marginal likelihood, i.e. the sum of all the calculated posteriors to ensure it is a valid probability distribution, and to find the multinomial distribution for the cluster label  $c_i$  of point  $i$ . If the point is assigned to a new cluster, a  $\mathcal{N}\mathcal{I}\mathcal{W}$  prior is used to initialise the cluster and the number of clusters are increased by one. This process is repeated until all the data in the window have been reassessed.

There are two main benefits of using this model. It does not require an operator to set the number of clusters (or damage states in this case) before using the algorithm, which eliminates the need for prior knowledge of the process that can be impossible to obtain. Secondly, the whole model is controlled by the hyperparameters; hence threshold tuning, and calibration is not required.

## 7.2 Feature selection for the DPMM

The success of the DP in this context is heavily reliant on the choice of input features. As the clustering algorithm is capable of identifying novelty, features that are sensitive to tool failure instead of monotonic tool wear are preferred here. Therefore, an entirely new set of signal features are required in this chapter to the ones selected in Chapter 6.

Referring back to Chapter 5, harmonics of the chip formation frequencies were identified to be stationary during the start of tool life then increased in energy prior to failure. Due to aliasing effects it is advised to operate well below Nyquist frequency, therefore, the harmonics in the vicinity of 435kHz were disregarded as input features, as the AE were sampled at 1MHz.

Ultimately, the DP would be applied in an industrial setting to cluster data automatically when running online. To achieve this, the input features should behave similarly across a homogeneous<sup>1</sup> set of tools. Prior knowledge of the process is therefore required to identify the features that increase in intensity during the end of tool life. As the process is run online, it is not possible to identify the amplitude of which frequency bins will exhibit this behaviour for a new tool. However, identifying bins that display the desired behaviour on all tools unanimously, it is possible to infer which features should be used for an unseen tool. Therefore, by studying the tools, the frequency bins associated with the harmonics of the chip formation frequencies are considered and the bins that all tools have in common are chosen as features in this work.

Unlike in Chapter 6 there is no need here to use averages of data as targets are not required in DPGMMs. However, this means that input data from multiple sources must be synchronised as otherwise the data will be misaligned. Therefore, it is not possible to use the force data as input features here because the data acquisition systems used in the experimental set up are not synchronised; it is not possible to align the AE data and force data without synchronisation. Furthermore, due to the large difference between sampling frequencies of the two DAQ systems it is also not possible to accurately synchronise the data during preprocessing. As discussed previously however, it is beneficial to automate the tool testing procedure in the absence of force measurements as attaching a dynamometer to each tool holder is cumbersome.

---

<sup>1</sup>In this case, homogeneous in the material composition and insert shape.

### 7.3 Results and discussions

The DP results from some of the tools have been presented in this section highlighting various interesting findings. A complete list of results for both grade A and B tools can be found in Appendix B.

The chosen features are fed into the DP and the hyperparameter  $\alpha$  is chosen as 20 using engineering judgement; low values of  $\alpha$  limits the number of clusters that are initiated. The effect of varying  $\alpha$  will be discussed in Section 7.3.4. The Gibbs sampler is set with a window length of 200 (200 seconds of data) to enable the convergence to the target distribution by the Markov Chain with the available computing power. A few examples of unsupervised clustering of AE energy for both grade A and B tools are presented in this section. As this is an unsupervised method and does not require training on similar data, the DP is restarted from cluster 1 for each new tool. In all cases, clusters are presented in different colours and the initiation of each cluster is demonstrated with a vertical line with the same colour as the cluster label, where the first cluster is initiated at the arrival of the first data point according to the threshold<sup>2</sup>. The clusters are formed due to change in mean and variance of the AE features, presumably caused by the onset of tool wear. Practically, these indicators could be a warning of tool deterioration to the operator; the appearance of a new cluster could be useful to trigger an inspection rather than inspecting after a set sliding distance. The cluster initiations can therefore be considered as a novel method of setting a detection threshold. The observations from tool examination can be used to decide whether the tool is damaged beyond a pre-determined tolerance threshold or not. In the latter scenario, the new cluster can be treated as an undamaged state, thus allowing future classification of similar observations to the same cluster.

For grade B tools, four AE frequency bins amplitudes are chosen as input features. These are 246kHz (feature 1), 250kHz (feature 2), 254kHz (feature 3) and 258kHz (feature 4). Grade B tool B6 is shown as an example clustering that could lead to fewer tool examinations. The clustering is presented in figure 7.2.

---

<sup>2</sup>The threshold value set here to be 1 specifies the amount of data a given cluster needs to have before it is triggered as a new cluster.

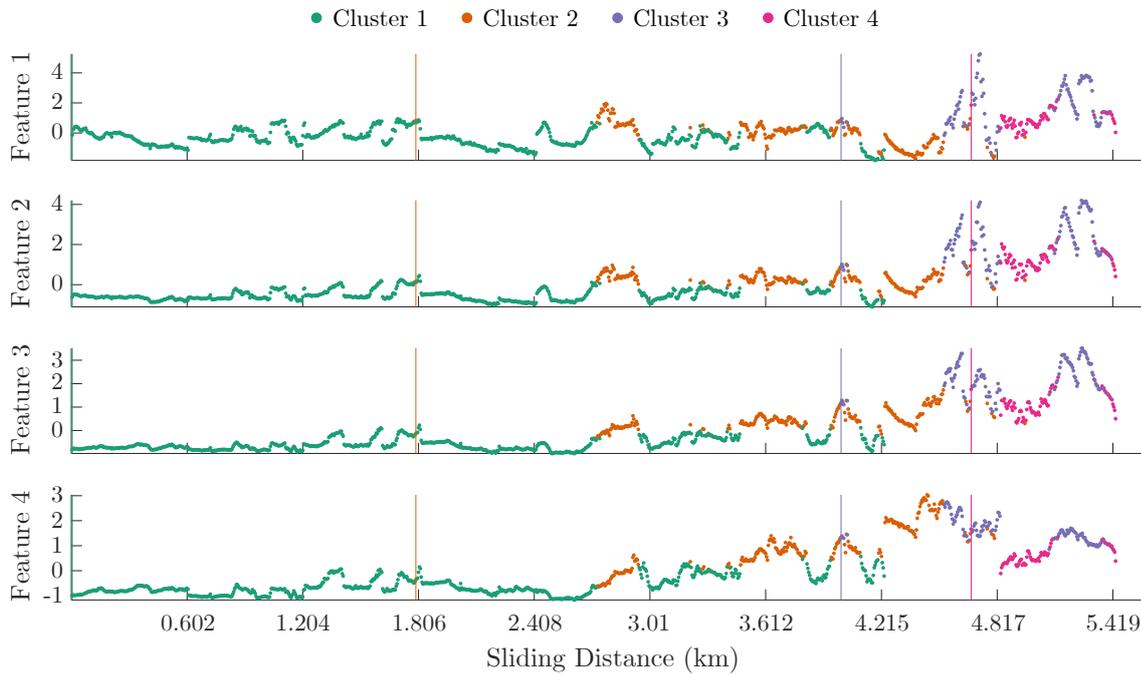


Figure 7.2: DP clustering results for tool B6 using harmonics of the chip formation frequency from AE as input features. Each cluster is displayed in a separate colour with initiation points shown with vertical lines. Four clusters are initiated here, three of which can be used as warnings to an operator.

As mentioned previously in this thesis, the tool is optically examined every four passes to study its condition, resulting in nine checks throughout tool life. On average, it takes over 8 minutes to machine and measure 4 passes. Over 42% of this time is spent on measuring the tool. Instead, if the operator is to only inspect the tool when prompted by the algorithm (3 times in this case), only around 14% of the time will be spent on measuring the tool in total. This is a substantial time saving.

The clustering presented in figure 7.2 depict promising results. The algorithm clusters almost all of the data up to sliding distance of 2.7km into cluster 1. Cluster 2 is initiated at 1.8km with one data point, possibly due to change in workpiece or reattachment of the tool, but does not sample more data until later. It would be possible to design an automated inspection system where the trigger for inspection is not initiated until a chosen amount of data is within the cluster to avoid unnecessary

inspection. However, the threshold value should be carefully considered as large threshold values could mean that the cluster initiation is less sensitive to tool damage; the cluster will not be initiated until the threshold is met, in which time, the tool could be damaged catastrophically. The large size of cluster 1 suggests that the energy of AE harmonics does not vary enough to warrant a new cluster under the current hyperparameters. Physically, this could mean that the early wear progression of the tool does not significantly affect the harmonics of AE. As the sliding distance increases beyond 2.7km, more data is assigned to cluster 2, indicating

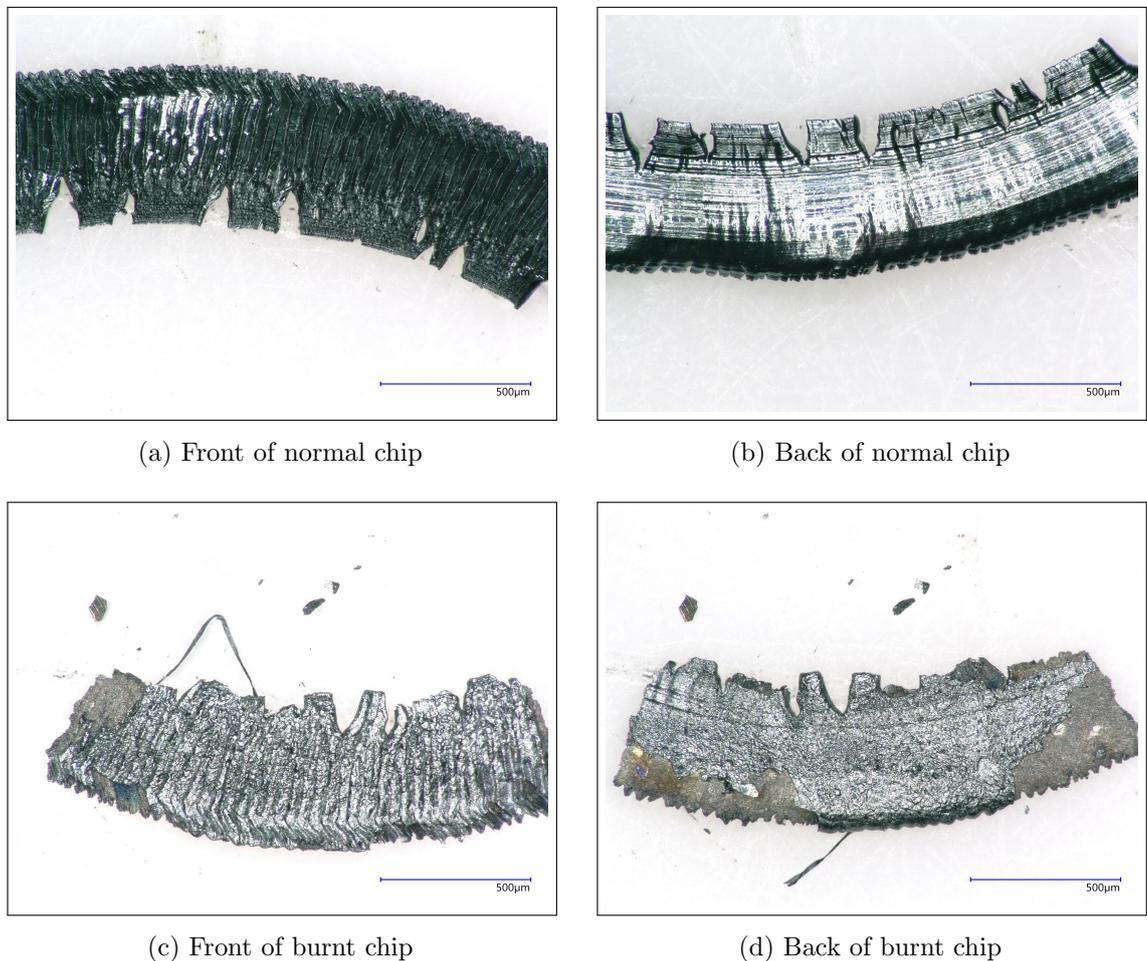


Figure 7.3: A sample of chips from tool B6 collected between sliding distances 3.612km and 4.215km.

a more permanent change of tool state, until a third cluster is initiated at around 4km.

At the initiation of cluster 4 around 4.7km, the energy of the harmonic frequencies increases rapidly, most likely due to tool wear. Due to the large forces acting on the tool, small dislocations and micro cracks in the grain boundaries within the tool material could release stress waves whilst also changing the wear profile of the tool. Consequently, the chip formation process may change permanently resulting in an increase in the energy of the AE signal. Upon inspection, it is likely that the operator retires the tool as the damage will be severe and the tool will be close to failure. Tool B6 catastrophically failed at around 5km. This means that if the tool was retired at the initiation of cluster 4, machining with a broken tool would be avoided whilst simultaneously reducing the number of tool inspections during the entire process.

It can be seen from figure 7.2 that after initiations of new clusters, the data is sometimes grouped into a previous cluster. For example, after the third cluster initiation at around 4km, some data is still allocated to clusters 1 and 2. One explanation for this behaviour could be the changing cutting edge. As the tool traverses along the workpiece, it is constantly losing material from its cutting edge. It may be possible that some of the particles ejected from the tool are bigger than others, leading to a temporary hole in the cutting edge, causing the chip formation characteristics to change momentarily. As a result, the DP clusters the data into a new class. However, once the tool traverses along the workpiece after this event, it may be possible for the tool to smooth over the hole and return to an edge similar to earlier. At this point, the AE data would behave similar to the previous cluster, resulting in the behaviours seen in the figure. Without collecting high speed videos of the process, however, it is challenging to verify this assumption.

Another reason for the aforementioned clustering may be due to increased temperature at the cutting edge. In some cases, the produced chips can wrap themselves around the workpiece and gather at the cutting edge. In this case, the temperature at the cutting edge increases rapidly, burning the chips surrounding the tool. As a result, the chip formation characteristics may alter (and consequently, the generated AE) as the workpiece and tool temperatures affects the way the materials behave.

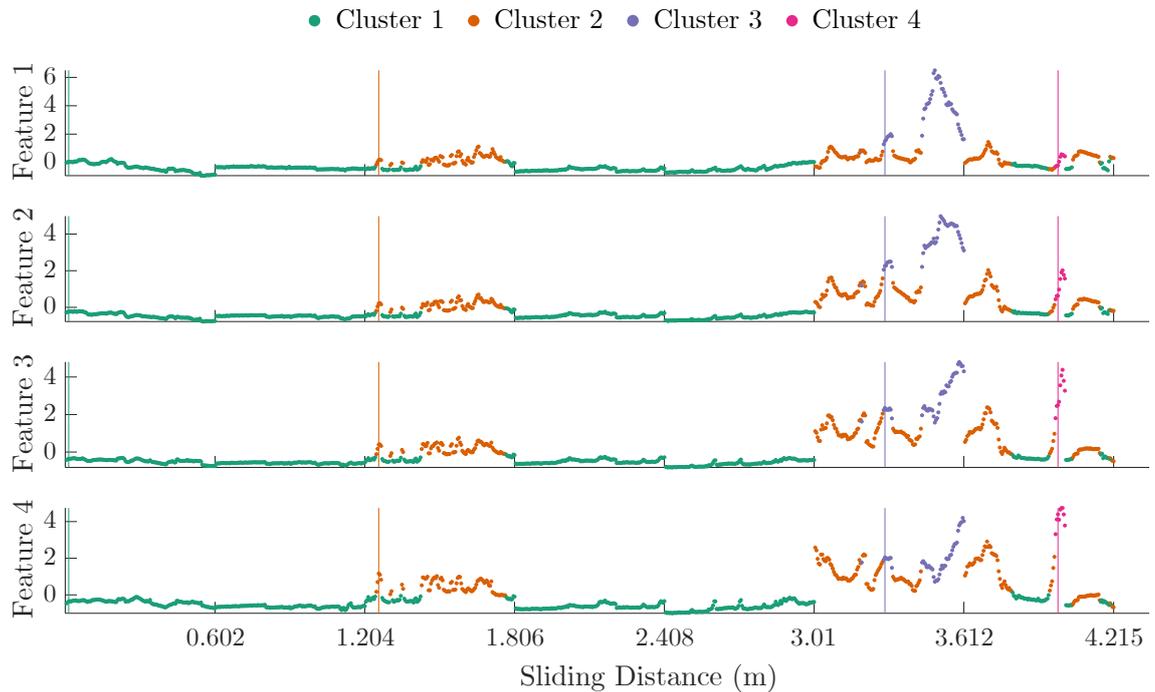


Figure 7.4: Results of tool A6. The DP is clustering AE harmonics between 250-262kHz.

Evidence of this may be seen in figures 7.3a to 7.3d showing images of chips collected between 3.612km and 4.215km where normal and burnt chips are presented. It may be possible that the data belonging to class 1 was a result of normal chips and the data belonging to cluster 2 (or 3) are due to burnt chips.

It was discussed in the previous chapter that Grade A tools did not correlate as well with tool features, making supervised learning a challenging task. Therefore, it is of much interest to investigate whether AE data from grade A tools can be clustered in an informative way to predict imminent tool failure. Here, slightly different input features were used according to the behaviour of grade A tools. These were 250kHz (feature 1), 254kHz (feature 2) and 258kHz (feature 3), 262kHz (feature 4) amplitudes. Tool A6 has been used as an example; the clustering is shown in figure 7.4.

The results are similar to the ones seen in figure 7.2 where most of the data up to around two thirds of the tool's life is clustered into clusters 1 and 2, suggesting that the AE features are not very sensitive to the increase in sliding distance before this

point, except for the data between around 1.2km to 1.8km. Then at around 3.3km a third cluster is formed which the operator can use as a warning of a change in tool behaviour, prompting a second investigation. At this point, it is possible that the tool is retired according to the tool wear measurement.

The DP results from tool A6 also exhibit similar clustering to tool B6 where data is clustered to earlier classes once a new cluster is initiated. Upon investigating the chips, evidence for burnt chips were found at sliding distances 1.806km and 3.612km. Furthermore, at 3.612km, chips that appear distinctly different were found, potentially supporting the theory of temporary edge chipping leading to changes in chip formation mechanisms and AE.

Though the return of the data to an earlier cluster may not be useful to an operator using the DP to monitor tools, this information can be very useful for research. This is especially true near the end of tool life where the data is clustered back to cluster 1 or 2 after tool failure, suggesting that the AE behave similarly to start of tool life. The reduced contact area between the tool and the workpiece may be a reason for this observation.

For tool A6, the current experimental procedure would entail seven tool measurements in total. When using DP cluster initiations as warnings for tool checks, only two checks may be required. Subsequently, the time taken to conduct an end of life test for this tool would be reduced from around 56 to 39 minutes, leading to a significant time saving of about 30%.

### 7.3.1 DP and confounding influences

DP clustering could lead to successfully avoiding failure whilst reducing the number of tool inspections in almost all of the tools that were tested in this work. However, unavoidable confounding influences such as the change in workpiece or the positioning of the tool post measurement at the specified intervals can cause variations in the AE signal that can be picked up by the DP and misjudged by the user as damage. Examples have been shown in figures 7.5a and 7.5b for tool B5 and B8 respectively,

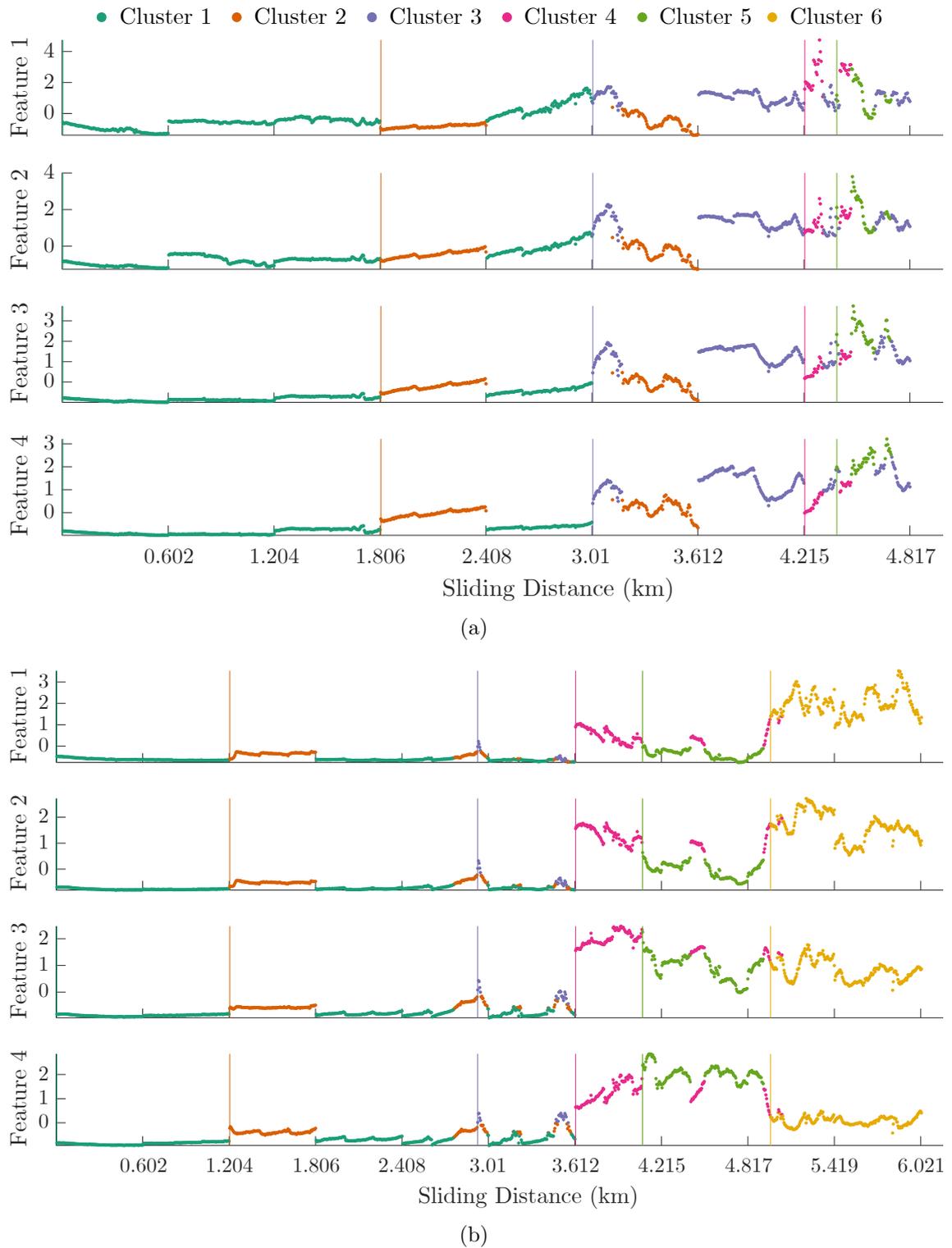


Figure 7.5: The DP clustering of AE harmonics data from (a) tool B5 and (b) tool B8. In both cases confounding influences have affected the clustering results.

where some cluster initiations occur at point of workpiece changes. The workpiece changes occur at the points where the sliding distances are indicated in the figures. Although the results for tool B5 and B8 could help save time spent on tool measurements, initiations for clusters 2, 3 and 4 are misleading if used as warnings of changes in tool behaviour as these clusters could be misinterpreted by the operator as damage; it is probable that the clusters are generated due to structural and conditional differences that present themselves as non-Gaussian clusters, prompting the algorithm to separate into smaller Gaussians. This information may not be very useful to an operator where the only interest lies in the state of the tool. It should be noted here that the cluster initiations at the start of passes may also be caused by the impact between the tool and the workpiece. In these instances, the AE signal may well be affected as the tool can be damaged. It is therefore important to differentiate between confounding influences and actual damage if cluster initiations occur at the start of passes.

When dealing with confounding influences, the first step should be to select input features that are robust against the changes in the workpieces and repositioning of the tool. Where this is not possible, the techniques suggested in Section 5.4 can be adopted. However, the methods that are suggested for dealing with confounding influences cannot be performed online without a bank of relevant training data.

In this work a relatively simple solution is suggested for dealing with confounding influences whilst also retaining sensitivity to possible tool damage at start of passes. It is proposed here to run an additional DP parallel to the original DP, where the clusters are reset at the start of each four passes of the workpiece, i.e. at the point where workpiece changes and tool positioning alterations occur. It is hoped that the resetting will essentially lead the second DP to only trigger clusters when a significant change is seen in the data during machining of each workpiece. It can therefore be assumed that the changes are not due to anything other than tool wear (and associated effects such as increased temperature and force), as all the other influences are kept constant. Two examples are presented to illustrate this method in figures 7.6 and 7.7 for tools B8 and B4.

In figures 7.6 and 7.7, the results are shown when the DP is run globally across all

tools at the top (figures 7.6a and 7.7a), and when the DP is reset every four passes of the workpiece at the bottom (figures 7.6b and 7.7b); the points at which the resetting occurs are shown in dashed lines. Here, the DPs are run with four features as before but only the first two are shown for ease of comparison. The colours of the clusters are kept the same in both cases for clarity. From now on, the clusters will be referred to as global clusters when the DP has been run across the entire tool life and local clusters when the clusters are local to each workpiece when the DP has been reset each time. In figure 7.6 where tool B8 is studied, some interesting observations can be made. Studying figure 7.6a, it is likely that the second cluster

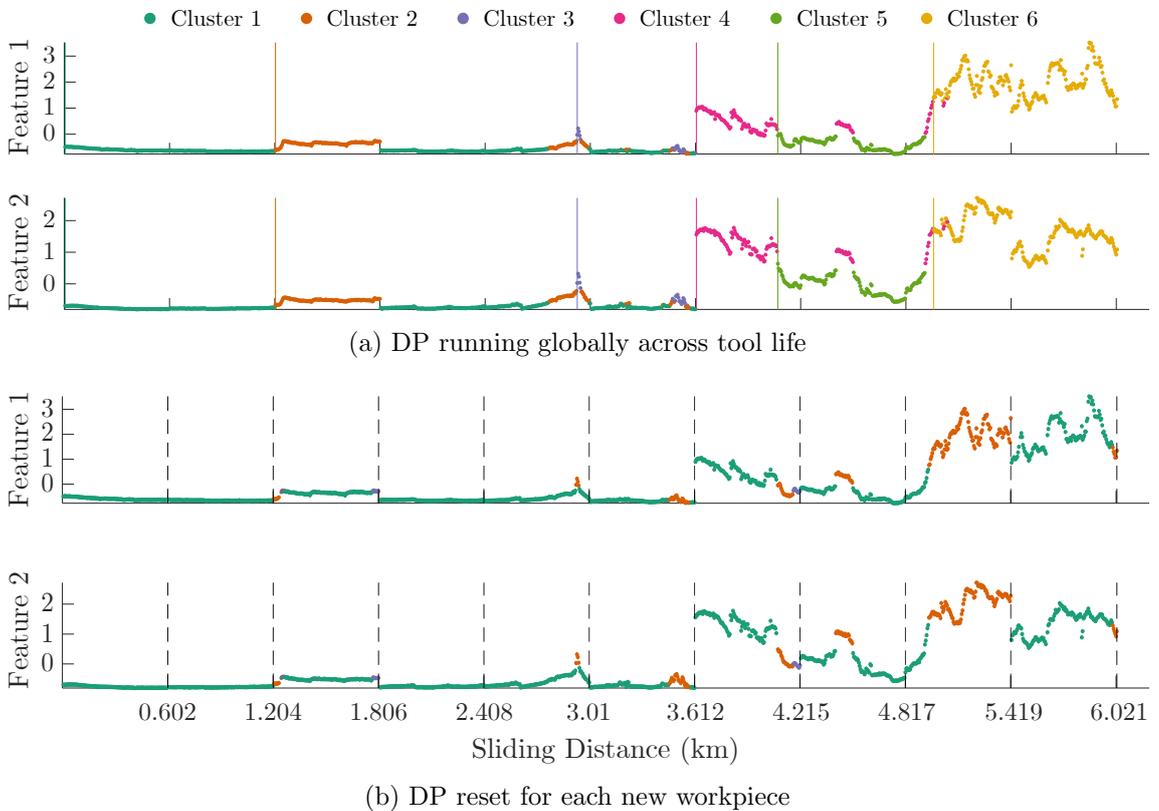


Figure 7.6: The DP clustering of AE data from tool B8 where (a) the DP is running globally and (b) the DP was reset at the point of workpiece change. The change of workpiece is indicated with the dashed black lines. Only features 1 and 2 are shown but the DP is run on all four features as before. The colour of the clusters are kept in the same order in (b) as they are in (a). Cluster initiation lines have not been shown in (b) for clarity.

has been caused by confounding influences; a change in workpiece. However, figure 7.6b shows that a 2nd cluster is also initiated in the second (local) DP at the start of the pass, suggesting that the data has been affected during the pass. This cluster initiation may be due to the tool's interaction with the workpiece as it comes into cut, where the sudden contact has affected the tool's chip formation characteristics. The initiation of cluster 4 is most likely due to confounding influences. Here, global clusters 3, 5 and 6 can be used as warnings to the operator to check for damage.

For tool B4, the results are somewhat different; in 7.7a a second cluster is not initiated

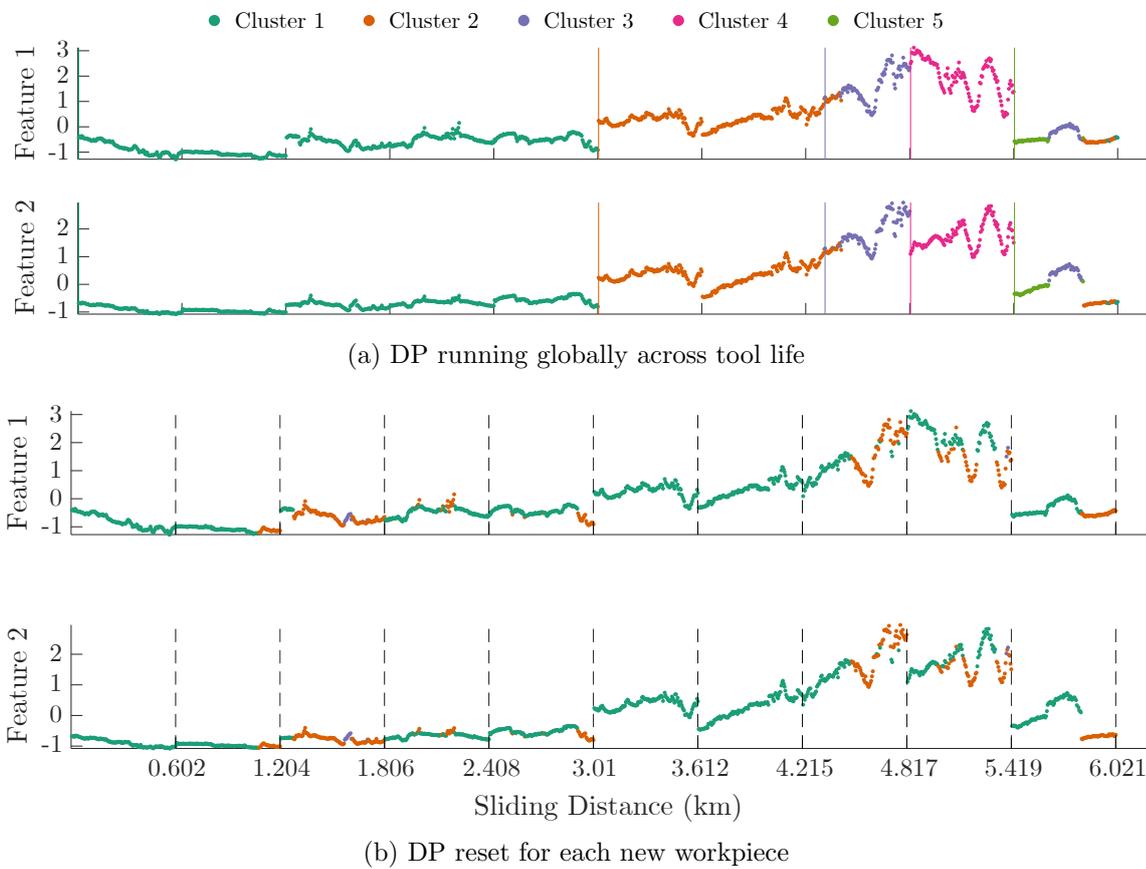


Figure 7.7: Parallel DP clustering of AE features from tool B4. (a) Global clustering. (b) Local clustering where the point of DP reset is indicated in vertical dashed lines and the cluster initiations are not displayed for clarity. The colour of the clusters are kept the same in both (a) and (b).

up till around 3km whereas figure 7.7b suggests that there may be changes in tool behaviour around 1-1.8km, indicating that globally, when considering the previous clusters, this local result is insignificant. The initiations of the global cluster may be used as a warning whereas the fourth global cluster is possibly due to confounding influences as there is no local cluster initiation aligned to it. However, at around 5.1km, a local cluster is initiated, a change that is not present in figure 7.7a. Again, as this local initiation occurs at the middle of the workpiece where a global cluster has also occurred at the start, the operator should halt machining and inspect the tool.

The application of DPs in parallel could help the operator stop machining before tool failure whilst being robust against confounding influences. Therefore it is beneficial to suggest some guidelines when using DPs in parallel in order to correctly identify when a tool may be close to failure.

- The global clustering may take priority over local clustering, unless a global cluster has initialised at the point of a workpiece change. This is because the global clustering takes the entire dataset into account and as a result, it will be sensitive to large changes in the dataset. As the workpiece change can cause new clusters, the clusters at workpiece changes may be treated with caution.
- If a global cluster has been initiated at the point of a workpiece change, and a local cluster gets initiated during that workpiece, the initiation of that local cluster may be used as a warning.
- If a global cluster has been initiated at the point of a workpiece change, but no local clusters are initiated during the workpiece, the global cluster initiation may be ignored as a warning.

As a result of using two DPs in parallel and following the guidelines stated above, it could be possible to reduce the time spent on periodically measuring the tool whilst also avoiding machining with a broken tool 5 out of 7 times for grade B tools and 6 out of 8 times for grade A tools. The complete set of results can be found in appendix B. As the risk of machining with a broken tool can be severe for the end user of these tools, in the future it is beneficial to undertake a cost analysis in order

to assess whether the risk of machining with a broken tool would outweigh the cost savings achieved by using parallel DPs in this way.

It should also be noted here that without validation it is not possible to find the link between these cluster initiations (global and local) and tool wear. As a result, currently this method can only be used as a technique for monitoring changes in the process observed through AE.

### 7.3.2 Comparing DP cluster initiations with tool wear curves

As the AE features chosen for the DP are not monotonic by nature, it may be challenging to validate this model against the wear observed on the tool. The in-depth feature selection process was conducted (for the DP) to find features that were sensitive to imminent failure; these are features that do not progress monotonically. As the DP is a clustering algorithm, it is also not suited for monotonically progressing features. Nevertheless, it is useful to investigate if significant changes in the DP results can be traced back to the tool wear curves.

Here, the DP results and some wear curves of tool A6 are investigated. The reader is referred to figure 7.4 for DP clustering of this tool and figures 7.8a to 7.8d for some of the corresponding wear curves used for this study. Here, one curve from each category of wear is presented. One of the interesting observations in figure 7.4 is that between sliding distances of 1.204km and 1.806km the data is clustered into Cluster 2. Additionally, between sliding distances 3.01km and 3.612km, the DP clusters the data into clusters 2 and 3. It is useful to explore if these changes are reflected in the wear curves. What is interesting here is that the changes seen in the DP clustering may be caused by the changes seen between the previously stated sliding distances in the cutting zone angle (figure 7.8d). The angle of the cutting zone changes rapidly as the crater and the flank wear combine to create the sharp cutting edge at the start of tool life, seen here possibly around 1.204km. Then near the end of tool life (3.612km), it is possible that the cutting edge is blunting, leading to an increased angle. However, these drastic changes are not observed in any of the other wear features, highlighting the difficulty of comparison between cluster initiations and characteristics of the tool wear curves.

However, it is assumed that the AE features reflect the chip formation process. If it was possible to link changes in chip formation to the tool wear, it may be possible to validate this model. Unfortunately, the current dataset used for this work does not contain information of the wear scar or swarf at the cluster initiation points to conclusively state whether the changes seen in the DP are due to chip formation; the reader may recall that the data was collected at the end of each four passes of the workpiece in the experimental procedure. Therefore, before use in an industrial setting another machining trial may be required to validate these results, where

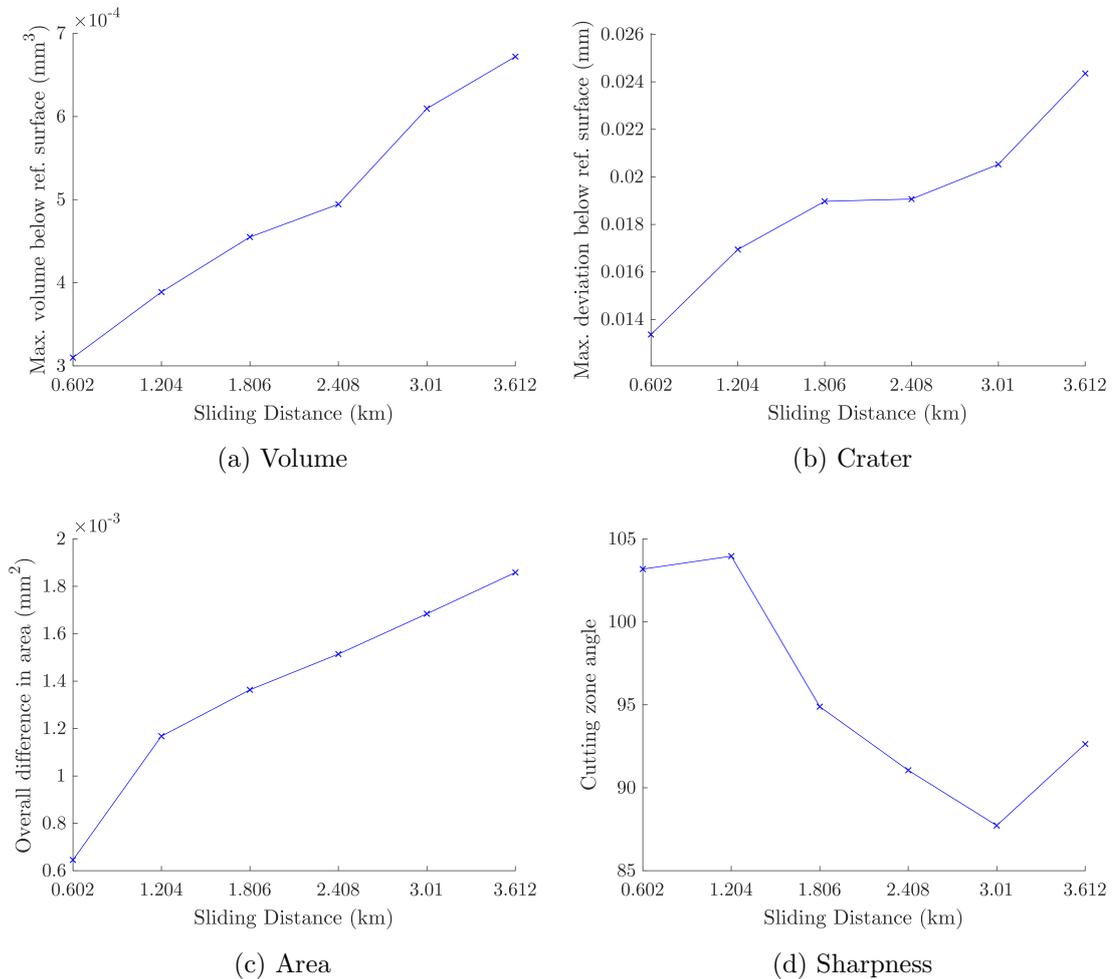


Figure 7.8: Figures showing some of the wear features for tool A6 collected as a result of feature selection discussed in Chapter 4.

machining would be halted at the cluster initiations to capture the tool wear as well as to collect swarf from that pass of the workpiece.

Given the lack of continuous wear measurements, it is difficult to directly validate that the DP is clustering at significant points. The previous discussion showed how behaviour highlighted by the DP correlates to changes in some of the tool wear observations but these are perhaps inconclusive. Without access to continuous wear measurements, another means of at least partially assessing the success of the DP is in considering the trials where anomalies or irregularities were noted. In the next section, clustering of dissimilar tools are explored to investigate this further.

### 7.3.3 Clustering of dissimilar tools

In Chapter 6, prediction of tool wear state for a tool dissimilar to others in the subset was attempted using supervised learning methods. Both RSL and GPR models failed to accurately predict wear on tool B2 as the training data could not successfully capture the underlying trend present in the dataset. As the DP does not require a training stage, the feasibility of detecting changes of AE for this tool using the DP is studied here.

Figures 7.9a and 7.9b show the results of the global and local DP results respectively. For comparison, the same features and hyperparameters are used for this tool as with all other grade B tools. The behaviour of the features and the clustering of the DP is similar to the other tools in the dataset up to around a sliding distance of 4.8km; the features are relatively ‘flat’ and increase in energy suddenly where the number of clusters formed increases as well (it is assumed that the initiations of local cluster 2 around 1km, 1.6km and 2.2km are due to increased temperature during cut or due to ejection of material from the tool. As mentioned before, these assumptions are challenging to verify without visual recordings of the tool passes. As these local initiations are not observed in the global clustering, they may be ignored by the operator as the changes in AE are not sensitive to global changes). For tool B2 however, machining is continued where the data is clustered back into earlier clusters after 4.8km. For an operator, the warnings at cluster initiations 3 and 4 may be used as before; the machining may be halted and tool may be inspected.

Upon inspection, the operator would observe the 3D models displayed in figure 7.10. The figure shows the wear on the tool at 4.215km. The highlighted red box on the right hand image of the side of the tool shows that the profile of the cutting edge is not smooth, but in fact, contains a small ridge. This is however not very clear on the image on the left hand side when the tool is viewed from the top. Therefore, sometimes, the machining is resumed with this tool following operator judgement. However, the surface roughness of the tool increases during continued machining as

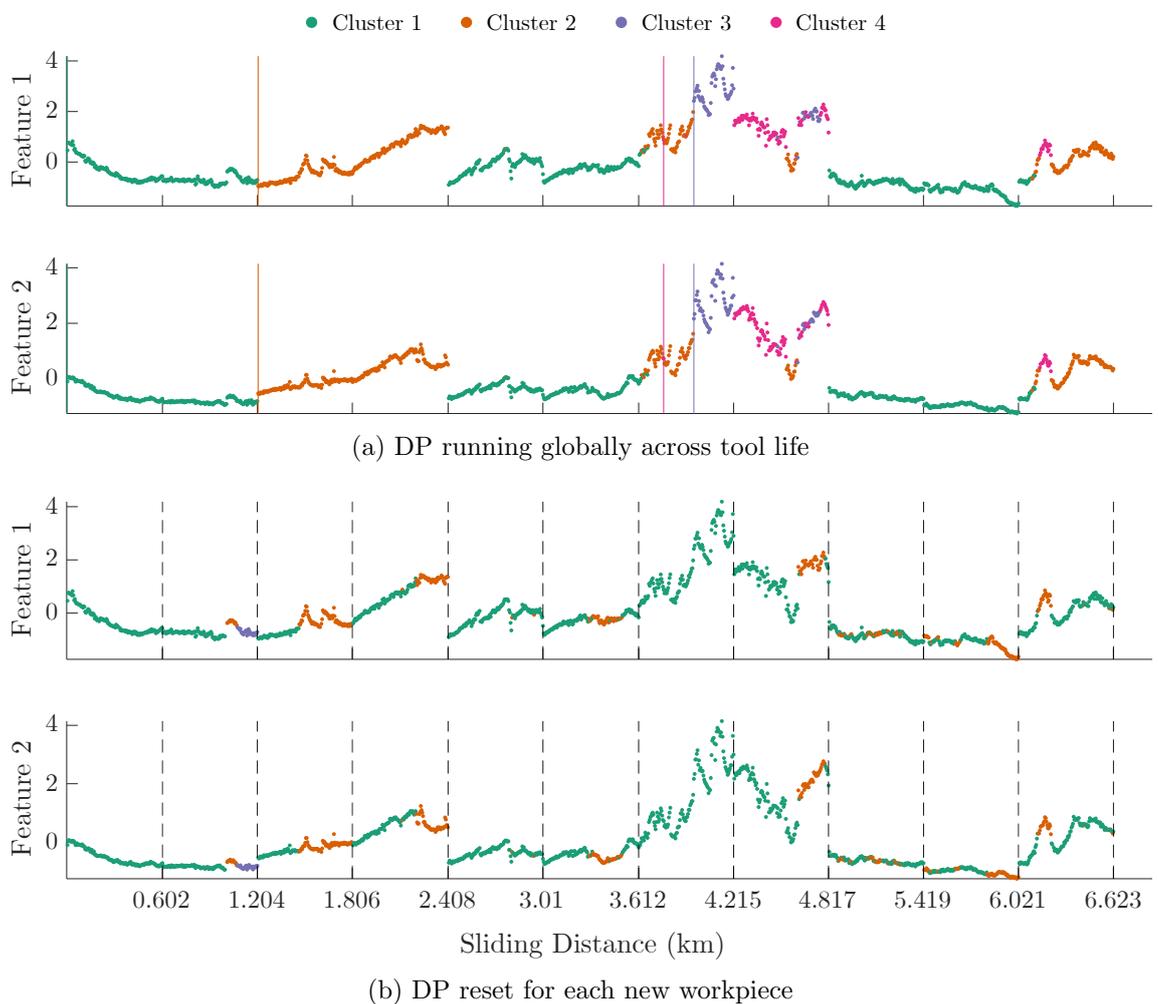


Figure 7.9: DP clustering of harmonics of AE collected from dissimilar tool B2. (a) Global DP clustering. (b) Local DP clustering where the DP is reset at every workpiece change.



Figure 7.10: 3D model of tool B2 at a sliding distance of 4.215km. Image on the left presents a view from the top of the tool and image on the right shows a side view. The arrows indicate the viewing directions for the adjacent image. The red box highlights the tool wear area of interest.

seen in figure 7.11. Therefore it is suggested that for tool B2, the operator should stop machining at the warning at the initiation of cluster 3 to avoid continuing with a chipped tool.

In this instance, the DP is able to cluster tools that are dissimilar to others in the subset. This is because it does not rely on data from other tools for training because the algorithm only considers variations within a dataset.

Another example is shown in figure 7.12 where the tool did not wear dissimilarly but the AE features behave differently to others in the dataset. In figure 7.12a

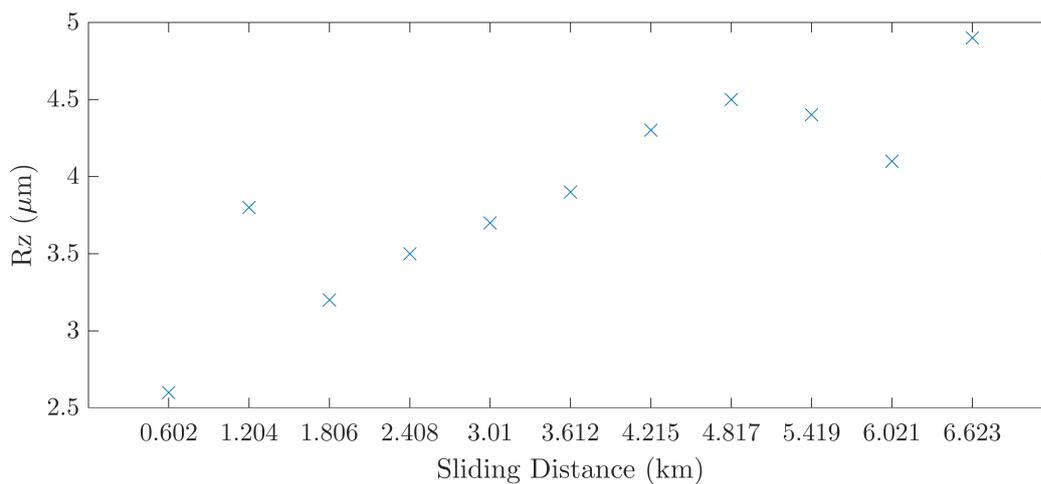


Figure 7.11: Surface roughness of tool B2 throughout tool life.

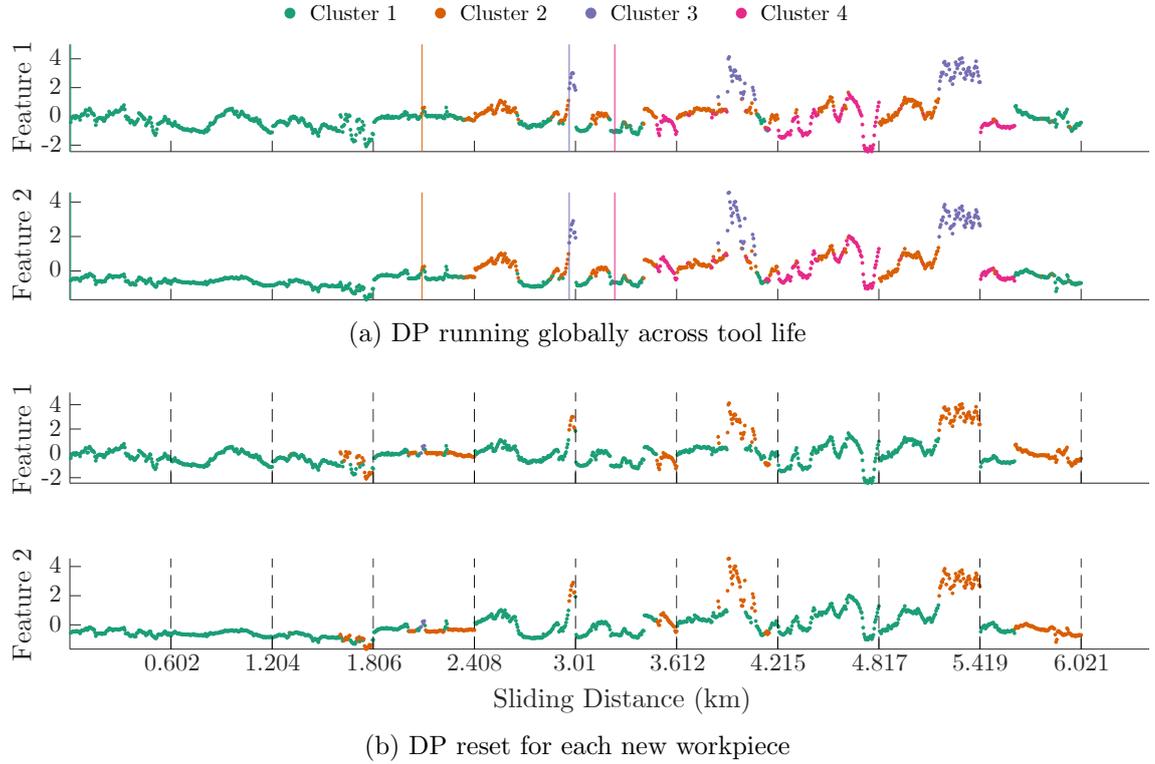


Figure 7.12: DP clustering of harmonics of AE collected from tool B3. (a) Global DP clustering. (b) Local DP clustering where the DP is reset at every workpiece change.

it can be seen that the DP detects cluster 3 at around 3km, due to an unknown behaviour causing AE energy to rise at these specified frequencies. Due to this increasing energy in the middle of tool life, the DP does not detect a new cluster around 5.2km even though the energy of the signal has increased. In figure 7.12b at this point, however, the local clustering has created a second cluster to reflect this energy increase, which could lead to avoiding machining with a broken tool. As the DP is able to cluster dissimilar data to the others within the dataset, this could be a step towards validating the model.

### 7.3.4 Effect of hyperparameters

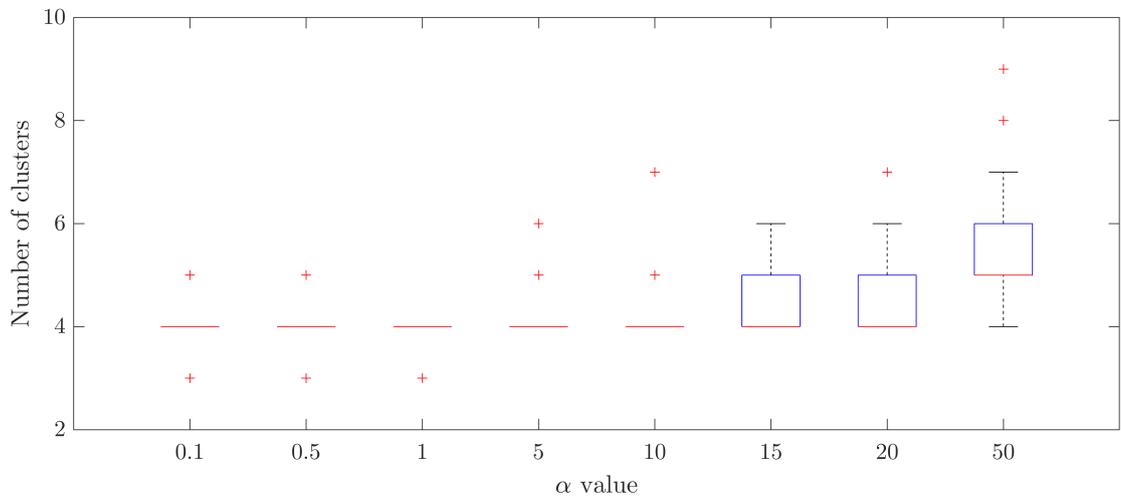
Tuning of the hyperparameter  $\alpha$  and threshold values can be used to alter the number of samples in each cluster. However, without knowing the exact nature of

the tool when new clusters are initiated, it is difficult/impossible to know which hyperparameters would be optimal. In generic SHM cases, the presence of a crack is treated as damage and cross-validation can be used with false negative rates for checking misclassification of the DP, aiding the tuning of hyperparameter,  $\alpha$ . If the clusters are well separated, then the number of clusters generated is insensitive to the  $\alpha$  parameter [131].

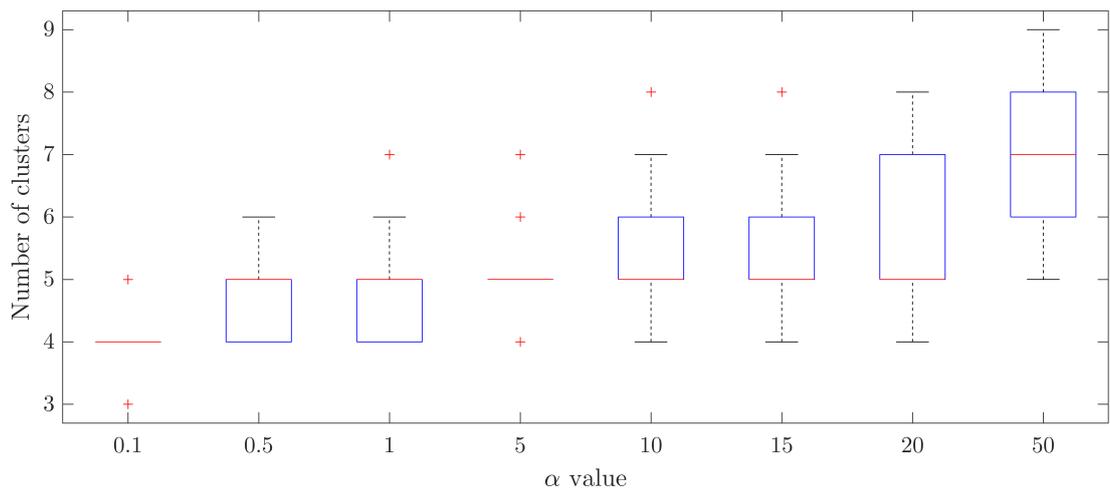
As the DP is used in this work to cluster data online, the value of  $\alpha$  cannot be learned *a priori* to find the optimum distribution. Instead,  $\alpha$  is picked through engineering judgement. To understand the effect of the  $\alpha$  parameter, the features used to obtain the result for tool B6 in figure 7.2 were fed through the algorithm for a number of  $\alpha$  values. In each case, the algorithm was repeated 100 times to obtain the average number of initiated clusters. The results have been presented using a box plot in figure 7.13. The average number of clusters across the 100 runs are shown with the red line, the outliers are shown in red crosses which stray beyond  $\pm 2.7\sigma$  (annotated by the whiskers). The top and the bottom of the boxes represent the 75th and 25th percentiles respectively.

It is clear from figure 7.13 that the number of clusters are independent of the  $\alpha$  value when  $\alpha$  is small. This means that the clusters are well separated. When considering  $\alpha$  values between 0.1 and 50, the mean number of clusters change very little. The recommendation here, therefore, is that any value of  $\alpha$  between 1 and 50 can be used without significantly affecting the formation of number of clusters.

The window length of the Gibbs sampler could also affect the results of the DP. Though it is not a hyperparameter, it is a value that can impact the number of clusters and the points of cluster initiations. This is because the window length governs the number of data points that are re-evaluated by the algorithm. Practically, it is not useful for the operator if the Gibbs sampler re-evaluates a data point from an earlier sliding distance and reallocates it to a different cluster than the one it is in at that time. However, the window length must be sufficiently large to ensure the Markov chain converges into the target distribution. However, increasing the window length can reduce the computational speed so again, engineering judgement should be used here to find a reasonable window length.



(a) Tool B6



(b) Tool A6

Figure 7.13: The number of clusters formed for a range of  $\alpha$  values for two example tools. A hundred repeats were conducted for each  $\alpha$  value.

## 7.4 Conclusions

The feasibility of applying DP to AE data in order to reduce the number of tool checks during machining and to predict imminent tool failure was explored in this chapter.

By using unsupervised cluster initiations as warnings/prompts to the operator to conduct tool state investigation, the DP may be successfully used to reduce the time taken for measuring tools and provide a point at which tools can be retired. This method of using cluster initiations as a detection threshold has not been observed in past literature.

By carefully analysing the data, it was possible to find a feature set for each tool grade that could be used for DP clustering. This type of feature has not been investigated in the past, possibly due to lack of available AE data. However, further work on this topic could be to extend the dataset to a much larger number of tools in order to find the most well suited features for the DP clustering as in some occasions, the behaviour of the features led to clustering that were not successful in avoiding tool failure.

Due to the confounding influences that affect this dataset, a combination of global and local DPs running in parallel were implemented here for the first time in literature; the global DP sampled from the entire dataset collected during the life of the tool whereas the local DP only sampled from the data collected during each workpiece. By using these parallel DPs, it could be possible to avoid machining with a broken tool in almost all the tools investigated. The models do, however, need to be validated with wear values at the points of cluster initiations. A step towards validation was taken by showing that the parallel DP can be used to cluster data of dissimilar tools in order to avoid tool failure.

A futuristic set up may entail a dedicated microscope that is capable of taking 3D images of the tool whilst still in the holder, negating the need to remove it. By doing so, it is possible to eliminate the positional change that acts as a confounding influence in this current dataset. As a result, the first of the four passes of the workpiece can be retained in the dataset, leading to a smoother transition between data from each workpiece, which may lead to reduction in the effects of confounding influences. Furthermore, a set up such as this could be programmed with tool wear tolerances and thresholds that can be used to assess the tools at cluster initiations, negating the need for intervention by an operator, leading to completely online tool wear tests.

The results from hyperparameter tuning suggested that for this dataset, a small  $\alpha$  value could be used as the clusters are well separated. This result is beneficial in the context of this work, as practically, it is disadvantageous if many clusters are formed as cluster initiations are warnings for tool checks. If a large number of clusters are initiated, then the number of tool checks would also increase, which would reduce the time savings that can be achieved by the DP.

As the DP does not rely on learning damage labels, the dataset can be any size. As larger datasets are more informative, the frequency at which the data arrives for clustering can be increased by taking averages of the AE data more often. Currently, a data point represents a second of data which comprise of 1 million raw AE data points. In the future, each second of data could be represented by a much larger number of data points, therefore producing a more informative trend or cloud of data for clustering. As more data becomes available, the DP may be able to initiate clusters earlier than currently as the clouds would be better defined, reducing the time taken for tool tests.

Here, the DP has been used as a unsupervised technique where data has been clustered without the need for target values. In the future, the DP could be used as a semi-supervised learning technique where damage labels could also be included in the model at the detection of change. By doing so, the third level of Rytter's Hierarchy could be achieved as the damage may be detected, and quantified. Interestingly in this work, the increase in AE energy (in the harmonics) leading to the final cluster initiation prior to failure usually occurs during the penultimate four passes of the workpiece. This information could be investigated further to build a system capable of prognosis; it may be possible to predict the remaining useful life of the tool by studying the nature of the final cluster initiation prior to catastrophic failure.



---

## CONCLUSIONS

The focus of this thesis was to explore tool condition monitoring techniques that may be capable of reducing or eliminating the number of time consuming tool wear measurements currently undertaken at Element Six Global Innovation Centre. Owing to the complexities of machining and the costs associated with collecting damage labels, the tool condition monitoring community has, thus far, struggled to find a solution that can be applied fully online. For the PcBN tools used this work, each of the constituent elements of the pattern recognition paradigm for condition monitoring were studied in detail. Firstly, Chapter 1 and 2 introduced the challenges in industry followed by a literature review in attempts to assess the operational evaluation. Then, Chapter 3 focused on data acquisition where the tool wear data was used in Chapter 4 in order to extract novel features to be used as damage labels. In Chapter 5, the signal data acquired from Chapter 3 were used to find the input features for feature discrimination. Chapter 6 and 7 used these features for prognosis and diagnosis respectively.

The experimental setup used to collect the data set for this work was described in Chapter 3. The main objective of the setup was to collect data to be used in tool condition monitoring algorithms later in the thesis that is representative of the real life machining process. The collected dataset was made up of AE and force data

from 17 tools across two tool grades; each tool was tested to end of life where damage labels were collected (at set intervals) using an optical and 3D scanning microscope. At these instances, samples of chips as well as surface roughness and depth of cut measurements were also collected for each tool. In its entirety, this dataset is one of the most comprehensive collected in literature to date.

In the experimental set up, two AE sensors were attached to the machine in order to investigate whether a network of sensors was required for monitoring (represented by Sensor 1), or whether one central sensor (Sensor 2) would suffice. Clearly, the set up with a central sensor would be preferred due to the lower associated costs and amount of equipment required. Upon investigating the data from Sensor 2, the conclusion was that it is likely that the signal to noise ratio was too small for any meaningful analysis. The distance the acoustic emissions have to travel whilst passing several material boundaries was assumed to be the reason for this loss of information. In the future, a sensor with a higher sensitivity rating could be used to test whether a central sensor could be used for this work. Using AE sensors with larger frequency ranges (upwards of 1MHz) in the setup may also aid better feature selection in the future; AE frequencies above 1MHz have been associated with tool wear in past literature, due to the transient nature of the bursts.

The next two chapters focused on the features that could be extracted from this dataset; Chapter 4 concentrated on tool wear, whereas Chapter 5 investigated the AE signal.

In general, flank wear width is used as the damage label in TCM of turning tools as it is relatively easy to obtain. Without having access to the cross section of the tool, information from the crater is challenging to access. Chapter 4 focused on using the 3D models of the tool to procure novel features that may be more descriptive of the wear scar compared to the average flank wear width. The work was conducted under the assumption that the new descriptive wear labels will correlate strongly to the signal features, resulting in higher prediction accuracies than if flank wear width was used. By manipulating the profiles provided by the 3D scanning software it was possible to study tool wear across the scar, as well as throughout the progression of sliding distance, for the first time. The wear scar was separated into three sections,

leading and trailing edge as well as the middle, according to the changing shape of the wear scar. In the future, it would be possible to increase the number of sections for a better localisation of damage. However, the final use of any damage labels should be taken into consideration here. For instance, there is a clear difference between the leading and trailing edge and the centre of the wear scar. As a result, it could be interesting to investigate whether the changes in shapes and contact area have an influence on the collected signals. Unfortunately, due to the size of the wear scar, it is likely that a tighter resolution would not be very informative, as this variation may not be detected in the sensor signal. The cross sections of the tool allowed the extraction of 21 features, 13 of which were novel. A pragmatic ranking criteria based on physical understanding of tool wear was used to distinguish between the highest and the lowest ranking features. The ranking criteria used here was subjective to engineering judgement and knowledge of the tool wear progression. In the absence of a single function that perfectly encapsulates the wear scar, it is challenging to rank the features objectively. However, it is interesting to investigate whether different features would have been chosen if a different ranking criteria was adopted. For example each feature could be ranked on reliability by accounting for measurement error. Additionally, the tool wear features could have been ranked on criteria such as ease of collection/ measurement which would include the time taken to measure the features as well as the skills and equipment required to carry out the process. This could be explored further in future work. The chapter concluded by choosing three features to represent the damage labels for each tool grade.

With a set of tool features finalised as damage labels, it was fitting to find features from the AE signal that could be used as inputs in a TCM algorithm. Chapter 5 evaluated the AE signal in the time and frequency domain with the help of numerous signal processing techniques. By analysing the samples of chips collected during machining, it was possible to identify what is assumed to be the chip formation frequency in the AE signal in oblique cutting with PcBN, possibly for the first time. To verify this assumption, further testing has been suggested. Due to the nature of oblique cutting, it was found that chip formation had a bandwidth of around 160kHz in this work. To visualise the AE data across the entire life of the tool, spectrograms

were used in this work. This led to the discovery of harmonics (of the chip formation frequency), that intensified close to tool failure, possibly owing to nonlinear effects. Presently, it is unclear where the exact origins of these harmonics may be. A set of controlled experiments could be conducted in order to investigate these harmonics in a laboratory setting.

By applying statistical analysis, dimensionality reduction, short-time Fourier transform and discrete wavelet transform, a set of signal features were identified. Due to the behaviour of these trends, it was assumed that some features would be better suited for diagnosis whereas others would aid with prognosis in later chapters. There are a number of other features from the AE signal that could be dissected in order to further understand AE generation in turning. For example, in the time domain, the time series could be smoothed or enveloped with signal processing techniques in order to find underlying trends in the data. A variety of parameters such as window length and the type of wavelet could be tested during signal processing to investigate whether more meaningful features could be extracted. The length of the time signal that is represented by each data point could be reduced in order to achieve a better time resolution; currently, each data point in Chapter 5 represents a second of data which is  $1 \times 10^6$  data points. However due to time constraints, only 21 features were chosen for Chapters 6 and 7.

Chapter 6 commenced with finding linear correlations between the chosen tool wear features (damage labels) from Chapter 4 and the signal features (inputs) from Chapter 5 to find the strongest existing relationships; it was hoped that these would encourage accurate predictions later in the chapter using supervised learning methods. Features encompassing standard deviation were strongly correlated to tool wear with sliding distance, owing to the increased activity and nonstationarity of the machining process. The strongly correlated features were then scrutinised against ease of collection and eligibility for online application. To test whether the inputs could be used for predicting tool wear with a simple model, RSL methodology was then utilised. It was found that flank area could be predicted with normalised mean square error of around 28% when using signal features that can be collected online. This was a promising result.

To explore whether this result could be improved via a nonlinear model, features with high mutual information were fed into a GPR algorithm. The model generalised well for grade B tools but did not predict wear accurately for grade A tools, possibly owing to the size of the dataset and variance in the data. Both methods were tested to see if they were capable of predicting wear on a dissimilar tool to the others in the feature set. As the training dataset was not diverse enough to capture the underlying trends in the dissimilar tool, both methods failed at predicting its wear. The shortcomings of both RSL methodology results and the GPR results were deemed to be a consequence of the unavoidable heavy averaging the feature sets, not capturing all operational conditions as well as features that describe wear well, in the training set. There are a number of adjustments that could be performed in attempts to improve the results of these learning methods. By using an alternate metric such as maximum values or only considering the signal values that are close to the point of a wear feature, stronger correlations may be found. Furthermore, each data point of the signal features was an average across numerous frequency bins that either represented the chip formation frequency or its harmonics. It would be interesting to observe the correlations between each frequency bin in the spectrogram, against the tool wear features, in hopes to identify specific bins that may present stronger correlations. With regards to RSL methodology, cross validation could be performed on the training set to find coefficients that may be more representative. The number of training examples could be increased to build a larger, more informative dataset that captures a wider range of tool wear trends. For this, further experiments are needed.

It would be interesting to explore whether a larger feature set would be able to predict wear on a tool made of a different material grade. A system that is capable of prediction across tool wear grades would be preferred in an industrial setting as fewer training phases would be required. To start with, further improvements to the GPR algorithm may be achieved by testing a variety of covariance functions. It is also advised to include features from dissimilar tools in the algorithm to improve its capability against environmental and operational variations.

Chapter 7 explored the idea of unsupervised learning where changes in data could

be detected with incomplete labels sets without the need for a training phase. The author introduced the DP and Gibbs sampling which was used in this work to cluster AE data. As this process may be considered as an infinite Gaussian mixture model, the number of clusters do not need to be set *a-priori*, leading to an entirely online implementation of the algorithm. The idea was to use the initiation of online DP clustering as an indicator for the operator to inspect the tool, at which point a damage label could be assigned to the cluster. At each cluster initiation, the operator should make a decision to whether continue machining or discard the tool. The harmonics of chip formation frequencies energy were used as the input features as they were found to lay relatively dormant for most of tool life and increase in intensity as a tool approaches failure. This type of feature is well suited for the DP as it clusters according to changes in mean and variance of the data. Promising initial results showed that tool measurement time could be reduced significantly by this method. However, confounding influences from unavoidable workpiece changes and tool repositioning also resulted in cluster initiations. As these clusters were not caused as a result of tool wear, a solution was needed to avoid unnecessary tool wear measurements. To address this issue, the idea of running two DPs in parallel was proposed. Here, the first clustered the data from start to end of tool life as before, meanwhile the second DP reset every time the structure of the process changed. At the instances where confounding influences take precedence, the cluster initiation was ignored unless a cluster was initiated in the second DP. This method of parallel DP was not only capable of reducing measurement time for almost all normal tools in the dataset but also performed well for dissimilar tools.

It is challenging to validate the DP model due to the lack of availability of data at the point of cluster initiations. Furthermore, the chosen features were specifically selected due to their sensitivity to imminent failure, therefore they are not monotonic. This means that the information from the current wear curves themselves is not sufficient to validate the models. Further experiments are needed to obtain information about the chips at cluster initiations in order to attempt to validate this model.

In Chapter 7, it was not possible to assess whether sensor fusion would improve the results of the unsupervised method as the dynamometer was not synchronised with

the AE sensors. In future work, all sensors in the setup should be synchronised by one clock in order to address this issue. There are numerous other adjustments that can also be made in attempt to improve the unsupervised clustering process. As this method relies heavily on the input features, first it would be interesting to increase the number of frequency bins when calculating the short-time Fourier transform. By doing so, better frequency resolution will be achieved, leading to, hopefully, features that are dormant and increase in intensity close to tool failure.

These results may be beneficial in industrial application for a number of reasons. Firstly, this method does not require training data that must encompass all operational conditions to safely predict unseen tools. Secondly, it reduces the number of regular tool wear measurements that are conducted currently to only collecting damage labels when prompted by the algorithm. By running the DP in parallel, the effect from confounding influences could be avoided. This method also works on AE features that can be used online without the need for preprocessing, such as dimensionality reduction. Furthermore, as the DP clusters according to changes in the features, it is possible to predict imminent failure on tools that are not correlated linearly to tool wear and also behaves dissimilar to others. In conclusion, this method could be applied to any tool as long as the input features intensify when approaching tool failure.

The main aims and the objectives of this thesis were to understand the tool wear and chip formation of turning in greater detail than previous work and to explore the feasibility of integrating automatic tool wear inspections for an unmanned tool testing procedure. By collecting the largest AE dataset generated during tool wear tests to date, it was possible not only to make links between chip formation mechanics and AE in oblique turning for the first time, but also to identify novel AE features that are sensitive to tool degradation and imminent tool failure. Furthermore, the detailed exploration of the wear scar using 3D scanning microscopy allowed the discovery of novel tool wear features that are more descriptive than the ones in use today. As a result, it was possible to find signal and tool feature sets that were better correlated than the ones previously used in literature, leading to good prediction accuracies when employing supervised learning algorithms. Finally, cluster initiations

of DPMMs were considered as a new method of setting threshold values for damage detection in TCM, an idea that has not been explored in past literature.

---

## FUTURE WORK

There are number of avenues that can be explored in the future to improve and implement the work presented in this thesis. The objective of this chapter is to detail the prospective possibilities of each aspect of the project. Therefore the layout of this chapter will follow the arrangement of the chapters, starting from Chapter 3 - experimental work and ending on Chapter 7.

### **9.1 Improvements to the experimental setup**

As the drive towards automation is well under way in the manufacturing industry, it would be useful to attempt to automate the experimental set up in the future. As discussed in Chapter 2, the tool condition monitoring community has a large uphill climb in order to achieve this goal. One of the many aspects of TCM research that has large room for improvements is the experimental set up.

For example, to reduce confounding influences on the collected data in this work, a futuristic setup could contain a dedicated robotic arm with a 3D scanning measurement head at each CNC machine. These robots are available for purchase in the current market, therefore this would not be a giant leap. A dedicated 3D

scanning microscope would eliminate the need to remove the tool from the machine for measurement. This would not only increase productivity but would also reduce uncertainties in the process. As a result, the first of the four passes of the workpiece in this work could be retained in the dataset.

Robotics could be used in other parts of the machining process too. For example, the loading and unloading of the workpieces could easily be automated by a robotic arm. Again, this technology exists today and therefore would not be an unrealistic aim. If integrated, the entire tool wear test procedure will be fully automated for an entire turret of tools.

To achieve these goals however, research must be conducted into integrating the TCM technology into the control system of the CNC lathe.

## 9.2 Exploring tool wear features

In a fully automated test setup with a dedicated measurement system, the feature extraction process from 2D profiles could be implemented relatively easily. There is scope in the future to investigate smoothing functions that may allow for a better selection method for tool wear features, allowing the process to be entirely automated. As there is no need to move the tool for measurement, the position and orientation of the scans will not vary, therefore alignment would not be necessary. As a result, a simple program could be installed in to the measurement software to extract the highest ranking features from Chapter 4.

It would also be beneficial to investigate other tool grades outside of the materials used in this project, in order to investigate whether the chosen features used in this work would behave similarly to others. By doing so it may be possible to identify a small number of features that are descriptive of the wear scar across a large number of tool materials.

### 9.3 Understanding acoustic emissions further

Chapter 5 explored the acoustic emissions generated in turning of case hardened steel with PcBN inserts. By analysing the chips formed during machining, it was possible to further understand the AE signal when converted to the frequency domain. By doing so, numerous useful features were found.

It may be possible to further the current knowledge of AE generation in turning by visually studying the chip formation process and validate the findings in Chapter 5; there exists cameras that are capable of matching frame rates to the sampling rates of AE, thus allowing the capture of chip formation in slow motion. Ideally, a machine with open access is required to build a rig with a high speed camera, as light sources that can produce large amount of light are required inside the machine due to short exposure times. As a result, a number of light sources must be focused on the tool, which is challenging to achieve inside a typical CNC machine. The videos may be analysed using image processing and pattern recognition techniques. By learning more about the AE signal during machining, it may be possible to pinpoint the aspects of tool wear that affect the AE features. For example, observing the change in curl of the chips or the direction of chip evacuation may provide a better insight into the shape of the wear scar that drives the behaviour. By doing so, it may be possible to link AE and the chip formation with that specific tool wear scar. This information would be useful for feature extraction as specific behaviours of the AE signal could be used to identify certain aspects of the wear scar. However, due to the set up requirements, it would not be possible to apply this method to a machining process that requires coolant.

### 9.4 Extending the supervised learning

In Chapter 6 supervised learning methods were used to predict wear on previously unseen tools. The RSL methodology was adopted here where the strongest AE features that were correlated to the chosen tool wear labels were used as inputs and

responses respectively. In attempts to assess the capability of a nonlinear model, GPR was applied to the same data.

In tool condition monitoring, predicting the surface finish is often just as important as tool wear, because these processes are intrinsically linked. It is possible that the AE features are correlated more strongly with surface roughness features in comparison to tool wear features. It would be interesting to test whether supervised learning of workpiece surface roughness could be performed with higher accuracy than predicting tool wear.

With the availability of the large data set comprising of AE and force features from three sensors and a variety of tool wear features in this work, the application of a genetic algorithm may provide a superior solution to the tool wear prediction problem compared to the RSL methodology and GPR. A genetic algorithm alongside a chosen regression model could be used to select the most suitable feature combinations, in order to yield the most accurate predictions. Due to constraints on time and computational power, it was not possible to apply this method in this work. However, in the future, a generic algorithm may be the best choice for a feature set of this size.

## 9.5 Improvements to the online clustering

In Chapter 7, Dirchlet process mixture models were used cluster the harmonic frequencies of chip formation, where the cluster initiations were used to predict imminent tool failure.

This Dirchlet process is versatile across two grades of PcBN and also across dissimilarly worn tools. It would be interesting to investigate whether this method is robust against different cutting conditions, insert shapes, workpiece materials and turning operations such as interrupted machining. If it is possible to isolate signal features that can successfully cluster data across numerous set ups, then this method could be applied across the entire shop floor in the future. However, further research is needed to find more reliable tool features that do not increase intensity during the middle of the tool life, leading to cluster initiations that are not helpful to the operator.

A futuristic set up taking advantage of robotics could benefit the Dirchlet process clustering. As discussed in detail in Chapter 7, confounding influences affect this process a great deal. Firstly, due to the repositioning of the tool after measurement, the data that is collected immediately after must be removed from analysis. Even with this solution, it is difficult to know whether depth of cut inconsistencies have affected the data in any way. By implementing a dedicated measurement arm to each machine, as discussed earlier in this chapter, it may be possible to use all the data produced by the tool, hopefully resulting in a smoother transition between clusters and initiating fewer clusters that are misconstrued as damage.

Secondly, and perhaps more interestingly, a dedicated measurement system could be programmed with tool wear thresholds in order to automatically accept or reject a tool at the warnings indicated by cluster initiations. This would negate the need for an operator to inspect the tool, leading to an entirely automated process.

By studying the point at which the final cluster before failure is initiated, it may be possible to predict when a tool may fail, to the nearest pass of the workpiece. For instance, if it was found that all final indicators before failure occurs around three passes of the workpiece, it may be possible to identify the point of failure in a probabilistic way.

This method does however require validation as it is not yet possible to link tool wear to AE features without further experimentation. It is assumed that chip formation is the key to make inferences about AE and tool wear. To validate the model, it is suggested that the machining process is halted at the points of cluster initiations to collect tool wear data as well as chips.

Transfer learning [135] may also be applied to this dataset; here clusters (along with known labels) are projected to a space that already contains clusters that are yet to be labelled. The idea is to minimise the distance between the cluster means. For the dataset collected here, inductive transfer learning could be applied, where the algorithm has information about the current cluster and requires a small amount of data from the new cluster to make judgements on similarity. However, it is not clear whether this method would be robust against tool breakage or tool wear.



---

## BIBLIOGRAPHY

- [1] A. Rytter, *Vibrational Based Inspection of Civil Engineering Structures*. PhD thesis, University of Aalborg, 1993.
- [2] C. M. Bishop, *Neural Networks for Pattern Recognition*. New York, NY, USA: Oxford University Press, Inc., 1995.
- [3] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.
- [4] H. V. Ravindra, Y. G. Srinivasa, and R. Krishnamurthy, “Acoustic emission for tool condition monitoring in metal cutting,” *Wear*, vol. 212, no. 1, pp. 78–84, 1997.
- [5] E. Emel and E. Kannatey-Asibu, “Acoustic emission and force sensor fusion for monitoring the cutting process,” *International Journal of Mechanical Sciences*, vol. 31, no. 11-12, pp. 795–809, 1989.
- [6] S. Dolinšek and J. Kopač, “Acoustic emission signals for tool wear identification,” *Wear*, vol. 225-229, no. I, pp. 295–303, 1999.
- [7] E. Trent and P. Wright, *Metal cutting*. Woburn: Butterworth-Heinemann, 4th ed., 2000.
- [8] A. Bhattacharyya and I. Ham, “Analysis of Tool Wear Part I: Theoretical Models of Flank Wear,” *Journal of Engineering for Industry*, vol. 91, no. 3, pp. 790–796, 1969.

- 
- [9] V. Šolaja, “Wear of carbide tools and surface finish generated in finish turning of steel,” *Wear*, vol. 2, pp. 40–58, aug 1958.
- [10] H. Wiklund, “Bayesian and regression approaches to on-line prediction of residual tool life,” *Quality and Reliability Engineering International*, vol. 14, pp. 303–309, sep 1998.
- [11] Sandvik Coromant, “Green Light Machining.”
- [12] Z. C. Lin and D. Y. Chen, “A study of cutting with a CBN tool,” *Journal of Materials Processing Tech.*, vol. 49, pp. 149–164, feb 1995.
- [13] H. K. Tonshoff, C. Arendt, and R. Ben Amor, “Cutting of hardened steel,” *CIRP Annals - Manufacturing Technology*, vol. 49, pp. 547–566, jan 2000.
- [14] C. M. Lahiff, B. Sc, S. S. Gordon, and P. Phelan, “An Investigation into the Effect of High Speed Cutting on the Wear Behaviour of PCBN Tools in Continuous Cutting of Hardened Steel Masters of Engineering by,” Tech. Rep. November, University of Limerick, 2008.
- [15] Elementsix, “PCBN Metalworking,” tech. rep., Element Six Ltd, Shannon, 2008.
- [16] C. R. Liu and S. Mittal, “Single-step superfinish hard machining: Feasibility and feasible cutting conditions,” *Robotics and Computer-Integrated Manufacturing*, vol. 12, pp. 15–27, mar 1996.
- [17] M. Dogra, V. S. Sharma, A. Sachdeva, N. M. Suri, and J. S. Dureja, “Tool wear, chip formation and workpiece surface issues in CBN hard turning: A review,” *International Journal of Precision Engineering and Manufacturing*, vol. 11, pp. 341–358, apr 2010.
- [18] D. A. Dornfeld, Y. Lee, and A. Chang, “Monitoring of ultraprecision machining processes,” *International Journal of Advanced Manufacturing Technology*, vol. 21, no. 8, pp. 571–578, 2003.
- [19] J. Barry and G. Byrne, “Study on acoustic emission in machining hardened steels Part 1: Acoustic emission during saw-tooth chip formation,” *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 215, no. 11, pp. 1549–1559, 2001.

- [20] L. Dan and J. Mathew, "Tool wear and failure monitoring techniques for turning-A review," *International Journal of Machine Tools and Manufacture*, vol. 30, no. 4, pp. 579–598, 1990.
- [21] G. Byrne, D. Dornfeld, I. Inasaki, G. Ketteler, W. König, and R. Teti, "Tool Condition Monitoring (TCM) - The Status of Research and Industrial Application," *CIRP Annals - Manufacturing Technology*, vol. 44, pp. 541–567, jan 1995.
- [22] D. E. Dimla, P. M. Lister, and N. J. Leighton, "Neural network solutions to the tool condition monitoring problem in metal cutting - A critical review of methods," *International Journal of Machine Tools and Manufacture*, vol. 37, pp. 1219–1241, sep 1997.
- [23] D. E. Dimla, "Sensor signals for tool-wear monitoring in metal cutting operations - a review of methods," *International Journal of Machine Tools and Manufacture*, vol. 40, no. 8, pp. 1073–1098, 2000.
- [24] B. Sick, "On-line and indirect tool wear monitoring in turning with artificial neural networks: A review of more than a decade of research," *Mechanical Systems and Signal Processing*, vol. 16, no. 4, pp. 487–546, 2002.
- [25] A. G. Rehorn, J. Jiang, and P. E. Orban, "State-of-the-art methods and results in tool condition monitoring: A review," *International Journal of Advanced Manufacturing Technology*, vol. 26, no. 7-8, pp. 693–710, 2005.
- [26] P. N. Botsaris and J. A. Tsanakas, "State-of-the-Art in Methods Applied To Tool Condition Monitoring (Tcm) in Unmanned Machining Operations: a Review," *The International Conference of COMADEM*, pp. 73–87, 2008.
- [27] J. V. Abellan-Nebot and F. Romero Subirón, "A review of machining monitoring systems based on artificial intelligence process models," *International Journal of Advanced Manufacturing Technology*, vol. 47, no. 1-4, pp. 237–257, 2010.
- [28] N. Ambhore, D. Kamble, S. Chinchankar, and V. Wayal, "Tool condition monitoring system: A review," *Materials Today: Proceedings*, vol. 2, no. 4-5, pp. 3419–3428, 2015.

- [29] Y. Zhou and W. Xue, "Review of tool condition monitoring methods in milling processes," *International Journal of Advanced Manufacturing Technology*, vol. 96, no. 5-8, pp. 2509–2523, 2018.
- [30] A. Siddhpura and R. Paurobally, "A review of flank wear prediction methods for tool condition monitoring in a turning process," *International Journal of Advanced Manufacturing Technology*, vol. 65, no. 1-4, pp. 371–393, 2013.
- [31] S. Das, R. Roy, and A. B. Chattopadhyay, "Evaluation of wear of turning carbide inserts using neural networks," *International Journal of Machine Tools and Manufacture*, vol. 36, pp. 789–797, jul 1996.
- [32] S. Das, P. P. Bandyopadhyay, and A. B. Chattopadhyay, "Neural-networks-based tool wear monitoring in turning medium carbon steel using a coated carbide tool," *Journal of Materials Processing Technology*, vol. 63, no. 1-3, pp. 187–192, 1997.
- [33] D. E. Dimla and P. M. Lister, "On-line metal cutting tool condition monitoring. II: tool-state classification using multi-layer perceptron neural networks," *International Journal of Machine Tools and Manufacture*, vol. 40, no. 5, pp. 769–781, 2000.
- [34] J. Sun, M. Rahman, Y. S. Wong, and G. S. Hong, "Multiclassification of tool wear with support vector machine by manufacturing loss consideration," *International Journal of Machine Tools and Manufacture*, vol. 44, no. 11, pp. 1179–1187, 2004.
- [35] S. Binsaeid, S. Asfour, S. Cho, and A. Onar, "Machine ensemble approach for simultaneous detection of transient and gradual abnormalities in end milling using multisensor fusion," *Journal of Materials Processing Technology*, vol. 209, pp. 4728–4738, jun 2009.
- [36] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Health assessment and life prediction of cutting tools based on support vector regression," *Journal of Intelligent Manufacturing*, vol. 26, no. 2, pp. 213–223, 2015.
- [37] G. Wang, L. Qian, and Z. Guo, "Continuous tool wear prediction based on Gaussian mixture regression model," *International Journal of Advanced Manufacturing Technology*, vol. 66, no. 9-12, pp. 1921–1929, 2013.

- [38] S. Dutta, S. K. Pal, and R. Sen, "On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression," *Precision Engineering*, vol. 43, pp. 34–42, jan 2016.
- [39] J. Karandikar, T. McLeay, S. Turner, and T. Schmitz, "Tool wear monitoring using naïve Bayes classifiers," *International Journal of Advanced Manufacturing Technology*, vol. 77, no. 9-12, pp. 1613–1626, 2015.
- [40] T. E. Mcleay, *Unsupervised Monitoring of Machining Processes*. PhD thesis, University of Sheffield, 2016.
- [41] S. Y. Liang and D. A. Dornfeld, "Tool Wear Detection Using Time Series Analysis of Acoustic Emission," *Journal of Engineering for Industry*, vol. 111, no. 3, p. 199, 1989.
- [42] J. Wang, P. Wang, and R. X. Gao, "Enhanced particle filter for tool wear prediction," *Journal of Manufacturing Systems*, vol. 36, pp. 35–45, 2015.
- [43] R. K. Miller, E. v. K. Hill, P. O. Moore, and American Society for Nondestructive Testing., *Acoustic emission testing*. American Society for Nondestructive Testing, 2005.
- [44] K. Uehara and Y. Kanda, "Identification of Chip Formation Mechanism through Acoustic Emission Measurements," *CIRP Annals - Manufacturing Technology*, vol. 33, no. 1, pp. 71–74, 1984.
- [45] E. Kannatey-Asibu and D. A. Dornfeld, "A study of tool wear using statistical analysis of metal-cutting acoustic emission," *Wear*, vol. 76, no. 2, pp. 247–261, 1982.
- [46] A. E. Diniz, J. J. Liu, and D. A. Dornfeld, "Correlating tool life, tool wear and surface roughness by monitoring acoustic emission in finish turning," *Wear*, vol. 152, no. 2, pp. 395–407, 1992.
- [47] X. Li, "A brief review: Acoustic emission method for tool wear monitoring during turning," *International Journal of Machine Tools and Manufacture*, vol. 42, no. 2, pp. 157–165, 2002.

- [48] T. Moriwaki and K. Okushima, "Detection for Cutting Tool Fracture by Acoustic Emission Measurement," *CIRP Annals - Manufacturing Technology*, vol. 29, no. 1, pp. 35–40, 1980.
- [49] Y. B. Guo and S. C. Ammala, "Real-time acoustic emission monitoring for surface damage in hard machining," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 14, pp. 1622–1627, 2005.
- [50] S. V. Kamarthi, S. R. T. Kumara, and P. H. Cohen, "Flank Wear Estimation in Turning Through Wavelet Representation of Acoustic Emission Signals," *Journal of Manufacturing Science and Engineering*, vol. 122, no. 1, p. 12, 2000.
- [51] K. Jemielniak, J. Kossakowska, and T. Urbański, "Application of wavelet transform of acoustic emission and cutting force signals for tool condition monitoring in rough turning of Inconel 625," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 225, no. 1, pp. 123–129, 2011.
- [52] N. Ray, *Correlation between machining monitoring signals, cutting tool wear and surface integrity on high strength titanium alloys*. PhD thesis, University of Sheffield, 2018.
- [53] J. Bhaskaran, M. Murugan, N. Balashanmugam, and M. Chellamalai, "Monitoring of hard turning using acoustic emission signal," *Journal of Mechanical Science and Technology*, vol. 26, no. 2, pp. 609–615, 2012.
- [54] A. S. More, W. Jiang, W. D. Brown, and A. P. Malshe, "Tool wear and machining performance of cBN-TiN coated carbide inserts and PCBN compact inserts in turning AISI 4340 hardened steel," *Journal of Materials Processing Technology*, vol. 180, no. 1-3, pp. 253–262, 2006.
- [55] R. T. Coelho, E. G. Ng, and M. A. Elbestawi, "Tool wear when turning hardened AISI 4340 with coated PCBN tools using finishing cutting conditions," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 2, pp. 263–272, 2007.
- [56] I. T. 29, "ISO 3685:1993 - Tool-life testing with single-point turning tools," tech. rep., International Organization for Standardization, Geneva, 1993.

- [57] G. Poulachon, B. P. Bandyopadhyay, I. S. Jawahir, S. Pheulpin, and E. Seguin, "Wear behavior of CBN tools while turning various hardened steels," *Wear*, vol. 256, no. 3-4, pp. 302–310, 2004.
- [58] N. Narukati and Y. Yamane, "Tool wear and cutting temperature of CBN tools in machining of hardened steels," *Annal CIRP*, vol. 45, no. (1), pp. 77–82, 1979.
- [59] C. Lahiff, S. Gordon, and P. Phelan, "PCBN tool wear modes and mechanisms in finish hard turning," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 6, pp. 638–644, 2007.
- [60] J. A. Arsecularatne, L. C. Zhang, and C. Montross, "Wear and tool life of tungsten carbide, PCBN and PCD cutting tools," *International Journal of Machine Tools and Manufacture*, vol. 46, pp. 482–491, apr 2006.
- [61] G. Poulachon, A. Moisan, and I. S. Jawahir, "Tool-wear mechanisms in hard turning with polycrystalline cubic boron nitride tools," *Wear*, vol. 250-251, no. 1-12, pp. 576–586, 2001.
- [62] J. C. Camargo, D. S. Dominguez, E. O. Ezugwu, and Á. R. Machado, "Wear model in turning of hardened steel with PCBN tool," *International Journal of Refractory Metals and Hard Materials*, vol. 47, pp. 61–70, 2014.
- [63] K. Komvopoulos and S. A. Erpenbeck, "Finite Element Modeling of Orthogonal Metal Cutting," *Journal of Engineering for Industry*, vol. 113, p. 253, aug 1991.
- [64] N. Ghosh, Y. B. Ravi, A. Patra, S. Mukhopadhyay, S. Paul, A. R. Mohanty, and A. B. Chattopadhyay, "Estimation of tool wear during CNC milling using neural network-based sensor fusion," *Mechanical Systems and Signal Processing*, vol. 21, pp. 466–479, jan 2007.
- [65] F. Mahfoudi, G. List, A. Molinari, A. Moufki, and L. Boulanouar, "High speed turning for hard material with PCBN inserts: tool wear analysis," *International Journal of Machining and Machinability of Materials*, vol. 3, no. 1/2, p. 62, 2008.

- [66] I. Durazo-Cardenas, P. Shore, X. Luo, T. Jacklin, S. A. Impey, and A. Cox, "3D characterisation of tool wear whilst diamond turning silicon," *Wear*, vol. 262, no. 3-4, pp. 340–349, 2007.
- [67] S. Thamizhmanii and S. Hasan, "Measurement of surface roughness and flank wear on hard martensitic stainless steel by CBN and PCBN cutting tools," *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 31, no. 2, pp. 415–421, 2008.
- [68] S. Kurada and C. Bradley, "A review of machine vision sensors for tool condition monitoring," *Computers in Industry*, vol. 34, no. 1, pp. 55–72, 1997.
- [69] S. Kurada and C. Bradley, "A machine vision system for tool wear assessment," *Tribology International*, vol. 30, no. 4, pp. 295–304, 1997.
- [70] D. Kerr, J. Pengilley, and R. Garwood, "Assessment and visualisation of machine tool wear using computer vision," *International Journal of Advanced Manufacturing Technology*, vol. 28, no. 7-8, pp. 781–791, 2006.
- [71] K. B. Pedersen, "Wear measurement of cutting tools by computer vision," *International Journal of Machine Tools and Manufacture*, vol. 30, no. 1, pp. 131–139, 1990.
- [72] J. Jurkovic, M. Korosec, and J. Kopac, "New approach in tool wear measuring technique using CCD vision system," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 9, pp. 1023–1030, 2005.
- [73] T. Pfeifer and L. Wieggers, "Reliable tool wear monitoring by optimized image and illumination control in machine vision," *Measurement: Journal of the International Measurement Confederation*, vol. 28, no. 3, pp. 209–218, 2000.
- [74] M. Sortino, "Application of statistical filtering for optical detection of tool wear," *International Journal of Machine Tools and Manufacture*, vol. 43, no. 5, pp. 493–497, 2003.
- [75] A. Faraz, D. Biermann, and K. Weinert, "Cutting edge rounding: An innovative tool wear criterion in drilling CFRP composite laminates," *International Journal of Machine Tools and Manufacture*, vol. 49, no. 15, pp. 1185–1196, 2009.

- [76] T. Zhao, J. M. Zhou, V. Bushlya, and J. E. Ståhl, “Effect of cutting edge radius on surface roughness and tool wear in hard turning of AISI 52100 steel,” *International Journal of Advanced Manufacturing Technology*, vol. 91, no. 9-12, pp. 3611–3618, 2017.
- [77] L. Saidi, J. Ben Ali, E. Bechhoefer, and M. Benbouzid, “Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR,” *Applied Acoustics*, vol. 120, pp. 1–8, 2017.
- [78] D. J. C. MacKay, *Information Theory and Learning Algorithms*. Cambridge University Press, 2003.
- [79] J. M. Zhou, H. Walter, M. Andersson, and J. E. Stahl, “Effect of chamfer angle on wear of PCBN cutting tool,” *International Journal of Machine Tools and Manufacture*, vol. 43, no. 3, pp. 301–305, 2003.
- [80] J. J. Hensman, “Novel techniques for acoustic emission monitoring of fatigue fractures in landing gear,” no. February, 2009.
- [81] E. Kannatey-Asibu and D. A. Dornfeld, “Quantitative Relationships for Acoustic Emission from Orthogonal Metal Cutting,” *Journal of Engineering for Industry*, vol. 103, no. 3, p. 330, 2009.
- [82] H. Dunegan and A. Green, “Factors Affecting Acoustic Emission Response From Materials,” *Acoustic Emission*, pp. 100–100–14, 2009.
- [83] P. Gillis, “Dislocation Motions and Acoustic Emissions,” *Acoustic Emission*, pp. 20–20–10, 2009.
- [84] D. A. Dornfeld and E. Kannatey-Asibu, “Acoustic emission during orthogonal metal cutting,” *International Journal of Mechanical Sciences*, vol. 22, no. 5, pp. 285–296, 1980.
- [85] J. Barry and G. Byrne, “The Mechanisms of Chip Formation in Machining Hardened Steels,” *Journal of Manufacturing Science and Engineering*, vol. 124, no. 3, p. 528, 2002.
- [86] J. Barry and G. Byrne, “Chip formation, acoustic emission and surface white layers in hard machining,” *CIRP Annals - Manufacturing Technology*, vol. 51, no. 1, pp. 65–70, 2002.

- [87] M. A. Davies, T. Burns, and C. J. Evans, "On the dynamics of chip formation in machining hard metals," *CIRP Annals - Manufacturing Technology*, vol. 46, no. 1, pp. 25–30, 1997.
- [88] R. H. Brown and E. J. Armarego, "Oblique machining with a single cutting edge," *International Journal of Machine Tool Design and Research*, vol. 4, pp. 9–25, jun 1964.
- [89] M. Neslušan, B. Mičieta, A. Mičietová, M. Čilliková, and I. Mrkvica, "Detection of tool breakage during hard turning through acoustic emission at low removal rates," *Measurement: Journal of the International Measurement Confederation*, vol. 70, pp. 1–13, 2015.
- [90] T. Blum and I. Inasaki, "A Study on Acoustic Emission from the Orthogonal Cutting Process," *Journal of Engineering for Industry*, vol. 112, no. 3, p. 203, 2008.
- [91] E. Kannatey-Asibu and E. Emel, "Linear discriminant function analysis of acoustic emission signals for cutting tool monitoring," *Mechanical Systems and Signal Processing*, vol. 1, no. 4, pp. 333–347, 1987.
- [92] X. Li, Y. Yao, and Z. Yuan, "On-line tool condition monitoring system with wavelet fuzzy neural network," *Journal of Intelligent Manufacturing*, vol. 8, no. 4, pp. 271–276, 1997.
- [93] I. Marinescu and D. A. Axinte, "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*, vol. 48, pp. 1148–1160, aug 2008.
- [94] O. Olufayo and K. Abou-El-Hossein, "Tool life estimation based on acoustic emission monitoring in end-milling of H13 mould-steel," *International Journal of Advanced Manufacturing Technology*, vol. 81, no. 1-4, pp. 39–51, 2015.
- [95] A. J. Mian, N. Driver, and P. T. Mativenga, "Estimation of minimum chip thickness in micro-milling using acoustic emission," in *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 225, pp. 1535–1551, 2011.

- [96] E. Hussein, “Acoustic Emission,” in *Non-Destructive Testing, Course 7.2*, ch. 6, pp. 125–144, Chulalongkorn University, 1996.
- [97] J. Barry, G. Byrne, and D. Lennon, “Observations on chip formation and acoustic emission in machining Ti-6Al-4V alloy,” *International Journal of Machine Tools and Manufacture*, vol. 41, no. 7, pp. 1055–1070, 2001.
- [98] E. J. Cross, K. Worden, and Q. Chen, “Cointegration: A novel approach for the removal of environmental trends in structural health monitoring data,” in *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, pp. 2712–2732, The Royal Society Publishing, sep 2011.
- [99] E. J. Cross, G. Manson, K. Worden, and S. G. Pierce, “Features for damage detection with insensitivity to environmental and operational variations,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 468, no. 2148, pp. 4098–4122, 2012.
- [100] D. Baccar and D. Söffker, “Wear detection by means of wavelet-based acoustic emission analysis,” *Mechanical Systems and Signal Processing*, vol. 60, pp. 198–207, aug 2015.
- [101] M. S. H. Bhuiyan, I. A. Choudhury, M. Dahari, Y. Nukman, and S. Z. Dawal, “Application of acoustic emission sensor to investigate the frequency of tool wear and plastic deformation in tool condition monitoring,” *Measurement: Journal of the International Measurement Confederation*, vol. 92, no. December 2017, pp. 208–217, 2016.
- [102] D. A. Axinte and N. Gindy, “Tool condition monitoring in broaching,” *Wear*, vol. 254, pp. 370–382, feb 2003.
- [103] O. Buck, “Harmonic Generation for Measurement of Internal Stresses as Produced by Dislocations,” *IEEE Transactions on Sonics and Ultrasonics*, vol. 23, no. 5, pp. 346–350, 1976.
- [104] W. T. Yost, M. Namkung, and S. G. Allison, “The Use of Ultrasonic Harmonic Generation to Determine Crack Opening Conditions in Compact Tension Specimens,” in *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 1525–1529, 2011.

- [105] X. Li, “Real-time detection of the breakage of small diameter drills with wavelet transform,” *International Journal of Advanced Manufacturing Technology*, vol. 14, no. 8, pp. 539–543, 1998.
- [106] G. E. P. Box and K. B. Wilson, “On the Experimental Attainment of Optimum Conditions,” Tech. Rep. 1, 1951.
- [107] A. J. Makadia and J. I. Nanavati, “Optimisation of machining parameters for turning operations based on response surface methodology,” *Measurement: Journal of the International Measurement Confederation*, vol. 46, no. 4, pp. 1521–1529, 2013.
- [108] R. Suresh, S. Basavarajappa, and G. L. Samuel, “Predictive modeling of cutting forces and tool wear in hard turning using response surface methodology,” *Procedia Engineering*, vol. 38, pp. 73–81, 2012.
- [109] S. Saini, I. S. Ahuja, and V. S. Sharma, “Influence of cutting parameters on tool wear and surface roughness in hard turning of AISI H11 tool steel using ceramic tools,” *International Journal of Precision Engineering and Manufacturing*, vol. 13, no. 8, pp. 1295–1302, 2012.
- [110] M. Seeman, G. Ganesan, R. Karthikeyan, and A. Velayudham, “Study on tool wear and surface roughness in machining of particulate aluminum metal matrix composite-response surface methodology approach,” *International Journal of Advanced Manufacturing Technology*, vol. 48, no. 5-8, pp. 613–624, 2010.
- [111] A. K. Gupta, “Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression,” *International Journal of Production Research*, vol. 48, no. 3, pp. 763–778, 2010.
- [112] E. Cross, *On Structural Health Monitoring in Changing Environmental and Operational Conditions*. PhD thesis, University of Sheffield, 2012.
- [113] T. J. Rogers, *Towards Bayesian System Identification : With Application to SHM of Offshore Structures*. PhD thesis, University of Sheffield, 2018.
- [114] R. Fuentes and E. Cross, “Aircraft parametric structural load monitoring using Gaussian regression techniques,” in *Proceedings of the 7th European Workshop on Structural Health Monitoring*, pp. 1933–1940, 2014.

- [115] K. Chandrasekhar, N. Stevanovic, E. J. Cross, N. Dervilis, and K. Worden, “Gaussian processes for structural health monitoring of wind turbine blades,” in *International Workshop for Structural Health Monitoring*, 2019.
- [116] J. Mahler, S. Krishnan, M. Laskey, S. Sen, A. Murali, B. Kehoe, S. Patil, J. Wang, M. Franklin, P. Abbeel, and K. Goldberg, “Learning accurate kinematic control of cable-driven surgical robots using data cleaning and Gaussian Process Regression,” in *IEEE International Conference on Automation Science and Engineering*, vol. 2014-Janua, pp. 532–539, 2014.
- [117] T. Koriyama and T. Kobayashi, “Prosody generation using frame-based Gaussian process regression and classification for statistical parametric speech synthesis,” in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2015-Augus, pp. 4929–4933, 2015.
- [118] D. Kong, Y. Chen, and N. Li, “Gaussian process regression for tool wear prediction,” *Mechanical Systems and Signal Processing*, vol. 104, pp. 556–574, may 2018.
- [119] M. Wang and J. Wang, “CHMM for tool condition monitoring and remaining useful life prediction,” *International Journal of Advanced Manufacturing Technology*, vol. 59, no. 5-8, pp. 463–471, 2012.
- [120] J. Hong, J. H. Zhou, H. L. Chan, C. Zhang, H. Xu, and G. S. Hong, “Tool condition monitoring in deep hole gun drilling: A data-driven approach,” in *IEEE International Conference on Industrial Engineering and Engineering Management*, vol. 2017-Decem, pp. 2148–2152, 2018.
- [121] G. Zhang, J. Li, Y. Chen, Y. Huang, X. Shao, and M. Li, “Prediction of surface roughness in end face milling based on Gaussian process regression and cause analysis considering tool vibration,” *International Journal of Advanced Manufacturing Technology*, vol. 75, no. 9-12, pp. 1357–1370, 2014.
- [122] D. S. Moore, G. P. McCabe, and C. A. Bruce, *Introduction to the Practice of Statistics*, vol. 79. WH Freeman, seventh ed., 2012.
- [123] M. Kutner, C. Nachtsheim, J. Neter, and W. Li, *Applied Linear Regression Models*, vol. 32 of *Irwin/McGraw-Hill series in operations and decision sciences*. McGraw-Hill/Irwin, 2005.

- 
- [124] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, vol. 2. MIT Press, 2006.
- [125] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012.
- [126] E. P. Xing, “Advanced Gaussian Processes,” 2015.
- [127] C. E. Rasmussen and Z. Ghahramani, “Infinite Mixtures of Gaussian Process Experts,” in *Advances in Neural Information Processing Systems 14*, vol. 14, pp. 881—888, MIT Press, 2001.
- [128] A. Vlachos, A. Korhonen, and Z. Ghahramani, “Unsupervised and constrained Dirichlet process mixture models for verb clustering,” in *GEMS '09 Proceedings of the Workshop on Geometrical Models of Natural Language Semantics*, pp. 74–82, 2009.
- [129] M. Dreyer and J. Eisner, “Discovering Morphological Paradigms from Plain Text Using a Dirichlet Process Mixture Model,” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, no. 0347822, pp. 616–627, 2011.
- [130] A. R. Ferreira da Silva, “A Dirichlet process mixture model for brain MRI tissue classification,” *Medical Image Analysis*, vol. 11, no. 2, pp. 169–182, 2007.
- [131] T. J. Rogers, K. Worden, R. Fuentes, N. Dervilis, U. T. Tygesen, and E. J. Cross, “A Bayesian non-parametric clustering approach for semi-supervised Structural Health Monitoring,” *Mechanical Systems and Signal Processing*, vol. 119, pp. 100–119, 2019.
- [132] S. Liang, J. Yu, D. Tang, and H. Liu, “A Bayesian nonparametric approach for tool condition monitoring,” *2016 UKACC International Conference on Control, UKACC Control 2016*, pp. 1–6, 2016.
- [133] C. E. Rasmussen, “The Infinite Gaussian Mixture Model,” in *Advances in Neural Information Processing Systems 12*, pp. 554–560, 2000.
- [134] T. J. Rogers, “A Bayesian non-parametric clustering approach for semi-supervised Structural Health Monitoring,” 2019.

- [135] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [136] S. Mallat, *A wavelet tour of signal processing - The sparse way*. Burlington, MA: Elsevier, 2009.

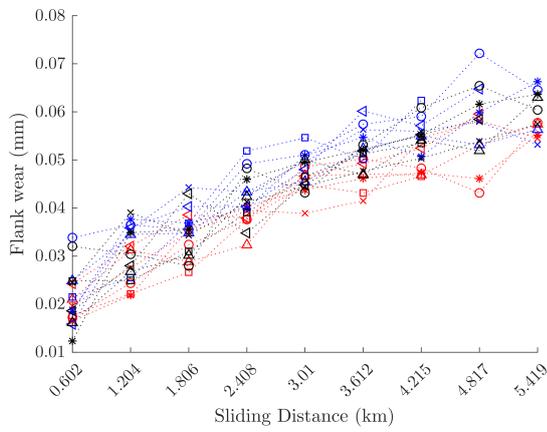


---

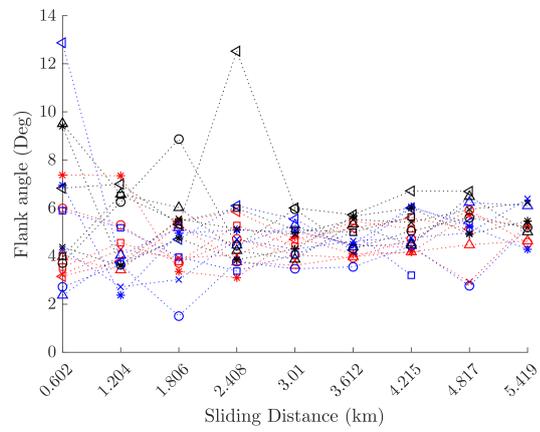
## TOOL WEAR FEATURES

This appendix displays all of the novel tool wear features collected in Section 4.4. In each plot, if relevant, the wear values from all three sections (leading edge - LE, centre - CE and trailing edge - TE) of the tool wear profile is shown on this page. The key to the plots have been presented here. Each plot is numbered according to the numbering system used in Section 4.4.

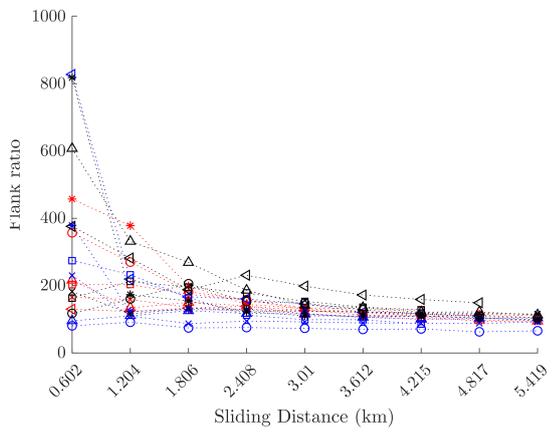
...x... B3 LE	...x... A3 LE
...o... B4 LE	...o... A4 LE
...□... B5 LE	...□... A5 LE
...◁... B6 LE	...◁... A6 LE
...*... B7 LE	...*... A7 LE
...△... B8 LE	...△... A8 LE
...x... B3 C	...x... A3 C
...o... B4 C	...o... A4 C
...□... B5 C	...□... A5 C
...◁... B6 C	...◁... A6 C
...*... B7 C	...*... A7 C
...△... B8 C	...△... A8 C
...x... B3 TE	...x... A3 TE
...o... B4 TE	...o... A4 TE
...□... B5 TE	...□... A5 TE
...◁... B6 TE	...◁... A6 TE
...*... B7 TE	...*... A7 TE
...△... B8 TE	...△... A8 TE
...x... B3 TE	...x... A9 TE
...o... B4 TE	...o... A9 TE
...□... B5 TE	...□... A9 TE
...◁... B6 TE	...◁... A9 TE
...*... B7 TE	...*... A9 TE
...△... B8 TE	...△... A9 TE



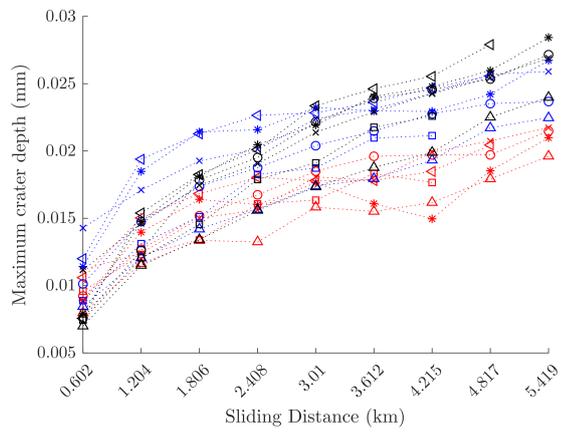
(a) Grade B - 1



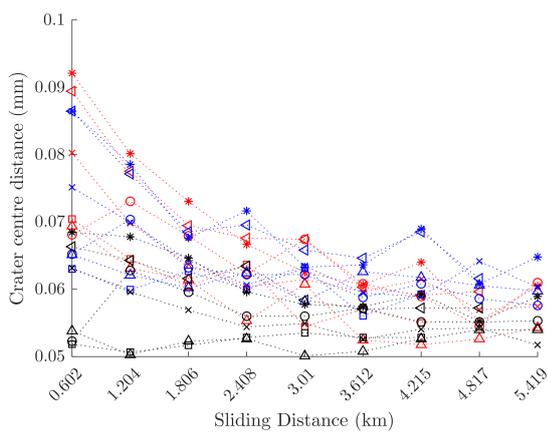
(b) Grade B - 2



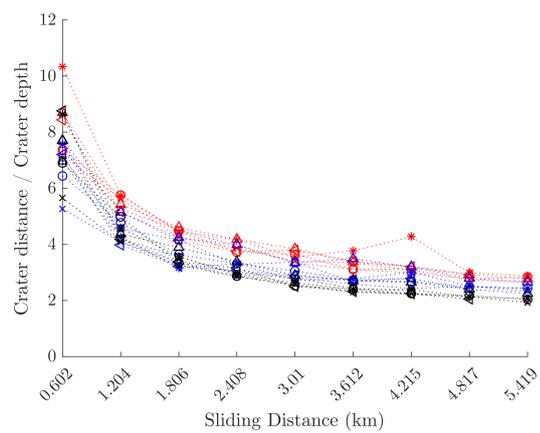
(c) Grade B - 3



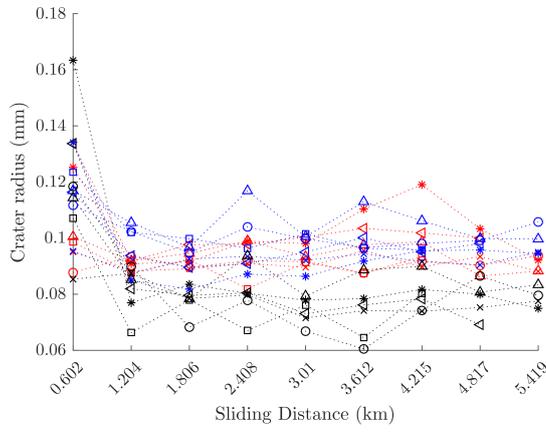
(d) Grade B - 4



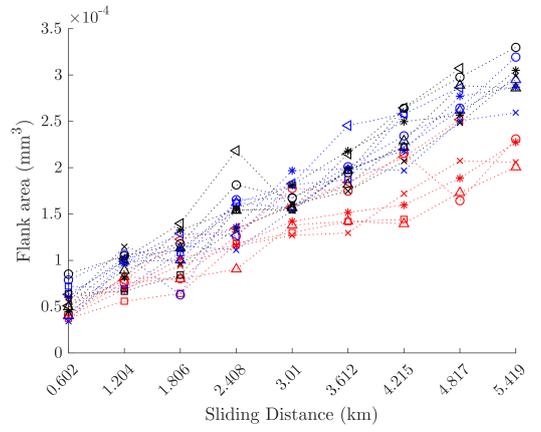
(e) Grade B - 5



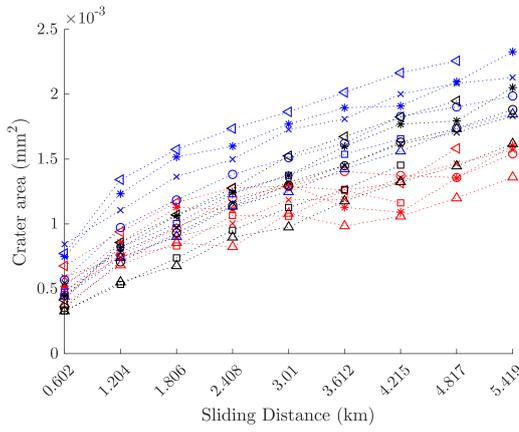
(f) Grade B - 6



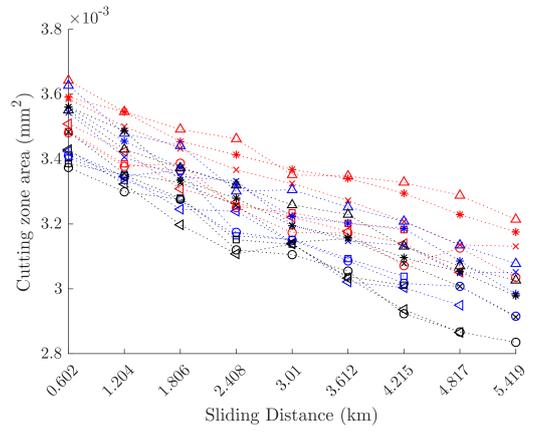
(a) Grade B - 7



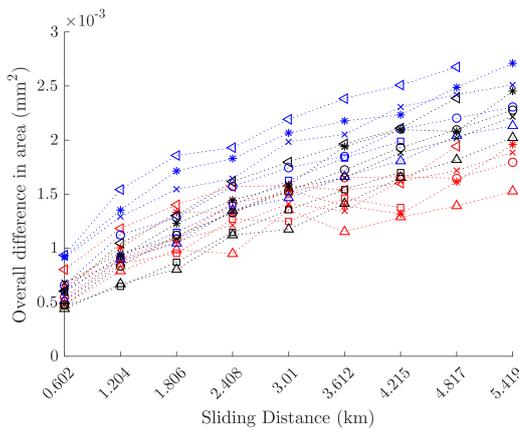
(b) Grade B - 8



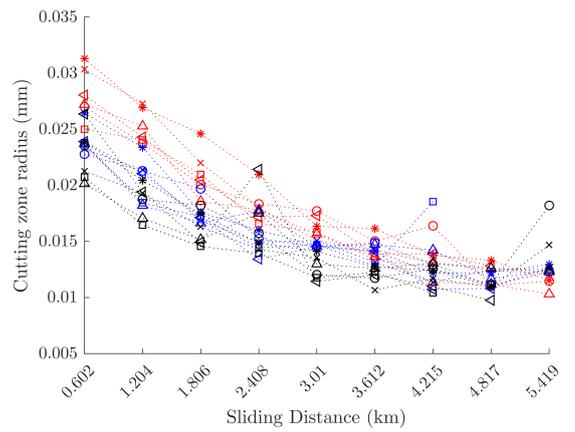
(c) Grade B - 9



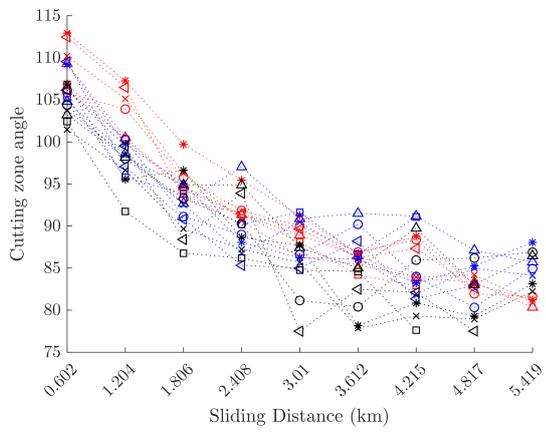
(d) Grade B - 10



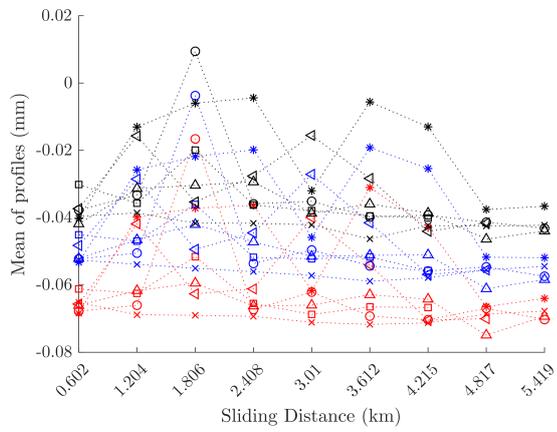
(e) Grade B - 11



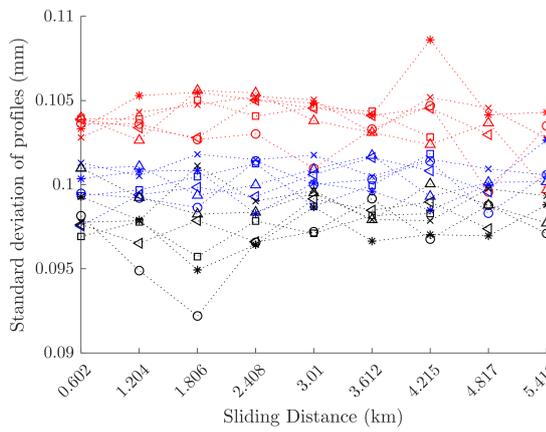
(f) Grade B - 12



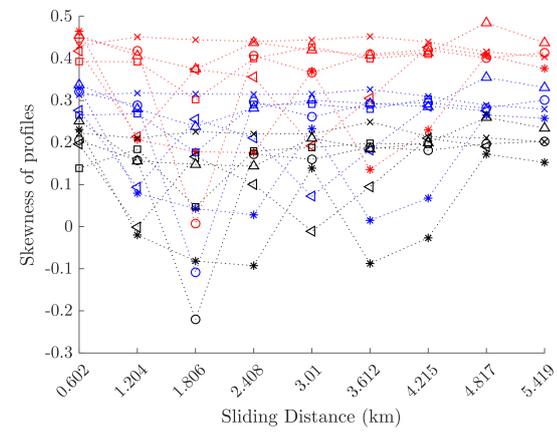
(a) Grade B - 13



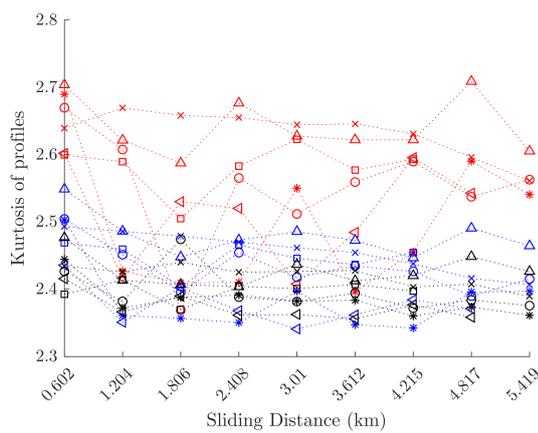
(b) Grade B - 14



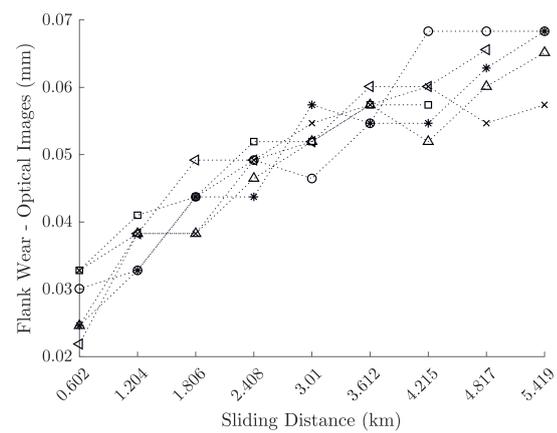
(c) Grade B - 15



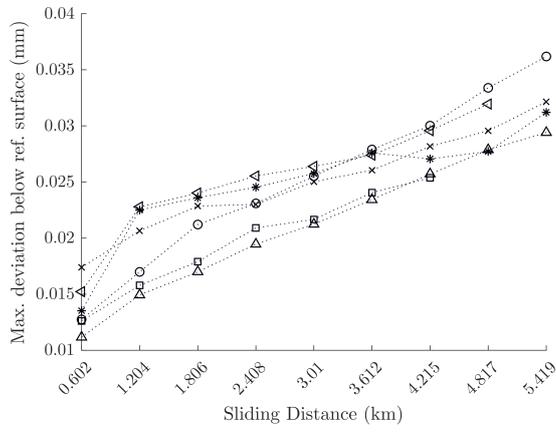
(d) Grade B - 16



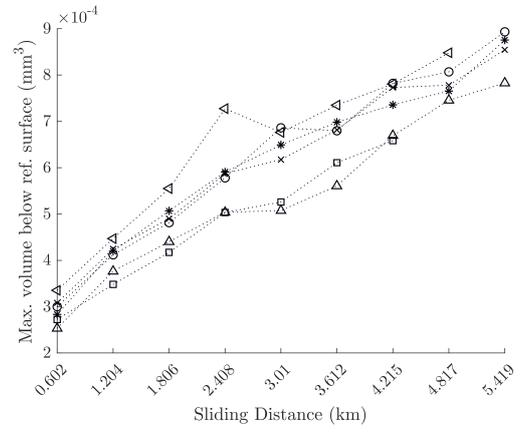
(e) Grade B - 17



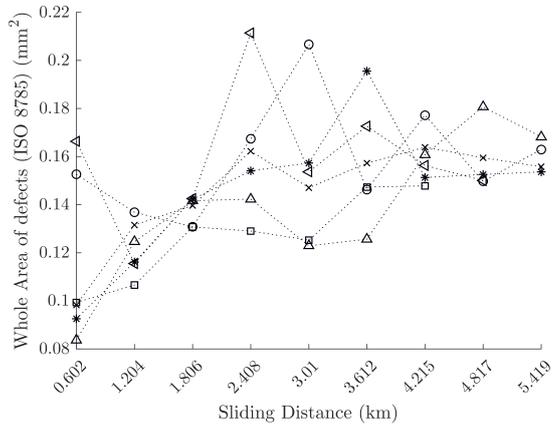
(f) Grade B - 18



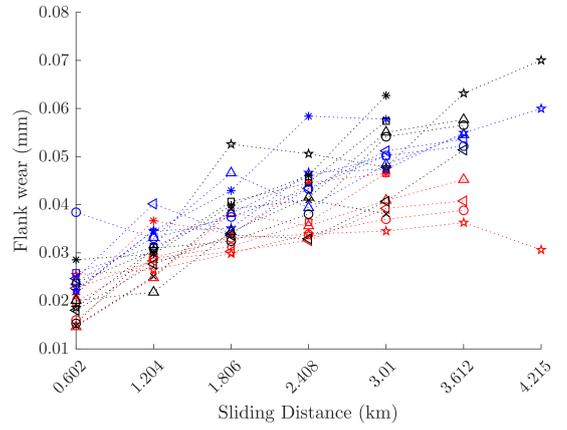
(a) Grade B - 19



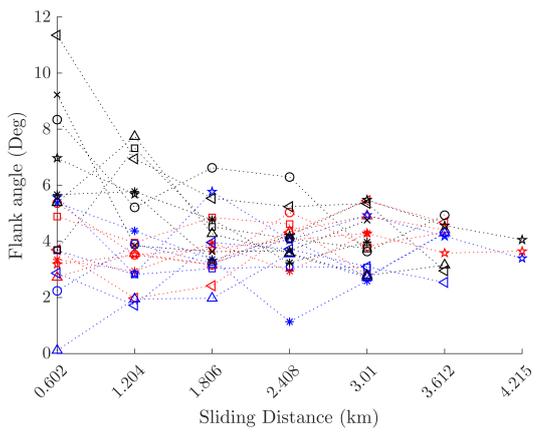
(b) Grade B - 20



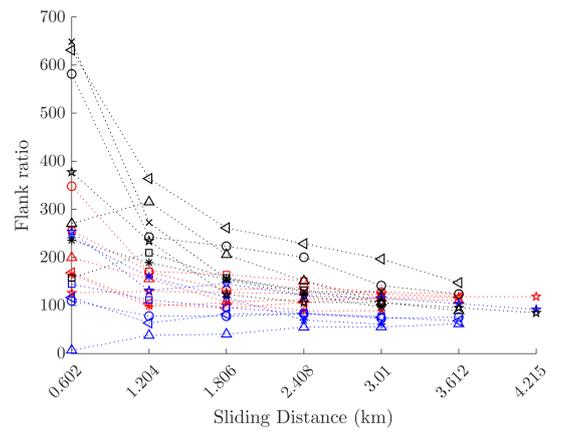
(c) Grade B - 21



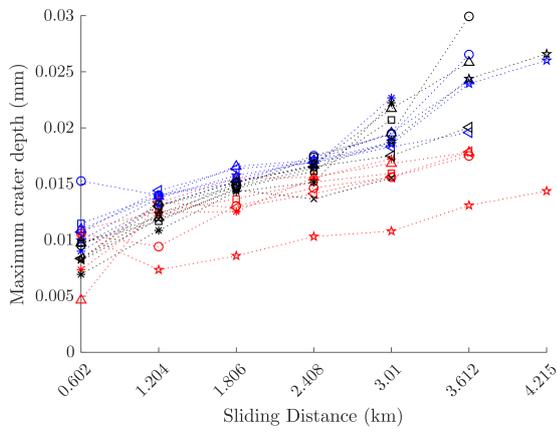
(d) Grade A - 1



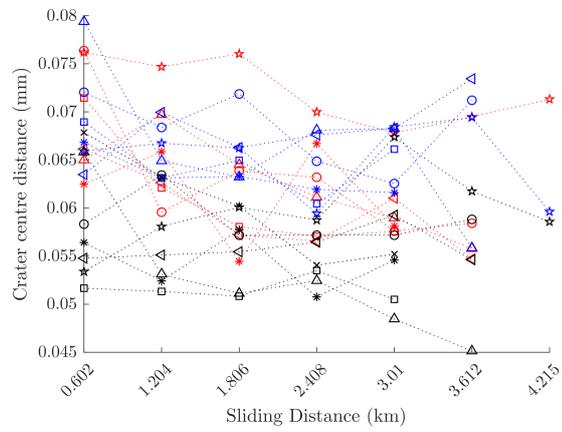
(e) Grade A - 2



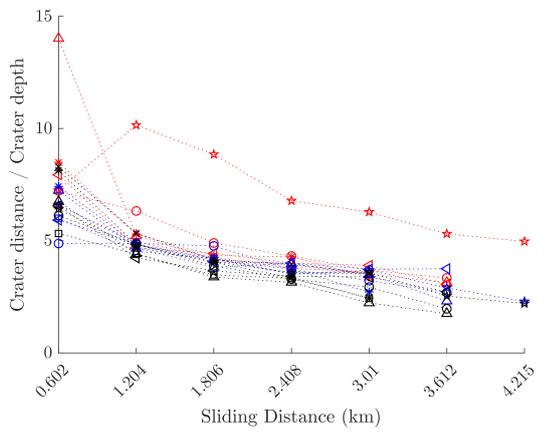
(f) Grade A - 3



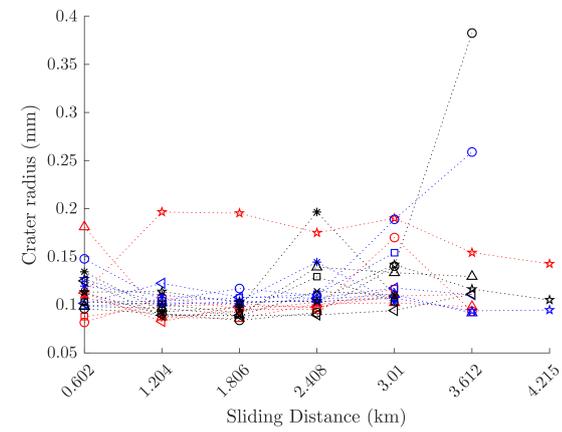
(a) Grade A - 4



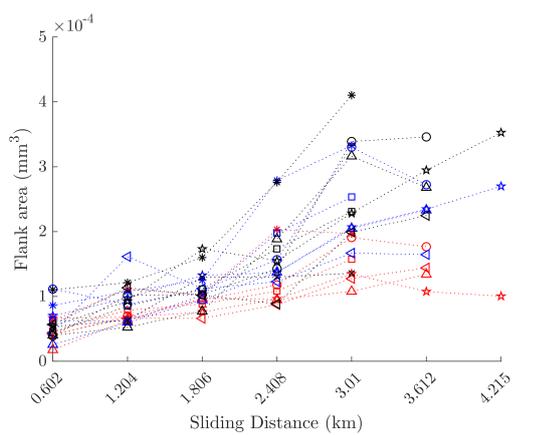
(b) Grade A - 5



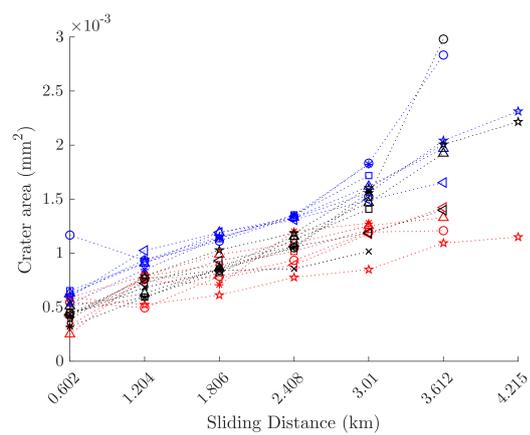
(c) Grade A - 6



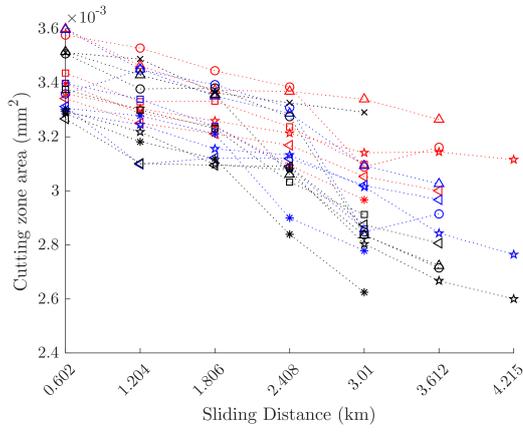
(d) Grade A - 7



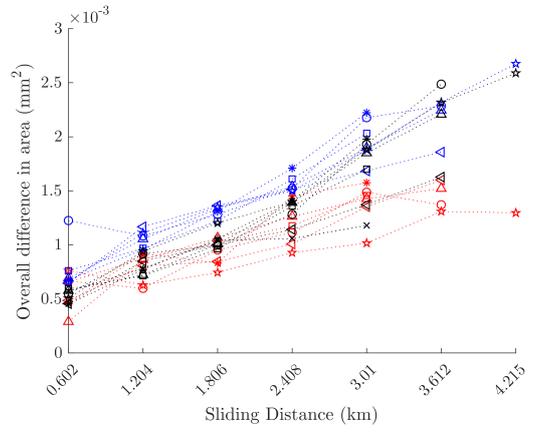
(e) Grade A - 8



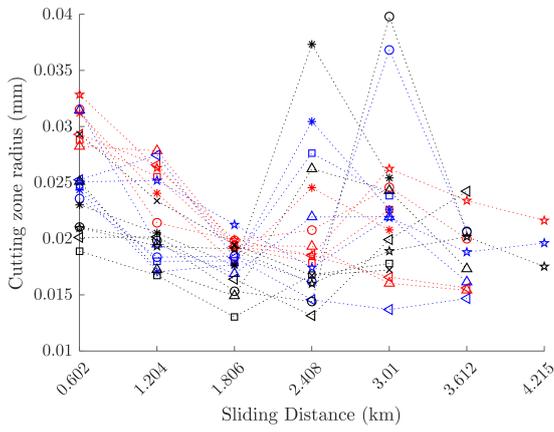
(f) Grade A - 9



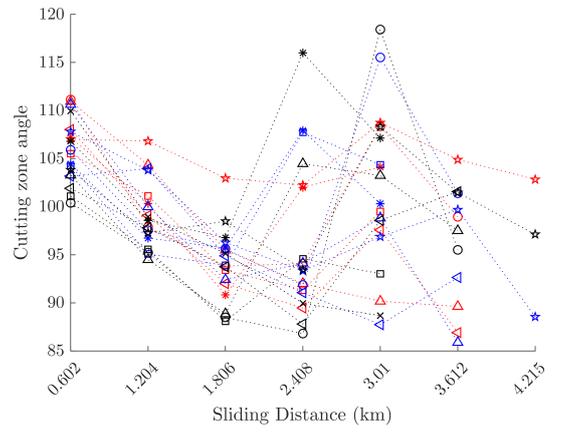
(a) Grade A - 10



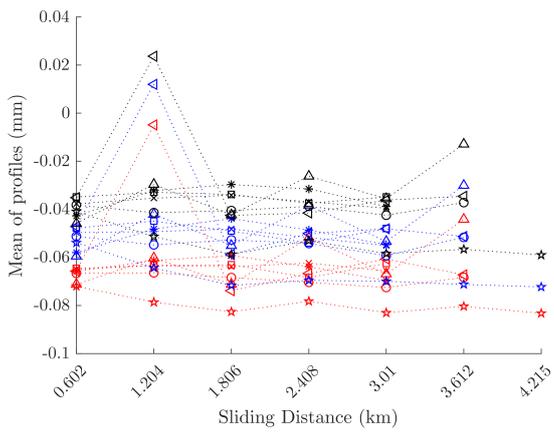
(b) Grade A - 11



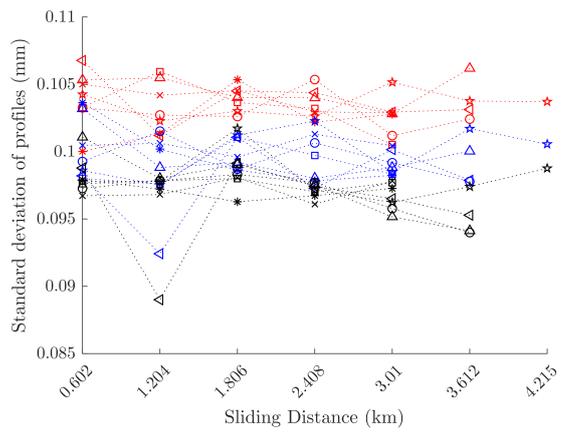
(c) Grade A - 12



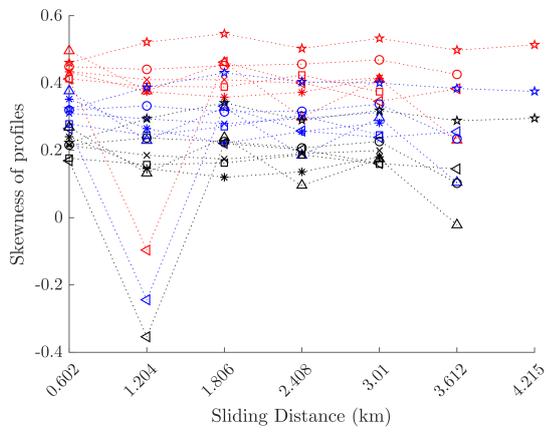
(d) Grade A - 13



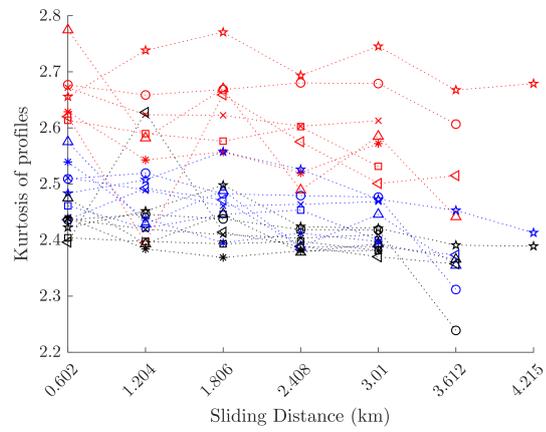
(e) Grade A - 14



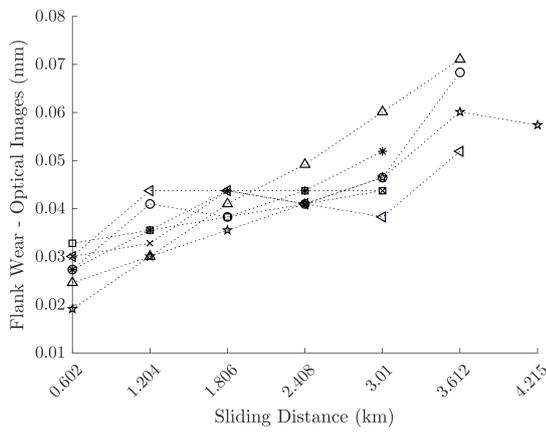
(f) Grade A - 15



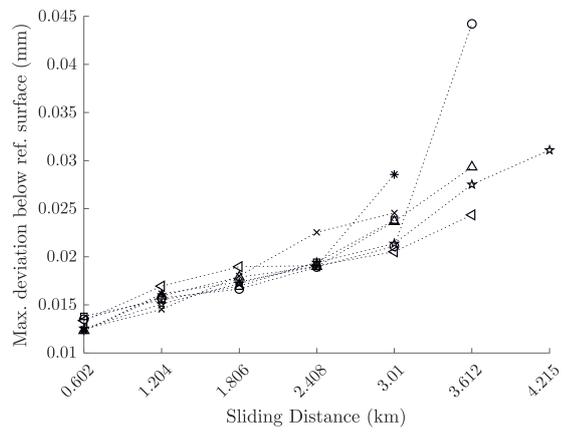
(a) Grade A - 16



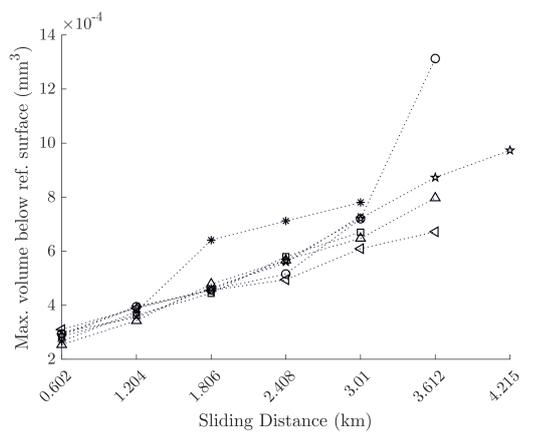
(b) Grade A - 17



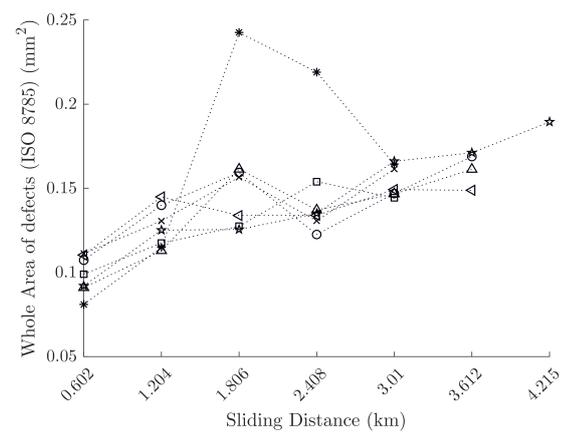
(c) Grade A - 18



(d) Grade A - 19



(e) Grade A - 20



(f) Grade A - 21

---

## RESULTS FROM THE DIRICHLET PROCESS

The results from the Dirichlet process clustering that were not presented in Section 7.3 have been presented here.

In all cases, the DP clustering of AE data where the DP is running globally is shown in the top part of the figure and where the DP was reset at the point of workpiece change has been shown by the bottom part of the figure. The change of workpiece has been indicated with the dashed black lines. Only features 1 and 2 have been shown for clarity.

The legend for all figures presented in this appendix is here:

● Cluster 1   ● Cluster 2   ● Cluster 3   ● Cluster 4   ● Cluster 5   ● Cluster 6

Figure B.1 shows the results for tool A2. Following the guidelines set in Section 7.3.1, this tool would be investigated at around 1km, then around 2.1km where the tool may be retired by the operator as it is approaching failure.

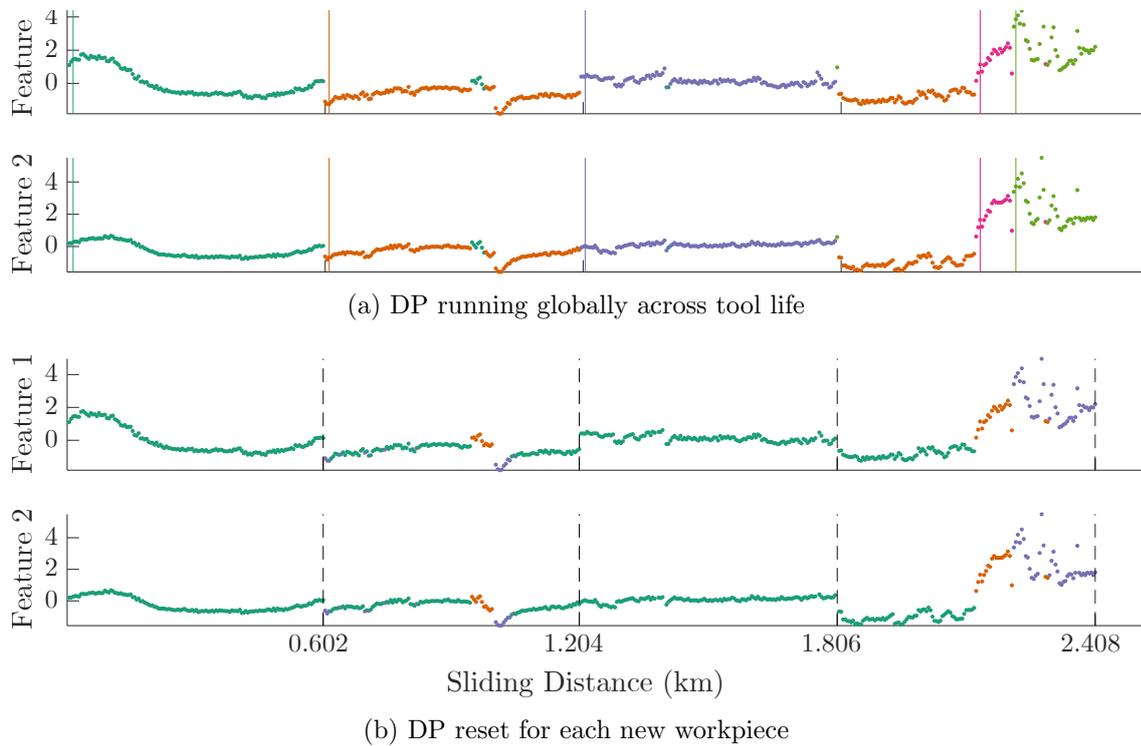


Figure B.1: DP results for tool A2.

Figure B.2 shows the results for tool A3. Again, following the guidelines specified in Chapter 7, this tool may be inspected at around 600m, and 3km, at which point, the tool would probably be retired.

The DP clustering results for tool A4 has been presented in figure B.3. Here, the tool will be investigated at sliding distances 200m, 2.35km, and finally retires after the inspection at around 3.1km.

For tool A5 in figure B.4, the cluster initiations at sliding distances 1.6km, possibly 2.2km (though there is no cluster initiation at 1.8km) and 2.6km will warn the operator to investigate the tool.

Figures B.5 and B.6 presents the DP clustering of tools A7 and A8 respectively. For tool A7, the clustering may result in avoiding tool failure if inspected at 3.4km according to the guidelines set in Section 7.3.1. For tool A8 however, there intensity of the features increase during the middle of tool leading to data near the end of tool

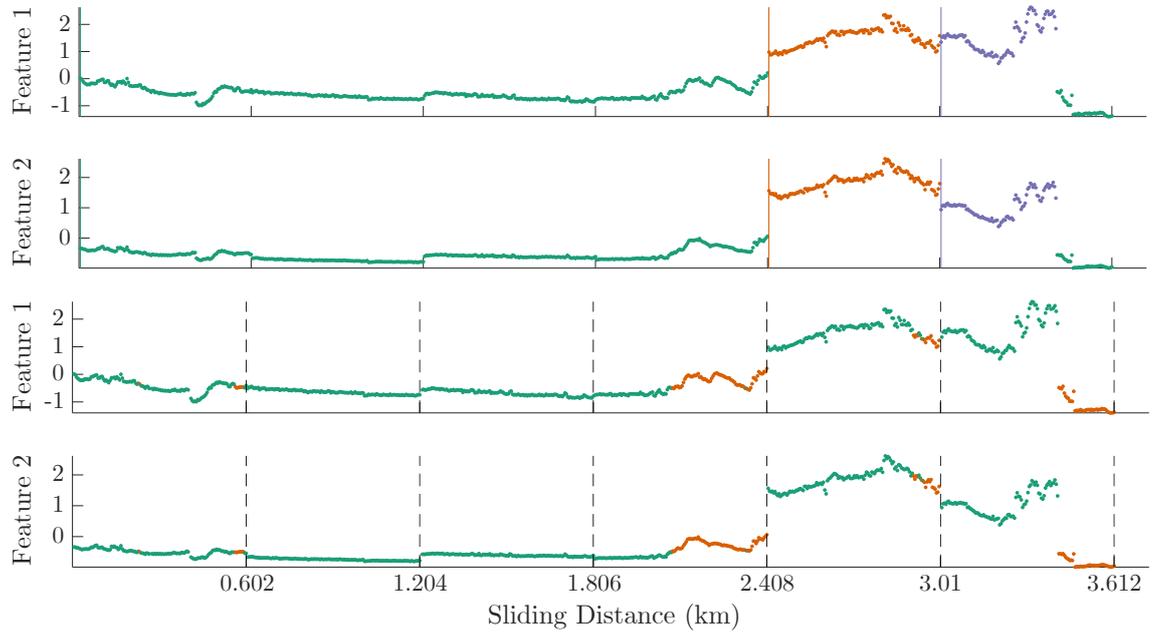


Figure B.2: DP results for tool A3.

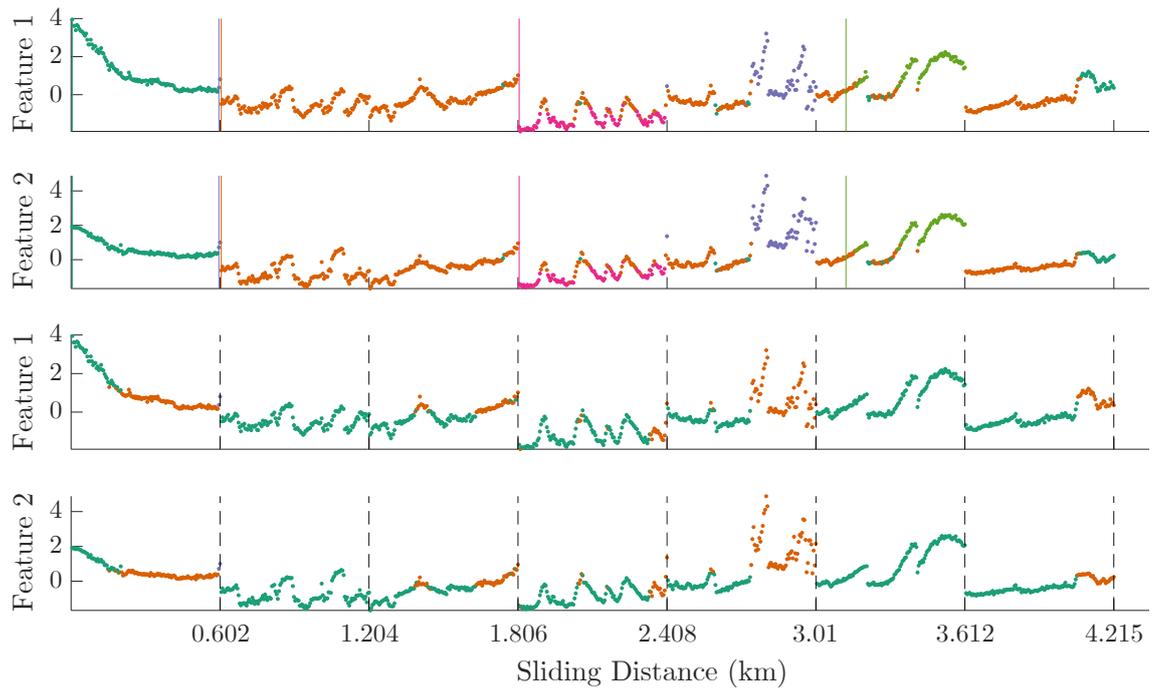


Figure B.3: DP results for tool A4.

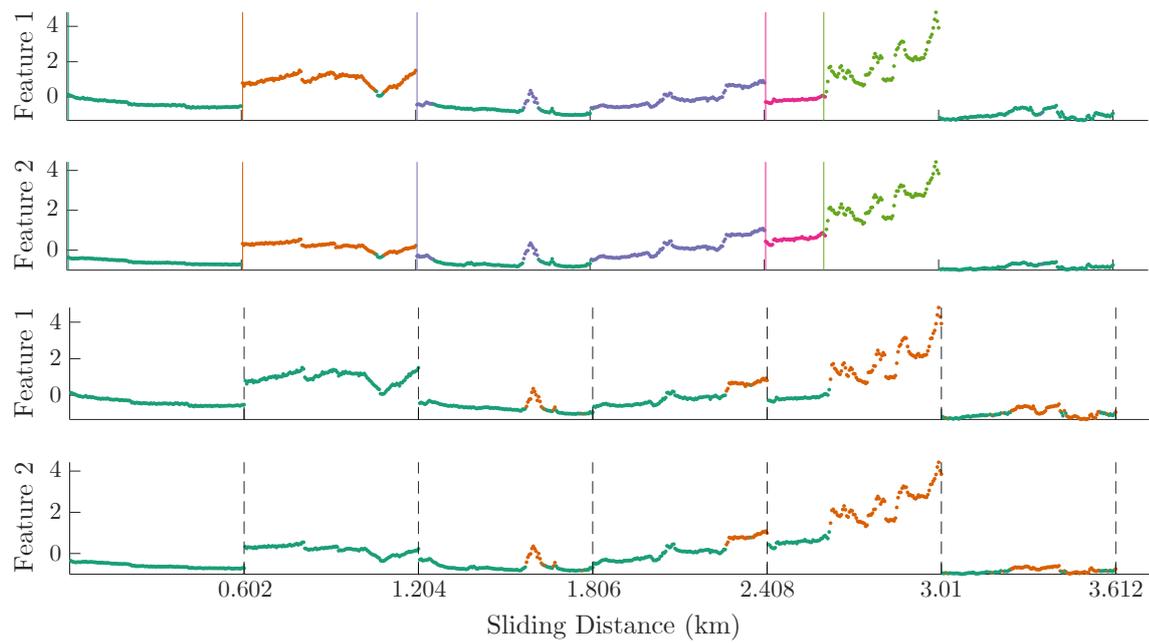


Figure B.4: DP results for tool A5.

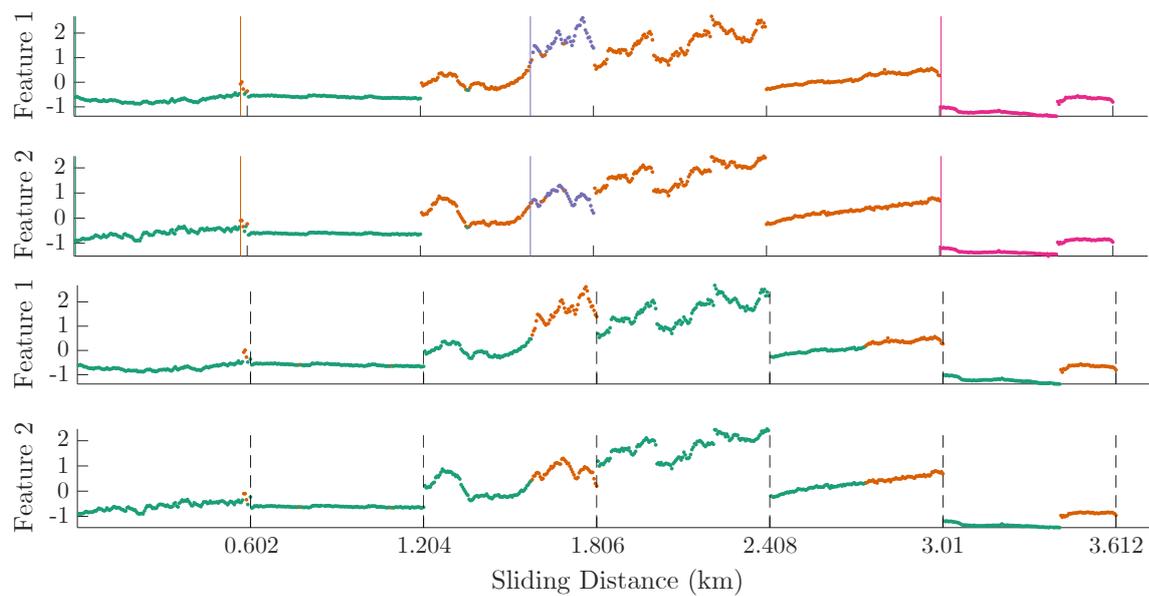


Figure B.5: DP results for tool A7.

life clustering into earlier clusters seen by the algorithm. According to the guidelines described in Section 7.3.1, this tool would not be retired in time.

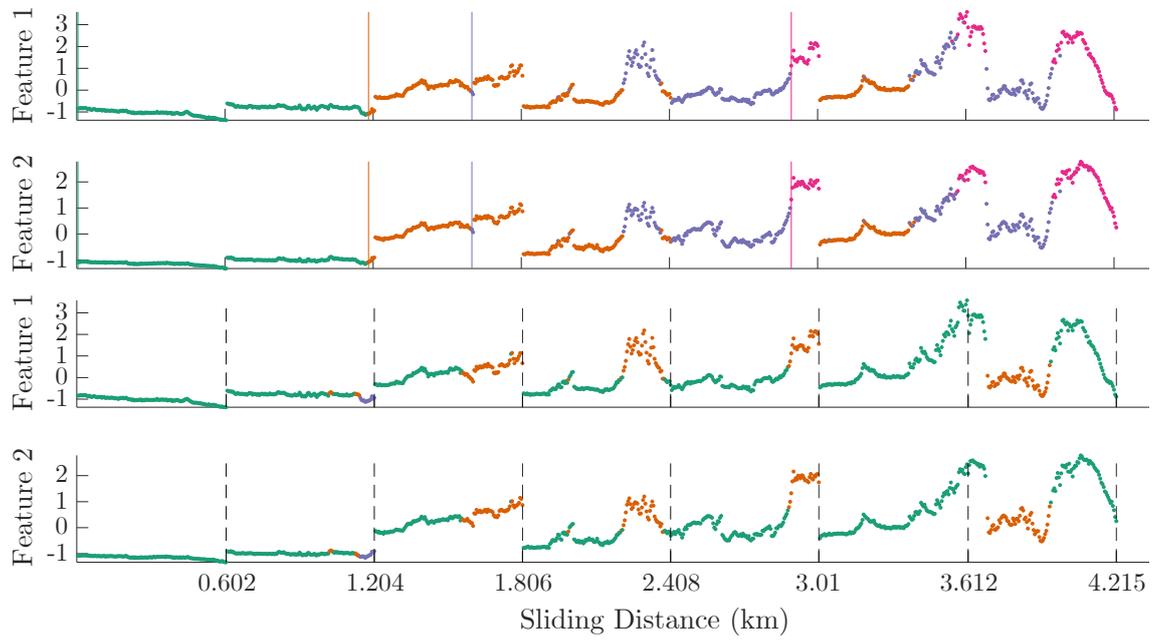


Figure B.6: DP results for tool A8.

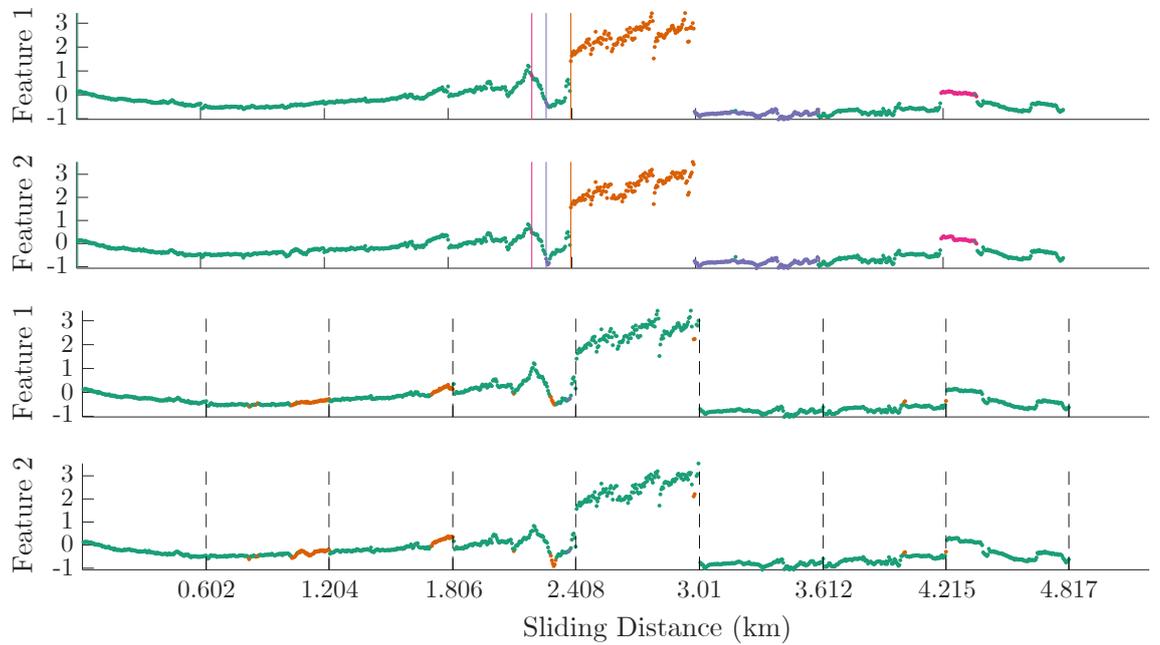


Figure B.7: DP results for tool A9.

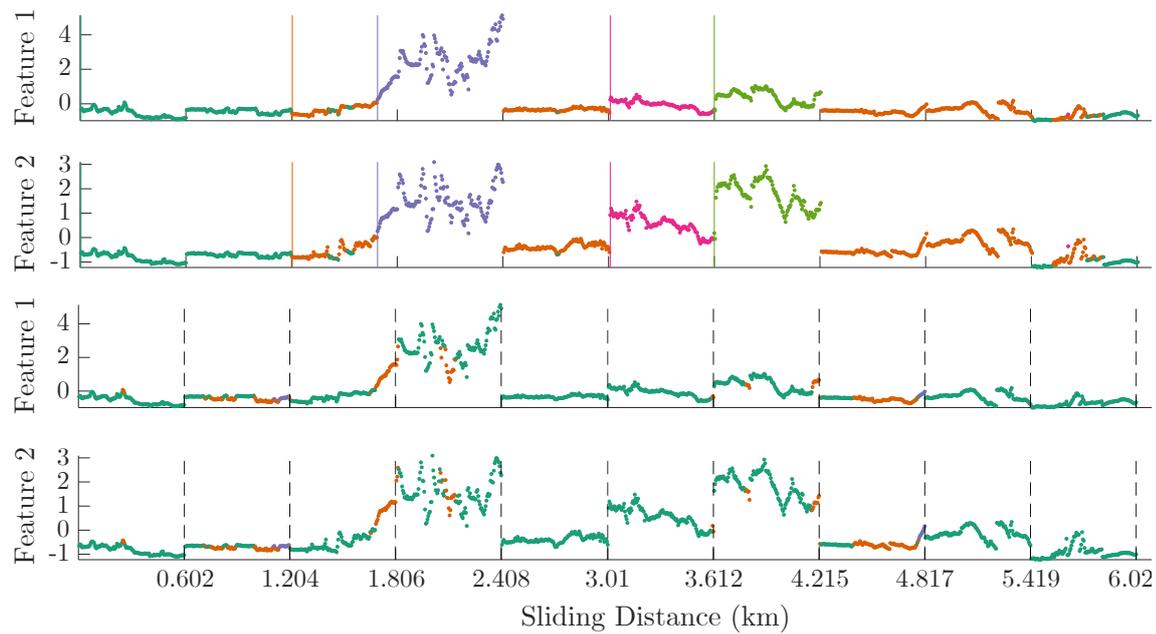


Figure B.8: DP results for tool B7.

Figure B.7 presents DP clustering results for tool A9. This was an unsuccessful clustering process due to the increased intensity of the signal features around 2.4km to 3km.

The clustering results for tool B7 is presented in figure B.8. This result was again not a successful, meaning that the operator is likely to machine with this tool until it reaches catastrophic failure if the guidelines in Section 7.3.1 is followed. This is because, once again, there is an increase in intensity in the features during the middle of the pass.

---

## DWT - A BRIEF INTRODUCTION

The Discrete wavelet transform (DWT) allows time series to be analysed and visualised in both time and frequency domain without having any data loss. When the data is transformed, the signal retains the number of data points that were in the original signal. As a result, the original data can be reconstructed from the wavelet coefficients.

Wavelet coefficients are found by using a series of high pass and low pass filters that split the data evenly into approximate and detailed coefficients; high pass filtering results in the detailed components and low pass filtering in approximate components. At each step of filtering, half of the data are removed due to Nyquist criteria. Depending on the level of decomposition, the approximate coefficients are put through these filters until the complete decomposition is achieved. Figure C.1 presents a visualisation of a level 3 DWT breakdown. Here if the signal was sampled at 1MHz and contained  $1 \times 10^6$  data points, D1 and A1 would contain  $5 \times 10^5$  data points each after the decomposition. A1 would contain frequencies from 0 - 250kHz whereas D1 would contain the information from 250 - 500kHz.

For more detail, the reader is referred to [136].

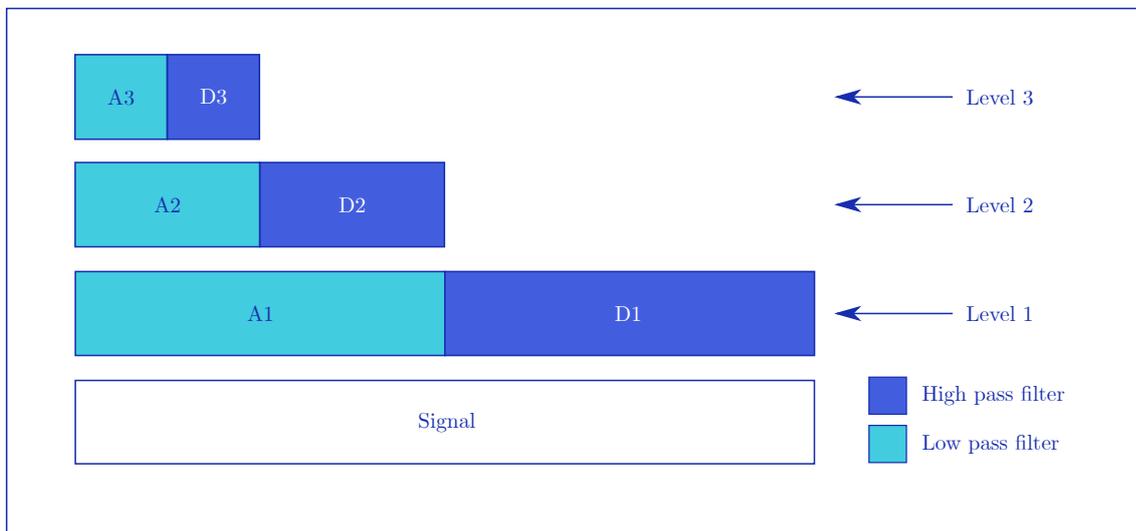


Figure C.1: Discrete wavelet decomposition presented visually.