

Application of machine learning for acoustic emissions waveform to classify galling wear on sheet metal stamping tools

James Marcus Griffin (✉ ac0393@coventry.ac.uk)

Coventry University <https://orcid.org/0000-0002-9179-5130>

Vignesh. V. Shanbhag

Norwegian Research Centre AS

Michael. P. Pereira

Deakin University

Bernard. F. Rolfe

Deakin University

Research Article

Keywords: Sheet metal stamping, Galling, Acoustic emissions, Mean frequency, and Machine leaning

Posted Date: May 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-186756/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Galling wear in sheet metal stamping processes can degrade the product quality and adversely affect mass production. Studies have shown that acoustic emission sensors can be used to measure galling. In the literature, attempts have been made to correlate the acoustic emission features and galling wear in the sheet metal stamping process. However, there is very little attempt made to implement machine learning techniques to detect acoustic emission features that can classify non-galling and galling wear as well as provide additional wear-state information in the form of strong visualisations. In the first part of the paper time domain and frequency domain analysis are used to determine the acoustic emission features that can be used for unsupervised classification. Due to galling wear progression on the stamping tools, the behaviour of acoustic emission waveform changes from stationary to a non-stationary state. The initial change in acoustic emission waveform behaviour due to galling wear initiation is very difficult to observe due to the ratio of change against the large data size of the waveform. Therefore, a time-frequency technique "Hilbert Huang Transform" is applied to the acoustic emission waveform as that is sensitive to change of wear state, and is used for the classification of 'non galling' and the 'transition of galling'. Also, the unsupervised learning algorithm fuzzy clustering is used as comparison against the supervised learning techniques. Despite not knowing *a priori* the wear state labels, fuzzy clustering is able to define three relatively accurate distinct classes: "unworn", "transition to galling", and "severe galling". In the second part of the paper, the acoustic emission features are used as an input to the supervised machine learning algorithms to classify acoustic emission features related to non-galling and galling wear. An accuracy of 96% was observed for the prediction of non-galling and galling wear using Classification, Regression Tree (CART) and Neural Network techniques. In the last part, a reduced Short Time Fourier Transform of top 10 absolute maximum component acoustic emission feature sets that correlates to wear measurement data "profile depth" is used to train and test supervised Neural Network and CART algorithms. The algorithms predicted the profile depth of 530 unseen parts (530 unseen cases), which did not have any associated labelled depth data. This shows the power of using machine learning techniques that can use a small data training set to provide additional predicted wear-state on a much larger data set. Furthermore, the machine learning techniques presented in this paper can be used further to develop a real-time measurement system to detect the transition of galling wear from measured acoustic emission features.

1. Introduction

In recent decades, condition-based maintenance has evolved from visual inspection methods to automated inspection methods. Automated based methods include use of advanced signal processing techniques and Machine Learning (ML) techniques. Automated based methods collect sensitive information from machines or tools [1], [2], whereas human inspections are sometimes prone to error. The sensitive information regarding health of machines or tools can assist in determining root cause of failure and reducing machine down time[3]. However, real-time tool monitoring systems need consistent data and can be subject to issues of sparse or missing data, or imbalanced data. These issues need to be

catered for otherwise, automation will fail [4]. This point is particularly important when dealing with large streams of information.

Every year, global automotive production exceeds more than 60 million vehicles. Each vehicle has hundreds of sheet metal components. Even a small improvement in maintenance strategy of sheet metal stamping process can increase the cost efficiency for the automotive industry. To manufacture light weight vehicles, there is also an increase in trend of using advanced high strength steels (AHSS) and ultra-high strength steels (UHSS). This has resulted in increased forming forces and galling wear on stamping tools causing premature failure of stamping tool. Visual inspection of stamping tools at regular intervals is both time consuming and impractical. Considering the huge cost involved in stamping tools, condition based maintenance is very much required for automotive industries to reduce machine down time and increase cost efficiency [5]. Therefore, attempts have been made in the literature to understand wear on the stamping tool using different sensors.

In the literature, sensors have been used for in-situ monitoring of the stamping tools or to distinguish wear profiles of the stamping part. In situ sensing applied to stamping tools/parts typically involves strain gauges to read either strain directly or different axial loads. Xu et al. [6] investigated strain experienced during the stamping process. Two strain gauges were used to obtain more uniform results. Daubechies Wavelets (Db 4 and Db 5) were used to provide time and frequency-based information to differentiate different states of process anomalies. Hidden Markov Models and probability density functions were used to predict anomalies. Hidden Markov Models and probability density functions have disadvantages in that they cannot express dependencies between hidden states which suggests they are poor for transparency and visualisation which is an important consideration for the work presented here. Bassiuny et al. and Ge et al. [7], [8] also carried out further investigations into using strain gauges for monitoring stamping processes. Bassiuny et al. [7] used frequency information to distinguish from normal state where higher frequencies are experienced to both mis-feed and the too thick state of workpiece. In addition, Hilbert Marginal Spectrum features obtained from analysing strain waveform was used as an input to the Learning Vector Quantisation Neural Network. Using this technique, it was possible to distinguish anomalies in the stamping process. Learning Vector Quantisation however usually require a pre-processing layer similar to a Self-organising map or k-means algorithm and therefore considered too complex when considering visualisation and transparency. Ge et al. [8] used Support Vector Machines (SVM) to distinguish features from strain waveform data. The SVM technique was preferred because it performs well when presented with low data sets. SVMs however are more complex and verbose when considering multiple outputs [9] which is why they are not appropriate for the work presented here. The use of a strain gauge however lacks resolution and is often difficult to distinguish between one anomaly from another. Hence the requirement for other sensing technologies especially those combined as a multispectral approach. Garcia [10] used another technique based on the use of digital camera and applied optimised wavelet to distinguish wrinkles and surface roughness by extracting 2D images. In summary, machining learning techniques have been applied to sensor data, such as strain data obtained from strain gauges, in sheet metal stamping. Also other machine learning techniques have been applied to distinguish the evolution of scratch forming with ball-on-disk sliding based on input parameters [11].

The machine learning technique used fuzzy clustering and quantum-behaved particle swarm optimisation to provide accurate and efficient predictions. It is clear there is research to predict scratch formations however not for the onset of galling where galling is a very minuet process and standard sensory quantities such as strain gauges will not have the resolution to see such effects and changing characteristics. Therefore, there is a need for machining learning to be applied to a more sensitive material measurement process.

Within the above mentioned works the sensors used to provide damage mechanisms lack in information accuracy, precision and resolution where there is a need for using more sensitive measurement technologies to provide more information and allow preventative maintenance as opposed to failure reactive maintenance. Having such capabilities help to provide extension of live where material strengthening or damage recovery can be carried out as part of maintenance. With total or near total failure, the part is simply swapped out. Such ideas for preventative maintenance fit in with our need to reduce energy and carbon emissions. One sensor that provides more information and has been used before in stamping tests is acoustic emission. To the knowledge of the authors there has not been any work in measuring acoustic emission from stamping tests and applying it machining learning techniques to gain visualisations, classifications and predictions of damage mechanisms such as the onset, and established wear. Furthermore there has been no attempts to use machine learning techniques to provide automation in detection and preventative maintenance for the stamping process using acoustic emission measurements. The reason behind this can be down to the fact that real time detection of tool wear in a slow mechanical process is very challenging, especially when using AE with wideband sensors where changes are very small in nature compared with the total amount of data extracted.

A large number of researchers have reported the application of Neural Network (NN) models for the tool condition monitoring data to classify tool wear in turning [12], [13]. Turning however has a lot of similarities with sheet metal stamping where scratches or galling can occur if the conditions are right. NNs are very good for low data sets as well as good and accurate visual output tools. Not to mention their prediction and classification capabilities which score fairly high when ranked against similar supervised methods.

In terms of AE being used in tensile tests (slow varying mechanical change), Godin et al. [14] looked at using k-means and self-organising maps in segregating different mechanisms of material failure through different AE waveform fingerprints. Such AE is reduced in n-dimensionality to give the rise time, peaks and counts [14]. By using these reduced AE parameters, only three values are required as opposed to a whole AE signature. These reduction techniques coupled with the Short Time Fourier Transform (STFT) of the AE signature has been seen in the AE literature before and specifically, applied to stamping tests [15], however the use of ML techniques have not been used to date. In comparing the two machine learning techniques, self-organising maps are less prone to local optima than k-means as k-means can suffer from premature convergence. That said, other unsupervised techniques such as fuzzy clustering take information of all the surrounding clusters to calculate the best distance for the point of interest [16].

More recently studies have focussed on wear mechanisms experienced in micro milling to quantify how AE can be used to distinguish such microscopic phenomena [16]. Ren et al. [17] also looks at other precision machining processes where fuzzy identification can accurately measure material removal rates using extended subtractive cluster analysis and adaptive filtering techniques, which when tuned, gives the process more accuracy against unwanted noise [17], [18]. This is another reason why fuzzy clustering is considered a better visualiser/classifier when compared with self-organising maps and k-means.

The tooling insert wear mechanism has a defined cutting edge and can partially represent the mechanics of single grit cutting seen in scratch tests replicating wear mechanisms as seen in tribology studies. Venkatesh et al. [19] predicts insert wear through NN models using the input of time, velocity, feed and cutting force. In other studies [20] ML techniques were applied to the AE data recorded from the scratch tests and is directly applicable to the work proposed in this paper. Moreover, the precision of AE technologies applied to wear can also be directly related to material removal mechanisms achieved during stamping. Based on the success of ML application for the AE data, in this work, ML techniques are used to classify AE data related to galling wear in sheet metal stamping process.

AE features have shown interesting tendency to identify galling wear at the very initial stages, that is, much prior to wear that is visually visible [15], [21]. Further work using AE identified the various stages of galling wear and focused on sensitivities of AE features that would indicate both the initiation of wear and severe wear on stamping tools [22], [23]. By applying ML techniques to AE features and understand the transition from non-galling to galling wear seems to afford many new insights not offered by previous means. This paper's work was inspired from previous works where AE sensors have been used to study wear mechanisms and source location without complex signal processing and data orientated algorithms [7], [8], [10], [12], [13], [16]–[18]. Time-frequency techniques to study acoustic emission waveforms for the fault diagnosis of machining processes and the in-service operation of bearings has been previously investigated [9–16], [22]–[29]. Condition monitoring of stamping processes using acoustic emissions discussed in [1], [15] is applied with the work presented in [17, 18] [30], [31] where Hilbert Huang Transform (HHT) provides a method to track the state from both stationary and non-stationary data. In displaying such information, it is possible to show a better understanding of the onset, transition and severe galling wear condition. This application of classifying non-galling and galling wear through using ML techniques applied to AE data is the identified knowledge gap in the literature and needs further research.

In this study, data obtained from AE sensors in previous work [15] is used to segregate unworn and worn stamped parts using ML techniques. AE waveforms were analysed using a number of time-frequency techniques to determine a suitable technique to study the wear behaviour. To ensure the data is more salient for automation of wear initiation; ML techniques were used. With a limited data set of destructive tests (namely depth profile measurements) it was possible to predict the remaining depth profile measurements from just AE signatures and non-tested cases. The non-tested cases are a much larger data set. This holistic approach conforms to a non-destructive testing technique. AE and ML were selected, AE because of the wide bandwidth data and is often difficult to identify the wear features of

interests that can correlate from one pattern to another. With ML techniques it is possible to segregate the different conditions of interest useful for maintenance control.

The ML techniques that are used for the classification in this work are Neural Networks (NNs) and Classification and Regression Trees (CART) as they both perform well when presented with small data sets and they are supervised classification techniques. The other technique to obtain good coverage of ML techniques is fuzzy clustering, which also responds fairly well to small data sets and is unsupervised learning in nature. The three techniques not only give a good coverage of ML techniques but also provide a comparison of supervised versus unsupervised learning, which is not common in literature especially when applied to tool wear. Apart from regression trees, both NNs and fuzzy clustering have already been discussed in the introduction where researchers applied this to Tool Condition Monitoring (TCM) and this is a further reason why they have been studied in this paper.

To distinguish different wear mechanisms which are non-linear in nature there is a need to use ML techniques that allow the visualization of such behaviour and this is another reason why these techniques have been chosen over others. To extract out the salient minut behaviour digital signal processing techniques such as HHT were applied to acoustic emission measurements before being input to the machine learning techniques. Using these transforms as a pre-processing layer is another unique method when introducing data to machine learning or cognitive layer. Finally, load has been introduced as a quantity to compare and contrast sensor technology sensitivities as well a known source to calibrate acoustic emission.

The rest of this paper is organised as follows: Chap. 2 discusses the experimental setup and algorithms, where both AE and ML are discussed in greater depth; Chap. 3 discusses the AE signal to physical data correlation; Chap. 4 discusses the classifier results applied to the signal and physical data.

2. Experiments And Methodology

2.1 Stamping setup and materials

The experiments were performed on a semi-industrial stamping setup [32], which closely replicates the continuous stamping process employed in the automotive industry. This semi-industrial setup uses a progressive die set that is typically installed in a mechanical press. Channel-shaped parts are produced from the tooling used in the stamping setup [15]. The AE sensors were mounted on the die inserts to study the change in AE signal energy due to change of wear state of the dies. Removable die corner inserts were used in this study to allow the visual examination of wear of the dies during the stamping tests. The accelerated tests were adopted in this study to allow examination of all the non-galling, transition of galling wear and galling wear on the stamping tool in short duration. The experimental process parameters used in this study were kept constant and are summarised in Table 1. It was considered too resource intensive to measure each individual stamped part, therefore every fifth part was measured,

where the sidewall surfaces of the stamped parts were examined to analyse the wear behaviour of the die radii surfaces.

Table 1: Process parameters

Lubrication	Anti-corrosive oil
Punch width	30 mm
Die to punch gap	2.35 mm
Die corner radius	5 mm
Punch radius	5 mm
Blank size (L × W × t)	150 × 26 × 1.6 mm
Draw depth	40 mm
Average blank holder force (h_f)	28 kN
Press stroke rate	32 strokes per minute
Number of parts formed	600
Die Corner insert materials	AISI D2 steel hardened to 60 HRC
Blank material	Bluescope Steel; grade: XF300; thickness: 1.6mm
Anti-corrosive oil	Supplier: Quaker Chemical; product label: Ferrocote 366 K2 50

2.2 AE data acquisition and analysis

According to Pereira et al. [32], wear of the stamping process is most severe at the die corner radii. The two wideband AE sensors used in this study has a frequency range of 20 kHz to 2.5 MHz (supplier: Vallen Systeme; model: AE2045S), however only 2 MHz was used for sampling as this was considered appropriate for recording the specific phenomena of interest. The AE sensors were connected to the data acquisition system (supplier: National Instruments, model: PXIe-1078) via a high speed digitiser (supplier: National Instruments, model: DCPL2) and an amplifier (supplier: Vallen Systeme; model: AEP3N) with a gain of 40 dB. The data was recorded in increments of every 5 parts until the final part.

AE analysis was performed on the AE waveforms collected for each part and for both sensors used in this study. This was mainly performed to ensure repeatability of time domain analysis, time-frequency analysis, frequency-based feature studies, and, analysis of whether the AE waveforms were stationary or non-stationary. To understand the frequency range associated with the wear mechanisms, the entire AE signal of the process was analysed rather than analysing only the burst AE signal. Segments of AE signal were then translated from time domain to time-frequency to study if the waveform has changed from stationary to non-stationary which gives us information regarding indicative of the transition of galling tending towards severe galling. In the literature, different time-frequency techniques have been used for fault diagnosis of stamping tools [5], [7], [22], [23] however none has been indentified before earlier work displayed in [15]. Already Wavelet Packet Transforms, Short time fourier transform, and HHT were discussed extensively in the literature [12], [26]–[31], [33], only the application as rich data summary technique providing ML input will be discussed here. Discussions were made in [15] to understand the compatibility of aforementioned techniques with understanding the wear condition of the stamping dies. To display verification of ML techniques where possible both classifications of Sensor 1 and Sensor 2 (which relates to Die 1 and Die 2) will be used to show similarites and correlations within the dataset therefore providing a high confidence from a data recording perspective. If similarities and correlarions

were not present it could be concluded the data recorded by each individual sensor has associated errors, such as poor attachment of the AE sensor with the dies when compared to the other sensor.

It is known that the performance of the time-frequency techniques is strongly dependant on whether the AE waveforms are stationary or non-stationary [34]. Therefore, using the statistical technique proposed by Papoulis et al. [35], the stationarity of the AE signal is first examined. This technique (HHT) is summarised in [15] as well as used to determine the non-stationary components within the extracted AE waveforms. Using HHT AE data is central to the ML technique to segregate no wear, on set of wear and, tending towards severe wear.

Also, load data was obtained in terms of downward force and the same setup and apparatus was obtained from [36]. For the work presented here, the load data gives a standardised view of waveforms for correlation with AE waveforms. The force, however, is less sensitive change and serves as useful waveform scale connector to compare against.

2.2.1 Data visualisation methods

There are many ways to connect sensor information with physical information however some methods are more sensitive to change than others. For example, mean frequency of the signal displays a more salient trend than peak voltage. Other techniques such as STFT [37] and HHT are applied to the raw signal and the resulting information is transformed from a time series waveforms to both time-frequency which provides more useful data to discern features of interest. For example, amplitude only gives a dynamic representation of the changing power whereas frequency also provides more information in terms of failure [15]. The higher the amplitude and frequency not only distinguishes failure but also what type of failure is taking place [15]. HHT however takes this further where stationary waveforms based on normal operating conditions remain constant and non-stationary waveforms significant of change related to failure and provides a clear distinction from the normal [15].

2.3 Machine learning techniques

Based on the discussed literature in the introduction section, three techniques that favour the recorded AE waveforms are fuzzy clustering, CARTs and NNs will be discussed in greater depth within this subsection. These techniques have been chosen on their support for small data sets as well as allowing a general comparison of supervised vs. unsupervised techniques.

2.3.1 Clustering techniques for segregating stationary and non-stationary data

Clustering techniques are unsupervised in terms of learning where the relationship is based on the data structure between all of the presented data. The clusters were assigned based on a fuzzy measure where the shortest distance measure found to a cluster centre places the data in that cluster compared to other cluster centres. Because there are no labels assigned with the data set, we have categorised our data by

arbitrary sequence numbers to provide three classes: "no galling", "transition to galling", and "severe galling". In the literature [38] the unsupervised technique appears to be more efficient and representative of the data structure when compared to its supervised counterpart.

Clustering techniques have emerged from work carried out in statistical probability [39]. When looking at real world phenomena most cases are not finite and instead, possess a lot of continuum values (soft sets). Fuzzy clustering-mean algorithm used around 170 iterations to find the optimised clusters for the data presented. This technique was used for grouping data and finding structures in data. Fuzzy clustering provides rules in the form of distance measurements that segregate the different cluster sets from each other, in this case; AE or force parameters.

After ranking the features in the order of similarity values, it is then possible to segregate these features using the closest cluster distance membership function and distinguish the AE or force data in terms of distance values (See Fig. 1, calculate centres and distances). The fuzzy algorithm iterates through Fig. 1 flow chart until it can no longer improve the separation of one cluster from another (optimisation).

2.3.2 Classification Tree techniques for segregating stationary and non-stationary data

The CART algorithm is another supervised classifier technique that is particularly useful in segregating n-dimensional data sets and it produces a transparent, easily readable set of classification rules.

CART builds classification and regression trees for predicting continuous dependent variables (regression) and categorical predictor variables (classification) [40]. Using Eq. (1) the feature space of AE waveforms is recursively split to find datasets of non-overlapping regions. This equation aims to give the most optimised tree for classification and prediction. Prediction is made from the rules obtained from recursively fitting the training data.

A good splitting criterion is the following:

$$PRE = \emptyset(s,t)$$

Misclassification error:

$$Q_m = \frac{1}{N_m} \sum_{x_i \in R^m} (y_i \neq k(m)) = 1 - \hat{P}_{mk(m)} \quad (1)$$

Where, y_i is the output of the individual under test and $k(m)$ is the class category under test.

PRE is the minimum production reduction in error and 's' is the split at any 't' node when applied to a tree structure, which is constructed based on the data presented (see CART rules 1 (Table 4) and 2 (Table 7)

for example output code and Fig. 2 for flow chart of CART process). The best purity measure looks at the best unique class classification where less impure looks at more multiclass representation. For the CART algorithm, the percentage accuracy of classifications is used as the best purity measure.

This method of classification is chosen because the tree fitting methods are actually closely related to cluster analysis [41]. This is where each node can be thought of as a cluster of objects, or cases, that are split by further branches in the tree. Note that the top node covers the whole sample amount, and each remaining node contains a sub amount of the original sample and so on as the split levels increase.

Looking at the results section of CART analysis, the pseudo code displays how the input data of AE (Max and Min parameters) are used to distinguish 0 no wear, 1 transition to wear and 2, severe wear. The values used to give 'if else' statements are to segregate the data accordingly to provide the correct classifications. From the introduction it should be clear why NN and fuzzy clustering techniques were used as classifiers. For CART however this is based on one of its features, robust to outliers [40]. All tests carried out using this technique were verified against test and verification unseen data sets. With high accuracy classifications, added confidence in terms of the accuracy was achieved.

2.3.3 Neural Networks used as prediction classifiers when using limited data

A large number of researchers have reported the application of using NN models for the classification of phenomena of interest for tool condition monitoring [13]. A feed-forward NN model was used with the back-propagation learning strategy to provide the segregation of data [42]. The parameters for the used neural network were as follows listed in Table 2:

Table 2
Neural Network Parameters

Learning rule	Backpropagation (trainrp & learnk)
Size of input layer	40 (tan-sigmoid)
Size of hidden layer	40 (tan-sigmoid) and 3 o/p layer (pure-linear)
Number of hidden layers	3 including o/p layer
Learning rate	0.1e10
Momentum	0.9
Transfer functions	I/P, H/L1 and 2 = Tan-sigmoid, O/P = pure-linear
Sum Squared Error	2.36 *e10 -25

This method segregates the different classes based on the supervised training data given to the NN. The summation of weights and bias values are multiplied by a differential transfer function to give a neuron output [42]. See Fig. 3 for flow chart for AE Max and Min HHT data applied to NN paradigm.

3. Signal To Physical Data Correlation

This section introduces some of the key AE parameters used to visualise and classify the different physical wear states which are defined from surface profile measurements.

3.1 Correlation between time domain features and profilometry measurements

It was noticed from literature [15] that the AE Root Mean Squared (RMS) tends to increase when the scratch increases more than 15 µm (Figs. 4a and 4b). In addition, it was noticed that both RMS (V) and Peak (V) were consistent in terms of the two sensor data sets where both the unworn parts started to increase from 170 parts onwards. Another useful view of AE waveforms is the Mean Frequency (MFreq) displayed on the second y axis of Fig. 4a. Mean Frequency gives more trend orientated data where values decrease as the wear state increases. In Die 1 the transition of galling was much earlier when compared with Die 2 (210 vs 280 parts respectively). The AE time and frequency domain information therefore captures such features (see Fig. 4a).

4. Classifier Results Connecting Signal To Physical Phenomena

This section looks at applying classifier techniques mentioned in Sect. 2.3 where the ‘unsupervised learning’ fuzzy clustering and, ‘supervised learning,’ CART and NN algorithms are used to segregate no galling, transition to galling and finally, severe galling. The accuracy of data segregation/classification will be compared to physical phenomena measurements to display how confident the findings are for potential commercial exploitation.

4.1 Fuzzy clustering analysis

Figure 5. shows that as the state of galling changes from no galling to the transition of galling, the max amplitude force increases and at the same time the mean frequency of the force time-series waveform slightly decreases. The unsupervised learning has created three clusters, (Green, Blue, and Red). The authors have then arbitrarily applied their knowledge of the system, where the part numbers equal and less than 260 show no wear, part numbers between 265 and 340 start to have wear effects, and those above 340 often exhibit severe wear. These values of force and force mean frequency are different to AE due being less sensitive to changing state of wear. The unsupervised green cluster would align with parts that display no galling, the blue cluster would appear to correspond to the parts in transition to galling, and the red cluster would appear to correspond to parts tending towards severe galling. Figure 5 also indicates that the observed force amount was more sensitive to a change in the state of wear than mean frequency of the extracted force signal. It can be concluded that as a differentiator there is change with the mean frequency, however it is only small in significance. Variance of the force signal with force amplitude gave a similar segregation as Fig. 5 and provided no additional information. When compared with physical measured data of Fig. 4b taken from [15] the max penetration depth is where part number

210 and below is consistent with no galling, above this and below part number 310 is the transition to galling and beyond part number 311, severe galling.

Figure 6. displays the fuzzy cluster distances of the three unsupervised learning clusters where this time, the green cluster refers to expected non galling, the blue cluster is the expected transition of galling and finally the red cluster is the expected parts tending towards severe galling. It should be noted that as well as outliers where such cases are closer to other cluster centres, the grouping of the distances are compact in nature and therefore with just force alone it is difficult to distinguish the transition of galling.

Table 3
Signal utilisation of clusters

Cluster Focus	Green: No galling (%)	Blue: transition of galling (%)	Red: Severe galling (%)	Transition number (210th part) and signal max and min values**	Accuracy to arbitrary predicted phenomena (%)
Force MFrEq. and Amplitude	280 (46.6)*	110 (18.3)*	210 (35)*	(265) 18.3/48.4	91
AE HHT Max and Min	255 (42.5)	230 (38.3)	115 (19.2)	(220) 0.56/-0.58	98.3
AE Max Min Amplitude	255 (42.5)	230 (38.3)	115 (19.2)	(220) 0.56/-0.58	98.3
Max AE and rise time	195 (32.5)	275 (45.8)	130 (21.6)	n/a	Low
Min AE and rise time	175 (29)	245 (41)	180 (30)	n/a	low
Key * : in terms of the clusters diagram, blue, green and red are the colours where blue is no galling and red is severe galling.					
** Transition number 210 is taken from actual Die set 1 data where the 210th part is displayed to be the part where it is estimated that the depth of penetration increases from non-galling type state (See Fig. 4a). In brackets is the part discriminator for that specific signal transition of wear and, the respective signal max and min values follow.					

Table 3. considers the comparison results of force and acoustic emission parameters. This information displays the correlation between both force and acoustic emission and its associated outputs which are indicative of the physical outputs. This is important when considering multi spectral approaches with more confidence due to multi-dimensional data describing non-physical phenomena. Observing Fig. 4a (Max AE HHT data), 210th part is consistent with the transition from no galling to galling for Die 1. As there were only 14 µm depth of profile measurements covering the 600 parts – only the boundary values of the clusters were compared. With respect to Table 3. the numbers for each material phenomena column (no galling, galling and severe

galling) display the amount of parts under that specific phenomena as well as its percentage utilisation from the full data set of 600 parts [15]). Another observation therefore is the AE clusters have a larger coverage for the transition of galling when compared with the same force clusters. This again shows the sensitivity to ‘pick-up’ transitional changes for AE compared with force. These clusters are obtained from unsupervised learning and the clusters best fit to a specific material phenomenon based on user best judgement. This best fit of clusters is then correlated against the physical phenomena of depth of penetration (with reference to Fig. 4b) for verification, this gives the accuracy to real phenomena. For example, a specific point will have an associated distance that is nearest to a specific cluster centre. The defined cluster groups are then compared against their true values based on signal intensity and profile depth measurements.

The AE Max amplitude and rise time had poor accuracy when compared to actual material phenomena and this was considered due to a too sensitive algorithm and requires a more trend-based approach over a longer period. The ‘accuracy to real phenomena’ of the final column of Table 3 is based on the error difference for the arbitrary transition of galling, which was estimated to be the 210th part. This gives a measure to how close the detection for the transition of galling for a specific sensing technology. It can be concluded AE had a wider coverage of transitional cases of wear as well as the closest to the actual measured transition point. Both Max and Min AE amplitude and Max and Min AE HHT achieved the same level of accuracy for Table 2 test, however the results of Max and Min AE HHT are more prominent in terms of different cluster densities when applied to different wear states. This is certainly more desired if larger data sets are required.

Figures 7a. and 7b. display the fuzzy cluster distances of the three unsupervised clusters, the green cluster is consistent with no galling, blue is correlated with the transition of galling and red is correlated with parts tending towards severe galling. What is very interesting here is the green cluster – that is correlated with no galling – is compact, while the blue cluster is more dispersed. To give greater confidence in the unsupervised clustered results, both Die 1 and Die 2 have very similar AE data segregations (Figs. 7a and 7b respectively) – this provides confidence in the unsupervised clustering technique. Using mean frequency parameters of HHT provides further clarity in terms of distinguishing galling from non-galling, as can be seen by the increase in accuracy in Table 3. The mean frequency is more of a differentiator for data segregation of AE when compared with that of force. It should be noted that force obtained a fairly constant response however there were ripples recorded in the extracted raw, time-based waveforms as they tend more, towards galling. This can be indicative of a cross coupling when measuring different states of material phenomena.

Figure 8. displays the unsupervised fuzzy cluster distances of the three clusters, green cluster is consistent with no galling, blue is the transition of galling and red is the parts tending towards severe galling. These results are very similar to Figs. 7a and 7b. where the green cluster of no galling is very compact and closed in, the blue cluster is almost like a dispersal effect from the green closed in cluster which further suggests HHT AE is also sensitive to non-stationary when compared to that of stationary waveforms. This supports the verification of HHT used to distinguish non-stationary from stationary

waveforms. In short, the clusters show there is a clear distinction between the non-stationary data, which shows the applicability of both the HHT transform and fuzzy clustering algorithm. Looking at the fuzzy distances, Fig. 8 transitional wear and severe wear are much greater than Fig. 7 which is Max AE and Mean Frequency AE.

4.2 CART Analysis

For CART analysis, the labels for each of the parts were included in the input data (supervised learning technique), where part sequence numbers less than 275 were attributed to class 0 – unworn or no galling parts; part numbers greater than 280 were attributed to class 1 – transition to galling; and part numbers with a greater Signal to Noise level of measured waveform were attributed to class 2 – severe galling.

Table 4
CART rules 1 for the decision tree classification

1 if $x_2 < -0.406854$ then node 2 elseif $x_2 \geq -0.406854$ then node 3 else 0
2 if $x_1 < 1.00458$ then node 4 elseif $x_1 \geq 1.00458$ then node 5 else 2
3 class = 0
4 class = 1
5 class = 2
x1 refers to maximum AE amplitude (HHT), x2 refers to minimum AE amplitude, class 0 refers to no galling,, class 1 transition to galling and class 2 tending towards severe galling.

The pseudo code for decision tree classification (above: CART rules 1) displays the CART classifications. Table 2 displays the transition signal AE IMFs min and max between 0.56/-0.58 and Max AE amplitude: < 1.00458 and Min AE amplitude: < -0.406854.

In this case, to test the validity of prediction, the AE Sensor 1 data was used to train the CART database and from that database, AE Sensor 2 data was used to predict the tooling state in terms of no galling, transition of galling and tending towards severe galling. The obtained accuracy was 97%. This should also be noted as a way of self-calibrating the sensors from the data extracted. If a huge difference existed, then one could conclude the setup needs modifying to ensure a good transfer of signal propagation where both sensors are picking up equal energy phenomena. This further backs up the calibration process of pencil lead break test which concluded consistently high results indicative of a valid, robust and rigid setup. From viewing the data, the AE sensor 1 was estimated as 210th part as the transition point and 280th part as the transition point for AE sensor 2. Looking at the results presented by CART prediction, the transition is at the 280th part for AE sensor 2 data and therefore it can be concluded CART rules are very sensitive to signal change when compared with fuzzy clustering. This is certainly true when sufficient training and test cases are present. Table 5 is a comparison where the AE HHT data was used as the input data of interest and this data was tested on the three classifiers, fuzzy clustering (unsupervised learning) and NNs and CART (supervised learning). To compare on equal terms the CART

and NN classifiers used the AE Sensor 1 data for training and AE Sensor 2 data for testing. The unsupervised method is tested just for AE Sensor 2 data.

Table 5
Comparison of supervised and unsupervised techniques

Data Focus	No galling (%)	Galling (%)	Severe galling (%)	Transition number 210th part for Die 1 and 280th for Die 2	Accuracy designated output (%)	RMSE (Root-Mean-Squared Error)	R ²
AE HHT Max and Min (Fuzzy Clustering)	255 (42.5)	230 (38.3)	115 (19.2)	220 & 305	74 (Die 1) 68 (Die 2) 70(Die 2)	54 (Die 1) 0.37 (Die 1) 0.35 (Die 2)	
AE HHT Max and Min (NN)	275 (46)	25 (4)	300 (50)	280	96 (Die 2)	13.4 (Die 2)	0.74 (Die 2)
AE HHT Max and Min (CART)	275 (46)	25 (4)	300 (50)	280	97 (Die 2)	3 (Die 2)	0.94 (Die 2)

Looking at the results presented in Table 5 it is possible to see that supervised learning outperforms unsupervised. The fuzzy clustering algorithm is certainly more efficient in its calculations which is consistent with [38], however, it is difficult to learn the input waveforms to output values where the intensities are intermittent, as galling occurs and then smooths out and then occurs again and carries on in this manner. There appears to be a discrepancy between the results of Table 3 compared with Table 5 when considering fuzzy clustering, this is due to the different tests that each table focus, where Table 3 is based on the class boundary accuracy and Table 5 the accuracy of class clusters.

Both the supervised techniques give good account when learning such phenomena. This is because unsupervised learning technique can only be tested against and not trained and tested like with the supervised techniques hence the two different values for Die 1 and Die 2. For the supervised techniques Die 1 used as the training data where all obtained 100% accuracy. On a final note, in regard to the results displayed in Table 3, for unsupervised learning of fuzzy clustering the main issue here was to show the transition of the wear state and the sensitivity of different sensing technologies/techniques. With more clusters the results would be more accurate however more difficult to display the distinguishing features. It is not surprising fuzzy clustering performs less than the other two, supervised techniques as other work uses optimisation algorithms to improve accuracy and efficiency of the fuzzy clustering technique [11].

To give more clarity in the results, the following metrics provided by equations. Here R² and Root Mean Squared Error (RMSE). The R² value gives a measure of goodness of fit where 0 corresponds to worse fit and 1 corresponds to the best fit. These metrics are used here to help make comparisons between the

three algorithms and supervised vs. unsupervised. These extra metrics are only used here as previous results displayed in Table 3 look at the transition of galling accuracy and beyond this work, in Sect. 4.3, the general behaviour of low data sets used to predict large data sets. The metrics used in Sect. 4.3 and Table 6 suffice for that study and R^2 and MSE are not used.

The following equations were used as extra metrics in this study where RMSE (2) and R^2 (3) statistical performance indicators are used as algorithm comparison differentiators:

$$RMSE = \sqrt{\frac{1}{N} \sum_i (a_i - p_i)^2} \quad (2)$$

$$R^2 = 1 - \left(\frac{\sum_i (a_i - p_i)^2}{\sum_i p_i^2} \right) \quad (3)$$

where a_i is the actual measured wear state, p_i the predicted wear state, and N the sample size.

Further confirming the other metrics in Table 5, CART performs very well with very near perfect fit in both metrics (RMSE and R^2 respectively), followed by NNs and significantly lower Fuzzy clustering where both data sets were measured albeit both data sets were classified in terms of cluster centres.

4.3 AE STFT 10 feature data correlated to depth of profile (Neural Networks Analysis)

Figure 9. displays STFT plots of parts 70, 220 and 585 respectively. The transition of galling displays an increasing energy spike from 0.02–0.35 MHz. As the extracted AE tends towards severe galling the prominent energy spike extends from 0.02 to 0.5 MHz. A less intense spike continues from 0.5 to 0.9 MHz. All of the amplitudes during the comparisons are all normalised at 40 dB for comparison purposes.

Figure 10. shows the parallel coordinates of the 10 Max frequency components of the HHT AE STFT plots (reference to Fig. 9). Here 10 components are shown for each individual case of a small data set of 14 cases which have associated measured profile depth data. Each measured case is representative of the following 4 channels in the stamping sequence – noting every fifth channel was saved for measuring, and therefore the labelled output results are made for 70 channels that are associated with the 14 measured channels. The processed sensor data could then be labelled with output observations associated with the wear state of the tooling. This small data set was then used to train a NN to predict

600 samples where 530 samples are unseen cases the remaining 70 cases are also used for training. These predictions can be further checked against Fig. 4b. which displays the actual max profile depth. In Figs. 8, 9 and 10, we can observe the pattern of wear phenomena. A ‘divide and conquer’ or ‘leave-one-out’ approach [43] was used to validate the small data set. This is where one case would be removed and the NN would learn the remaining 13 cases from a ‘reset weights state’ with no prior knowledge. After this initial learning, the NN would then predict the missing 1 case. This would be carried out 14 times for the total small data set and finally, a total distance error would be calculated (see Table 6 for more information).

The distance error was calculated from Eq. (4) and the total distance error was calculated from the following Eq.

$$D_i = (T_{Di} - T_{Ai}) \quad (4)$$

$$D_T = \frac{1}{14} \sum_{i=1}^{14} (D_i - \bar{D}_i)^2 \quad (5)$$

Where T_A = Actual Target, T_D = Desired Target, D_T = Total Distance Error and D_i = Distance Error for Target i

Table 6
Optimal Neural Network Architectures

Varying parameters	Total Distance Error	Total Distance Error ²	Sum Squared Error
'trainrp','learnk', Change 2 hidden layers = 40 baseline network architecture	75.72	4045	2.36 * E10 -25
'trainrp','learnk', Change 2 hidden layers = 80	-15.1	228.01	7.04 * E10 -30
'trainrp','learnk', 2 hidden layers = 40 and lr = 0.1 *E10 -9 (10% reduction)	-7.5	56.25	9.71 * E10 -31
'trainrp','learnk', 2 hidden layer = 80 and lr = 0.1 *E10 -9 (10% reduction)	-3.98	15.8404	9.02 * E10 -31
Change to 'trainlm','learnk',	-2.76	7.6176	1.03 * E10 -28
Change to 'traingdx','learnk'	4.7	22.09	0.000150

Table 6. lists the NN comparisons for the most optimal NNs. The highlighted green NN algorithms display that the Levenberg-Marquardt training rule and back propagation training rule with reduced learning rate and increased hidden layers both outperformed the others with very low total distance and sum squared error values obtained. These optimal NNs were used to predict the 600 cases (the 600 cases would consist of 10 Max AE frequency components correlated to measured profile depth of cut, see Fig. 11).

Figures 11 and 12 give a good account for profile depth prediction. When comparing back to Fig. 4b. the majority of the cases with no galling are above 8 μm and below 10 μm , transition is between 10 μm and 20 μm and severe galling, is greater than 20 μm which is consistent with Figs. 11 and 12. Using AE STFT and HHT data it is possible to correlate with profile depth based on both the intensity and dominant frequency bands. It should be noted the backpropagation learning rule NN appears to resemble the actual measurements where there is little change in profile depth with respect to the early parts and as galling tends towards more severe, there is only one outlier. This is both consistent with physical measurements as well as visually inspecting STFT AE images. Reference to Fig. 11, the second data cases relate to 600 cases for AE Sensor 2 where again, the main patterns are captured. Although the smoothness of AE responses of the non-galling is not as smooth as with AE Sensor 1, the transition however of galling is much further along which is consistent with physical measurements such as those displayed in Fig. 4b. AE Sensor 2 are totally unseen, hence the prediction is less smooth and there are more outliers mixed in with this data which relate to transition but is not considered as the main transition where this occurs around 295th part.

This holistic approach has high confidence of accuracy and especially applicable to industry as it allows a non-destructive testing approach where initial destructive tests (calibrate signals to physical phenomena) allows, follow on non-destructive tests (prediction based solely on signals) which is rewarding in terms of time and effort.

Table 7
CART rules output 2

CART Results – tree formed on 14 known data set (AE STFT correlated to actual measured profile depth):
CART Decision tree for classification
1 if $x_1 < 21.2502$ then node 2 elseif $x_1 \geq 21.2502$ then node 3 else -37.804
2 class* = -8.9039
3 if $x_1 < 23.3281$ then node 4 elseif $x_1 \geq 23.3281$ then node 5 else -8.824
4 class = -8.6211
5 if $x_1 < 25.7719$ then node 6 elseif $x_1 \geq 25.7719$ then node 7 else -37.804
6 class = -8.824
7 if $x_1 < 30.2589$ then node 8 elseif $x_1 \geq 30.2589$ then node 9 else -37.804
8 class = -9.8667
9 if $x_1 < 34.5643$ then node 10 elseif $x_1 \geq 34.5643$ then node 11 else -37.804
10 class = -15.765
11 class = -37.804

*class in this case refers directly to profile depth in μm and x_1 is the Max AE FFT component (1st out 10 STFT signal components which distinguishes different galling states)

4.4 CART prediction of profile depth (600 parts for both AE Sensor 1 and Sensor 2)

CART is considered as a similar method to fuzzy clustering where segregating data centroids are based around best variables tending towards the most optimal position significant of best segregation [36]. The only major difference being that CART provides its predictions in a supervised manner, while fuzzy clustering is an unsupervised technique.

Figure 13. displays the output profile depth prediction of 600 parts which again is sensitive to the transition from no galling to galling. This is true as two predictions are made one with AE Sensor 1 and the other with AE Sensor 2. To verify the correctness, look at Fig. 4b and compare the depth of penetration with Fig. 13 there is certainly a correlation between the measurements and the predictions. That said, this method is considered less sensitive than NNs as the optimal NNs of Figs. 11 and 12 give a better account of the transition zone tending towards the transition of galling from no galling. It can be concluded with a limited data set, the optimised NNs perform better than the CART and fuzzy clustering algorithms. In fact, it is true to say when observing AE Sensor 1 (Die 1) the classification distinction is from no wear to severe wear and there is little to no sensitivity with transitional wear. AE sensor 2 however has more distinction within the transitional zone. This is due to a low data set however slightly more coverage for transitional

zone with sensor 2 when compared with sensor 1. With more data, both CART and fuzzy clustering algorithms should perform better. It is fair to conclude with limited data sets NNs are better at gaining higher resolution around boundary conditions when compared with CART. All of these techniques perform well when presented with multi-dimensional data as presented by the max STFT frequency components (see Figs. 9 and 10).

5. Discussion Of Results

Table 3 displays the coverage of fuzzy clusters which relates to different levels of wear mechanisms. These wear mechanisms are evaluated in terms of their accuracy from comparing known measurement outputs of wear with input signal phenomena. The outputs are already known in this case and therefore a blind distance evaluation is made to segregate these different wear states. For clarity, the amount of parts with signals correlated to no galling are 280 for the Force mean frequency and force peak amplitude. In brackets (46.6% gives the overall percentage utilisation between the three wear states: no galling, transition of galling and severe galling). These states are then evaluated with actual physical wear measurements in terms of the galling transition and which part. How close gives a higher a prediction percentage. Here it is clear to see force is less accurate when compared with both HHT Max/Min AE Amplitude and standard Max/Min AE Amplitude in terms of blindly mapping, different wear mechanisms.

Figure 7. extends the work displayed by Table 3. where the AE Mfreq vs. the Max AE amplitude provides good segregations of the different wear states. Two sets of acoustic emission wear data is used where one sensor is attached to Die 1 and the other, Die 2. Both show similarities however galling transition with Die 1 has a higher % utilisation than Die 2.

Figure 8. Displays the use of pre-processing techniques such as HHT which visually ‘blows out’ the wear states from no galling (where the signals are considered stationary) and tend to stay in localized clusters with much smaller distances, however in the case of the other wear mechanisms where the signals become non stationary the clusters open out with much greater distances significant to change. This increases more for severe galling cases. With reference to Table 3. HHT Max/Min accuracy is the same as standard Max/Min accuracy however the former is more visually observable than the latter and an important consideration when promoting visual detection. In terms of results, Table 5 is different to Table 3 where Table 3 only considers the accuracy of the transition of galling and how close classifications/predictions are to the physical transition of material change, Table 5 looks at this and then compares the accurate utilization of the different wear states. With unsupervised techniques it is difficult to use a training and test set separately. Instead both data sets were tested individually. Here both the supervised techniques outperform the unsupervised technique. The accuracy of output evaluates each wear classification to known physical material measurements. Both the CART and NNs have almost the same accuracy where classifications lie in-between tolerances and with more accuracy required as these tests aim to go beyond general behaviour and visualisation, extra metrics are required. Extra metrics of RMSE and R^2 statistical measures are used to clarify the different algorithms. CART performs the best we near perfect results all round, followed by NNs with promising results. Fuzzy Clustering however scored

much lower and this was based on a poor fit optimisation function, where the algorithm started off at 81.6 and achieved a minimum of 17.4 which is still fairly high for termination criteria. This is where the intra-cluster variance is minimised until it cannot go any further and therefore achieves maximised intra-cluster similarity. With better optimisation functions that are not prone to local minima through too many constraints will avoid this.

The results provided by Table 6 and Figs. 11 and 12 display the use of using high dimensional data that has been transformed from the time-frequency domain and used to train and then predict unseen test cases. The training data in this case was very small and confirms the resilience of Neural Networks when faced with such constraints. The visualization outputs of Figs. 11 and 12 gave good clarity when compared to actual physical measured data of Fig. 4b. Looking also at Table 6 the chosen NNs obtain very low sum squared error and total distance error which suggests the NN fit the low data set very well to make good predictions with larger unseen data sets. Figure 13 displays the same low data set using CART. Comparing the visual outputs, NN provides a better coverage in terms of accuracy and higher precision.

Previous work [15] has detailed information in terms of the acoustic emission measurements obtained in stamping and used here. The data however taken from stamping tests has been processed in this work where HHT transforms, mean frequency and rise time have been applied as a pre-processing element for further application to a ML technique. By extracting salient signal parameters as provided by the pre-processing layer there is less work to do when distinguishing classes or providing prediction in ML layer. This method is already obtains its data from a semi-industrial stamping setup. With signal extracted parameters the time penalties are fairly small and therefore classification for automated intervention would be provided in near real-time.

There is very little work where machining learning is used to distinguish different material characteristics and predict the onset and increased wear states. The two main sources [19], [20] were picked to highlight similarities however the results displayed here give somewhat higher resolution and information than results displayed by the previous work. For example, [19] uses data from literature to build a model to then predict tool wear. These sensing inputs however are time, velocity, feed and cutting force – which are useful for detecting failure but less so, the onset of failure. The sensing capabilities with the work presented here look at load as well for comparison and calibration, however the acoustic emission gives far greater insight into material process and is more sensitive to change, that said, it can also correlate to load (cutting force) as well as other effects such as strain and temperature. The results presented in this work display more information providing a more accurate picture. For accuracy, these results are all checked against actual measurements for verification (see Fig. 4b). The other work [20] is much closer to the work provided here where ML is used to segregate different material removal mechanisms using acoustic emission measurements. The material removal mechanisms were then checked against material interferometer measurements which gives verification of the results. Methods carried in this work were followed very closely here as this appears the best way to display results and results verification when linking the non-physical signal with the physical material measurements.

Both sets of work used neural networks, CART and fuzzy clustering to show visualisations and predictions which is why they are used here. The neural network for [19] used a delay function to feedback information in the form of historical sensor data. This is very useful and an advancement for future work. That said, there were no studies looking into reduced data sets from empirical results which is often the case when gathering data from industrial processes. Here both NNs and CART are used to predict wear based a very limited data sets. Another reason why these techniques were chosen over others.

6. Conclusions

A new approach to classify wear using acoustic emission sensors and machine learning (ML) applied to stamping processes was defined. This work reinforces previous work where classifiers were used to exploit digital signal processing techniques that highlight non-stationary over stationary waveforms. Using ML as a form to segregate different states promotes an autonomous model ready for industrial exploitation. Based on the results and discussion, the following conclusions can be made:

- A number of different data representations were made to both force and AE signal waveforms. Where parameters and transforms provided rich summaries as well as salient features for both the said signal waveforms. AE HHT waveforms gave the best separation between non-galling and galling for the unsupervised fuzzy clustering technique. Force was used to highlight the sensitivity of AE and at the same time, verify to known quantity.
- The non-stationary AE waveforms with fuzzy clustering has a much greater separation in terms of distance classification output from fuzzy centres than the stationary AE counterparts which are bunched up, much closer and significant of no galling. This is due to the use of adding a pre-processing layer in the form of HHT that increases the distance for data points which were non-stationary in nature compared to stationary. The clustering technique using fuzzy distance measures to quantify a calculated point space and displayed this phenomena very well where visualisation for greater insight was one of the main aims for carrying out this work.
- ML techniques namely CART and fuzzy clustering distinguished different states: no galling, transition from no galling to galling, and galling tending towards severe galling. As well as AE HHT waveforms both AE minimum/maximum amplitude and force maximum amplitude with mean frequency showed good separation (force being the less sensitive to the transition of galling). This is important to show the sensitivity comparison between load (force) and acoustic emission and why such sensing technologies are chosen over others.
- It was found that supervised learning techniques provided a better level of accuracy prediction than unsupervised learning technique 96% (NN) and 97% (CART) compared with 68% (fuzzy clustering). Extra metrics of RMSE and R^2 further clarified the above results. The reasons for these differences are likely down to the low data set where fuzzy clustering needs more data to be able to generalise the data structure more, separating the different wear states. In addition, the optimisation function for best intra-cluster variance minimisation was found to be poor and with a more randomised

optimisation function with less constraints should prove better. Both NN and CART can work with much lower data sets in comparison.

- Prediction of profile depth using STFT maximum frequency amplitude information from a limited data set correlated to physical profilometer measurements were also carried out. This is very useful as it's often the case with industrial tests where low data is captured instead of high data amounts. This was also greater n-dimensional data than other mentioned tests and suggests the robustness of both NNs and CART.
- Using NN 'divide and conquer' method as well as CART, both data representations verified each other where there was significant correlation with predictions.
- Predictions of unseen AE waveforms (not correlated with profile depth measurements) to the ML choice gave good account when compared to AE waveforms correlated with profile depth measurements. This is a very important verification as it allows a method to measure profile depths using non-destructive testing (NDT) methods calibrated to actual physical measurements. The practical success here means only small portion of the data has to be checked physically. NDT can save a lot of time as the measurements can be taken in real time and in difficult to get to locations.

With considerations for future work, it was noticed that AE Max amplitude vs AE Max amplitude rise time resulted in poor results. This needs to be looked at over longer durations to give more accurate account of time. Another aspect to look into the future is investigate time delay neural networks to evaluate different wear mechanisms with dynamic historical data.

Declarations

6.1 Ethical Approval

Not applicable

7.2 Consent to Participate

Not applicable

7.3 Consent to Publish

Not applicable

7.4 Authors Contributions

Authors' contributions: J.M.G. and V.V.S. conceived of the presented idea. J.M.G developed the theory and performed the computations. J.M.G and B.F.R. verified the analytical methods. B.F.R encouraged J.M.G to investigate divide and conquer method of neural networks with very low data set and supervised the findings of this work. V.V.S. contributed to the flow, style and quality of manuscript.

Also V.V.S. provided the data from experiments. M.P.P. provided contributed on the materials and mechanical aspects of the work. All authors discussed the results and contributed to the final manuscript.

7.5 Funding

The authors have no relevant financial or non-financial interests to disclose.

7.6 Competing Interests

The authors did not receive support from any organization for the submitted work.

7.7 Availability of data and materials

Data will be made available on request.

References

1. Lu B, Zhou X, "Quality and reliability oriented maintenance for multistage manufacturing systems subject to condition monitoring," *J. Manuf. Syst.*, vol. 2019, no. Part A, pp. 76–85, 2019
2. Dong Q, Kontar R, Min L, Gang X, Xu J (2019) A simple approach to multivariate monitoring of production processes with non-Gaussian data. *J Manuf Syst* 53:291–304
3. Paolanti M, Romeo L, Felicetti A, Mancini A, Frontoni E, Loncarski J, "Machine learning approach for predictive maintenance in industry 4.0," (2018) *14th IEEE/ASME Int. Conf. Mechatron. Embed. Syst. Appl.*, pp. 1–6, 2018
4. Lee D-H, Yang J-K, Cho-Heui L, Kim K-J, "A data-driven approach to selection of critical process steps in the semiconductor manufacturing process considering missing and imbalanced data," *J. Manuf. Syst.*, vol. 52, no. Part A, pp. 146–156, 2019
5. Ubhayaratne I, Pereira MP, Xiang Y, Rolfe BF (2017) Audio signal analysis for tool wear monitoring in sheet metal stamping. *Mech Syst Signal Process*. doi:10.1016/j.ymssp.2016.09.014
6. Xu Y, Ge M (2004) Hidden Markov model-based process monitoring system. *J Intell Manuf.* doi:10.1023/B:JIMS.0000026572.03164.64
7. Bassiuny AM, Li X, Du R (2007) Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization. *Int J Mach Tools Manuf.* doi:10.1016/j.ijmachtools.2007.06.006
8. Ge M, Du R, Zhang G, Xu Y (2004) Fault diagnosis using support vector machine with an application in sheet metal stamping operations. *Mech Syst Signal Process*. doi:10.1016/S0888-3270(03)00071-2
9. Cervantes J, Garcia-Lamont F, Rodriguez-Mazahua L, Lopez A (2020) A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing* 408:189–215
10. García C (2005) Artificial intelligence applied to automatic supervision, diagnosis and control in sheet metal stamping processes. *J Mater Process Technol*. doi:10.1016/j.jmatprotec.2005.02.031

11. Wei Li., Zhang L, Chen X, Wu C, Cui Z, Niu C, "Predicting the evolution of sheet metal surface scratching by the technique of artificial intelligence.", *Int J Adv Manuf Technol*, 1–132020, doi:10.1007/s00170-020-06394-4
12. Li X (2002) A brief review: Acoustic emission method for tool wear monitoring during turning. *Int J Mach Tools Manuf.* doi:10.1016/S0890-6955(01)00108-0
13. Sick B (2002) On-line and indirect tool wear monitoring in turning with artificial neural networks: A review of more than a decade of research. *Mech Syst Signal Process.* doi:10.1006/mssp.2001.1460
14. G. and Godin SL (2004) N., Huguet. S., "Clustering of acoustic emission signals collected during tensile tests on unidirectional glass/polyester composite using supervised and unsupervised classifiers. *NDT&E Int* 37:253–264
15. Shanbhag VV, Rolfe BF, Arunachalam N, Pereira MP (2018) Investigating galling wear behaviour in sheet metal stamping using acoustic emissions. *Wear.* doi:10.1016/j.wear.2018.07.003
16. Ren Q, Balazinski M, Jemielniak K, Baron L, Achiche S (2013) Experimental and fuzzy modelling analysis on dynamic cutting force in micro milling. *Soft Comput.* doi:10.1007/s00500-013-0983-0
17. Ren Q, Baron L, Balazinski M (2012) Fuzzy identification of cutting acoustic emission with extended subtractive cluster analysis. *Nonlinear Dyn.* doi:10.1007/s11071-011-0173-5
18. Ren Q, Balazinski M, Baron L, Jemielniak K, Botez R, Achiche S (2014) Type-2 fuzzy tool condition monitoring system based on acoustic emission in micromilling. *Inf Sci (Ny).* doi:10.1016/j.ins.2013.06.010
19. Venkatesh K, Zhou M, Caudill RJ (1997) Design of artificial neural networks for tool wear monitoring. *J Intell Manuf.* doi:10.1023/A:1018573224739
20. Griffin J, Chen X (2014) Real-time fuzzy-clustering and CART rules classification of the characteristics of emitted acoustic emission during horizontal single-grit scratch tests. *Int J Adv Manuf Technol.* doi:10.1007/s00170-014-5959-4
21. Shanbhag VV, Rolfe BF, Pereira MP (2020) Investigation of galling wear using acoustic emission frequency characteristics. *Lubricants.* doi:10.3390/lubricants8030025
22. Teymuri Sindi C, Ahmadi Najafabadi M, Salehi M (2013) Tribological behavior of sheet metal forming process using acoustic emission characteristics. *Tribol Lett.* doi:10.1007/s11249-013-0193-z
23. Hase A, Mishina H, Wada M (2012) Correlation between features of acoustic emission signals and mechanical wear mechanisms. *Wear.* doi:10.1016/j.wear.2012.05.019
24. Yang M, Manabe K, Hayashi K, Miyazaki M, Aikawa N (2003) Data fusion of distributed AE sensors for the detection of friction sources during press forming. *J Mater Process Technol.* doi:10.1016/S0924-0136(03)00533-8
25. Nivesrangsang P, Steel JA, Reuben RL (2007) Source location of acoustic emission in diesel engines. *Mech Syst Signal Process.* doi:10.1016/j.ymssp.2005.12.010
26. Chen GX, Zhou ZR (2007) Time-frequency analysis of friction-induced vibration under reciprocating sliding conditions. *Wear.* doi:10.1016/j.wear.2006.03.055

27. Kaewkongka B, Joe T, Rakowski YH, R.T. and Jones (2003) A comparative study of short time Fourier transform and continuous wavelet transform for bearing condition monitoring. *Int J COMADEM* 6(1):41–48
28. Wang WJ, McFadden PD (1996) Application of wavelets to gear box vibration signals for fault detection. *J Sound Vib.* doi:10.1006/jsvi.1996.0226
29. Fan X, Zuo MJ (2006) Gearbox fault detection using Hilbert and wavelet packet transform. *Mech Syst Signal Process.* doi:10.1016/j.ymssp.2005.08.032
30. Yan R, Gao RX (2006) Hilbert-huang transform-based vibration signal analysis for machine health monitoring. *IEEE Trans Instrum Meas.* doi:10.1109/TIM.2006.887042
31. Hamdi SE, Le Duff A, Simon L, Plantier G, Source A, Feuilloy M (2013) Acoustic emission pattern recognition approach based on Hilbert-Huang transform for structural health monitoring in polymer-composite materials. *Appl Acoust.* doi:10.1016/j.apacoust.2012.11.018
32. Pereira MP, Weiss M, Rolfe BF, Hilditch TB (2013) The effect of the die radius profile accuracy on wear in sheet metal stamping. *Int J Mach Tools Manuf.* doi:10.1016/j.ijmachtools.2012.11.001
33. Tse PW, Chu FL, Peng ZK, "A comparison study of improved Hilbert–Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing," *Mechanical Systems and Signal Processing*. 2005
34. Yang Z, Yu Z, Xie C, Huang Y (2014) Application of Hilbert-Huang Transform to acoustic emission signal for burn feature extraction in surface grinding process. *Meas J Int Meas Confed.* doi:10.1016/j.measurement.2013.08.036
35. Miller I, Papoulis A (1966) Probability, Random Variables, and Stochastic Processes. *Technometrics*. doi:10.2307/1266379
36. Voss BM, Pereira MP, Rolfe BF, Doolan MC, "Using stamping punch force variation for the identification of changes in lubrication and wear mechanism," *IOP Conf. Ser. J. Phys.*, vol. 896, p. 12028, 2017, [Online]. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/896/1/012028/pdf>
37. Strang T, Nguyen G (1996) Wavelets and Filter Banks. Wesley, Cambridge
38. Sathya R, Abraham A (2013) Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *Int J Adv Res Artif Intell.* doi:10.14569/ijarai.2013.020206
39. Cuevas A, Febrero M, Fraiman R (2001) Cluster analysis: A further approach based on density estimation. *Comput Stat Data Anal.* doi:10.1016/S0167-9473(00)00052-9
40. Gordon AD, Breiman L, Friedman JH, Olshen RA, Stone CJ, "Classification and Regression Trees.,," *Biometrics*, 1984, doi: 10.2307/2530946
41. Lawrence RL, Wright A, "Rule-based classification systems using classification and regression tree (CART) analysis," *Photogramm. Eng. Remote Sensing*, 2001
42. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature*. doi:10.1038/323533a0

Figures

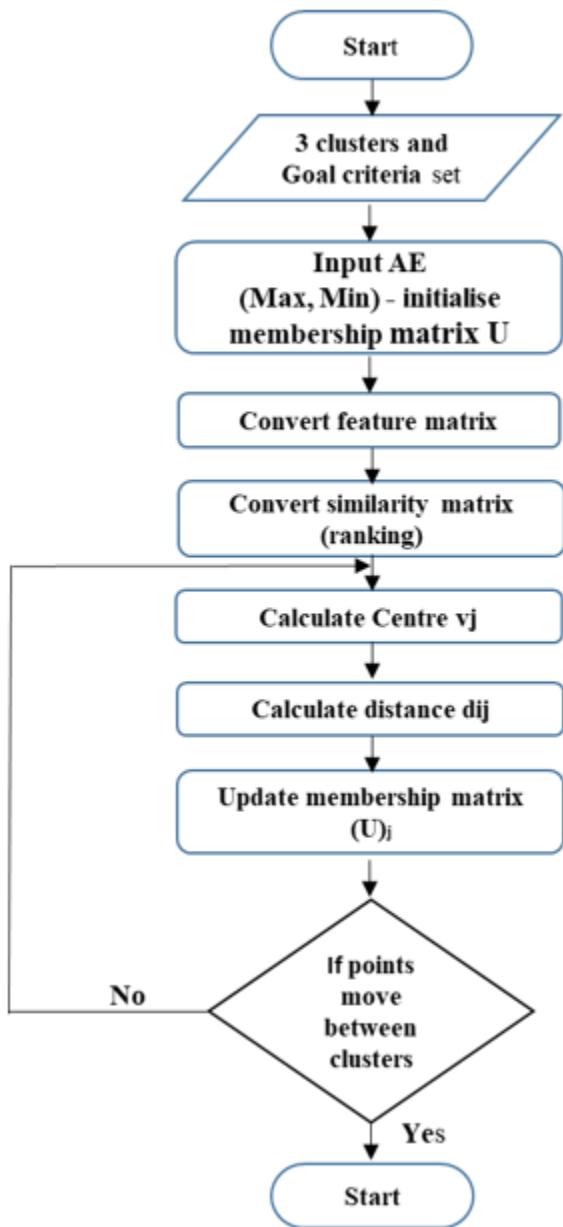


Figure 1

Flow Chart of fuzzy clustering example with AE Max and Min inputs.

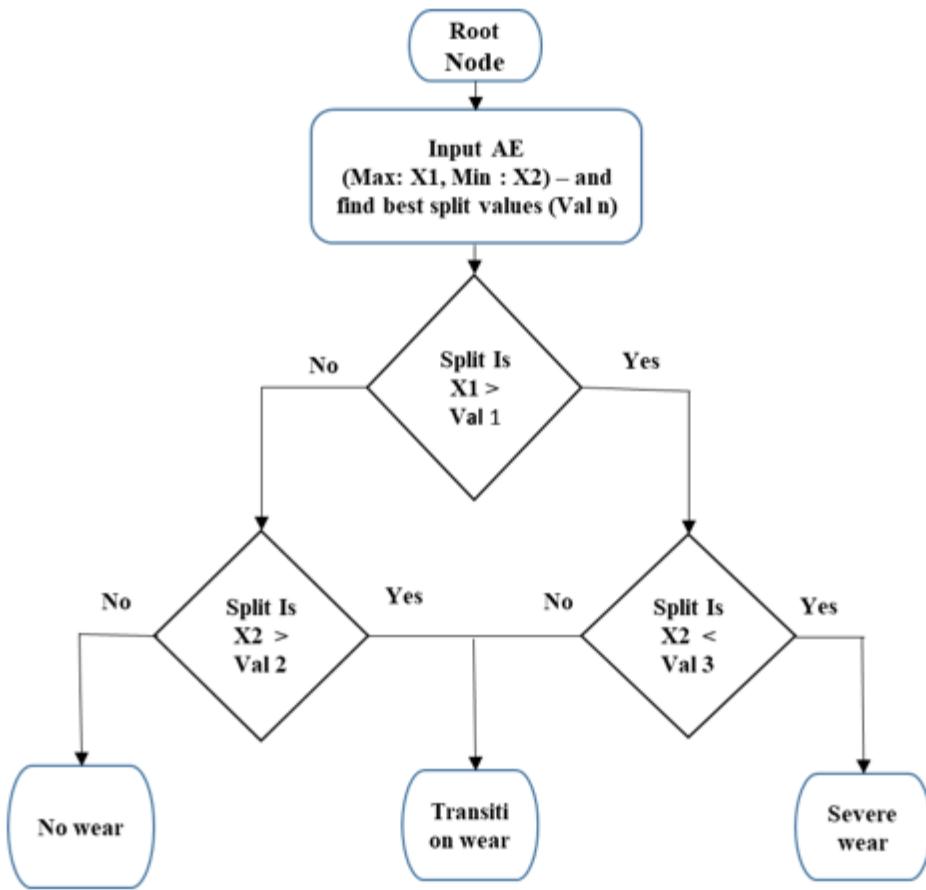


Figure 2

Flow Chart of CART Example with AE Max and Min inputs

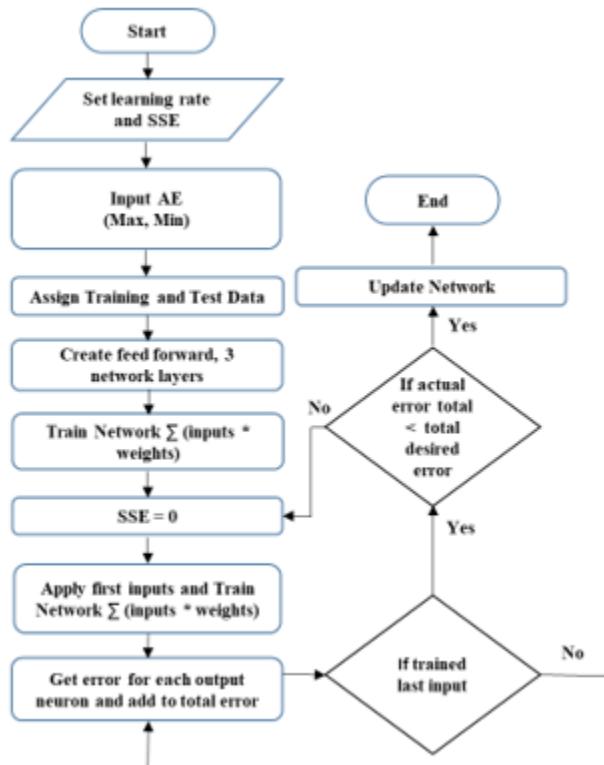


Figure 3

Flow Chart of Neural Network Example with AE Max and Min inputs.

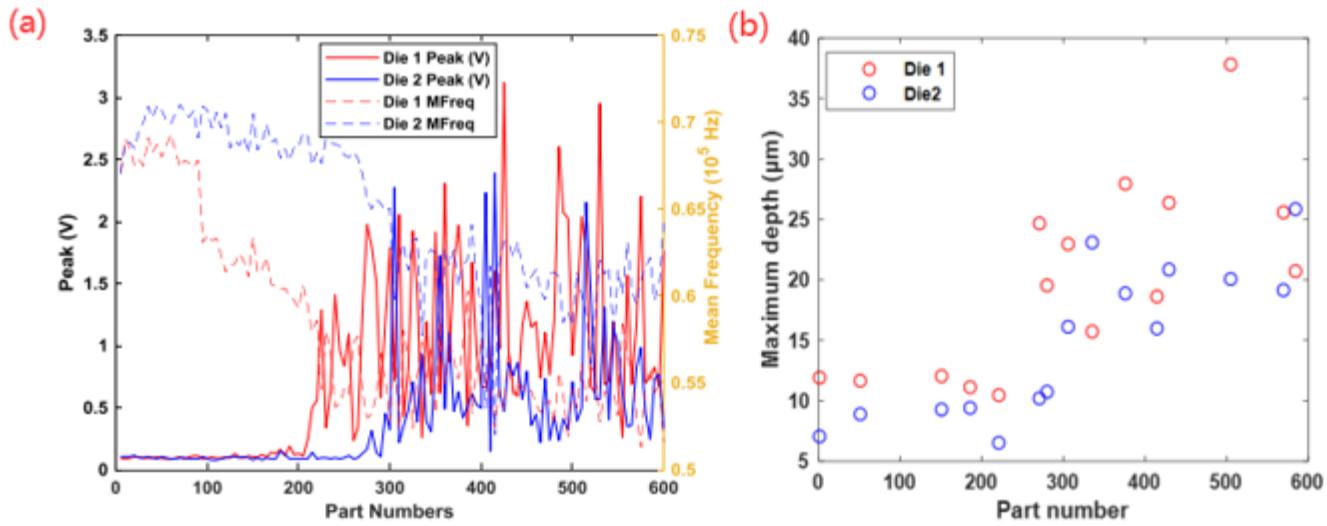


Figure 4

a Peak voltage and mean frequency of AE for both Die 1 and Die 2. b Maximum depth of the surface profile measurement of the stamped part.

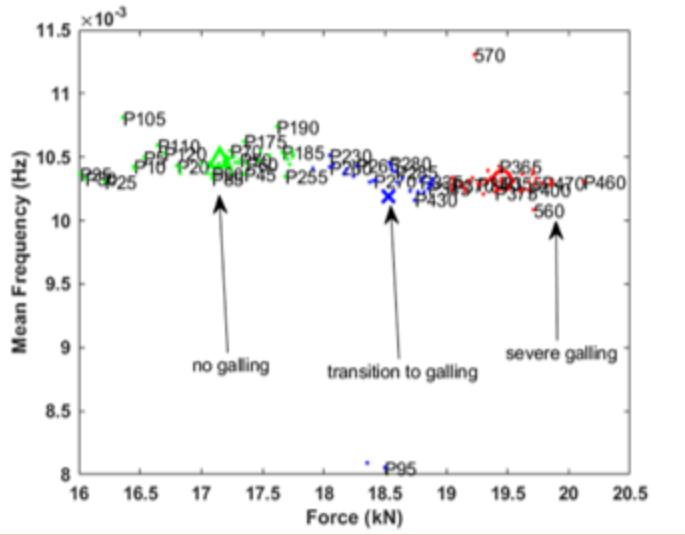


Figure 5

The unsupervised learning of fuzzy clustering to investigate the relationship between the forces mean frequency and max amplitude. The labels are the authors arbitrary understanding of the state of the tooling during the experimental sequence (part number (P), P<260 = no galling, P≥260 & <340 = transition, P≥340 = severe galling)

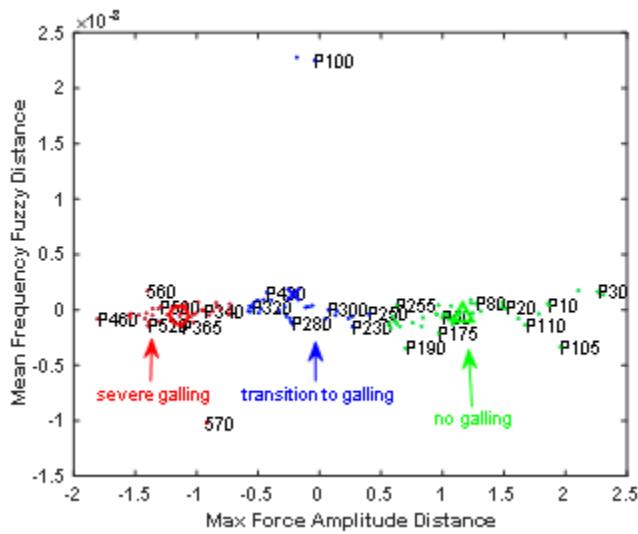


Figure 6

Unsupervised learning fuzzy cluster distances between the two force attributes as displayed in Fig. 5.

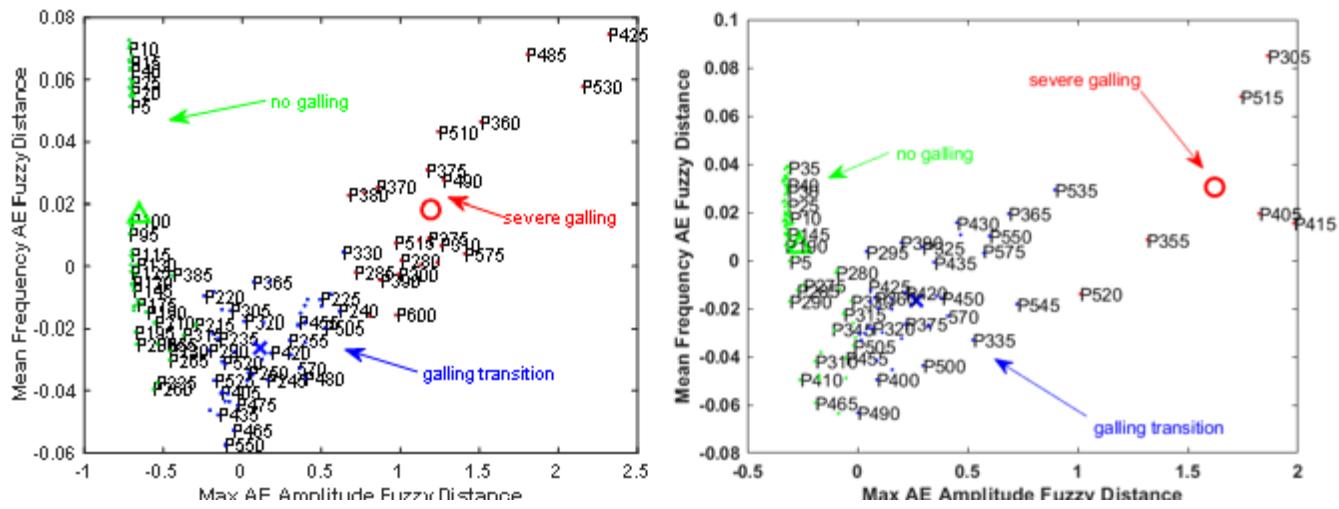


Figure 7

a. Die 1 and b. Die 2 fuzzy cluster distances between the AE Max and Mean Frequency parameters of HHT AE IMFs.

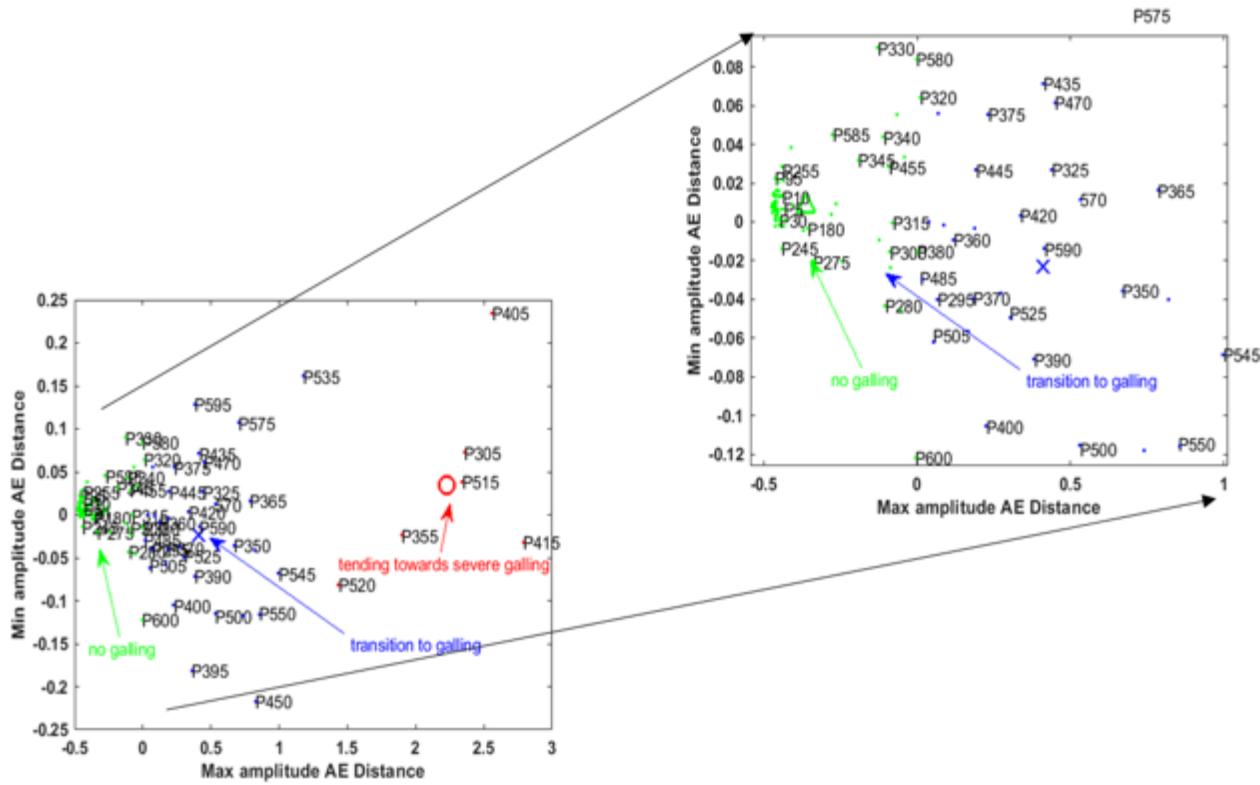


Figure 8

Fuzzy cluster distances between the Max and Min amplitudes of HHT AE IMFs.

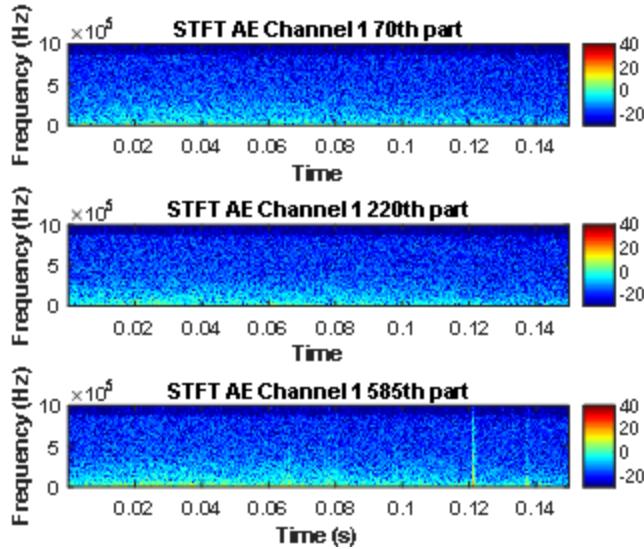


Figure 9

STFT AE Sensor 1 of Die 1 for a) 70 b) 220 and c) 585 parts

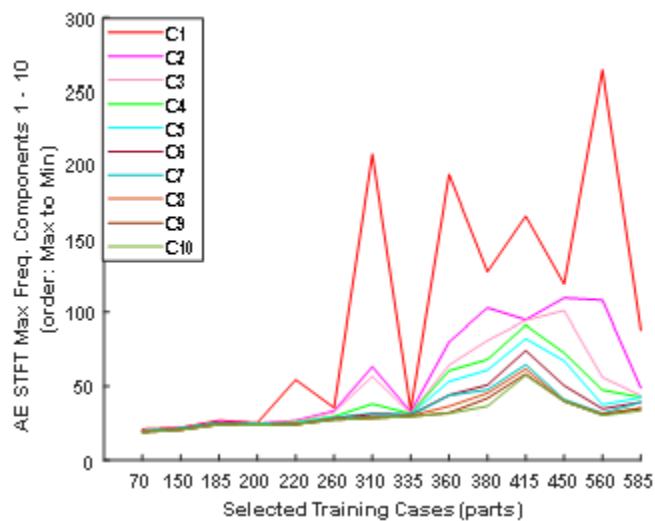


Figure 10

Parallel Coordinates of 10 STFT Max Frequency components for 14 selected training parts

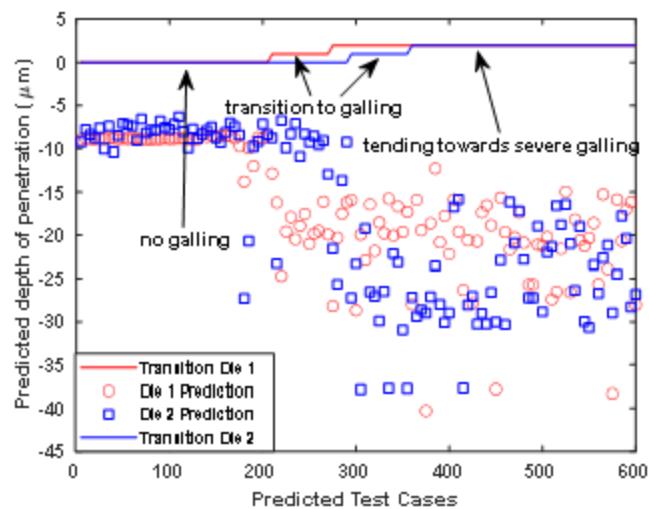


Figure 11

Default NN with increased H/Ls and reduced learning rate displaying predicted profile depth

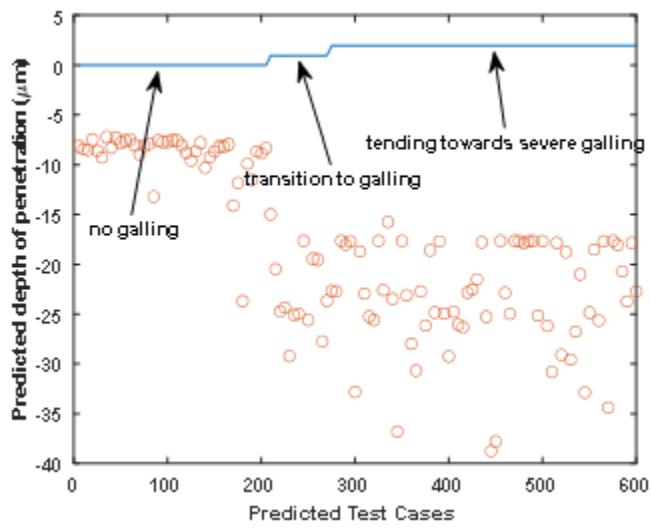


Figure 12

Levenberg-Marquardt NN displaying predicted profile depth (Die 1)

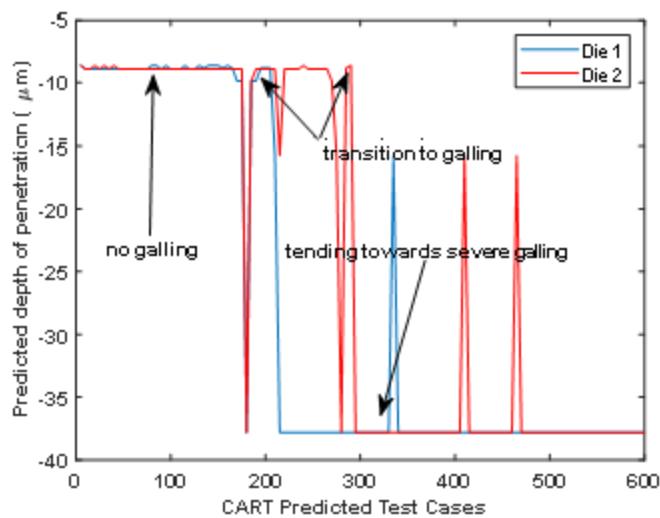


Figure 13

CART predicted profile depth