Query Sheet		
Journal Name: IJAT	Volume No. X	Issue No. Y

Article Name: Characteristics of the acoustic emission during horizontal single grit scratch tests: Part 2 classification and grinding tests

Author Name: J. Griffin and X. Chen

Chapter No. Two (Part 2) Paginated S/W: Word

S.No.	Page No.	Query
1	9	Author: Please check the sentence 'The supervised training method uses three times'.
2	11	Author: Please provide footnote links '*' for Table 2.
3	16 (Ref)	Author: Please provide the forename for authors Li and Xiaoli for reference Li and Xiaoli (1998).
4	17 (Ref)	Author: Please provide the forename for authors Peterson and Gerald for reference Peterson and Gerald (1992).

Characteristics of the acoustic emission during horizontal single grit scratch tests: Part 2 classification and grinding tests

James Griffin and Xun Chen*

School of Mechanical, Materials and Manufacturing,
University of Nottingham,
Nottingham NG7 2RD, UK
E-mail: epzjg3@ad.nottingham.ac.uk
E-mail: x.chen@hud.ac.uk
*Corresponding author

Abstract: The second part of this work follows on the work carried out in Part 1 where the investigations were made between the grinding phenomena: cutting, ploughing and rubbing. The demarcation between each of the phenomenon was identified from Acoustic Emission (AE) signals being converted to the frequency-time domains using Short-Time Fourier Transforms (STFTs). Other digital signal processing techniques were used and discussed; however, the more update and successful tests only required STFTs. This part of the paper looks at the classification using both Neural Networks (NNs) and fuzzy-c clustering/Genetic Algorithm (GA) techniques. After the cutting, ploughing and rubbing gave a high confidence in terms of classification accuracy, 1 μm and 0.1 mm grinding test data were applied to the classifiers. Interesting output results sufficed from both classifiers signifying a distinction that there is more cutting utilisation than both ploughing and rubbing as the interaction between grit and workpiece become more in contact with one another (measured depth of cut increases).

Keywords: single grit scratch; AE; acoustic emission; feature extraction; NNs; neural networks; fuzzy-c clustering and genetic algorithms.

Reference to this paper should be made as follows: Griffin, J. and Chen, X. (XXXX) 'Characteristics of the acoustic emission during horizontal single grit scratch tests: Part 2 classification and grinding tests', *Int. J. Abrasive Technology,* Vol. X, No. Y, pp.XXX–XXX.

Biographical notes: James Griffin completed his undergraduate studies in 1995 with a BE from The University of Wales, College Cardiff. Between 1996 and 2003, he was a Member of the Defence Evaluation Research Agency (DERA) on numerous projects. One of that was significant used of Genetic Algorithms/Neural Networks to control an Unmanned Aircraft Vehicle (UAV). Whilst working for DERA/QinetiQ he completed a second Degree BSc from Open University (1996–2001) in Technology biased towards Artificial Intelligence. He is now a PhD candidate within the Advanced Abrasive Technology Group at the University of Nottingham applying intelligent concepts to identify different grinding phenomena.

Xun Chen specialises in Advanced Manufacturing Technology including Application of Computer Science, Mechatronics and Artificial Intelligence to Broad Engineering Application. He has published more than 100 research papers. He is a founder Member of the International Committee of Abrasive

Technology. Before his employment at Nottingham, he was a Lecturer of Mechanical Engineering at the University of Dundee. He received his BEng from Fuzhou University. He received his MSc from Zhejiang University and his PhD from Liverpool John Moores University. He has been a Visiting Professor to Fuzhou University since 2001.

1 Introduction

As mentioned in Part 1, elastic waves such as that emitted by Acoustic Emission (AE) can be used for monitoring many machining processes and/or material non-destruction tests (Chen and Xue, 1999; Coman et al., 1999; Holford, 2000; Liu et al., 2005; Webster et al., 1994). Once the raw extracted AE has been transformed into the time-frequency domain it can then be presented to the classifier. Here the work looks at the Single Grit (SG) scratch classification of cutting, rubbing and ploughing using two classifiers: Neural Networks (NNs) and fuzzy-c clustering/Genetic Algorithm (GA). The characteristics of cutting, ploughing and rubbing through both material profile measurements and digital signal processing of AE extracted signals were discussed in Part 1 of this paper. It was found in the Part 1 that the Short-Time Fourier Transforms (STFTs) results of AE analysis can represent different characteristics of cutting, ploughing and rubbing in grinding, which may be used as input parameters for the classification. The classification of the three phenomenons is of particular importance to the fundamental understanding of grinding mechanics. From the accurate classification of cutting, ploughing and rubbing for SG tests it was then possible to classify 1 μm and 0.1 mm grinding wheel cuts in terms of the three phenomena.

The main investigation objectives of this part of the paper are:

- classify cutting, ploughing and rubbing in grinding with 1 μm and 0.1 mm cutting depths
- verify the classification process using both training, test and verification data
- compare two classification techniques (NNs and fuzzy-c clustering/GA).

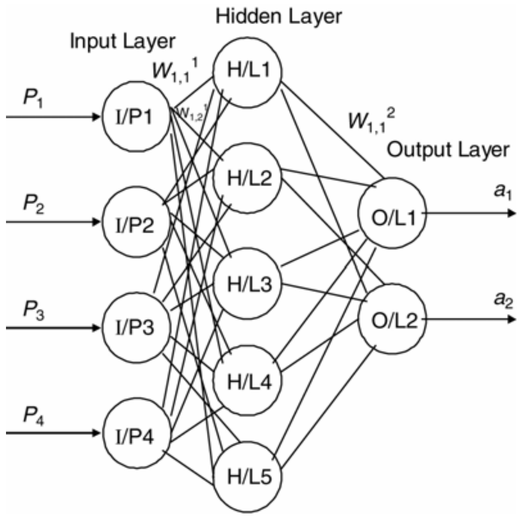
The work carried out in this paper is the first of its kind in that AE signal extracted data is classified in terms of cutting, ploughing and rubbing for single grinding pass. In addition by using an optimised GA fuzzy clustering algorithm for non-linear classification is the first seen in literature and therefore considered a novel classification approach.

2 NNs for classification of cutting, ploughing and rubbing in grinding

A large number of researchers have reported the application of using NN models for the classification of phenomena of interest when applied to tool condition monitoring (Ozel and Karpat, 2005; Sick, 2002). A feed-forward NN model was used with the back-propagation learning strategy to provide the segregation of data (Rumelhart et al., 1986). Commonly, NNs are used for pattern recognition in image analysis or sound waves in signal analysis. The NN consists of a complex interconnection of units which are otherwise known as nodes or neurons. The general layout for a NN consists of a set of neuron layers connected together through complex connections; this layout and features is known as the network architecture.

A two-layer NN model can map the basic logic functions of OR, AND and NOT however, a hidden layer is required when mapping non-linear functions such as that of exclusive-OR or the much complex functions such as the data presented by STFT signal processing techniques. This type of data is not only non-linear but also n-dimensional. The basic logic function network classifiers use a linear data separation approach however with the separation of much larger data sets there is need for a more dynamic learning system that takes all the information into consideration and maps the data in both parallel and gradient descent segregation fashion such as that seen by the back-propagation feed-forward network (McCulloch and Pitts, 1943). A Multilayer Perceptron (MLP) utilising the back-propagation learning rule is presented in Figure 1 for illustration purposes.

Figure 1 A four input NN with one hidden layer



As displayed in Figure 1 each of the inputs P_1–P_4 are multiplied by a changing weight function and are associated to a target vector, in this example a_1–a_2, respectively. This is called the associations of input–output pairs and provides the supervised training data (test and verification data set have both data that has been seen by the network in training (supervised) and data that has not been seen (testing the generalisation of the network). Each neuron has a summation function which sums up the weighted (e.g. $w^1_{1,1}$ and $w^1_{1,2}$ to $w^2_{n,n}$ reference to Figure 1) and input bias (bias input variable) connections. The transfer function (for non-linear problems a differential transfer function; such as Tan-sigmoid is used) is required to map the non-linear input–output relations which are obtained for each neuron and updated in an iterative fashion towards the desired target set. Back-propagation is so-called as the weights are updated from the error between the actual output and the desired output which in short is from the back to the front. 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.

The output of each neuron is a function of its inputs. Specifically, output of the jth neuron is any layer is described by the following equations:

$$U_j = \sum\left(Pi \times w_{ij}\right) \tag{1}$$

$$a_i = F\left(U_j + t_j\right) \tag{2}$$

For every neuron, j, in a layer, each of the i inputs, X_i to that layer is multiplied by a previously established weight, w_{ij}. These are all summed together, resulting in the internal value of the operation, U_j. This value is then biased by a previously established threshold value t_j, and sent through an activation function, F (Sigmoid or Linear) giving the NN output; a_i.

Equation (3) describes the output error obtained from each neuron.

$$\mathrm{ME} = \frac{1}{\Omega}\sum_{i=1}^{\Omega}\left(t_i - a_i\right)^2 \tag{3}$$

where ME is the mean squared error, a_i (a_1 and a_2 in the example displayed by Figure 1) is the output of the network corresponding to ith input P_1–P_4 looking at the example displayed in Figure 1. The error term of network is given from $(t_i - a_i)$ where t_i is the target vector or the desired value for given input vectors P_1–P_4. The T is used to transpose the matrix to ensure both matrices are multiplied together to get a Sum Squared Error (SSE) output. The error function can be applied to the NN in a batch training fashion at the end of data presentation or sequentially after each input–output pair.

For the back-propagation algorithm the weight and bias update equations are as follows:

$$\Delta w_{ij}^{k} = -\alpha \frac{\partial \mathrm{ME}}{\partial w_{ij}^{k}} \tag{4}$$

$$\Delta b_{i}^{k} = -\alpha \frac{\partial \mathrm{ME}}{\partial b_{i}^{k}} \tag{5}$$

where α is the learning rate, which has a trade-off in value to ensure it is small enough to gain a true convergence but large enough to separate the data space in adequate time. Equations (4) and (5) are iteratively changed across the network along with other functions to provide learning sensitivity. This process of weight and input, and error calculation propagates through the NN to provide the segregation rules which separates the data according to class (target vector). The b is a bias term used to influence the training weights and for NN training.

3 Clustering method for classification of cutting, ploughing and rubbing in grinding

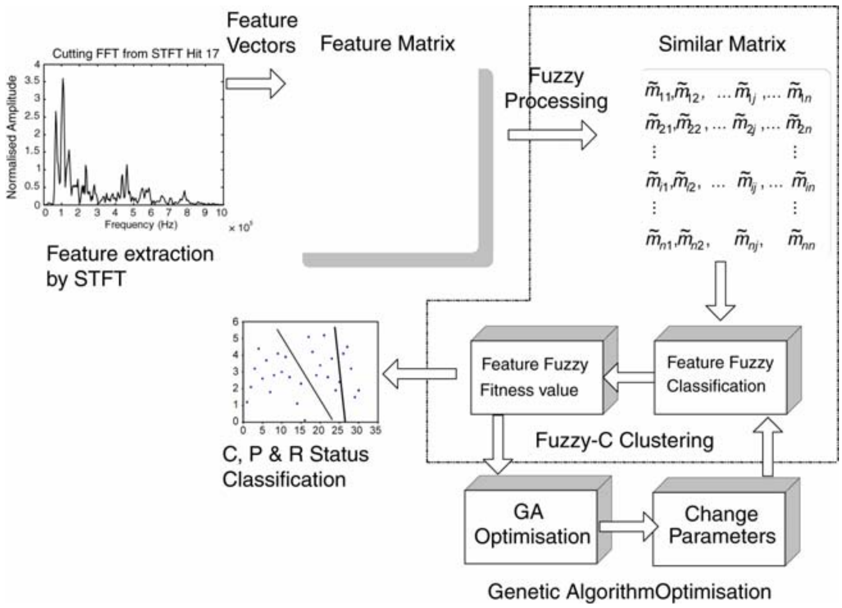
Another pattern recognition technique used for classification is through fuzzy-c clustering (Li and Xiaoli, 1998; Liu et al., 2005). Pattern recognition can select features of interest based on all data features. The technique of fuzzy clustering provides rules in the form of distance measurements that segregate the different cluster sets from each other, in this case; the cutting (*C*), ploughing (*P*) and rubbing (*R*).

Clustering techniques has emerged from work carried out in statistical probability (Cuevas et al., 2001; Hartigan, 1985). When looking at real world phenomena most cases are not finite and instead possess a lot of in between values such as that seen with fuzzy sets. Fuzzy-c mean algorithm is an iterative technique for clustering data sets in a soft rule set fashion. It is a technique for grouping data and finding structures in data. Essentially clustering techniques use a distance measure to segregate like data from other presented data into classes or sets (clusters).

There are two main types of clustering techniques; the conventional way of clustering is through hard clustering where partitions are formed representing each pattern similar to a threshold measure used for preprocessing the NN outputs. The difference in hard clustering is the data belongs to only one cluster. With non-linear data there is no sharp classification between clusters, especially at the boundaries; this is why fuzzy clustering is better as it can assign clustering between 0 and 1 and not just one fixed value, therefore giving cases for both cluster sets. The classification here is based on the set the input data has more similarities to when compared with the other and is known as soft clustering.

Figure 2 displays a block diagram of the fuzzy-c/GA is based on work carried out by (Liu et al., 2005) clustering of the cutting, ploughing and rubbing. The first step in this process is to convert the STFT vectors into a fuzzy similarity matrix defining the relations of similarity. The next process is to then define the fuzzy variable similarity matrix which evaluates each coefficients $\tilde{r}_{ij}$ degree of membership between the element *i* and *j*. Following on, the cluster centres are determined, the centres segregate and categorise one cluster from another. The centre cluster is the representing function of a particular cluster. Input test data that has a membership close to a particular centre than other centres mean that input data belongs to that centre.

Figure 2 Block diagram of the C, P and R classification



Let $X = \{X_1, X_2, \ldots, X_m\}, X \subset V,$ where X_1–X_m are feature vectors make up the total feature matrix set and $X_i = (X_{il}, X_{i2}, \ldots, X_{in}) \in V$ is a feature vector (element of total feature matrix set V); X_{ij} is the jth feature of individual X_i the feature matrix are made up from 1 to n feature vectors. To ensure there is normalisation across the feature matrix, Equation (6) is used which calculates the normalised mean for each input value divided by the variance.

$$x'_{ij} = m_j X_{ij} \text{ and } x''_{ij} = \frac{\left(x'_{ij} - \overline{x}'_{ij}\right)}{\sigma_i} \tag{6}$$

where

$$m_j = \max \frac{\left(X_{11}, X_{12}, \ldots, X_{1n}\right)}{X_{1j}}, \quad \overline{x}'_i = \frac{1}{n}\sum_{j=1}^{n} x'_{ij} \text{ and } \sigma_i = \left|\frac{1}{n}\sum_{j=1}^{n}\left(x'_{ij} - \overline{x}'_i\right)^2\right|^{1/2}$$

The normalised feature matrix is then represented by the feature matrix below in Equation (7).

$$X(m \times n) = \begin{bmatrix} x''_{11}, x''_{12}, \ldots, x''_{1j}, \ldots, x''_{1n} \\ x''_{21}, x''_{22}, \ldots, x''_{2j}, \ldots, x''_{2n} \\ \vdots \qquad\qquad \vdots \\ x''_{i1}, x''_{i2}, \ldots, x''_{ij}, \ldots, x''_{in} \\ \vdots \qquad\qquad \vdots \\ x''_{m1}, x''_{m2}, \ldots, x''_{mj}, \ldots, x''_{mn} \end{bmatrix} \tag{7}$$

The fuzzy similarity matrix is the next calculation required for the fuzzy clustering of the input data set. The similarity matrix uses a distance measure to show similarities within the matrix set. There are many distance functions available; however, fuzzy-c clustering uses Equation (8). The index of similarity is based on the minimum distance that equates to the maximum similarity.

$$m_{ij} = \frac{\sum_{k=1}^{n}\left|\left(x_{ik} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_j\right)\right|}{\left\{\left[\sum_{k=1}^{n}\left(x_{ik} - \overline{x}_i\right)^2\right]\left[\sum_{k=1}^{n}\left(x_{kj} - \overline{x}_j\right)^2\right]\right\}^{1/2}} \tag{8}$$

By using the correlation coefficient Equation (8) the normalised feature matrix is converted into a fuzzy proximity matrix M:

$$M = \begin{bmatrix} m_{11}, m_{12}, \ldots, m_{1j}, \ldots, m_{1n} \\ m_{21}, m_{22}, \ldots, m_{2j}, \ldots, m_{2n} \\ \vdots \qquad\qquad \vdots \\ m_{i1}, m_{i2}, \ldots, m_{ij}, \ldots, m_{in} \\ \vdots \qquad\qquad \vdots \\ m_{m1}, m_{m2}, \ldots, m_{mj}, \ldots, m_{mn} \end{bmatrix} \tag{9}$$

The fuzzy proximity matrix *M* is then converted into a fuzzy similarity matrix M^K as the proximity relationship does not have enough similarities for fuzzy clustering to be carried out. From the using the fuzzy algorithm such as transitive closure, the fuzzy matrix *M* can be converted into the fuzzy similarity matrix M^K.

$$M^K = \begin{bmatrix} \tilde{m}_{11}, \tilde{m}_{12}, \ldots, \tilde{m}_{1j}, \ldots, \tilde{m}_{1n} \\ \tilde{m}_{21}, \tilde{m}_{22}, \ldots, \tilde{m}_{2j}, \ldots, \tilde{m}_{2n} \\ \vdots \qquad\qquad \vdots \\ \tilde{m}_{i1}, \tilde{m}_{i2}, \ldots, \tilde{m}_{ij}, \ldots, \tilde{m}_{in} \\ \vdots \qquad\qquad \vdots \\ \tilde{m}_{m1}, \tilde{m}_{m2}, \ldots, \tilde{m}_{mj}, \ldots, \tilde{m}_{mn} \end{bmatrix} \tag{10}$$

$$\min\left\{ J_m = \sum_{j=1}^{m} \sum_{i=1}^{n} \left| \mu_j(x_i) \right|^b \left\| x_i - c_j \right\|^2 \right\} \tag{11}$$

Looking at Equation (10), $\tilde{m}_{ij}$ in the matrix M^k is the similarity between feature *i* and *j*. The maximum value of similarity is when $i = j$ and the feature itself equates to 1. 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 STFT data in terms of cutting, ploughing and rubbing phenomenon. The closest distance membership function of fuzzy-c clustering is based on the squared loss cost function (see Equation (11)) for each point. For Equation (11), x_i is the samples ($i = 1, 2, \ldots, n$), *m* is the number of known clusters, c_j is the cluster centre point where ($j = 1, 2, \ldots, m$), $\mu_j(x_i)$ is the fuzzy membership of sample x_i to cluster *j*. The *b*, term if equals 1, tends to more towards *k*-means clustering, similar to city-block distance statistical measure and if *b* tends towards ∞ it becomes completely fuzzy similar to Chebyshev maximum distance clustering. If however the term *b* takes the value of 2 it is similar to the Euclidean distance technique which was used in this work. The fuzzy algorithm iterates through Equation (11) until it can no longer best fit the separation of one cluster from another.

4 GA for classification of cutting, ploughing and rubbing in grinding

The GA is also a biologically inspired technique that optimises the evolutionary process of living organisms. They were invented specifically to avoid getting a solution stuck in a local minimum and to cover as much of the solution space as possible. The essential feature of a GA (Johnson and Picton, 1995) is the group of chromosomes that contain the genetic information. The genetic information is in the form of strings which define a particular solution. For instance a six bit binary number would be represented by the following string and stored as a chromosome: 1 0 1 0 0 1. The GA randomly produces strings forming a population. Once there are a significant number of strings in the population, a fitness function is used to test whether a particular chromosome is used to influence a new population, or whether it is to be discarded. The better of these chromosomes is kept in the population for the next generation. Thus each successive population of chromosomes will have a greater cumulative fitness compared with its predecessor. The fit chromosomes are then chosen for breeding; this then allows the fit chromosomes to influence the next population.

The breeding mechanism is known as cross-over and can be seen from the example below:

Parent 1	**1 0 1 | 0 1 0**	→	offspring 1	**1 0 1** | 0 1 1
Parent 2	1 1 0 | 0 1 1	→	offspring 2	1 1 0 | **0 1 0**

The result is two offspring chromosomes; combining the digits of the parents according to the cross-over point chosen (middle of the chromosome was chosen here).

Another influencing factor on a population is mutation, this is where a point is chosen on the chromosome and that genetic material is changed, in this case inverted. Mutation only occurs on randomly selected chromosomes, this random feature is dependant on the mutation rate (higher the rate the more random mutation becomes). The example below will show mutation be applied to the 5th bit of offspring 1.

Offspring 1	1 0 1 0 1 1	
	↓	
Mutated offspring 1	1 0 1 0 0 1	(5th bit inverted)

The GA was used to interact with the fuzzy-c clustering set by searching the solution space with a Darwinian fitness approach. The GA would convert the best individual from a population (genotype function) into two variables (phenotype functions) being: the cluster number and the number of iterations. From the returned fitness value of the fuzzy-c clustering algorithm the GA can evaluate the best individual. The fuzzy-c clustering algorithm is executed for each best individual presented. If the fitness value is less than the fitness function returned it is discarded as genotype material for the next GA pass, otherwise it used in the next GA pass. By continually simulating the fuzzy-c clustering algorithm the best fitness gained would result as the given classifier. If the GA was not used the fuzzy-c clustering algorithm is less likely to gain the optimised cluster set.

5 Experimental setup of horizontal SG scratch tests and grinding wheel tests

The scratch test simulates the grinding chip formation where the associated extracted AE of grinding chip formation may then be investigated. The scratch tests can be carried out by feeding a rotating Al_2O_3 grit towards a flat horizontally placed workpiece (CMSX4) as described in the Part 1 of this paper. With a micron incremental grit in and out stroke, a scratch groove will be formed on the surface of the flat sample. The average scratch depth is about 1 μm, which is a typical value of grinding chip in high efficiency grinding. The scratching wheel rotational speed is 4000 rpm with a feed rate of 4000 mm/min. The depths of cut of grinding tests were 1 μm and 0.1 mm under down grinding condition. The depths of cut were measured from the height change of the workpiece before and after grinding, in addition both workpieces were measured with a Fogale Photomap inference microscope to confirm the interactions. For the grinding tests, the wheel rotational speed was maintained at 4000 RPM with a federate of 4000 mm/min to compare like with like interaction. Both classification analysis and digital signal processing were carried out under the Matlab computing environment.

6 NN classification of cutting, ploughing and rubbing

The supervised training method uses three times larger than that of the examples used in both the test and verification data sets (in terms of vector numbers; 180 for training and 60, 60 vectors for training and verification data sets, respectively). The network verification test set similar to the test set albeit slightly different to ensure the network provides a good confidence when presenting its classifications. The classification accuracy is based on correct classification against misclassification data. The network momentum value was set to a high value to ensure the NN can search beyond localised minimum and stick to the minimum post near global minima. Table 1 lists the parameters used in the NN to classify the cutting, ploughing and rubbing phenomenon

Table 1 NN parameters for NN classifications

NN Parameters	*Value*
hidden layers	2
Input size	STFT processed data: 207 neurons
Transfer function for layer 1,2,3	Tan-sigmoid
Transfer function for output layer	Pure-linear
Epochs	10,000 for (1) Time: 37 min
Learning rule	Back-propagation
Learning rate	0.1^{-9}
Momentum	0.9
NN training performance	5.2423×10^{-31}
Training	STFT: 148 equally applied as phenomenon cases

Figures 3 and 4 represent both the NN test and test verification data sets. The STFT signal NN test results are very encouraging with an 87% unseen classification (out of total 60 test vector set) and overall network classification of 93%. To give further confidence (James, 1994; Peterson and Gerald, 1992), the STFT signal NN test verification results give an 83% unseen classification accuracy and total network accuracy of 92%. These two sets of results are conclusive when classifying SG cutting, ploughing and rubbing phenomenon; two sets of similar test data is used to check consistency of results. The training data was 148 cases which were considered sufficient for data generalisation of the NN.

The next results present Hit 4, Hit 14 and Hit 15 of 1 μm grinding cuts and Hit 20 of the 0.1 mm grinding cut, respectively. In addition to the above, the third hit taken from the rubbing signal before Test 212 (Test 211) is also tabulated to show the classification of pure rubbing. Here a STFT of the hit data was taken from start of the hit AE profile to the finish. Table 2 represents the amount of hit data classified as cutting, ploughing or rubbing and the percentage amounts of cutting, ploughing and rubbing. Figure 5 displays the NN outputs starting with the top left as Hit 4, top right Hit 14 and bottom left Hit 15 for a 1 μm grinding wheel cut. Bottom right displays the NN output for Hit 20 of a 0.1 mm grinding cut. All grinding cuts had no coolant present. It is possible to see the profile of cut from the NN outputs of Figure 5. Table 2 classifications used a rule based threshold system to gain a crisp cutting, ploughing and rubbing output the rule being; if greater than 2.5 then rubbing, if less than 2.5 and greater than 1.5 then ploughing and if less than 1.5 then its cutting phenomenon. T212 Hit 2 data is total hit data and not segments used for training and test (top row) 1 μm to 0.1 mm cuts are for grinding wheel passes and not SG scratch experiments (both NN and fuzzy-c/GA results).

Figure 3 Displays the NN results for cutting, ploughing and rubbing test set

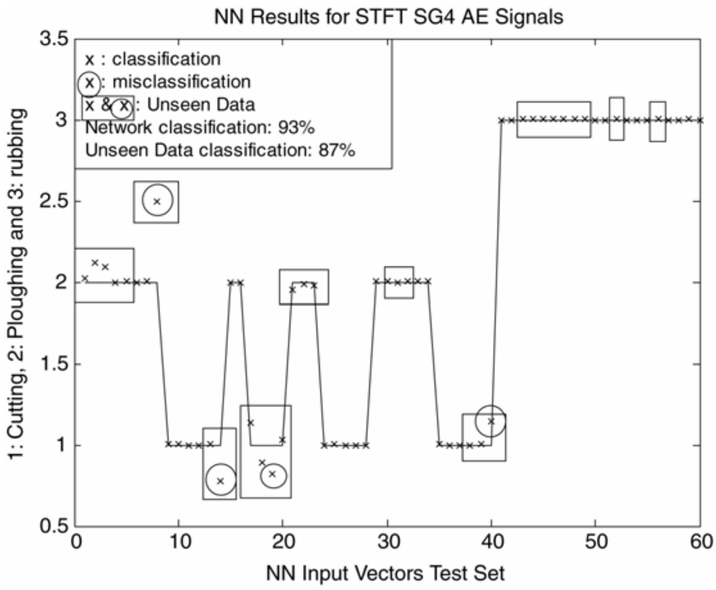


Figure 4 Displays the NN results for cutting, ploughing and rubbing verification set

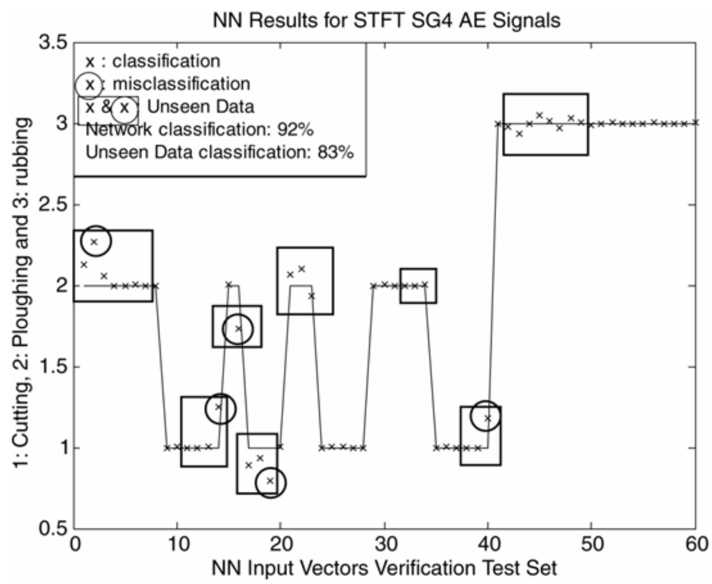
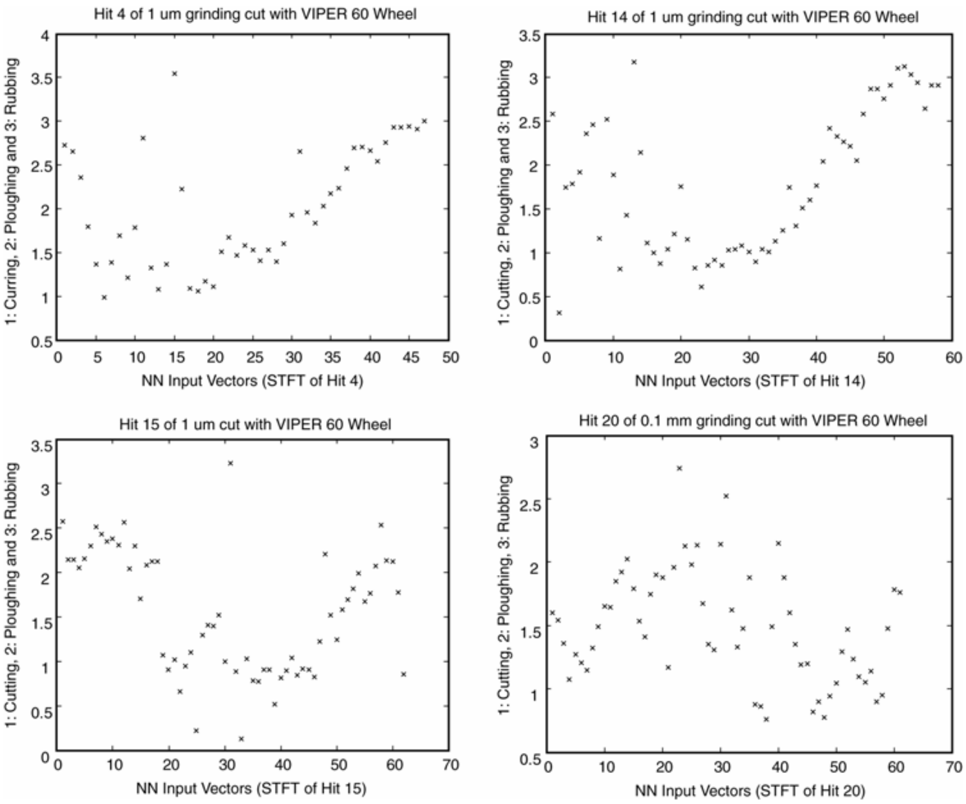


Table 2 Displays NN test vector outputs for grinding wheel cuts

Test set	*NN cutting (C)*	*NN ploughing (P)*	*NN rubbing (R)*	*Total vectors*	*C % correct/ total*	*P % correct/ total*	*R % correct/ total*
Hits 1–23 T211/T212	17/20	18/19	21/21	60	27/33	30/32	35/35
Hit 3 T211 (rubbing)	0	0	21	21	0	0	100
Hit 2 T212	13	20	5	38	34	53	13
Hit 4* (*1 μm cut)	14	17	16	47	30	36	34
Hit 14*	26	17	15	58	45	29	26
Hit 15*	28	29	5	62	45	47	8
Hit 20 0.1 mm cut	34	24	2	60	57	40	3

Figure 5 Top left Hit 4, top right Hit 14, bottom left Hit 15 (1 μm cuts) and bottom right Hit 20 (0.1 mm cut)



Looking at Figure 5 and Table 2 further there appears to be more of a percentage of cutting for the more interaction between workpiece and grit. There is a higher percentage of cutting compared with ploughing and rubbing when the depth of cut is increased for grinding wheel passes (middle parts of the 1 μm cuts had greater depth of cuts when measured against the beginning and end grit hits – this is due to a greater surface area being present during the cut rather than at the start or end). With pure rubbing extracted signal there is hardly any interaction between grit and workpiece which is also displayed by Table 2 results. The NN SSE for the cutting, ploughing and rubbing tests was 5.24×10^{-31} and the number of training epochs was 10,000. This NN engine was used throughout the applied hit data tests of 1 μm cuts and 0.1 mm cuts. The hit tests were applied to the NN engine that had already gained 93% classification accuracy from cutting, ploughing and rubbing tests and was considered a good score for applying the hit data tests. For instance; as the hits are classified from the 1st interactions (Hit 4) to the mid interactions (Hit 14 and Hit 15) appears to have a higher percentage of cutting. Certainly with a mid hit of a 0.1 mm depth cut there is a lot more cutting interaction and a lot less rubbing when compared to the 1 μm cuts (more so with Hit 4 than Hit 14 and Hit 15 as it is at the beginning of grit/scratch interaction). This signifies that cutting, ploughing and rubbing changes in ratio as there is more interaction between workpiece and grinding wheel (depth of cut increases therefore more grit interaction between material surface and grit). Note that AE of all depth cut hits were normalised to a 1 μm cut to compare the cutting, ploughing and rubbing training/test set. The normalisation is used to distinguish between different deforming phenomena. Looking at the lower right of Figure 5 the 0.1 mm cut has more crisp cutting classifications than with other displayed grinding pass cuts this is due to the training data and 0.1 mm grinding cut pass data being very similar in characteristics. With the other grinding pass cuts the NN output displays linear regression with respect to the groove being cut.

7 Classification of cutting, ploughing and rubbing using fuzzy-c clustering/GA

Figure 6 displays the equivalent fuzzy-c/GA cluster outputs to the NN outputs of Figure 5. Figure 6 displays the first two principle components (which contain 65% of the total variance of the 300 cases of 205 elements of data). The fuzzy-c/GA outputs are mapped onto these two principle components which are based on the highest evaluated data cluster centre membership. For cutting data two cluster centres were output based on the optimised criteria between fuzzy-c algorithm and GA.

Table 3 displays the known classification accuracy of fuzzy-c clustering. Here the correct/incorrect clusters were checked against the known phenomenon and a percentage of classification was determined. Table 4 displays the percentage of cutting, ploughing and rubbing phenomenon for Hit 4, Hit 14 and Hit 15, 0.1 μm cuts and Hit 20, 0.1 mm cut.

From looking at Table 4 it is again possible to see more utilisation of cutting, then ploughing and rubbing when the depth of cut increases leading to more interaction between grit and workpiece. The fuzzy-c/GA classifier was first tested for the SG4 data (T212 and T211 SG4 full cutting, ploughing and rubbing data) and the classifier produced 90% classification accuracy. Then the output of the hit data (for 1 μm and

0.1 mm cuts, respectively) using the fuzzy-c/GA clustering classifier is displayed in Figure 7, this is in comparison with the NN output displayed in Figure 5. For each individual hit the classification consisted of both the training and test SG4 data used in the NN experiments.

Figure 6 Fuzzy-c clustering, top left: Hit 4, top right: Hit 14, bottom left: Hit 15 of 0.1 μm cut and bottom right: Hit 20 of 0.1 mm cut

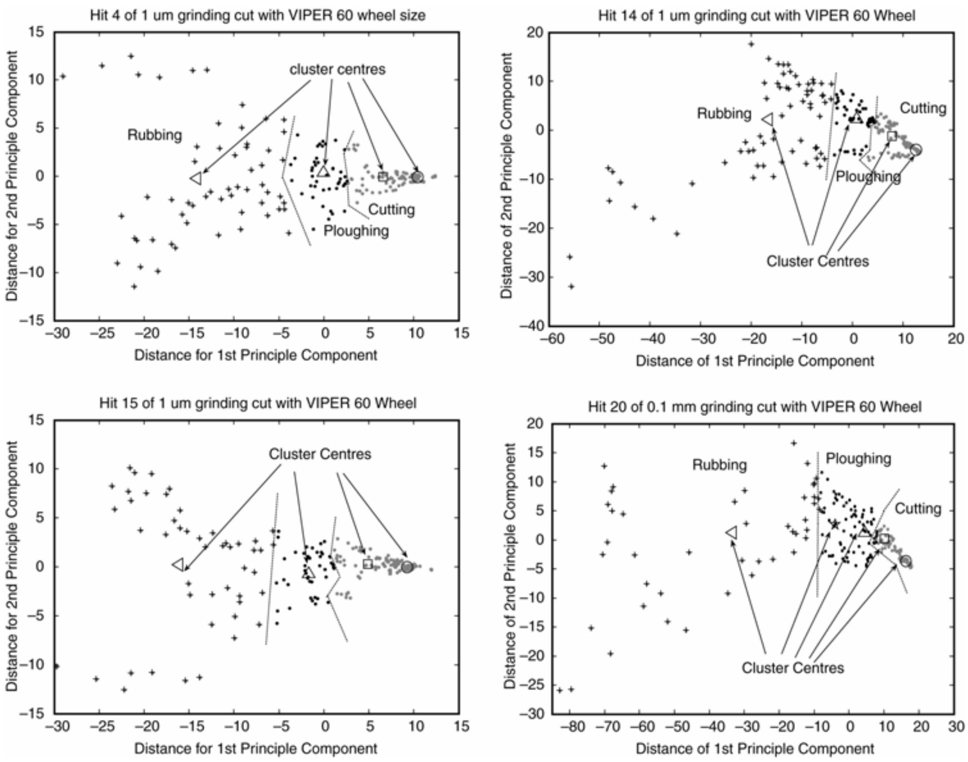


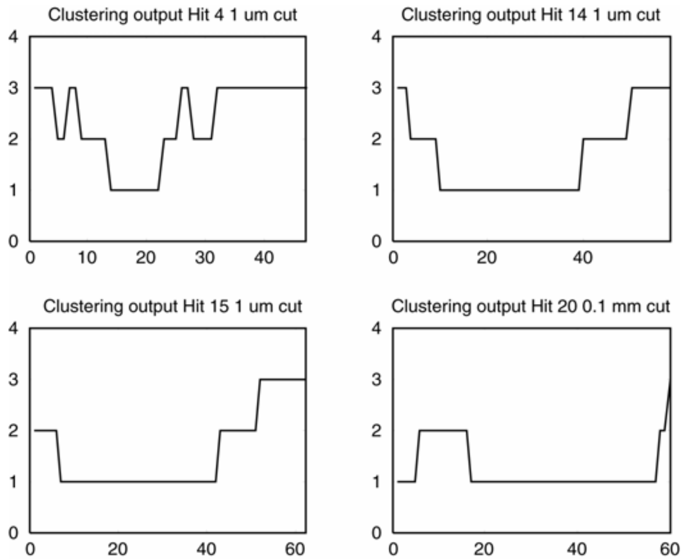
Table 3 Fuzzy-c clustering results for 0.1 μm cuts and 0.1 mm cut

Test set	*FuzzyGA cutting (C)*	*FuzzyGA ploughing (P)*	*FuzzyGA rubbing (R)*	*Classification accuracy %*
Test SG4	61/62	43/47	81/97	90 (185/206)
Hit 3 T211 (rubbing)	53/55	51/66	84/85	91 (188/206)
Hit 2 T212	52/55	60/66	84/85	95 (196/206)
Hit 4* (*1 μm cut)	69/69	41/50	71/87	88 (181/206)
Hit 14*	42/48	58/59	68/99	82 (167/206)
Hit 15*	46/46	72/96	55/64	84 (173/206)
Hit 20 0.1 mm cut	39/44	82/85	77/77	96 (198/206)

Table 4 The fuzzy-c cluster for the percentage of C, P and R phenomenon

Test set	*C (%)*	*P (%)*	*R (%)*	*Iterations and clusters*	*Fitness*
Test SG4	33 (20/60)	23 (19/60)	44 (21/60)	78/4	0.7245
Hit 3 T211 (rubbing)	0 (0/21)	0 (0/21)	100 (21/21)	190/6	0.822
Hit 2 T212	32 (12/38)	61 (23/38)	8 (3/38)	153/5	0.1305
Hit 4 (1 µm cut)	19 (9/47)	26 (12/47)	55 (26/47)	135/6	1.659
Hit 14 (1 µm cut)	52 (30/58)	23 (13/58)	21 (12/58)	81/4	0.8862
Hit 15 (1 µm cut)	58 (36/62)	24 (15/62)	18 (11/62)	165/6	1.85
Hit 20 0.1 mm Cut	76 (48/60)	22 (13/60)	2 (1/60)	120/5	0.968

Figure 7 Displays the fuzzy-c clustering and GA classification for grinding wheel cuts



When the associated hit data for 1 µm and 0.1 mm cuts were concatenated with the training data, the classification had to be recalculated each time the different hit was presented this is due to the training data being of known phenomenon distinction and the 1 µm and 0.1 mm being of unknown phenomenon distinction. This creates a problem for verification purposes in that the more unknown data the less to check the recently added test data. By calculating each individual test case to the known cluster memberships the more accuracy and confidence is gained in phenomenon distinction. It would appear that both the NN and fuzzy-c/GA clustering techniques have given similar results and therefore the findings in this paper are conclusive of the cutting, ploughing and rubbing phenomenon distinguished by the energy released from the workpiece and grit interaction in the form of an AE signal. That said where the results are slightly in favour of one phenomenon than another when compared between the two classification techniques this can be attributed to incorrectly classifying a cluster set (hence no hard separation rule with fuzzy-c clustering) or a low returned fitness value. In short if the classifier produces 90% or more for the known data set then there is more confidence when the rubbing, 1 μm and 0.1 mm cuts are introduced to the classifier.

8 Discussion of classification results

From comparing the results between the two classifiers there are differences between the given outputs. These differences can be attributed to the different methods of classification. For instance; looking at Figure 6 the NN outputs for the rubbing, 1 μm and 0.1 mm cuts have levels between 1 and 3 and were not identified as crisp integers. This is due to the presented data being similar albeit different to the training and test data of the original NN data set. The varying levels display linear patterns with reference to the interaction of between the workpiece and grit. A threshold rule based system was used to classify which was cutting, rubbing and ploughing. The thresholds used were any input vector less than 1.5; cutting, greater than 1.5 but less than or equal to 2.5; ploughing and above classified as rubbing. This hard classification rule based system segregates the boundaries based on absolute outputs. The fuzzy-c clustering/GA classification system looks at all the data points and works out a crisp output albeit the identified cluster may have the same distance as another cluster or be very near to another cluster and therefore has attributes for both classes. It was noted that some of the outputs for the fuzzy-c clustering/GA classification system had data points belonging to two cluster pairs however the one absolute cluster was chosen based on its position in the identified fuzzy matrix. The output differences have correlations with each other in that the more workpiece and grit interaction through measured depth of cut the more cutting phenomenon occurs followed by ploughing then lastly, rubbing. Both classifiers had a high confidence in terms of classifying both the training and test data with 93% and 90%, respectively for the NN and fuzzy-c clustering/GA classifiers. That said, the data hit data presented to the NN had close boundary conditions and with the hard rule based threshold postprocessor some that was applied to the NN for 1 μm and 0.1 mm cuts, output cases could have been cutting when instead they were ploughing and the same for ploughing with rubbing. With fuzzy-c/GA the results that were significantly different to the NN results had a higher fitness function, when a lower fitness function returns a higher classification accuracy. In addition, more clusters are returned which can easily be classified incorrectly giving an incorrect classification. Even with some similar but slightly different results gained by the comparison of the two classifiers a pattern can be scene in the classification of the three phenomenons. This pattern is summarised as greater intensities towards rubbing when grinding wheel passes with slight touch are made and, greater intensities towards cutting when operational grinding wheel passes are made.

9 Conclusions

The 1st part of this work concerned identifying the fundamental signatures of SG phenomenon which is the key to the process of monitoring grinding. This part of the work concludes the investigation of horizontal SG scratch cutting, ploughing and rubbing by classifying the three different phenomenon using NNs and fuzzy-c clustering/GA classification techniques. This paper has demonstrated that STFTs as a useful technique to distinguish the frequency bands occupied by cutting, ploughing and rubbing phenomena.

Both the NN and fuzzy-c clustering/GA classifiers were found to have a high confidence of distinguishing cutting, ploughing and rubbing phenomenon. The NN classified the phenomena at 87% for the unseen test data set classification accuracy and, 93% for the total test classification accuracy. The fuzzy-c clustering GA had a 90% of total data set classification accuracy (both training and test sets was presented to the classifier for this accuracy result).

The further tests looked at hit data taken from grindings with 1 μm and 0.1 mm depth cuts, both classifiers indicated more percentage of cutting utilisation when that the process had more interaction between workpiece and grit (i.e. measured increased actual depth cut). The rubbing (Test 211) and 1st hit for 1 μm scratch had more rubbing percentage utilisation when compared with both ploughing and cutting phenomenon. Looking at the general patterns from the three different grinding passes both classifiers provided some encouraging results with differences attributed to classification accuracy and increased cluster sets for the fuzzy-c GA classifier where the NN engine remained the same throughout the applied grinding pass data. There will always be slight differences between the two classifiers as fuzzy-C/GA is a non-supervised technique and NN is a supervised technique.

References

Chen, M. and Xue, B.Y. (1999) 'Study on acoustic emission in the grinding process automation', *American Society of Mechanical Engineers, Manufacturing Engineering Division, MED Manufacturing Science and Engineering – (The ASME International Mechanical Engineering Congress and Exhibition)*, 14–19 November, Nashville, TN, USA, ASME, Fairfield, NJ, USA.

Coman, R., et al. (1999) 'Acoustic emission signal – an effective tool for monitoring the grinding process', *Abrasives*, December/January, p.5.

Cuevas, A., Febrero, M. and Fraiman, R. (2001) 'Cluster analysis: a further approach based on density estimation', *Computational Statistics and Data Analysis*, Vol. 36, pp.441–459.

Hartigan, J. (1985) 'Statistical theory in clustering', *Journal of Classification*, Vol. 2, pp.63–76.

Holford, K.M. (2000) 'Acoustic emission – basic principles and future directions', *Strain*, Vol. 36, No. 2, pp.51–54.

James, S.D. (1994) 'Verification and validation of neural networks', in P.E. Keller (Ed). *Proceedings of the Neural Network Workshop for the Hanford Community*, Pacific Northwest National Laboratory, Richland, WA, USA, pp.76–80.

Johnson, J. and Picton, P. (1995) *Designing Intelligent Machines*, Open University, Vol. 2, pp.85–94.

Li and Xiaoli (1998) 'Tool wear monitoring with wavelet packet transform-fuzzy clustering method', *Wear*, Vol. 219, No. 2, pp.145–154.

Liu, Q., Chen, X. and Gindy, N. (2005) 'Fuzzy pattern recognition of AE signals for grinding burn', *International Journal of Machine Tools and Manufacture*, Vol. 45, Nos. 7–8, pp.811–818.

MATLAB Neural Network Toolbox (XXXX) Available at: http://www.mathworks.com/products/neuralnet/.

McCulloch, W. and Pitts, W. (1943) 'A logical calculus of the ideas immanent in nervous activity', *Bulletin of Mathematical Biophysics*, Vol. 7, pp.115–133.

Ozel, T. and Karpat, Y. (2005) 'Predictive, predictive modelling of surface roughness and tool wear in hard turning using regression and neural networks', *International Journal of Machine Tools and Manufacture*, Vol. 45, pp.467–479.

Peterson and Gerald (1992) 'A framework for neural network evaluation', *McDonnel Douglas Technical Report MDRL TN-92-02*, McDonnell Douglas Research Laboratories, St. Louis, Missouri.

Rumelhart, D.D., Hinton, G.E. and Williams, R.J. (1986) 'Learning representations by back-propagating errors', *Nature*, Vol. 323, pp.533–536.

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', *Mechanical Systems and Signal Processing*, Vol. 16, pp.487–546.

Webster, J., Marinescu, I. and Bennett, R. (1994) 'Acoustic emission for process control and monitoring of surface integrity during grinding', *Annals of the CIRP*, Vol. 43, No. 1, pp.299–304.