CIVIL AVIATION AUTHORITY

CAA PAPER 99006

INTELLIGENT MANAGEMENT OF HELICOPTER HUMS DATA

Study I **A Demonstration of the Feasibility and Performance of an Intelligent Management System Operating on HUMS In-Service Data**

Study II **Intelligent Management of HUMS Data: The use of Artificial Intelligence Techniques to Detect Main Rotor Gearbox Faults**

CIVIL AVIATION AUTHORITY, LONDON

Price £17.00

CAA PAPER 99006

INTELLIGENT MANAGEMENT OF HELICOPTER HUMS DATA

Study I **A Demonstration of the Feasibility and Performance of an Intelligent Management System Operating on HUMS In-Service Data**

Study II **Intelligent Management of HUMS Data: The use of Artificial Intelligence Techniques to Detect Main Rotor Gearbox Faults**

Study I Prepared and approved by Dr Hesham Azzam

Study II Prepared by N Harrison
Approved by Dr N C Baines

REPORT PREPARED BY MJA DYNAMICS LTD, HAMBLE, HAMPSHIRE AND PUBLISHED BY
CIVIL AVIATION AUTHORITY, LONDON, SEPTEMBER 1999

ISBN 0 86039 760 2

First published September 1999
Reprinted December 1999

Printed and distributed by
Westward Digital Limited, 37 Windsor Street, Cheltenham, England

General Foreword

The research reported in this paper was funded by the Safety Regulation Group of the UK Civil Aviation Authority and Shell Aircraft Limited, and was performed by MJA Dynamics Limited. The work forms part of the Authority's helicopter HUMS research programme which was instigated in response to the recommendations of the Review of Helicopter Airworthiness (HARP Report – CAP 491), and the subsequent recommendations of the Report of the Working Group on Helicopter Health Monitoring (CAA Paper 85012).

The issue of the intelligent management of HUMS data was identified for attention in response to concerns over the depth and effectiveness with which individuals are able to examine the large quantities of HUMS data being generated by fleets of helicopters in-service. Successes achieved in other applications had demonstrated the potential of artificial intelligence techniques to assist with such problems, and the Authority wished to establish the effectiveness of the technology in relation to HUMS data. Accordingly, the following two studies were instigated:

I An initial study, started in 1993, was performed to demonstrate the feasibility and performance of a computer-based Intelligent Data Management process, comprising unsupervised machine learning, supervised machine learning, and data pre-processing. The demonstration was performed using Spectrometric Oil Analysis, HUMS and Flight Data Recorder data from in-service helicopters.

II An additional study was subsequently commissioned in 1995 to demonstrate the performance of Artificial Intelligence fault detection techniques using data from two Sikorsky S61 main rotor gearbox seeded defect tests.

This paper contains unabridged versions of the corresponding MJA Dynamics Ltd reports MJAD/R/224/98 and MJAD/R/219/97 respectively.

It is the view of the CAA that the above studies have clearly demonstrated the potentially significant benefits of the application of advanced analysis techniques to HUMS data. The vast quantities of data available to characterise serviceable components and/or systems should enable unsupervised machine learning to be used to particularly good effect. While success with supervised machine learning was also achieved, the absence of large numbers of examples of all possible failure conditions is likely to always limit effectiveness.

The Authority believes that, initially at least, this technology should only be used in combination with existing analysis techniques, making use of unsupervised machine learning to identify data warranting detailed investigation by the analyst. It is anticipated that further development of the system to address the practical issues associated with the routine every day use of the technology will be required. The Authority considers that this would best be achieved by the contractor working in conjunction with one or more helicopter operators having significant experience of the use of HUMS.

Study I

A Demonstration of the Feasibility and Performance of an Intelligent Management System Operating on HUMS In-Service Data

Summary

The work presented in this report was performed under the Civil Aviation Authority (CAA) contract 7D/S/1128. The aim of the contract was to demonstrate the feasibility and performance of a computer based Intelligent Data Management (IDM) process. The IDM process should detect, by *unsupervised learning* methods, abnormal patterns in large volumes of HUMS data even when the underlying cause is unknown. The process should also assimilate the relationships between mechanical faults and abnormalities in the data using *supervised learning* methods. Other factors such as atypical operational conditions or equipment noise, can also induce abnormal data patterns which can trigger false alarms. By using data pre-processing mechanisms, the IDM process attempts to discriminate between the abnormal data patterns induced by such factors and those induced by faults. A HUMS IDM system should provide a framework that integrates these three intelligent processes, namely unsupervised learning, supervised learning and data pre-processing.

The benefits of **unsupervised learning** were demonstrated using CAA supplied Spectrometric Oil Analysis Programme (SOAP) data. The demonstration adhered to the following generic steps, which closely reflect what would happen in-service:

Step 1: Identify and apply initial data pre-processing requirements, i.e. how the data is conditioned before being operated upon by the unsupervised machine learning process.

Step 2: Identify atypical data patterns using the unsupervised machine learning core algorithm – cause may be unknown.

Step 3: If the atypical data patterns persist, trigger an inspection in order to identify the cause.

Step 4: Utilise the identified cause and unsupervised machine learning techniques to determine optimum data pre-processing requirements which consistently highlight the underlying cause.

For Step 1, element concentration levels were selected which were consistently present in the measurement samples and exhibited variability with time. The elements which met these criteria were Iron (FE), Silver (AG), Aluminium (AL) and Magnesium (MG). Step 2 identified two gearboxes which were associated with atypical measurement samples, namely serial numbers M186 and M361. Whilst nothing untoward was reported on M361, gearbox M186 had suffered spalling of a mast bearing (Step 3). The data pre-processing refinement (Step 4) revealed that correlated noise within a measurement sample could be mitigated by dividing concentration levels by one selected concentration level, and uncorrelated noise between samples better managed by the implementation of a moving average. Employing the refined pre-processing, two further gearboxes were consistently exhibiting atypical characteristics, namely serial numbers M468 and M285. Whilst the former gearbox was associated with a high SOAP sampling frequency (indicative of engineering staff concern), a strip report on M285 revealed that spalling of an epicyclic bearing had occurred. The demonstration of the four steps in the machine learning process has shown that atypical measurement patterns may be identified even when the cause is unknown and, once the measurement/cause relationship is isolated, can be used to refine the selection of measurements and their conditioning. Significantly, the analysis process presented is independent of what is being measured and therefore will be of fundamental importance in the interpretation of HUMS data, which contains many different types of measurements from a variety of sensors.

The importance of **data pre-processing** was demonstrated by utilising Flight Data Recorder (FDR) parameters in the context of helicopter HUMS data management and analysis. Model-based and statistical data pre-processing were used to mitigate operational influences and noise effects on measured airframe vibration. This process, which has been called normalisation, produced vibration signatures which only reflected the mechanical health of the helicopter. The demonstration adhered to the following generic steps:

Step 1: Define a method to filter out the noise which can be correlated with factors such as individual helicopter effects or age but can not be definitely estimated. This type of noise is called the correlated noise.

Step 2: Establish a math-model to pre-process the FDR parameters so that the vibration induced by operational conditions can be adequately simulated and thus eliminated.

Step 3: Train a simulation system so that the system can predict the vibration amplitude given the flight parameters.

Step 4: Three criteria for success are monitored and, if satisfied, the system will be accepted; otherwise, the normalisation process will be refined by re-working steps 1, 2 and/or 3. The criteria are 1. the correlation between the measured and predicted vibration is high and 2. the probability distribution of the normalised vibration amplitude of healthy helicopters is peaky at a central value defining the normal vibration and 3. the system can predict the vibration of test data from FDR parameters with the same accuracy achieved during training.

Step 5: Filter the random noise.

For Step 1, a simple statistical model was established to filter the correlated noise. The statistical model was based on temporal monitoring of the average vibration levels of each helicopter. For Step 2, a simplified math-model which related the vibration to the Mach number at the tip of the advancing rotor blade and the air density was initially considered. For Step 3, a Multi-Variate Regression (MVR) analysis was used to test the non-linear modelling capability of the model-based pre-processing of Step 2. For Step 4, the correlation coefficient between the measured and predicted 4R lateral vibration of three Super Puma MK I helicopters was evaluated and found to be low (0.17). This triggered the refinement process which started by establishing a detailed math model for effective pre-processing. The model utilised 12 FDR parameters to simulate the vibration. The correlation coefficients in this case were 0.78 and 0.67 for lateral and vertical vibration respectively. The probability distributions as well as the generalisation capability of the system were found to be satisfactory. A further normalisation refinement was achieved by filtering the uncorrelated noise. For Step 5, a formula that can attenuate the random noise was presented and it was established that the method can be optimised such that only noise and not health states are attenuated.

The benefits of **supervised machine learning** were demonstrated using 133 SOAP samples taken from 10 gearboxes fitted on 6 Super Puma helicopters. Eight samples were related to a mast bearing spalling fault. This fault had been indicated by a magnetic chip detector prior to the head replacement. Nine samples were related to an epicyclic bearing spalling. With the cause known, a Multi-Variate Regression (MVR) system and an Artificial Neural Network (ANN) were trained using 58, 4 and 5 SOAP samples relating to normal bearings, mast bearing spalling and epicyclic bearing spalling respectively. The diagnostic capabilities of the systems were tested by presenting the effects (measurements) which were not used for training to the systems and comparing the predictions (mechanical state) with the actual causes; 100% of the normal bearing samples, 100% of the epicyclic bearing fault samples and 75% of the mast bearing fault samples were correctly identified. Two mast bearing samples were classified as normal. Nevertheless, the correctly identified samples

indicated that the supervised systems could have detected the mast bearing fault 170 hours prior to the magnetic plug.

The performance of an IDM system was demonstrated by identifying its ability to recognise previously reported and unreported features in HUMS data. For the purpose of this demonstration, 5585 data downloads from Bristow Helicopters Limited HUMS were analysed. The data consisted of vertical and lateral airframe vibration harmonics along with main rotor track and lag measurements from 23 Super Puma MK I helicopters. The vibration harmonics were extracted at multiples of the main rotor frequency. The data was used to establish and refine an IDM process. The process was targeted at main rotor non-adjustable faults with particular reference to frequency adaptor faults. These faults manifest themselves through features evaluated from high and low loading conditions. Only climb, descent, cruise and Minimum Pitch on Ground (MPOG) measurements were therefore considered. The data were filtered to attenuate the random noise, and the influence of main rotor adjustable faults was removed. Diagnostic features were extracted from the residual vibration and blade displacements. The features were ratios between vibration amplitudes, differences between phase angles and predicted stiffness values of blades' frequency adaptors. The IDM process was partially refined using a mathematical model, features selection and supervised classification.

Another data set was pre-processed and used to test the performance of the refined process. The data consisted of 1812 downloads of which 1412 downloads did not overlap with the first data set. The overlapped data was only included to ensure smooth filtered results and was not used in testing. As a result of clustering the test data, an alarm was triggered regarding the G-BLXR helicopter and a query was raised regarding the G-TIGT helicopter. The IDM process suggested that G-BLXR had a main rotor non-adjustable fault which was highly likely to be a frequency adaptor fault. The IDM analysis also reported a signature which had not been reported before. This signature indicated that the effect of the fault was equivalent to a difference in the flap-wise stiffness between two opposite blades. Regarding G-TIGT, the results of the IDM analysis indicated that maintenance actions could have been carried out and not recorded in the HUMS database. The observations of the line engineer regarding the two helicopters substantiated, to some extent, the IDM analysis. G-BLXR was under heavy maintenance and the Data Acquisition and Processing Unit (DAPU) of G-TIGT had been replaced. The line engineer also indicated his concern regarding G-TIGO. The cluster analysis did not identify this as an atypical helicopter. This was attributed to the absence of valid data records; a valid data record is a record that contains information regarding the four operational conditions required for diagnosis.

This section of the project demonstrated the importance of an IDM framework which would facilitate interactions between pre-processing, unsupervised learning and supervised learning. Central to the success of an IDM system are the refinement processes which should be based on realistic mathematical models, representative extracted features, enhanced supervised and unsupervised algorithms, and engineering knowledge from in-service experience and HUMS data. It is important to appreciate that the IDM process described in this report was developed for demonstration purposes. A practical IDM system for main rotor non-adjustable faults should be based on the framework described in this report. The system should not only possess the benefits of the IDM processes, but the software should also be capable of accommodating the refinements without the need for re-designing or re-certifying the system. In order to maximise the benefits of the system, a dedicated programme should consider using a large data set and concentrate on the refinement issues.

Contents

Glossary

Ag	Silver
Al	Aluminium
ANN	Artificial Neural Network
ART	Adaptive Resonance Theory
DAPU	Data Acquisition and Processing Unit
CAA	The Civil Aviation Authority
Cr	Chromium
Cu	Copper
FDR	Flight Data Recorder
Fe	Iron
HUMS	Helicopter Health and Usage Monitoring Systems
IDM	Intelligent Data Management
LMS	Least Mean Square
Mg	Magnesium
MPOG	Minimum Pitch angle On Ground
MVR	Multi Variate Regression
Ni	Nickel
PCR	Principal Component Regression
ppm	parts per million
RTB	Rotor Track and Balance
Si	Silicon
SOAP	Spectrometric Oil Analysis Programme
Zi	Zinc

1 INTRODUCTION

Helicopter Health and Usage Monitoring Systems (HUMS) will typically generate in excess of 1 megabyte of data per flight, which will be down-loaded to ground station computers for further analysis. Such vast quantities of data cannot be effectively examined by individuals; symptoms associated with developing mechanical faults could be overlooked. A need therefore exists to automate an intelligent scanning process of the data and, in particular, report pattern abnormalities even when the underlying cause is unknown. An intelligent process must be able to discriminate between the vibration signatures induced by atypical operational conditions and those induced by faults. In addition, a supervised process is required to assimilate the relationships between causes (mechanical defects) and effects (symptoms as indicated by HUMS measurements).

The work presented in this report is performed under the Civil Aviation Authority (CAA) contract 7D/S/1128, and is driven by five objectives, namely:

objective 1: demonstrate the benefits of unsupervised machine learning techniques by searching for patterns associated with bearing defects in spectrometric oil analysis SOAP data.

objective 2: demonstrate the diagnostic importance of data pre-processing by utilising FDR data combined with HUMS vibration data.

objective 3: demonstrate the benefits of supervised machine learning techniques by making use of an Artificial Neural Network (ANN) operating on HUMS data.

objective 4: demonstrate the feasibility of an Intelligent Data Management (IDM) system by integrating data pre-processing with unsupervised and supervised machine learning facilities, and interfacing the system directly to a HUMS ground station database.

objective 5: demonstrate the performance of an IDM system by establishing its ability to recognise previously reported and unreported features in HUMS data, which may be used to diagnose mechanical defects.

The ultimate objective of the contract is therefore to demonstrate the feasibility and performance of a computer based Intelligent Data Management (IDM) process operating on HUMS data.

1.1 The Intelligent Data Management Process

A practical IDM system needs to offer a framework that is capable of organising intelligent interactions between data pre-processing, unsupervised learning and supervised learning.

1.1.1 *Unsupervised Learning*

Unsupervised learning is a process that classifies objects by natural association according to some similarity measure. If the attributes of objects (e.g. between heavy and light, far and near) are quantified and geometrically represented, the points close to each other are grouped together into a cluster. The process is independent

of any prior knowledge or data type. Unsupervised identification of clusters through hierarchical trees, density search and/or fuzzy rules can identify the structures within a data set. A hierarchical tree is formed such that each cluster in the top layer contains one object (one entity) and the root cluster contains all entities. A hierarchical agglomerative technique starts from the top clusters and fuses the two clusters which are most similar. Starting from the root of the tree, a hierarchical divisive method splits the root cluster into two subsets. Density search techniques form clusters by identifying dense concentrations of objects (data points). Fuzzy techniques utilise a set of fuzzy rules which can generally be considered as non-linear surfaces (equations) that partition the entities into clusters. More than one technique can be used and optimisation can be performed in order to enhance the unsupervised analysis. A cluster in a data set can have a compact or a connected (a segregate) form. A compact cluster is a collection of similar objects (objects close to each other in the geometrical representation). Whilst the neighbours in a connected cluster are similar, there exists two objects in the cluster which are relatively dissimilar. An example of a connected cluster is a collection of dense points along a straight line where the first point and the last point are far apart. A natural cluster is either a compact or a connected cluster that contains a relatively high density collection of objects and is separated from other clusters by a relatively low density collection of objects. It is possible that two clusters can overlap.

The results of a clustering algorithm can be very sensitive to the initial data statistics. In other words, the boundary of the clusters can be significantly modified by adding data points that have different statistical properties to those of the initial data set. Most of the unsupervised learning techniques can uncover compact, spherical clusters and fail to uncover connected clusters. Nevertheless, to a large extent, a cluster technique can be optimised to unveil a required cluster type, but this requires a prior knowledge of objects' structures in the data. In practice it is difficult to determine the optimum number of clusters within a data set. The results of the hierarchical algorithms can be sensitive to noise and more than one solution can be identified by a density seeking method; in other words, the solution obtained can be a local optimum and not a global optimum. The initial choice of objects' attributes defines the frame of reference within which the clusters relevant to these attributes can only be identified.

1.1.2 *Supervised Learning*

Supervised learning relies on a priori knowledge and uses a pre-defined set of data to solve a specific problem; this process is termed 'training'. The training data consists of input-output pairs; input patterns and associated desired output patterns. The learning capability of any supervised system is accomplished through its adaptive weights (coefficients). Training entails allowing the supervised system to adjust its weights such that the required mapping between the input and output can be reproduced. Only after training can the supervised process solve the required problem. Supervised learning can establish relationships between output causes (mechanical faults) and input effects (measurements), even when these relationships can not be expressed explicitly. Most of the learning algorithms which are used to adjust the weights stem from the well recognised error minimisation rule known as the Least Mean Square (LMS). Multi Variate Regression (MVR) analysis tools are linear computational tools that perform a single-shot error minimisation. Artificial Neural Networks (ANNs) can be regarded as non-linear computational tools that perform recursive error minimisation.

1.1.3 *Data Pre-processing*

Central to the success of the IDM system is data pre-processing. An established definition of this task is not explicitly available in the published literature known to the author. This can be attributed to the fact that pre-processing is largely dependent on the problem under consideration. In engineering applications, the task is often driven by the laws of physics. Nevertheless, pre-processing can be classified under two main headings: application-specific and generic.

The former concentrates on extracting features from measurements. For example, Fourier analysis can be applied to extract a feature such as a frequency amplitude of interest. Another feature can be extracted by combining a set of measurements (or a set of features) using the underlying laws of physics. A required signature in a feature can be isolated and extracted by implementing appropriate mathematical models where required. For example the influence of operational effects can be removed from a set of features. In this case, the clusters signify other aspects such as mechanical health or age effects. The extracted features must provide an adequate description of the required phenomena. The selection of a set of features appropriate to a particular phenomenon is an application specific process. For example, the colour of a vehicle can only be a significant feature to some of the marketing aspects and, the harmonics extracted at multiples of the tail rotor frequency are generally irrelevant to classifying main rotor faults.

IDM systems can also significantly benefit from generic pre-processors. Examples of such pre-processors are as follows:

- Identify each individual object which is described by the measurements (e.g. identify each individual helicopter or helicopter component).
- Evaluate the probability distribution of each extracted feature.
- Apply linear filters (e.g. calculate moving averages to attenuate noise effects).
- Monitor the trend of the probability distribution with time for each individual object.
- When appropriate, normalise individual object effects.
- Evaluate the correlation between various features.
- Identify a set of independent features.
- Identify the principal, orthogonal components of the independent features.

The ranges of two features can be significantly different because of the units (or scales) which are used to describe each feature. Whilst generic pre-processing can be targeted at removing scale effects and fusing two different data types, application dependent feature weights can be introduced to emphasise the importance of a particular feature.

The selection of features and their weights are largely application dependent processes. However, generic parts of these processes can be identified and

implemented through statistics, optimisation and cluster analysis under the following assumptions:

- There exists a large set of features that cover all aspects of the problem under consideration and this set includes a number of features that depend on another number of features.

- There exists an unsupervised learning algorithm which is insensitive to the number of features relevant to the problem under consideration and is insensitive to the units of features.

In a practical situation, these two assumptions can never be fully satisfied.

1.2 The Organisation of the Report

Section 2 reports the demonstration of the benefits of unsupervised learning techniques and, therefore covers the first contract objective. The engineering data used for demonstration purposes are CAA supplied Spectrometric Oil Analysis Programme (SOAP) data.

Section 3 concentrates on the second objective of the CAA contract. It demonstrates the importance of data pre-processing. Vibration and FDR measurements from 23 Super Puma MK I helicopters are used for the purpose of this demonstration.

Section 4 addresses the third objective of the contract and demonstrates the benefits of supervised learning techniques. The demonstration data are the SOAP data along with information indicating two mechanical defects.

Section 5 concentrates on the fifth objective of the contract. It demonstrates the performance of an IDM system by identifying its ability to recognise previously reported and unreported features in HUMS data.

It should be noted that the fourth objective of the contract comprised a physical demonstration of the feasibility of the IDM system to the CAA using PC-based integrated processes interfacing to a HUMS database.

Whilst the above sections are essentially self contained, Section 6 summarises the conclusions of the work performed under the CAA contract.

2 UNSUPERVISED MACHINE LEARNING

2.1 General

This section concentrates on the 1st objective of the CAA contract 7D/S/1128, namely demonstrating the benefits of unsupervised machine learning.

2.1.1 *CAA Supplied SOAP Data*

The Spectrometric Oil Analysis Programme (SOAP) data supplied to MJAD was obtained from a CAA HUMS trial which, in part, involved Super Puma Helicopters. During this trial, oil samples were taken within 20 minutes of shut-down from specified Super Puma main rotor gearboxes at a sample rate of approximately 50

flying hours. The oil samples were taken from a self sealing chamber which allows for removal and inspection of a magnetic chip detector. The spectrometric analysis was performed by Spectro Laboratories, which delivers a breakdown of the elements in the oil to less than one part per million. In all, 344 samples where analysed. MJAD was supplied with records describing the analysis of 138 oil samples, covering 11 gearboxes. Tables 2.1 and 2.2 present the details supplied.

2.1.2 *Unsupervised Machine Learning*

Unsupervised machine learning can take many forms. For this report, emphasis is placed on a geometrical analysis approach. In essence, each measurement is considered as a dimension in space. A familiar representation is the two dimensional (xy) graph. By plotting points on such a graph, measurement patterns or clusters may be discernible. From a mathematical view point, although totally abstract to the reader, the number of dimensions is unlimited. For n measurements an equivalent n dimensional measurement space can be constructed. The thrust of unsupervised machine learning is to isolate discernible data clusters in this multidimensional space.

Generally the population of measurement samples in each data cluster will be different. Clusters with the majority of measurement samples are considered 'typical'. From a mechanical diagnostic viewpoint, clusters with smaller populations merit particular attention as they are in some way atypical from the norm.

2.1.3 *Unsupervised Machine Learning Process*

Figure 2.1 presents the process (i.e. independent of the measurements) which closely reflects what would happen in-service [Reference 1].

Section 2.2 describes each step of this process.

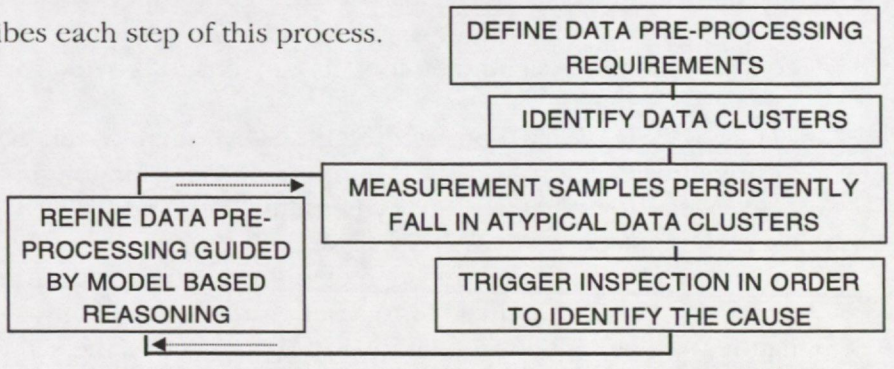
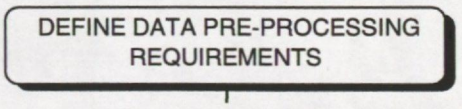


Fig.2.1 Unsupervised machine learning process

2.2 Demonstration of Unsupervised Machine Learning

2.2.1 *Defining Data Pre-processing Requirements*



The first step in the unsupervised machine learning process is to establish how the measurements will be conditioned. For example, a vibration signature may be signal averaged, Fourier Transformed to the frequency domain, particular spectral lines removed, Inverse Fourier Transformed back to the time domain and/or statistical parameters such as kurtosis determined. In practice, any number of parameters or even the whole signature may be presented as input to the subsequent clustering analysis.

The pre-processing criteria are typically driven by extracting features in the measurements which are sensitive to fault conditions and, with reference to the machine learning process, offer maximum separability from a healthy mechanical state.

Often prior knowledge of exactly what fault conditions or failure modes that can develop is not available, so the pre-processing concentrates on estimating which features in the measurements will offer the potential of separating mechanical fault symptoms from typical measurement patterns.

The particular measurement set used for demonstration purposes was SOAP data. Typically, the quantity of each measurable element in the oil is expressed in parts per million (ppm), where one ppm is equivalent to one milligram per litre of oil. The measured concentration may be described as being composed of the following terms:

- The actual concentration at any time t (the sample point), assuming no oil leakage, oil addition, contamination and so on. Generally, an upward trend with time would be expected.

- A non-deterministic error (stochastic noise) on the measured concentration, which can be introduced by contamination, spectrometer inaccuracies and so on.

- A term which describes the effect on the measurement caused by changes in oil volume – oil top-ups, drain and flush etc.

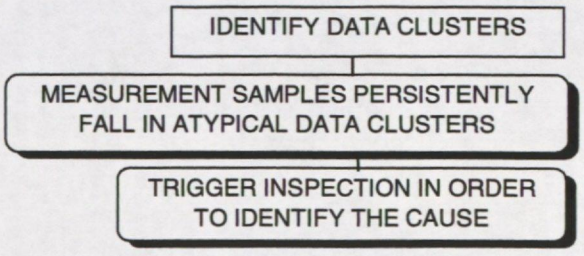
A further consideration is: 'do the measurements offer visibility of developing mechanical defects'? In the case of the CAA supplied SOAP data, information on changes to the oil volume is absent and therefore casts doubt on whether raw concentration levels could be used as a means to provide robust diagnostics.

The actual SOAP data consisted of 138 SOAP samples taken from 11 gearboxes. The ppm values covered Iron (Fe), Chromium (Cr), Aluminium (Al), Copper (Cu), Silver (Ag), Nickel (Ni), Silicon (Si), Magnesium (Mg) and Zinc (Zi). For the data supplied, four elements namely Cr, Ni, Si and Zi in the majority of cases could not be quantified in ppm (ie, the elements were effectively absent in the oil). A fifth element, namely Cu had a ppm measurement which was essentially 'binary' in nature, oscillating between either 1 or 2 ppm across the samples. With a pronounced absence of measurement patterns in these elements, it would suffice to set simple threshold criteria as a means to trigger a warning.

Of the remaining elements, Fe is the main constituent of gears and bearings. Al is the main constituent of the gearbox casing and could indicate the occurrence of bearing track slip or casing corrosion. Ag is used in plating material on bearing cages and could therefore give the earliest sign of wear or damage to bearings, in particular the cage.

Initial pre-processing was therefore limited to element selection, which included Fe, Al, Ag and Mg. The last element, Magnesium, was included because it exhibited pattern changes and therefore may contain useful diagnostic information.

2.2.2 *Cluster analysis*

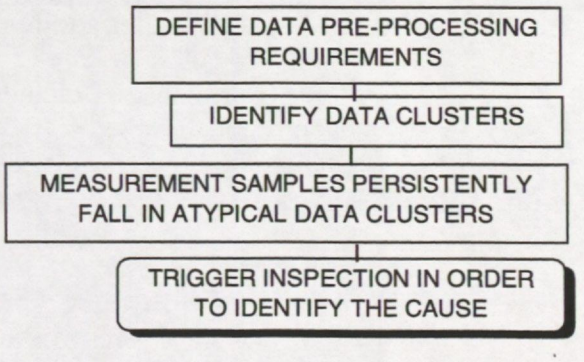


Cluster analysis is performed by searching for regions in the measurement space where data 'groupings' reside. In essence, the analysis locates an axis which is directed along the line of greatest variance in the data. By establishing a further axis at right angles to the axis of 'greatest variance' at the geometrical mid-point in the data, two regions or clusters are isolated. The variance in each cluster is next checked and, if greater than a prescribed threshold, further clusters are isolated by following the same procedure. For each cluster the geometrical centre is determined and, if the distance between two cluster centres is less than a prescribed value, the clusters are merged. The variance threshold and distance term were set based on experience, and remained constant throughout the analysis.

Cluster analysis of the raw concentration levels of Fe, Mg, Al and Ag revealed a number of small clusters. Ordering the clusters with reference to the number of SOAP samples within each cluster (i.e. the highest priority given to the cluster with the smallest number of samples), gearbox M186 fell into the first three clusters, each containing one sample. Gearbox M361 occupied the forth cluster, which contained 2 samples. Further samples from M361 also fell into the fifth cluster (see Table 2.3 for details).

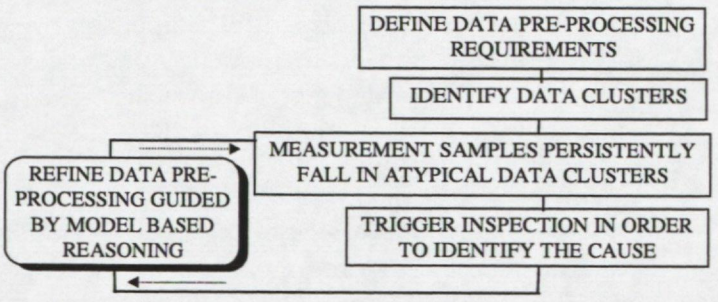
Samples which persistently fall into minority clusters will trigger investigative actions. The number of samples that constitute persistent presence in minority clusters can be optimised by considering the sampling rate, the total number of samples and the failure mode (specially the growth rates of possible defects). Persistent presence can be also evaluated by monitoring the ratio between the number of gearbox samples within minority clusters and the total number of samples. Nevertheless, at this stage of the process, two simple criteria were used as a means to decide on what action to take. First, at least three samples from a gearbox need to fall into minority clusters before any action is taken. Second, if no fault is associated with samples for such a gearbox falling into minority clusters, the refinement stage will be initiated (see Section 2.2.4).

2.2.3 *Identify Causes of Atypical Measurements*



A strip report of gearbox M186 revealed spalling of a mast bearing, which constitutes a potential cause for SOAP samples from this gearbox falling into minority clusters. Nothing untoward could be associated with gearbox M361 and, following the criteria cited in the previous section, no further investigative actions would therefore be triggered.

2.2.4 *Refining the Data Pre-processing Requirements*



Armed with a known mechanical defect, the cluster analysis can be repeated with different data pre-processing options in order to isolate the optimum measurement features which relate to that cause. For the more mathematically orientated reader, Appendix A presents salient math formulations which were used to establish data pre-processing refinements, the descriptions of which follow.

Returning to the problems of noise contamination, first consider any correlated noise that may exist between measurements within a sample. This, for example, may be caused by unreported changes to the oil volume. Such a change can be assumed to effect each element in a similar way and therefore, dividing one element by another should mitigate this effect. The 'normalising' element should be selected with some care as it should largely reflect changes in oil volume and not other 'contaminating' influences (e.g., Fe from other tools, Si from sand ingestion etc.). Of the four elements used in the initial analysis, Mg was chosen as the normalising element.

Second, suppression of further uncorrelated noise between samples could be realised by implementing a moving average rather than using individual measurements.

Combined with these pre-processing techniques, additional features such as 'wear' rates were considered worth adding to the analysis. The wear rate definition for an element is given by the difference between the present and previous concentration levels, divided by the elapsed time in flight hours. Without information on changes to the oil it was assumed that concentration levels would not go down and therefore it was proposed that a 'corrected' wear rate could also be computed. That is, negative or zero wear rates, if found, were set to the previously calculated value.

In summary therefore, the following pre-processing options were considered in order to establish whether additional discriminatory capability could be realised:

- Divide (normalise) element concentrations by a reference element concentration.
- Implement a moving average on the measured concentration levels .
- Compute wear rates.
- Correct wear rates such that negative or zero values are set to the previously calculated value.

In addition, statistical pre-processing options were also considered. These were scaling of data, mean centring and unit vectoring. Scaling of the data simply divides a series of elemental measurements by their respective standard deviations, whilst mean centring subtracts the mean of an element's measurements from every measurement sample for that element. Finally, unit vectoring reduces every measurement in a sample to a component of a unit vector – the direction of the vector now being the discriminant.

Whilst numerous runs with different pre-processing options were undertaken, the best two runs used the following as input parameters to the cluster analysis routine:

- Average concentration levels of Fe, Al and Ag, normalised by Mg.
- Average wear rates of Fe and Mg

Note that average values are based on a simple moving average calculation, where the present plus four previous measurements are used to compute the average.

Notwithstanding the quality of the data and the limited information available, the second run was notable because repeated SOAP samples from particular gearboxes fell into a minority cluster. In particular, the first 5 minority clusters contained 15 successive samples from gearbox M468. Samples from no other gearbox fell into these clusters. Further information on this gearbox was requested. No mechanical defects were reported, although the SOAP sampling rate had increased which may indicate engineering staff concern over this gearbox. The next minority cluster contained 7 samples from one gearbox, namely M285. Request for information on this gearbox revealed that an epicyclic bearing had spalled. This minority cluster also contained 6 and 5 samples from gearboxes M186 and M361 respectively.

Using the two known fault conditions associated with gearboxes M186 and M285, a full series of different pre-processing options were analysed in a quantifiable manner. This is achieved by using quality (*of discrimination*) curves. Figures 2.2 to 2.5 are typical examples of quality curves. The 'x' axis represents the number of SOAP samples, where the samples nearest the origin correspond to samples associated with the smallest (in population terms) clusters. The 'y' axis represents the number of fault cases found |Reference 2|. If the assumption that measurement samples which fall into minority clusters have a higher probability of being associated with mechanical defects, a one to one correspondence between the x and y values close to the origin would be expected.

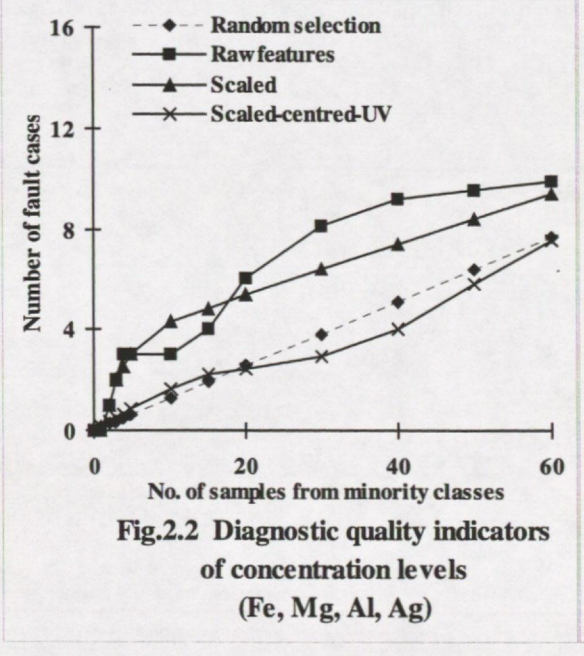


Fig.2.2 Diagnostic quality indicators of concentration levels (Fe, Mg, Al, Ag)

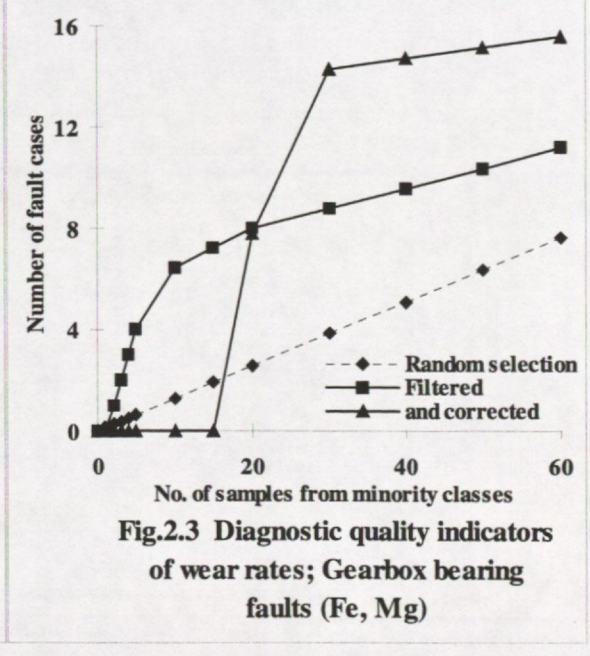


Fig.2.3 Diagnostic quality indicators of wear rates; Gearbox bearing faults (Fe, Mg)

The magnetic chip detector indicated the M186 fault a few hours prior to 6794.35 flying hours. With reference to 'fault' measurement samples it has been assumed that

the SOAP measurements from the gearbox M186 prior to 6794.35 flying hours and all available measurements from the gearbox M285 fall into this category. In all, eight and nine SOAP samples were associated with gearboxes M186 and M285 respectively. Referring to Table 2.4 and the flying hours (TIME) of Table 2.1, it can be seen that the cluster analysis indicated the M186 fault about 150 to 230 hours before the magnetic chip detector. The cluster analysis also indicated the M285 fault at its assumed early stage of development.

Figure 2.2 presents quality curves where Fe, Mg, Al and Ag were used as input. A reference line (annotated 'random selection') is plotted for comparative purposes, which assumes that no relevance is attached to cluster analysis, and therefore fault related measurement samples will be randomly distributed throughout all clusters. Quality curves which lie above the reference line and exhibit a steeper gradient are diagnostically significant. When the quality curve lies parallel to or has a shallower gradient than the reference line, its diagnostic impact is lost. Typically, this would be expected away from the origin, since measurement samples in this region (as indicated on the 'x' axis) fall into majority clusters which should represent mechanical health. Note that raw input (concentrations without pre-processing) and scaled (concentrations normalised by standard deviations) showed some promise in the smallest clusters only.

Figure 2.3 presents quality curves associated with wear rates for Fe and Mg. In particular, two curves are displayed; the filtered curve exploits a moving average (average based on present concentration plus four previous values), whilst the second curve also applies a correction factor. This correction factor attempts to mitigate the effect of changes to the oil condition/volume, details of which were not recorded. The correction factor assumed that some positive change in an element's concentration level will occur between SOAP samples; negative or zero wear rates being reset to their previously calculated value. These curves, when compared with the curves in Figure 2.2, show additional promise.

Figure 2.4 details 'normalised' concentration levels – concentrations are divided by the concentration level of a selected element. In this case, Fe, Al and Ag are normalised by Mg. The results of the application of the moving average is annotated 'filtered-normalised'. Significant improvements over the 'random selection' line indicate that normalisation has particular merit.

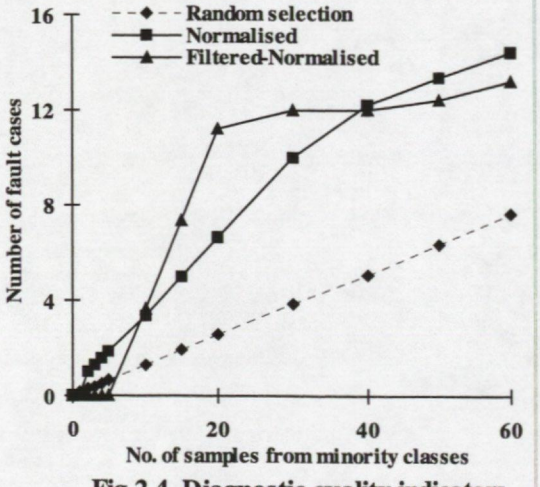


Fig.2.4 Diagnostic quality indicators of normalised concentration levels (Fe, Al, Ag normalised by Mg)

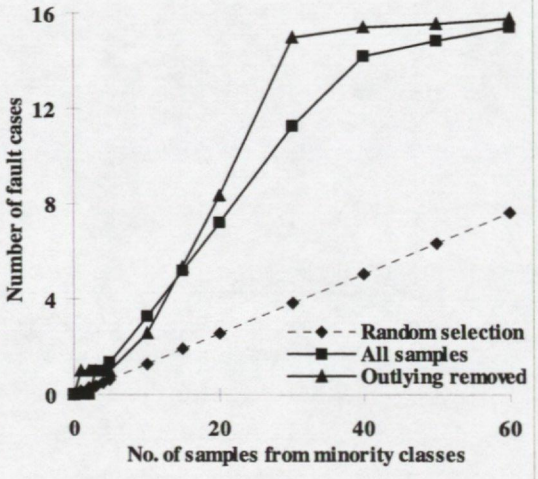


Fig.2.5 The effect of removing 3 outlying samples; Fe, Al and Ag normalised by Mg, and Fe, Al and Ag centred w.r.t. Mg

Finally, Figure 2.5 presents Fe, Al and Ag normalised by Mg and the same elements centred with respect to Mg (e.g. Fe-Mg etc.). It will be noted that in all cases the curves indicate that measurement samples are present which are not known to be fault related. The reasons for these occurrences can include measurement error associated with the spectrometer, sample contamination and so on. As a final refinement, the 'outlying removed' curve in Figure 2.5 is based on a cluster run where assumed non-fault related samples which fell into the 3 smallest clusters were removed.

2.3 Discussion

The previous section presented the elemental steps in the unsupervised machine learning process. It can be seen that this process is independent of the measurements and therefore is applicable to a wide range of HUMS data. A particular advantage of unsupervised machine learning is that it can highlight abnormalities in the data even when the underlying cause is unknown.

The start point in the process is to define what data will be analysed and how it will be pre-processed. For the particular demonstration described in Section 2, raw SOAP concentration levels were used. This data highlighted two 'abnormal' gearboxes, one of which was found to have a mast bearing which had spalled. Nothing untoward could be established for the second gearbox. This has to be expected since the pre-processing will be far from optimised in the initial stages of development.

Once a fault has been isolated, the measurements can be repeatedly re-analysed as a means to refine the data pre-processing. The objective is to drive the system such that measurements from the fault related gearbox consistently fall into minority data clusters. This process revealed two further gearboxes, namely M468 and M285 as 'abnormal'. Nothing was reported on M468 which could be identified as a fault although the SOAP sampling rate had been increased, possibly indicating engineering staff concern. However it was found that M285 was suffering spalling of an epicyclic bearing.

With two known faults in the measurement data, a quantifiable performance indicator was introduced, called the quality curve. These curves quickly reveal whether the measurements combined with the cluster analysis are diagnostically meaningful. For the SOAP demonstration it was established that correlated noise within a measurement sample could be mitigated by normalising (dividing) concentration levels by one selected concentration level. For the demonstration Fe, Ag and Al were normalised by Mg. The attenuation of uncorrelated noise between measurement samples was achieved by implementing a moving average, where the 'latest' concentration is given by the average of the present plus the four previous values. Further diagnostic performance was also realised by including corrected wear rates.

2.4 Conclusions

There is much still to be gained from the patterns concealed in helicopter HUMS measurements. It is important that measurement patterns which are abnormal are reported even when the cause of the abnormality is unknown. It has been demonstrated that unsupervised machine learning is one mechanism by which this need can be delivered.

The essence of the machine learning technique proposed was that each measurement may be considered as one dimension (axis). Therefore n measurements would be mapped into n-dimensional measurement space. By 'plotting' measurements in this space, data clusters may become discernible. The machine learning process actively searches for such clusters. Further, it has been demonstrated that clusters which have a small number of measurement samples are diagnostically significant in that such measurements can be related to the development of mechanical faults.

Directly associated with the success of detecting mechanical faults is the selection of appropriate measurements and how they are conditioned. Initially, it is doubtful that the optimum data pre-processing requirements will be known. The demonstration presented in this section identified a process by which various pre-processing options could be assessed. In particular, quantifiable checks are proposed in the form of quality (*of discrimination*) curves. It has been demonstrated that these curves allow rapid quantifiable assessments to be realised, and emphasises which data selection and pre-processing will deliver the best fault discrimination capability.

Whilst the demonstration of the machine learning process has used CAA supplied Spectrometric Oil Analysis (SOAP) data, it has been demonstrated that a process may be followed which is measurement independent. This generic process will be of fundamental importance in the analysis and interpretation of HUMS data, which contains many different types of measurements from a variety of sensors.

Table 2.1 CAA supplied SOAP data

ID	AC_REG	SER	TIME	S_DATE	Fe	Cr	Al	Cu	Ag	Ni	Si	Mg	Zn
1	G_TIGE	M285	10852.25	27/02/89	0.6	0.0	0.4	0.1	0.0	0.0	0.0	0.1	0.0
2	G_TIGE	M285	10900.00	18/03/89	0.8	0.0	0.6	0.1	0.0	0.0	0.0	0.1	0.0
3	G_TIGE	M285	11003.05	17/04/89	1.2	0.1	0.0	0.1	0.2	0.1	0.0	0.1	0.0
4	G_TIGE	M285	10947.00	04/04/89	0.4	0.0	0.6	0.1	0.0	0.1	0.0	0.1	0.0
5	G_TIGE	M285	11049.20	26/04/89	1.3	0.0	1.0	0.1	0.1	0.0	0.0	0.2	0.0
6	G_TIGE	M285	11103.30	08/05/89	0.8	0.1	1.0	0.2	0.2	0.1	0.0	0.1	0.3
7	G_TIGE	M285	11019.00	14/05/89	0.7	0.0	0.2	0.1	0.0	0.1	0.0	0.1	0.0
8	G_TIGE	M285	11152.19	23/05/89	0.7	0.0	0.5	0.1	0.1	0.2	0.0	0.1	0.0
9	G_TIGE	M285	11207.00	04/06/89	0.5	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
10	G_TIGE	M468	11250.00		0.6	0.0	0.8	0.1	0.1	0.1	0.0	0.3	0.0
11	G_TIGE	M468	11295.00	26/06/89	0.4	0.0	0.0	0.1	0.1	0.1	2.0	0.1	0.0
12	G_TIGE	M468	11362.00	06/07/89	0.3	0.0	0.0	0.1	0.1	0.0	2.0	0.1	0.0
13	G_TIGE	M468	11394.00	17/07/89	0.4	0.0	0.0	0.1	0.0	0.0	2.0	0.2	0.0
14	G_TIGE	M468	11468.00	13/08/89	0.6	0.0	0.0	0.1	0.1	0.0	2.0	0.3	0.0
15	G_TIGE	M468	11495.00	22/08/89	0.8	0.0	0.0	0.1	0.1	0.0	2.0	0.4	0.0
16	G_TIGE	M468	11609.00	01/11/89	0.5	0.0	0.3	0.1	0.0	0.0	0.0	0.2	0.0
17	G_TIGE	M468	11746.00	18/11/89	0.9	0.0	1.0	0.1	0.2	0.0	0.0	0.5	0.0
18	G_TIGE	M468	11798.00	27/11/89	1.0	0.0	1.0	0.1	0.3	0.0	0.0	0.6	0.0
19	G_TIGE	M468	11946.00	22/12/89	1.5	0.0	1.4	0.1	1.6	0.0	0.0	0.6	0.0
20	G_TIGE	M468	12014.88	12/01/90	0.8	0.0	0.7	0.1	0.2	0.0	1.0	0.3	0.0
21	G_TIGE	M468	12034.56	18/01/90	0.8	0.0	0.4	0.1	0.2	0.0	0.0	0.3	0.0
22	G_TIGE	M468	12047.68	22/01/90	0.8	0.0	0.2	0.1	0.1	0.0	0.0	0.3	0.0
23	G_TIGE	M468	12051.00	23/01/90	0.8	0.0	0.5	0.1	0.2	0.0	0.0	0.3	0.0
24	G_TIGE	M468	12054.20	24/01/90	0.9	0.0	0.5	0.1	0.2	0.0	0.0	0.3	0.0
25	G_TIGE	M468	12057.40	25/01/90	0.9	0.0	0.5	0.1	0.2	0.0	0.0	0.3	0.0
26	G_TIGE	M468	12063.80	27/01/90	1.0	0.0	0.6	0.1	0.2	0.0	0.0	0.3	0.0
27	G_TIGE	M468	12073.40	30/01/90	1.0	0.0	0.3	0.1	0.1	0.0	0.0	0.3	0.0
28	G_TIGE	M468	12076.60	31/01/90	1.1	0.0	0.2	0.1	0.1	0.0	0.0	0.4	0.0
29	G_TIGE	M468	12083.00	02/02/90	1.0	0.0	0.0	0.1	0.1	0.0	0.0	0.3	0.0
30	G_TIGE	M468	12115.00	12/02/90	0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
31	G_TIGE	M468	12118.20	13/02/90	0.5	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
32	G_TIGE	M468	12121.40	14/02/90	0.5	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
33	G_TIGE	M468	12124.60	15/02/90	0.5	0.0	0.4	0.1	0.0	0.0	0.0	0.2	0.0
34	G_TIGE	M468	12137.40	19/02/90	0.7	0.0	0.4	0.1	0.1	0.0	1.0	0.3	0.0
35	G_TIGE	M468	12137.40	19/02/90	0.6	0.0	0.4	0.1	0.1	0.0	1.0	0.2	0.0
36	G_TIGE	M468	12124.60	15/02/90	0.5	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
37	G_TIGE	M468	12143.60	21/02/90	0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
38	G_TIGE	M468	12192.00	08/03/90	0.3	0.0	0.2	0.1	0.0	0.0	1.0	0.1	0.0
39	G_TIGE	M468	12248.00	21/03/90	0.5	0.0	0.0	0.1	0.1	0.0	0.0	0.2	0.0
40	G_TIGE	M468	12347.16	17/04/90	0.7	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
41	G_TIGH	M306	7895.30	30/01/89	0.8	0.0	0.4	0.1	0.4	0.1	0.0	0.2	0.0
42	G_TIGH	M306	7968.05	14/02/89	1.2	0.0	0.8	0.1	0.5	0.1	0.0	0.4	0.0
43	G_TIGH	M306	8013.35	28/02/89	1.5	0.0	0.9	0.1	0.5	0.0	0.0	0.5	0.0
44	G_TIGH	M306	8049.20	06/03/89	2.0	0.0	1.2	0.1	0.7	0.0	0.0	0.7	0.3
45	G_TIGH	M306	8119.55	22/03/89	0.6	0.0	0.5	0.1	0.0	0.1	0.0	0.2	0.0
46	G_TIGH	M180	8246.45	04/05/89	1.1	0.0	0.6	0.1	0.0	0.0	0.0	0.2	0.0
47	G_TIGH	M180	8300.00	18/05/89	1.0	0.0	0.6	0.1	0.0	0.1	2.0	0.2	0.0
48	G_TIGH	M180	8348.00	01/06/89	0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
49	G_TIGH	M180	8397.00	16/06/89	0.5	0.0	0.2	0.1	0.0	0.1	0.0	0.1	0.0
50	G_TIGH	M180	8453.00	29/06/89	0.8	0.0	0.6	0.1	0.1	0.1	0.0	0.2	0.0
51	G_TIGH	M180	8505.58	09/07/89	0.9	0.1	1.6	0.1	0.0	0.0	0.0	0.3	0.0
52	G_TIGH	M180	8543.00	17/07/89	1.1	0.0	0.0	0.1	0.0	0.0	0.0	0.4	0.2
53	G_TIGH	M180	8690.00	14/08/89	1.4	0.0	1.5	0.1	0.2	0.0	1.0	0.5	0.0
54	G_TIGH	M180	8749.00	25/08/89	1.4	0.0	2.0	0.1	0.3	0.0	0.0	0.5	0.0
55	G_TIGH	M180	8809.23	15/09/89	1.1	0.0	1.3	0.1	0.2	0.2	0.0	0.4	0.0

ID	AC_REG	SER	TIME	S_DATE	Fe	Cr	Al	Cu	Ag	Ni	Si	Mg	Zn
56	G_TIGH	M180	8985.90	24/10/89	0.5	0.0	0.4	0.1	0.0	0.0	0.0	0.1	0.0
57	G_TIGH	M180	9051.36	22/11/89	0.9	0.0	1.0	0.1	0.0	0.0	0.0	0.2	0.0
58	G_TIGH	M180	9107.39	05/12/89	1.0	0.0	1.3	0.2	0.3	0.1	0.0	0.3	0.0
59	G_TIGK	M246	6620.00	26/06/89	0.4	0.0	0.5	0.2	0.2	0.1	2.0	0.1	0.0
60	G_TIGU	M186	6505.35	24/05/89	2.1	0.0	0.5	0.1	0.9	0.0	1.0	0.5	0.2
61	G_TIGU	M186	6549.00	08/06/89	3.0	0.0	1.5	0.2	1.0	0.0	0.0	1.0	0.2
62	G_TIGU	M186	6555.00	09/06/89	2.3	0.1	1.8	0.2	0.8	0.1	0.0	0.8	0.2
63	G_TIGU	M186	6624.00	22/06/89	1.6	0.1	1.0	0.1	0.3	0.1	0.0	0.4	0.0
64	G_TIGU	M186	6640.00	26/06/89	1.7	0.1	0.4	0.2	0.6	0.2	2.0	0.4	0.2
65	G_TIGU	M186	6743.00	17/07/89	2.6	0.0	2.9	0.1	0.7	0.0	1.0	0.6	0.0
66	G_TIGU	M186	6758.00	22/07/89	0.9	0.0	1.0	0.1	0.2	0.1	0.0	0.2	0.0
67	G_TIGU	M186	6786.00	27/07/89	0.9	0.0	0.5	0.1	0.0	0.1	0.0	0.2	0.0
68	G_TIGU	M186	6799.00	01/08/89	1.2	0.0	0.0	0.1	0.0	0.0	0.0	0.3	0.0
69	G_TIGU	M186	6872.00	14/08/89	0.7	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
70	G_TIGU	M186	6892.00	25/08/89	0.8	0.0	0.5	0.1	0.1	0.0	0.0	0.2	0.0
71	G_TIGU	M186	6943.20	22/09/89	0.3	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
72	G_TIGU	M186	6984.55	03/10/89	0.5	0.0	0.3	0.1	0.0	0.1	0.0	0.1	0.0
73	G_TIGU	M186	7002.00	06/10/89	0.4	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0
74	G_TIGU	M186	7035.10	25/10/89	0.4	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0
75	G_TIGU	M186	7145.35	09/11/89	0.5	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0
76	G_TIGU	M186	7299.30	11/12/89	1.2	0.0	0.4	0.1	0.0	0.0	0.0	0.3	0.0
77	G_TIGV	M306			0.4	0.0	0.2	0.1	0.1	0.0	0.0	0.1	0.0
78	G_TIGV	M306	6226.25	16/08/88	0.7	0.0	0.3	0.1	0.0	0.0	2.0	0.1	0.0
79	G_TIGV	M306	6294.55	13/09/88	1.0	0.0	0.4	0.1	0.0	0.0	0.0	0.2	0.0
80	G_TIGV	M306	6397.00	25/10/88	0.7	0.0	0.4	0.1	0.1	0.0	3.0	0.2	0.0
81	G_TIGV	M127		01/11/89	0.5	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0
82	G_TIGW	M132	4810.00	14/09/88	0.6	0.0	0.4	0.1	0.0	0.0	2.0	0.1	0.0
83	G_TIGW	M132	4892.60	11/10/88	0.6	0.0	0.5	0.1	0.0	0.0	2.0	0.1	0.0
84	G_TIGW	M132	4957.30	01/11/88	1.0	0.0	0.8	0.1	0.0	0.0	2.0	0.3	0.0
85	G_TIGW	M132	4998.55	15/11/88	0.6	0.0	0.4	0.1	0.0	0.0	2.0	0.2	0.0
86	G_TIGW	M132	5057.00	02/12/88	0.5	0.0	0.0	0.2	0.2	0.0	2.0	0.1	0.0
87	G_TIGW	M132	5057.00	02/12/88	0.4	0.0	0.0	0.2	0.2	0.0	2.0	0.1	0.0
88	G_TIGW	M132	5149.25	24/01/89	0.4	0.0	0.6	0.2	0.1	0.0	2.0	0.1	0.0
89	G_TIGW	M132	5216.25	18/02/89	0.5	0.0	0.5	0.1	0.0	0.0	0.0	0.1	0.0
90	G_TIGW	M132	5250.45	01/03/89	0.6	0.0	0.7	0.1	0.0	0.0	0.0	0.1	0.0
91	G_TIGW	M132	5396.05	20/04/89	0.6	0.0	0.0	0.1	0.1	0.1	0.0	0.2	0.0
92	G_TIGW	M132	5486.10	09/05/89	1.0	0.0	1.5	0.2	0.2	0.2	0.0	0.2	0.2
93	G_TIGW	M132	5500.00	14/05/89	0.4	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0
94	G_TIGW	M132	5602.00	21/06/89	0.4	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0
95	G_TIGW	M132	5764.00	27/07/89	0.3	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
96	G_TIGW	M132	5802.00	02/08/89	0.3	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
97	G_TIGW	M132	5841.00	14/08/89	0.3	0.0	0.0	0.1	0.0	0.0	1.0	0.1	0.0
98	G_TIGW	M132	5900.00	30/08/89	0.6	0.0	0.2	0.1	0.1	0.0	0.0	0.2	0.0
99	G_TIGW	M132	5985.00	18/09/89	0.7	0.0	0.6	0.1	0.0	0.0	0.0	0.2	0.2
100	G_TIGW	M132	6016.31	28/09/89	0.4	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
101	G_TIGW	M132	6048.20	03/10/89	0.4	0.0	0.0	0.1	0.0	0.1	0.0	0.2	0.0
102	G_TIGW	M132	6101.35	25/10/89	0.4	0.0	0.3	0.1	0.0	0.0	0.0	0.1	0.0
103	G_TIGW	M453	6345.00	09/02/90	0.6	0.0	0.3	0.1	0.0	0.0	1.0	0.1	0.0
104	G_TIGW	M453	6447.00	03/03/90	0.8	0.0	0.2	0.1	0.0	0.0	0.0	0.2	0.0
105	G_TIGW	M453	6521.00	19/03/90	0.5	0.0	0.0	0.1	0.0	0.0	1.0	0.1	0.0
106	G_TIGW	M453	6557.00	26/03/90	0.6	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0
107	G_TIGW	M453	6592.00	09/04/90	0.9	0.0	0.2	0.1	0.0	0.0	0.0	0.2	0.0
108	G_TIGW	M453	6676.00	11/05/90	1.0	0.0	0.0	0.1	0.0	0.0	0.0	0.3	0.3
109	G_TIGW	M453	6747.00	07/06/90	0.5	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
110	G_TIGW	M453	6928.00	15/08/90	2.0	0.0	0.0	0.1	0.8	0.0	0.0	0.3	0.2
111	G_BMCX	M149	2405.15	01/09/88	0.5	0.0	0.7	0.1	0.0	0.0	0.0	0.1	0.0
112	G_BMCX	M149	2497.00	27/09/88	0.7	0.0	0.5	0.1	0.0	0.0	0.0	0.2	0.0

ID	AC_REG	SER	TIME	S_DATE	Fe	Cr	Al	Cu	Ag	Ni	Si	Mg	Zn
113	G_BMCX	M149	2557.50	19/10/88	0.4	0.0	0.4	0.1	0.0	0.0	3.0	0.1	0.0
114	G_BMCX	M149	2613.00	07/11/88	0.7	0.0	0.5	0.1	0.0	0.0	2.0	0.2	0.0
115	G_BMCX	M149	2701.00	09/12/88	1.0	0.0	0.5	0.2	0.3	0.1	2.0	0.2	0.0
116	G_BMCX	M149	2746.00	21/12/88	1.2	0.0	0.5	0.2	0.3	0.1	2.0	0.3	0.0
117	G_BMCX	M149	2801.30	06/01/89	1.3	0.0	1.0	0.1	0.1	0.1	2.0	0.5	0.0
118	G_BMCX	M149	2853.25	24/01/89	0.6	0.0	0.6	0.1	0.0	0.1	2.0	0.2	0.0
119	G_BMCX	M149	2902.05	08/02/89	0.9	0.0	0.4	0.1	0.0	0.0	0.0	0.2	0.0
120	G_BMCX	M149	2957.10	23/02/89	1.0	0.0	0.5	0.1	0.1	0.0	0.0	0.2	0.0
121	G_BMCX	M149	2987.00		1.1	0.0	0.4	0.1	0.1	0.0	0.0	0.3	0.0
122	G_BMCX	M361	3146.27	16/04/89	1.8	0.1	0.4	0.2	0.2	0.1	0.0	0.3	0.3
123	G_BMCX	M361	3041.50	28/03/89	0.9	0.0	0.5	0.1	0.0	0.1	0.0	0.1	0.0
124	G_BMCX	M361	3195.58	26/04/89	2.4	0.0	1.2	0.1	0.4	0.1	0.0	0.4	0.2
125	G_BMCX	M361	3247.40	10/05/89	1.9	0.0	1.6	0.1	0.5	0.1	0.0	0.5	0.2
126	G_BMCX	M361	3274.05	23/05/89	2.2	0.0	1.0	0.1	0.5	0.1	0.0	0.7	0.3
127	G_BMCX	M361	3333.00	01/06/89	2.2	0.0	0.5	0.1	0.2	0.1	0.0	0.7	0.2
128	G_BMCX	M361	3353.34	09/06/89	1.8	0.0	1.4	0.1	0.5	0.1	0.0	0.6	0.2
129	G_BMCX	M361	3420.00	22/06/89	1.6	0.0	0.8	0.1	0.5	0.1	0.0	0.6	0.0
130	G_BMCX	M361	3449.00	29/06/89	1.6	0.0	0.8	0.1	0.4	0.1	0.0	0.6	0.2
131	G_BMCX	M361	3475.00	06/07/89	1.4	0.1	1.0	0.2	0.4	0.1	0.0	0.5	0.0
132	G_BMCX	M361	3554.00	27/07/89	0.9	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
133	G_BMCX	M361	3595.00	12/08/89	1.4	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0
134	G_BMCX	M361	3648.00	20/08/89	1.0	0.0	0.0	0.1	0.0	0.0	0.0	0.3	0.0
135	G_BMCX	M361	3745.22	14/09/89	0.8	0.0	0.0	0.1	0.1	0.1	0.0	0.3	0.0
136	G_BMCX	M361	3793.00	18/09/89	1.0	0.0	0.0	0.1	0.1	0.0	0.0	0.5	0.0
137	G_BMCX	M361	3849.00	02/10/89	1.1	0.0	0.2	0.1	0.1	0.2	0.0	0.5	0.0
138	G_BMCX	M361	4046.21	13/12/89	0.7	0.0	0.4	0.1	0.0	0.0	0.0	0.1	0.0

Table 2.2 CAA supplied SOAP data

ID	TIME	T_DATE	L_REF	REMOVAL	REASON	OIL_CH	PQ
1	10852.25	31/03/89	C1990				10
2	10900.00	31/03/89	C1991				10
3	11003.05	20/04/89	D1391				12
4	10947.00	21/04/89	D1401				8
5	11049.20	03/05/89	E224				13
6	11103.30	12/05/89	E864				11
7	11019.00	20/05/89	E1373				12
8	11152.19	26/05/89	E1667				10
9	11207.00	10/06/89	F674	REMOVED	OVERHAUL		11
10	11250.00	17/06/89	F1099				10
11	11295.00	28/06/89	F1834				9
12	11362.00	10/07/89	G559				9
13	11394.00	20/07/89	G1249				9
14	11468.00	16/08/89	H1132				7
15	11495.00	24/08/89	H1675				8
16	11609.00	02/11/89	L114				9
17	11746.00	05/12/89	M204				8
18	11798.00	05/12/89	M205				8
19	11946.00	11/01/90	A638				8
20	12021.60	17/01/90	A1088				
21	12043.20	26/01/90	A1657				
22	12050.60	26/01/90	A1658				
23	12051.00	01/02/90	B004				9
24	12054.60	01/02/90	B005				
25	12058.20	01/02/90	B006				

ID	TIME	T_DATE	L_REF	REMOVAL	REASON	OIL_CH	PQ
26	12065.40	01/02/90	B007				
27	12076.20	03/02/90	B126				
28	12079.80	03/02/90	B127				
29	12087.00	16/02/90	B920				
30	12123.00	16/02/90	B921				
31	12126.60	16/02/90	B922				
32	12130.20	16/02/90	B923				
33	12133.80	17/02/90	B1063				
34	12148.20	22/02/90	B1312				
35	12148.20	22/02/90	B1313				
36	12133.80	23/02/90	B1398				
37	12155.40	23/02/90	B1399				
38	12192.00	14/03/90	C814			OIL	10
39	12248.00	23/03/90	C1432				7
40	12347.16	20/04/90	D1164				8
41	7895.30	08/03/89	C488				8
42	7968.05	08/03/89	C489				10
43	8013.35	31/03/89	C1992				11
44	8049.20	31/03/89	C1993				10
45	8119.55	21/04/89	D1402	REMOVED	METAL CONTAMINATION		
46	8246.45	09/05/89	E504				14
47	8300.00	20/05/89	E1375				14
48	8348.00	05/06/89	F223				15
49	8397.00	22/06/89	F1455				8
50	8453.00	30/06/89	F2100				9
51	8505.58	12/07/89	G713				9
52	8543.00	20/07/89	G1250				8
53	8690.00	16/08/89	H1133				7
54	8749.00	29/08/89	H1935				9
55	8809.23	16/09/89	J1070	REMOVED	EPICYCLIC CHANGE	OIL	9
56	8985.90	02/11/89	L112				
57	9051.36	05/12/89	M206				9
58	9107.39	06/12/89	M292				8
59	6620.00	28/06/89	F1835				10
60	6505.35	30/05/89	E1922				13
61	6549.00	10/06/89	F673				11
62	6555.00	16/06/89	F1060				
63	6624.00	23/06/89	F1497				13
64	6640.00	28/06/89	F1836				10
65	6743.00	20/07/89	G1251				10
66	6758.00	27/07/89	G1761				9
67	6786.00	31/07/89	G1964				9
68	6799.00	03/08/89	H238	REMOVED	SEAL CHANGE	OIL	11
69	6872.00	16/08/89	H1134				6
70	6892.00	29/08/89	H1936				10
71	6943.20	07/10/89	K427				8
72	6984.55	30/10/89	K1871				9
73	7002.00	30/10/89	K1872				8
74	7035.10	30/10/89	K1873	REMOVED	SEAL CHANGE	OIL	9
75	7145.35	16/11/89	L940				11
76	7299.30	11/01/90	A639	REMOVED	EPICYCLIC CHANGE	OIL	9
77		06/08/88	H289				
78	6226.25	19/10/88	K1157				6
79	6294.55	19/10/88	K1158				4
80	6397.00	22/11/88	L1296	REMOVED	OPERATION REQUIREMEN		4
81		02/11/89	L113				
82	4810.00	19/10/88	K1159				4

ID	TIME	T_DATE	L_REF	REMOVAL	REASON	OIL_CH	PQ
83	4889.50	19/10/88	K1160				
84	4957.30	22/11/88	L1297				4
85	4998.55	22/11/88	L1298				4
86	5057.00	12/01/89	A757				8
87	5057.00	12/01/89	A758				8
88	5149.25	10/02/89	B663				9
89	5216.25	31/03/89	C1994				14
90	5250.45	31/03/89	C1995				14
91	5396.05	24/04/89	D1579				11
92	5486.10	12/05/89	E865				10
93	5500.00	20/05/89	E1374				10
94	5602.00	22/06/89	F1456				8
95	5764.00	31/07/89	G1968				8
96	5802.00	03/08/89	H239				8
97	5841.00	16/08/89	H1135				8
98	5900.00	07/09/89	J422				7
99	5985.00	23/09/89	J1544				8
100	6016.31	30/09/89	J2108				9
101	6048.20	07/10/89	K429				9
102	6101.35	30/10/89	K1874	REMOVED	METAL CONTAMINATION		10
103	6364.70	14/02/90	B804				9
104	6447.00	06/03/90	C259				8
105	6521.00	21/03/90	C1302				7
106	6557.00	29/03/90	C1838				6
107	6592.00	12/04/90	D756				8
108	6717.40	12/05/90	E702				
109	6747.00	08/06/90	F495				8
110	7017.40	17/08/90	H1044				
111	2405.15	19/10/88	K1161				5
112	2497.00	19/10/88	K1162				4
113	2557.50	22/11/88	L1294				
114	2613.00	22/11/88	L1293				4
115	2701.00	12/01/89	A762				8
116	2746.00	12/01/89	A763				7
117	2801.30	08/03/89	C486				7
118	2853.25	08/03/89	C487				9
119	2902.05	31/03/89	C1986				12
120	2957.10	31/03/89	C1987				11
121	2987.00	31/05/89	C1988	REMOVED	OVERHAUL		10
122	3146.27	20/04/89	D1390				10
123	3041.50	21/04/89	D1403				10
124	3195.58	03/05/89	E223				14
125	3247.40	18/05/89	E934				16
126	3274.05	25/05/89	E1666				14
127	3333.00	05/06/89	F222			OIL	14
128	3353.34	16/06/89	F1059				11
129	3420.00	23/06/89	F1496				8
130	3449.00	30/06/89	F2099				8
131	3475.00	10/07/89	G558				10
132	3554.00	31/07/89	G1965				7
133	3595.00	16/08/89	H1136				7
134	3648.00	24/08/89	H1674				9
135	3745.22	16/09/89	J1071				11
136	3793.00	19/09/89	J1214				13
137	3849.00	07/10/89	K428			OIL	9
138	4046.21	11/01/90	A637				8

Table 2.3 Cluster analysis based on Fe, Mg, Al and Ag

ID	Cluster	SER_NO	ID	Cluster	SER_NO	ID	Cluster	SER_NO
65	1	M186	3	10	M285	46	10	M180
61	2	M186	21	10	M468	104	10	M453
62	3	M186	23	10	M468	106	10	M453
124	4	M361	25	10	M468	108	10	M453
126	4	M361	27	10	M468	107	10	M453
44	4	M306	118	10	M149	103	10	M453
19	5	M468	116	10	M149	45	10	M306
125	6	M361	114	10	M149	59	10	M246
128	6	M361	113	11	M149	56	10	M180
54	6	M180	73	11	M186	41	10	M306
53	6	M180	74	11	M186	47	10	M180
129	7	M361	93	11	M132	52	10	M180
43	7	M306	94	11	M132	50	10	M180
63	7	M186	96	11	M132	24	10	M468
130	7	M361	100	11	M132	22	10	M468
122	8	M361	101	11	M132	20	10	M468
60	8	M186	102	11	M132	15	10	M468
64	8	M186	109	11	M453	16	10	M468
127	8	M361	105	11	M453	2	10	M285
110	8	M453	97	11	M132	136	10	M361
117	9	M149	95	11	M132	134	10	M361
51	9	M180	91	11	M132	132	10	M361
6	9	M285	29	10	M468	115	10	M149
5	9	M285	28	10	M468	119	10	M149
58	9	M180	33	10	M468	121	10	M149
92	9	M132	34	10	M468	120	10	M149
42	9	M306	26	10	M468	123	10	M361
57	9	M180	40	10	M468	75	11	M186
66	9	M186	67	10	M186	71	11	M186
55	9	M180	69	10	M186	48	11	M180
131	9	M361	70	10	M186	49	11	M180
17	9	M468	72	10	M186	11	11	M468
18	9	M468	68	10	M186	12	11	M468
84	9	M132	76	10	M186	31	11	M468
111	10	M149	78	10	M306	32	11	M468
112	10	M149	80	10	M306	37	11	M468
133	10	M361	79	10	M306	38	11	M468
135	10	M361	85	10	M132	39	11	M468
137	10	M361	88	10	M132	30	11	M468
138	10	M361	90	10	M132	13	11	M468
1	10	M285	89	10	M132	14	11	M468
4	10	M285	83	10	M132	9	11	M285
8	10	M285	98	10	M132	86	11	M132
7	10	M285	99	10	M132			
10	10	M468	82	10	M132			

Table 2.4 Cluster analysis based on Fe, Al and Ag normalised by Mg

ID	Cluster	SER_NO	ID	Cluster	SER_NO	ID	Cluster	SER_NO
138	1	M361	51	7	M180	46	7	M180
1	1	M285	49	7	M180	47	7	M180
4	1	M285	123	7	M361	50	7	M180
2	1	M285	116	7	M149	52	7	M180
6	1	M285	114	7	M149	65	7	M186
5	1	M285	112	7	M149	66	7	M186
7	1	M285	121	8	M149	54	7	M180
3	1	M285	134	8	M361	45	7	M306
88	1	M132	9	8	M285	56	7	M180
59	2	M246	135	8	M361	57	7	M180
110	3	M453	27	8	M468	71	8	M186
58	4	M180	28	8	M468	73	8	M186
61	4	M186	29	8	M468	74	8	M186
41	4	M306	31	8	M468	75	8	M186
63	4	M186	32	8	M468	76	8	M186
64	4	M186	33	8	M468	92	8	M132
62	4	M186	37	8	M468	93	8	M132
60	5	M186	38	8	M468	94	8	M132
107	5	M453	39	8	M468	96	8	M132
108	5	M453	40	8	M468	97	8	M132
109	5	M453	67	8	M186	98	8	M132
122	6	M361	69	8	M186	100	8	M132
124	6	M361	70	8	M186	101	8	M132
126	6	M361	19	6	M468	99	8	M132
11	6	M468	42	6	M306	103	8	M453
17	6	M468	44	6	M306	104	8	M453
21	6	M468	80	6	M306	106	8	M453
25	6	M468	43	6	M306	105	8	M453
24	6	M468	18	6	M468	95	8	M132
23	6	M468	16	6	M468	91	8	M132
22	6	M468	15	6	M468	72	8	M186
20	6	M468	10	6	M468	68	8	M186
55	7	M180	125	6	M361	48	8	M180
78	7	M306	127	6	M361	34	8	M468
79	7	M306	129	6	M361	30	8	M468
83	7	M132	128	6	M361	26	8	M468
85	7	M132	111	7	M149	12	8	M468
90	7	M132	113	7	M149	13	8	M468
89	7	M132	115	7	M149	14	8	M468
86	7	M132	117	7	M149	136	8	M361
84	7	M132	119	7	M149	130	8	M361
82	7	M132	8	7	M285	131	8	M361
53	7	M180	120	7	M149	133	8	M361
102	7	M132	118	7	M149	132	8	M361
51	7	M180	137	7	M361			

3 DATA PRE-PROCESSING

3.1 General

This section concentrates on the **2nd objective** of the CAA contract 7D/S/1128, namely demonstrating the diagnostic importance of data pre-processing by utilising Flight Data Recorder (FDR) data combined with HUMS vibration data.

3.1.1 *The FDR and HUMS Vibration Data*

Vibration and FDR data from 23 Super Puma MK I helicopters were used for the purpose of this demonstration. The data was acquired in 1993 during more than 1645 revenue earning flights and was stored in a database comprising 3137 records. The FDR parameters include pressure·altitude, indicated airspeed, magnetic heading, pitch attitude, roll attitude, engine 1 torque, engine 2 torque, main rotor RPM, tail rotor RPM, collective pitch angle, longitudinal cyclic control, lateral cyclic control, tail rotor position, outside air temperature, yaw rate, altitude rate, normal acceleration. The HUMS data comprises a signal divided into a number of cycles having the same period (frequency) of the main rotor. About 100 cycles (22.6 seconds worth of data) are averaged. The vibration harmonics are calculated from the average cycle. The FDR parameters are captured at the beginning of the acquisition frame. The fourth harmonics (4R) of the vertical and lateral vibration (from the RTB sensors) were considered in this investigation.

3.1.2 *Vibration Data Pre-processing*

The first function of pre-processing is to convert the analogue signals to digital signals which are suitable for computer analysis. HUMS performs this function along with two other pre-processing functions, namely, the signal averaging and the Fourier analysis required to extract the vibration harmonics.

Further pre-processing is targeted at extracting a set of features from HUMS data that yield an optimum identification of fault patterns. Here, normalised vibration is derived from the measured vibration. The measured vibration V_m can be considered to consist of three components:

$$V_m = V_{op} + V_n + V_h \tag{3.1}$$

V_{op} is the vibration induced by operational conditions such as speed and weather.
V_n is the vibration induced by factors such as equipment noise or different pilots operating the same vehicle.
V_h is the vibration induced by mechanical faults.

It is clear that diagnosis based on V_m can trigger false alarms; a car moving on a rough road experiences high vibration which may be confused with fault induced vibration. Normalisation of helicopter vibration is required to remove V_{op} and filter V_n from the measured vibration. The normalised vibration V_{norm} can be expressed as the sum of V_h and a constant value V_c:

$$V_{norm} = V_h + V_c \tag{3.2}$$

V_c can be chosen to be the average value of the fleet vibration at standard operational conditions. The pre-processed vibration V_{norm} ensures robust diagnostics. The probability distribution of V_{norm} for healthy helicopters is expected to be a peaky distribution with a distinct central value at V_c .

V_h can be estimated by subtracting V_{op} from V_m and filtering out V_n. A suite of FDR parameters and associated V_m measurements of healthy helicopters can be used to train a system so that the system can predict V_{op} given the FDR parameters. The system prediction (generalisation) capability can be tested by another set of FDR parameters and associated V_m measurements. If the system predicts V_{op} for the test data with the same accuracy as obtained during training, the system will be adequately trained. The system performance can be enhanced by pre-processing the FDR parameters using model-based reasoning. If the test data and the training data do not cover the required operational range, the outputs of the system will be questionable.

3.1.3 *The Normalisation Process*

Figure 3.1 describes the normalisation process and closely reflects what would happen in-service.

Section 3.2 expands on each normalisation step.

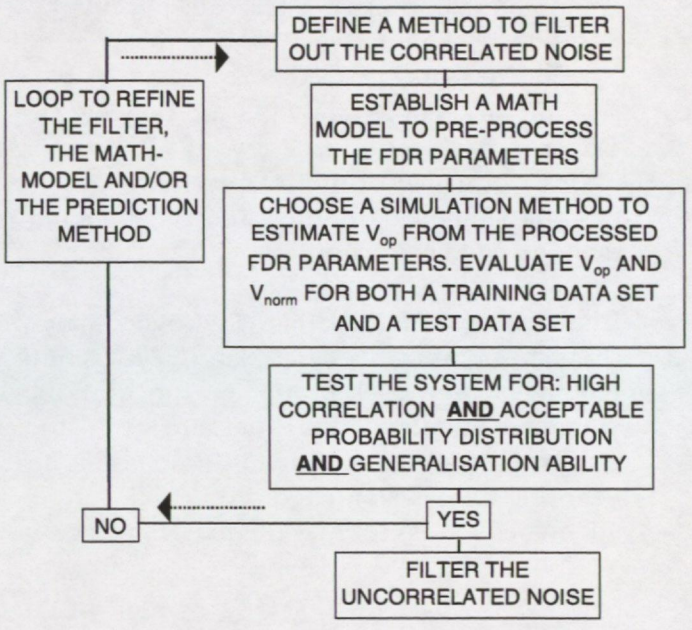


Figure 3.1 The normalisation process

3.2 Demonstration of Data Pre-processing

DEFINE A METHOD TO FILTER OUT THE CORRELATED NOISE

3.2.1 *The Correlated Noise*

Whilst the noise in general can not be definitely predicted by a system or a math model, its statistics can be adequately estimated. The correlated noise is induced by factors that may influence the vibration in a consistent manner. For example, the vibration levels of a helicopter may change with age. It is also possible that for a fleet, the average vibration levels of individual helicopters may not be identical. For these two particular cases, the correlated noise can be filtered out as follows:

$$V_{i)filtered} = V_i - V_{i)average} + V_{c2} \qquad (3.3)$$

V_i indicates the vibration of the i^{th} helicopter and $V_{i)average}$ is the average value of V_i during specified time period. V_{c2} is a constant value which can be chosen as the average value of the fleet vibration. Variations in the value of $V_{i)average}$ between specified time periods will trigger diagnostic alarms, if there is no noted mechanical modification or observed age effects that can explain these variations.

In practice, it will not be known in advance whether the individual helicopter effects are present. Therefore, this step may not be initially implemented.

3.2.2 *The Preliminary Math-Model*

ESTABLISH A MATH MODEL TO PRE-PROCESS THE FDR PARAMETERS

The preliminary math model assumes that the vibration is proportional to two derived parameters, namely the Mach number at the tip of the advancing blade and the air density.

3.2.3 *The Simulation Method*

CHOOSE A SIMULATION METHOD TO ESTIMATE V_{op} FROM THE PROCESSED FDR PARAMETERS. EVALUATE V_{op} AND V_{norm} FOR BOTH A TRAINING DATA SET AND A TEST DATA SET

Neural network technology offers computational tools which can perform non-linear simulations. An Artificial Neural Network (ANN) is applied to a specific problem by allowing the network to learn from a set of data how to solve the problem; this process is termed 'training'. The performance of the trained network is tested by using a set of test data which has not been used during training. The potential of the neural network technology is demonstrated in Section 4.

Unlike ANNs, Multi-Variate Regression analysis (MVR) relies on a math model to simulate non-linear effects. During a training session, the MVR evaluates a weight (coefficient) for each non-linear function obtained from the math model. Thereafter, weighted non-linear functions are evaluated and combined in a linear manner to estimate the vibration. The MVR method is a suitable candidate to reveal whether the math model has simulated the non-linearities and described the underlying physics adequately. Therefore, MVR analysis will be used for this demonstration.

3.2.4 *The Process Effectiveness Test*

TEST THE SYSTEM FOR: HIGH CORRELATION **AND** ACCEPTABLE PROBABILITY DISTRIBUTION **AND** GENERALISATION ABILITY

Having fixed the simulation method, the effectiveness of data pre-processing can be assessed by evaluating the following:

- *The correlation between the measured and predicted vibration as well as the slope and the y-axis intersect of the line that best fits the measured and predicted data.* A model having a correlation coefficient of 1.0, slope of 1.0 and zero y-axis intersect corresponds to 100% normalisation.

- *The probability distribution of the normalised vibration.* For healthy helicopters, a peaky distribution with a distinct central value is expected.

- *The generalisation capability.* The math model must simulate the test data with the same accuracy obtained during training.

The results obtained by analysing data, sampled with the aircraft in stable flight at an airspeed of 125 knots, from 457 flights involving three helicopters (G-TIGB, G-TIGC, and G-TIGE, arbitrarily selected) indicated poor correlation between the measured and predicted vibration; for example the correlation coefficient of the 4R lateral vibration was 0.17. It was therefore concluded that the mathematical simulation based on rotor blade tip Mach number and air density was inadequate.

3.2.5 *The Process Refinements*

LOOP TO REFINE THE FILTER, THE MATH-MODEL AND/OR THE PREDICTION METHOD

The refinement process was initially targeted at the math model to establish improved relationships between the FDR parameters and the vibration. Appendix B presents the math formulations which were used for data pre-processing as a means to enhance the non-linear modelling capability [Reference 3].

Considering the 457 flights of the three helicopters, the refined process produced the following results:

Table 3.1

	4R Vertical	4R Lateral
Correlation coefficient	0.67	0.78
Best line slope	0.98	1.0
Best line y-intersect	0.066	-0.015
SD ÷ Mean of measured	0.23	0.516
SD ÷ Mean of normalised	0.17	0.324
Spread reduction	25.8%	37.2%

The vibration traces shown in Figures 3.2 and 3.3 indicate that the normalisation has reduced the spread of vibration between flights significantly. A successful normalisation is expected to move flight records from the tails of the probability distribution of the measured vibration towards the central value. Considering the vertical vibration and referring to the frequency chart of Figure 3.4 the normalisation moved the sample at 1.2 ips, the sample at 0.5 ips and all samples at 0.4 ips towards the central value at about 0.2 ips. The desirable normalisation effect on the frequency distribution is evident as can be seen in Figure 3.4.

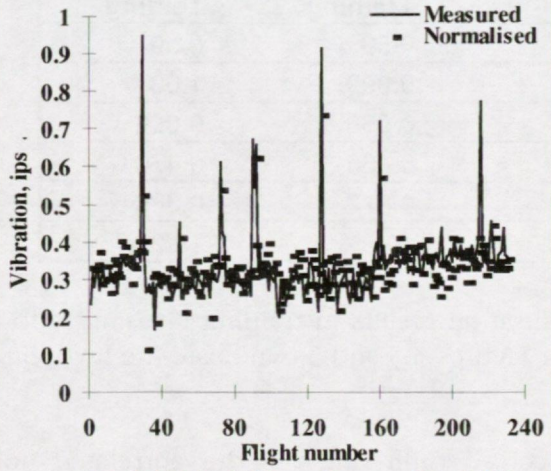


Fig.3.2 Model-based normalised 4R vertical vibration, G-TIGB, G-TIGC, G-TIGE at 125 knots

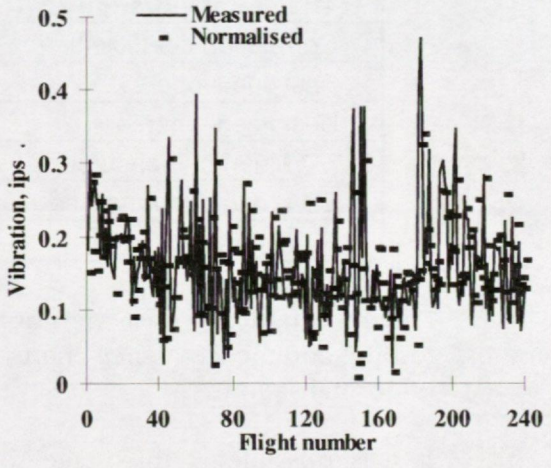


Fig.3.3 Model-based normalised 4R lateral vibration, G-TIGB, G-TIGC, G-TIGE at 125 knots

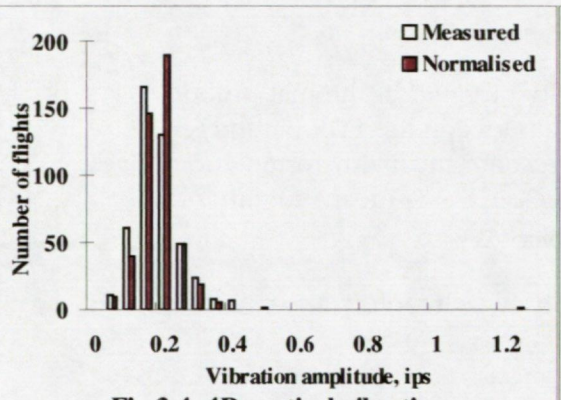


Fig.3.4 4R vertical vibration; a frequency chart, G-TIGB, G-TIGC, G-TIGE at 125 knots

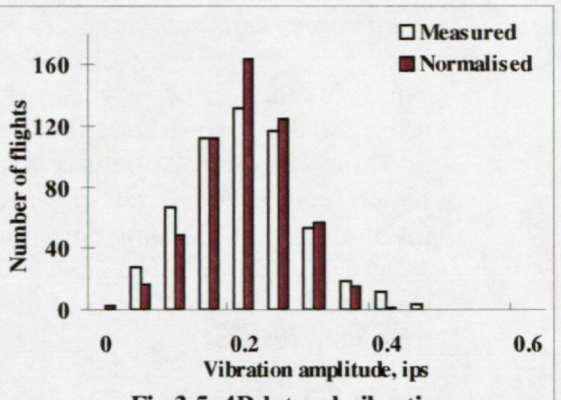


Fig.3.5 4R lateral vibration; a frequency chart, G-TIGB, G-TIGC, G-TIGE, 125 knots, training

The data set considered above is relatively small to be randomly divided into training and test data sets. The random division of data ensures the absence of any subjective bias. The size of each data set must be sufficient for the operational conditions covered by one set to be similar to those covered by the other set. A sub-set of the original data set representing all 23 Super Puma MK1 helicopters was therefore constructed. This subset comprised records for which the 4R lateral vibration was less than 0.7 ips, i.e. below the threshold value which would trigger engineering concern. Only 4R lateral vibration was considered since its variance is far greater than the variance of the 4R vertical vibration (see the SD ÷ Mean values of Table 3.1). The data set was divided randomly into two equal sets. The first set was used for training and the second set was used for testing. The normalisation results are presented in Table 3.2.

Table 3.2

4R Lateral vibration	Training	Testing
Correlation coefficient	0.56	0.59
Best line slope	0.999	1.03
Best line y-intersect	0.056	0.006
SD/Mean of measured	0.453	0.470
SD/Mean of normalised	0.377	0.386
Spread reduction	16.8%	19.2%

Table 3.2 indicates that the normalisation results in training are similar to those in testing, and the frequency charts of Figures 3.5 and 3.6 indicate the favourable effect of normalisation.

It was possible at this stage to consider filtering out the correlated noise. The individual helicopter effects were removed as described in Section 3.2.1. Referring to Figure 3.7, the enhanced normalisation through this process can be observed as an increase in the peak value of the distribution.

Further refinements are possible by revisiting the math model and the correlated noise model. While the above discussion suffices for the demonstration, it is important to appreciate that the in-service system must be sufficiently flexible that the refinement processes do not require the system to be redesigned.

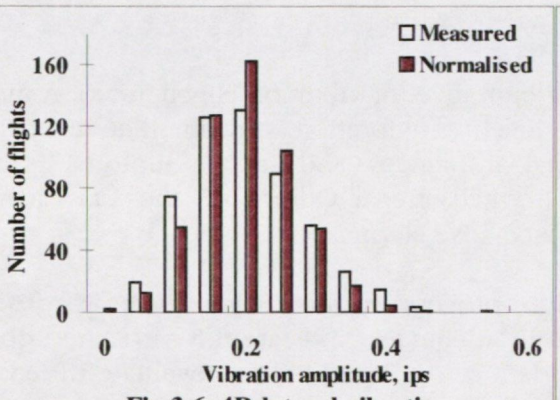


Fig.3.6 4R lateral vibration; a frequency chart, 23 helicopters, 125 knots, testing

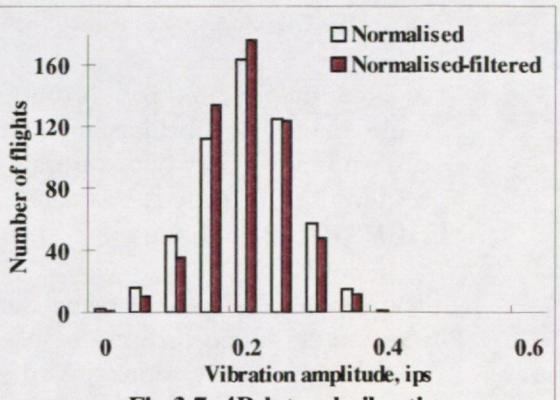


Fig.3.7 4R lateral vibration; a frequency chart, 23 helicopters, 125 knots, training

3.2.6 *The Uncorrelated Noise*

FILTER THE UNCORRELATED NOISE

Having largely removed the operational influences on the measured airframe vibration, and attenuated the correlated noise, the normalised vibration reflects the mechanical health state combined with random noise. The term uncorrelated is used to emphasise that this type of noise can not be correlated with available FDR parameters, individual helicopters or age. It can not be definitely estimated by a system, but its statistics can be determined. This type of noise is attributed to factors such as:

- Measurements' accuracy and equipment's noise.
- Different pilots flying the same helicopter.
- The flight parameters and the measured vibration are not simultaneously acquired; the vibration of the helicopter has been evaluated from an average over 100 main rotor cycles and the FDR parameters measured at the start of the first cycle. This is a constraint associated with the HUMS data and not a deliberate choice.

If the noise level is high relative to the mechanical health contribution (high noise to signal ratio), the noise must be suppressed, e.g. by implementing a simple moving average calculation, where the present plus a number of previous measurements are used to compute the average. Nevertheless, care must be taken not to diminish the health contribution. Suppression of fault induced vibration can be avoided by evaluating a moving weighted average, where the present has more weight than the past. Generally the noise to signal ratio can be reduced significantly by appropriate choice of the weights in the following moving average equation:

$$V_{FL} = \left(\sum_{j=L}^{L-N+1} V_j W_j \right) / \left(\sum_{j=L}^{L-N+1} W_j \right) \qquad (3.4)$$

V_{FL} is the L^{th} filtered value.
V_j is the j^{th} value of vibration.
W_j is the weight of the j^{th} vibration value.
L indicates the present sample.
N is an arbitrary window width.

3.3 Conclusions

As a car moves on a rough road, a high level of vibration is generated. A mechanical fault may also induce a high level of vibration. An intelligent process that discriminates between operational influences and fault symptoms is therefore required (this process is called normalisation). Otherwise, the vast quantities of HUMS data is highly likely to produce false alarms.

The essence of the normalisation process is that Flight Data Recorder (FDR) parameters, once pre-processed, can adequately indicate the associated operational conditions. Furthermore, given a data set, it is possible to evaluate the statistics of the noise which may have contaminated the data. The demonstration presented in this section has highlighted the importance of the model-based data pre-processing where non-linear functions of FDR parameters have been deduced through a helicopter math model. These functions have been used to mitigate the operational influences on the measured airframe vibration. Also, the normalisation process has implemented statistical calculations to attenuate the noise effects which can not be consistently simulated by a model or a system.

It is important to realise that our engineering knowledge will be continuously improved as a direct consequence of the patterns and information within the vast amount of HUMS data. The related mathematical and statistical models will be also enhanced. Therefore, the HUMS intelligent system must be flexible enough to incorporate the refined knowledge without re-building the system. In other words, the refinement processes are required to be part of the system. This report has presented a typical example of a normalisation refinement analysis.

4 SUPERVISED MACHINE LEARNING

4.1 General

This section concentrates on the 3rd objective of the CAA contract 7D/S/1128, namely demonstrating the diagnostic benefits of supervised machine learning techniques. The demonstration involves training a Neural Network to recognise bearing defects.

4.1.1 *The HUMS SOAP Data*

The Spectrometric Oil Analysis Programme (SOAP) data used for the purpose of this demonstration was obtained from a CAA HUMS trial which, in part, involved Super Puma Helicopters (see Sections 2.1.1 and 2.2.1). During the trial, two bearing faults were reported. The first fault was a mast bearing spalling on the M186 gearbox of the G-TIGU helicopter. The second fault was an epicyclic bearing spalling on the M285 gearbox of the G-TIGE helicopter. Eight samples from the M186 gearbox and nine samples from the M285 gearbox were considered to be fault related samples, a total of 17 fault cases.

Armed with the known mechanical defects, the cluster analysis of Section 2 has indicated that raw measurements can trigger false alarms and provide poor visibility of faults. It has been also concluded that enhanced discriminatory capability can be realised through the following pre-processing steps:

- Normalise element concentrations by a reference element concentration; filter the correlated noise and operational effects.

- Compute and correct wear rates. If the operational information required for correction is not available, correct the wear rates such that a negative or zero value is set to the previously calculated value.

- Implement a moving average on the normalised concentration levels and corrected wear rates; filter the random noise.

As indicated above, the supervised learning process determines the relationship between causes and effects. In this section, the effects will be considered to be filtered pre-processed measurements: concentration levels of Fe and Al normalised by Mg along with corrected wear rates for Fe and Mg (see Section 2.2.4). The causes of these effects will be three gearbox bearing states: a healthy or normal bearing (116 samples), a gearbox with a mast bearing spalling fault (8 samples) and a gearbox with an epicyclic bearing spalling fault (9 samples). Only 133 samples were considered since the pre-processed measurements of the remaining five samples could not be evaluated.

4.1.2 *Supervised Machine Learning*

Supervised learning relies on a priori knowledge and uses a pre-defined set of data to solve a specific problem; this process is termed 'training'. The training data consists of input-output pairs; input patterns (effects) and associated desired output patterns (causes). The learning capability of any supervised system is accomplished through its adaptive weights (coefficients). Training means allowing the supervised system to adjust its weights such that the required mapping between the input and output can be reproduced. For example, for each input pattern in the training set, the difference (error) between the actual output and the desired output is evaluated. Then, the weights are adjusted to minimise the mean of the squared errors. Usually, another set of data is used to test whether the training is successful. Only after training, can the supervised process solve the required problem.

According to the above definition, a successful supervised process delivers successful **approximation** and **generalisation**. Approximation may be regarded as learning a smooth mapping or a smooth surface construction from sparse data points. Generalisation means estimating the height of the surface where the locations (e.g. x, y) are not included in the training data. Interpolation is the limit of approximation where there is no noise in the data. The information in the training data can be insufficient to construct uniquely the mapping in regions where training data are not available. Also, extrapolation cannot happen in any predictable way.

Traditionally, supervised learning has been based on linear Multi-Variate Regression (MVR) analysis. The past few years however have seen a rapid growth of interest in Artificial Neural Networks (ANNs) which can offer powerful recursive, non-linear MVR processes.

4.1.2.1 Multi-Variate Regression (MVR) Analysis

MVR is a single-shot squared errors' minimisation process. To illustrate this, consider the two input measurements X and Y along with the corresponding desired output Z

of Table 4.1 (a). An MVR process is required to predict the output Z given the inputs X and Y.

The first step in the MVR analysis is to choose a model that contemplates how the output Z can be linearly composed from non-linear functions of input parameters. In other words, prior engineering knowledge of the model is essential; a badly chosen model can have a very poor generalisation capability. For the data shown in Table 4.1 (a), the exact model is $Z = w_1 + w_2 \sqrt{X} + w_3 Y^2$.

The second step is to evaluate the non-linear functions $X_1=\sqrt{X}$ and $Y_1= Y^2$ for each input pair (X,Y) as shown in Table 4.1 (b).

Table 4.1 (a)

$Z = w_1 + w_2 \sqrt{X} + w_3 Y^2$

X	Y	Z
1	1	9
4	2	18
9	3	31
16	4	48
25	5	69
1	5	57
4	4	42
9	3	31
16	2	24
25	1	21

Table 4.1 (b)

$Z = 4 + 3 X_1 + 2 Y_1$

$X_1=\sqrt{X}$	$Y_1= Y^2$	Z
1	1	9
2	4	18
3	9	31
4	16	48
5	25	69
1	25	57
2	16	42
3	9	31
4	4	24
5	1	21

The third step is the training step where a Least Mean Square (LMS) algorithm is implemented to estimate the weights w_1, w_2 and w_3 such that the sum of squares of differences between the desired Z values and the Z values as calculated from the model is a minimum. The weights are evaluated through single-shot matrix algebra and not through an iterative process.

Practical problems are often described by a large number of input measurements and hence, it is possible that a number of input measurements can be linearly dependent on other inputs. In this situation, direct LMS algorithms can give misleading results. Singular Value Decomposition methods (SVD) are available to rectify such a problem. It is also possible that only a sub-set of the inputs is significant, and that the remaining measurements do not influence the desired output. In this case, Principal Component Regression (PCR) is available to identify the principal, significant features in the input measurements.

4.1.2.2 Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are commonly described as being analogous to the physiology of the brain more than other kinds of information processing methods. The brain consists of a large number of nerve cells (neurons) which send very simple messages to each other via connections (synapses). The power of the mind probably stems from complicated sequential and massive parallel processing carried out by the neurons. The neurons themselves are slow processors.

Any ANN consists of a number of processing units referred to as neurons, units or cells. The neurons receive input values through connections. Each connection has a specific strength referred to as its weight. The input values to a neuron are referred to as the neuron input pattern (or input vector). The weights of the connections that deliver the input pattern are referred to as the weight pattern (or weight vector). The basic building block of the ANNs is the neuron and its connections; a detailed description is presented in Appendix C. Generally, each neuron performs three computational tasks:

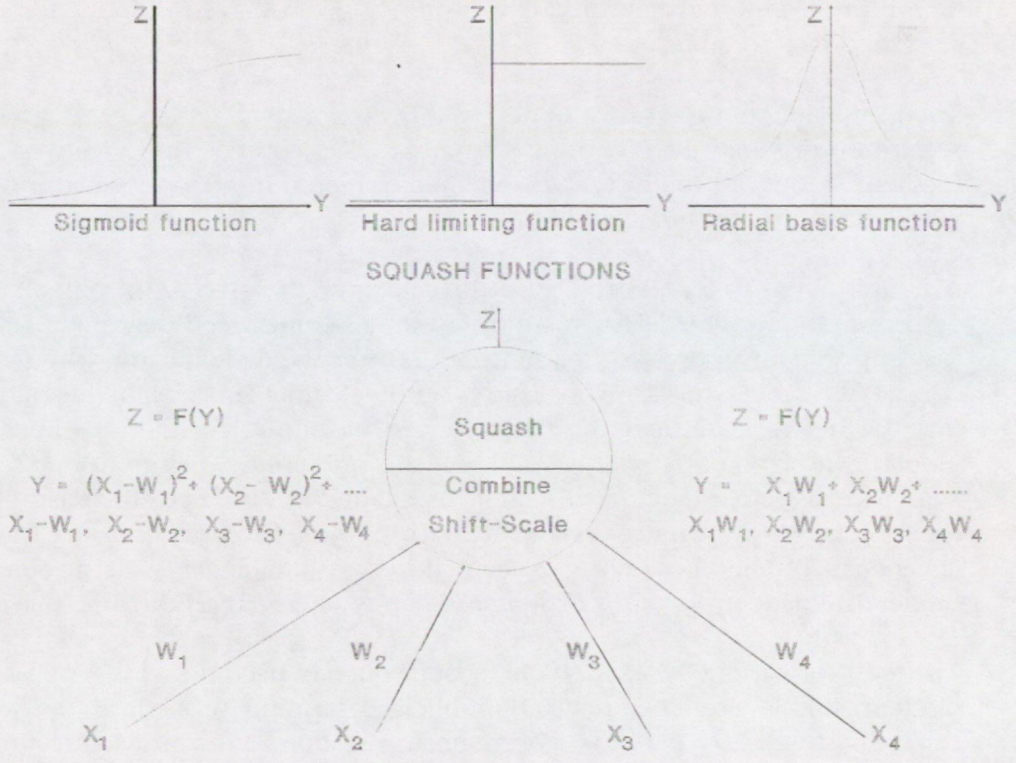


Fig.4.1 A neural network neuron

- Scale (or shift) the neuron input values using the connection weights which deliver these values. Often, this is achieved by either multiplying each input by the weight of its connection or by evaluating the difference between the input and the weight.

- Combine the scaled input values into a single value.

- Apply a non-linear function (activation function) to the combined value to produce the neuron output (activation). The non-linear function squashes the combined input, which can vary from $-\infty$ to $+\infty$, to finite values (e.g. values between 0 and 1).

Figure 4.1 summaries the general computational tasks of the neuron. The MVR is a special case of this general case where only one unit without a squashing function is used to combine its inputs linearly.

The neuron receives four types of inputs:

- Inputs from external sources. For example, the inputs may be N numbers that represent the concentration of elements in SOAP samples or amplitudes and phases of vibration signatures.

- A constant internal input referred to as the bias.

- The feedback from the neuron to itself. The current and previous (delayed) outputs may be feedback to the unit.

- Input from other neurons. The current and/or delayed outputs of a unit may be feedback to other units.

The inputs to a neuron may not necessarily comprise all types of input. For example, some neurons may receive inputs from external sources only. Networks that have feedback loops are referred to as dynamic or recurrent networks as opposed to static networks which are those without feedback.

A neural network learns from a set of data how to solve the problem, typically by adjusting its weights. The learning is either supervised learning, reinforcement learning or unsupervised learning. In **supervised learning**, the training data consists of input and 'known desired output' pairs. The weights are adjusted such that the network predicts the output given the input. In reinforcement learning, a global reinforcement signal is used and the probability distribution associated with each local variable (such as a weight) is changed in order to increase the expected reinforcement. In **unsupervised learning** the input patterns are grouped into clusters such that the patterns in each cluster are similar. The learning may be based on deterministic procedures or stochastic ones such as the Boltzman machine.

The feed-forward networks are the most frequently used networks, perhaps, because of their simple architectures and publicised training procedures as well as their capability to classify patterns, approximate functions and simulate arbitrary Boolean expressions such as AND, OR and XOR. The external input (X1, X2, X3, ...) is passed forward to the neurons of the first hidden layer through unidirectional connections. The name 'first hidden layer' is used because this layer is the first active layer that performs computational tasks and it is hidden from the external environment; the layer does not receive input from external sensors or send output out of the system. Each hidden layer feeds its output to the following layer. The last hidden layer feeds its output to an output layer, which may consist of more than one neuron. The network has no feedback connections from a unit to itself or from a unit to another unit on a preceding layer. Most of the learning algorithms of this type of networks utilise a supervised LMS learning technique which starts with an initial guess of weights and then searches for optimum weights using a recursive update. For example the back-propagation algorithm of Reference 4 is based on this technique.

Other types of neural networks include self organising maps and Adaptive Resonance Theory (ART) networks which belong to a category of learning called competitive, unsupervised, or self-organising. The ART networks prevent previously learned knowledge from being washed away by new learning, and enables new learning to be automatically incorporated into the total knowledge base of the system in a globally self-consistent way. Interested readers can refer to Reference 4 and the literature cited in Reference 5.

4.2 Demonstration of the Benefits of Supervised Learning

In Section 2, four pre-processed features were evaluated from three elements' concentration levels (Fe, Al, Mg). The possible random noise in these features were filtered using a moving average. The four features are as follows:

- Fe normalised by Mg.
- Al normalised by Mg.
- Fe corrected wear rate.
- Mg corrected wear rate.

The unsupervised process of Section 2 has indicated that these features can yield clear fault visibility.

The four features were evaluated for 133 samples relating to three bearing states; healthy state (116 samples denoted by H), an epicyclic bearing spalling (9 samples denoted by F) and a mast bearing spalling (8 samples denoted by G). For the purpose of this demonstration, the available data are split uniformly into a training data set and a test data set.

4.2.1 *MVR Analysis*

The use of MVR analysis for supervised classification purposes is not a straight forward task and requires a carefully chosen strategy. The strategy used in this section considers three MVR units and a logical gate.

The task of the first MVR unit is to discriminate between the two bearing faults F and G. The task of the second MVR unit is to discriminate between H and F. The third unit discriminates between H and G. The logical gate interprets the output of the three units and issues diagnostic reports.

Five F samples and four G samples are considered and the associated four pre-processed features are presented to the first MVR unit with the known causes for training. Successful training implies that the weights are adjusted to describe a plane that separates the F and G states. For two inputs the plane is reduced to a straight line, and for more than three inputs, the plane is called a hyper plane. The output of the MVR unit indicates the distance of a sample from the hyper plane. The trained MVR unit was tested using the remaining four F samples and four G samples and found to discriminate successfully between the two states. Figure 4.2 shows the output of the unit (training and test results). It can be seen that the points at one side of the hyper plane (above 2 on the y axis) represent the epicyclic bearing fault and those on the other side (below 2 on the y axis) represent the mast bearing fault indicating 100% success rate in both training and test.

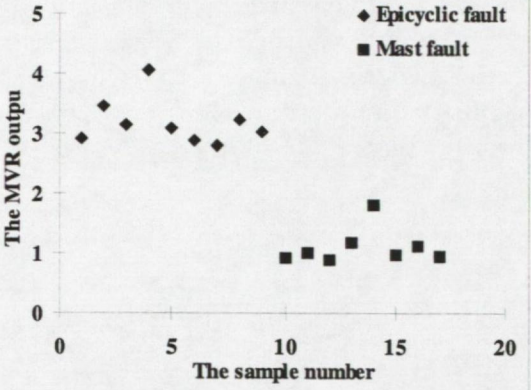


FIG.4.2 Two gearbox bearing faults as identified by a linear MVR from pre-processed FE, MG and AL (Super Puma)

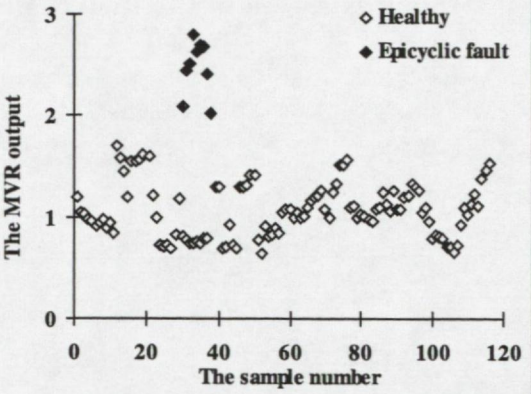


FIG.4.3 An epicyclic bearing fault as identified by a linear MVR from pre-processed FE, MG and AL (Super Puma)

Similarly, the second MVR unit is trained using 5 F samples and 58 H samples and tested using 4 F samples and 58 H samples. Figure 4.3 shows the output of this unit and indicates 100% success rate in training and test. The F samples are above 2 and the H samples are below this level. It is worth mentioning that the samples are presented in a chronological order and, generally, each training sample is followed by a test sample. Referring to Figure 4.3, the response of the MVR unit to the F samples starts with a value just above 2, peaks at the fourth sample to a value of 2.8 and returns to a value above 2 at the 9^{th} sample. This can be attributed to both the moving average effects and the fault mode. It is anticipated that the wear mode of a bearing fault can change with time; for example, the wear mode can change after the erosion of the plating material on bearing cages.

The third MVR unit is trained using 4 G samples and 58 H samples and tested using 4 G samples and 58 H samples. The unit output shown in Figure 4.4 indicates that the hyper plane can not separate the two states. The output levels of both the H and G samples are below 2 and overlap for all samples with the exception of one G sample. It seems that the mast bearing fault samples occupy a part of the feature space surrounded by healthy samples. In such a situation, a hyper plane is not the right choice and a hyper sphere or ellipsoid can offer better discrimination. Therefore, an MVR is allowed to adjust its weights using the training set described above such that the optimum size and location of the hyper ellipsoid that can separate the two states is determined. In training, 2 G samples and 58 H samples are correctly identified. In testing, 2 G samples and 57 H samples are correctly identified. One H sample is classified as a G sample which is one false alarm in 116 H samples. Figure 4.5 shows the output of this non-linear MVR unit. The term non-linear is used to indicate that the target is to define a non-linear separator (ellipsoid) not a linear separator (plane). The output of the unit for 4 G samples is definitely above 2. The other four samples are on the border between the two states. The degree of membership of these samples to each state can be determined and used for prognosis.

In the training strategy described above, samples describing only two bearing states are presented to each unit. The question now is what is the response of a unit trained to discriminate between F and H to G samples? In this case, it is not known where the G samples will be relative to the hyper separator. Thus, if an unknown sample is presented to the unit, the output can only indicate that the sample is F or G on one side of the hyper plane or, H or G on the other side. In other words, the output of each MVR unit indicates not a single state but two states as shown in the

following table (the symbol U is the union or the OR operator. O_1, O_2 and O_3 refers to the three unit outputs):

	trained on		O_1, O_2 and O_3 indicates either	or
Unit 1	F	G	F U H	G U H
Unit 2	F	H	F U G	H U G
Unit 3	G	H	G U F	H U F

The logical gate examines the output of the three units and evaluates the diagnostic report as follows:

The bearing state = $O_1 \cap O_2 \cap O_3$

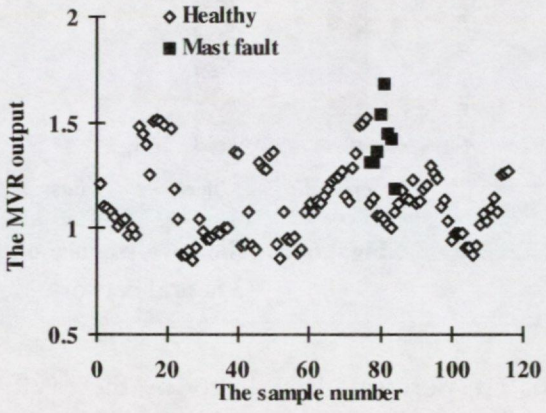


Fig. 4.4 A mast bearing fault as identified by a linear MVR from pre processed Fe, Mg and Al (Super Puma) linear MVR

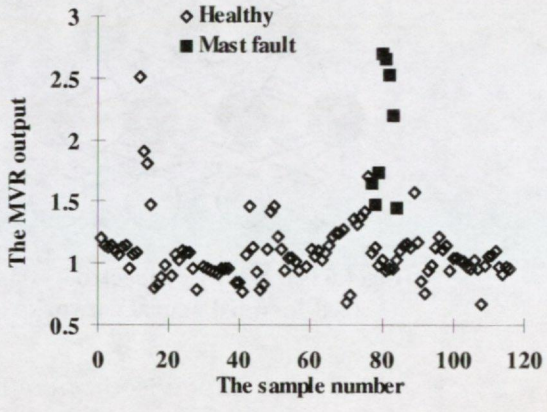


Fig. 4.5 A mast bearing fault as identified by a non-linear MVR from pre processed Fe, Mg and Al (Super Puma) linear MVR

The symbol $\cap$ is the intersection operator or the AND operator. For example, if the output of the three units is G U H, F U G and G U F, then the output of the logical gate will indicate the G fault. The collective results of the MVR system (three units and a logical gate) indicate 100% detection rate of the F state, 50% detection rate of the G state and one false alarm in 116 samples. It is worth mentioning that a magnetic chip detector indicated the mast bearing fault a few hours prior to the head replacement at 6794.35 flying hours. In Figure 4.5, the maximum response to the G state occurs at the fourth, fifth and sixth samples which were taken at 6624, 6640 and 6743 flying hours. This suggests that if the techniques described in this report had been available, pre-processed SOAP measurements would have indicated the mast bearing spalling 170 to 51 hours before the magnetic plug.

The above discussion not only demonstrates the benefits of supervised machine learning in the form of MVR systems but also provides a generic analysis for a supervised classification process. It breaks down the process to three operations:

- The construction of a series of hyper surfaces using computational units.
- The use of a learning rule (LMS rule) to optimise the size and location of these surfaces.
- The implementation of logical gates to evaluate the output of the computational units and issue the classification report.

4.2.2 *Supervised Artificial Neural Networks*

A neural network classifier implicitly combines the above three operations in a single architecture. Figure 4.6 (a) shows an ANN of 4 input units and 3 output units (white circle), and 7 computational units in two hidden layers (black circles). The 7 computational units can implicitly (and perhaps collectively) construct the hyper separators and act as logical gates. The analysis of the previous section is therefore very relevant to neural networks. Nevertheless, it is important to appreciate the basic differences between MVR and ANN based systems:

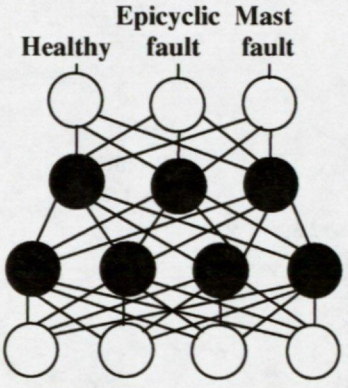


Fig.4.6 (a) A fully connected 4-4-3-3 feed-forward neural network

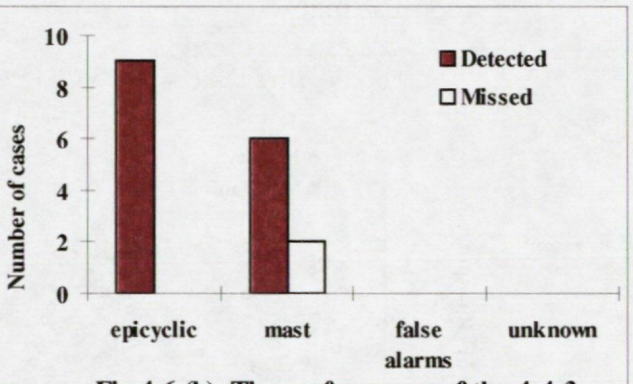


Fig.4.6 (b) The performance of the 4-4-3-3 neural network

- The three supervised operations are performed explicitly by the MVR system, and hence, the analysis of the system results is straight forward. At present, it is not possible to rigorously identify ***how*** an ANN has solved a particular problem, and hence it is difficult to analyse the results and systematically judge whether the trained network has been well trained or not. Nevertheless, this is currently achieved by testing the performance of the trained network using the test data which has not been used during training. If this data and the training data do not cover the required operational range however, the results of the test will be questionable.

- The guidelines for ANN design and training are much simpler than those for an MVR. For example, the number of units of an MVR system is a function of the number of classes and different training and test data sets used for the various units. The construction of non-linear hyper separators is not an easy task. For the ANN system, it suffices to specify a sufficient number of units and only one training data set and one test data set are used.

- The learning rule of the MVR is a single-shot rule as opposed to the recursive rule of the ANN and hence, the time required to train the MVR can be much less than that required to train the ANN. Nevertheless, the memory requirement of the recursive algorithm is much less than that of the MVR and hence, a large data set can be used for training.

- The ANN can be regarded as a generic recursive, non-linear mean square errors' minimisation tool. The ANN automatically defines the appropriate degrees of the various non-linear hyper separators.

The 4-4-3-3 network shown in Figure 4.6 (a) was trained using 5 F samples, 4 G samples and 58 H samples. After 4000 iterations the network was found to classify all F and H samples correctly. Three G samples were correctly classified and the fourth G sample is classified as H. The ANN was tested using 66 samples (4 F, 4 G and 58 H) of which 64 samples (4F, 57 H and 3 G) were correctly classified. One H sample was classified as G; one false alarm in 116 H samples. Also one G sample was classified as an H sample. It is apparent that the hyper non-linear separator of the ANN is more accurate than the hyper ellipsoid of the MVR; 6 G samples were correctly classified by the ANN as opposed to 4 samples identified by the MVR. By increasing the number of iterations to 5000, the performance of the network was improved and the false alarm was eliminated. Figure 4.6 (b) summarises the training and test results and indicates 100% detection rate of the F state, 75% detection rate of the G state along with 0% false alarm.

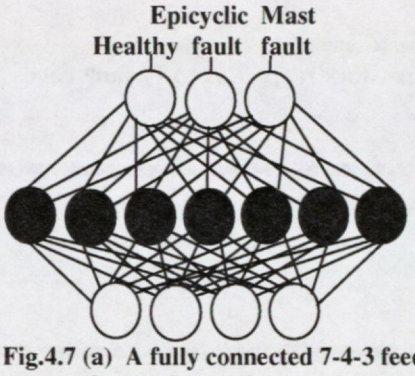


Fig.4.7 (a) A fully connected 7-4-3 feed-forward neural network

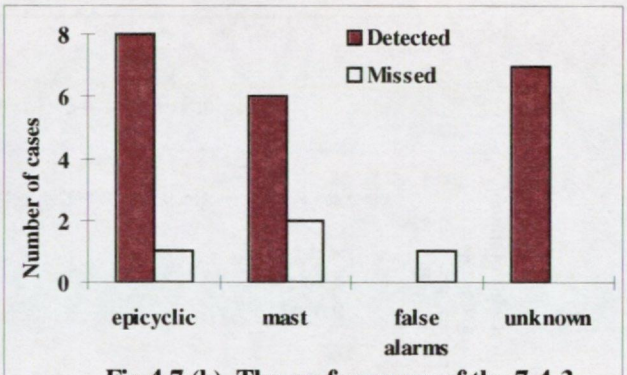


Fig.4.7 (b) The performance of the 7-4-3 neural network

Although the literature survey of Reference 5 has indicated a customary use of a single hidden layer architecture, the two hidden layer was found to provide more accurate results and better generalisation. The second layer condenses the output of the first layer further which makes the various classes more distinct and the output more accurate. The second hidden layer can also be considered as an independent logical gate acting on the hyper surfaces described by the first hidden layer. Combining the two operations in one layer seems to be inefficient. To illustrate this, an ANN of a single hidden layer is considered (a 4-7-3 network). The number of computational units is chosen to be the same as that of the two hidden layer network (7 units). At first sight, it seems that the computational power of the 4-7-3 is more than that of the 4-4-3-3 network; the number of unidirectional connections of the 4-7-3 ANN is 49 connections as opposed to 37 in the case of the 4-4-3-3 ANN. Nevertheless, the training and test of the 4-7-3 ANN using the same data sets and the same number of iterations (5000 iterations) indicate that the efficiency and generalisation capability of the 4-4-3-3 ANN is better than that of the 4-7-3 ANN. Figure 4.7 (b) shows the results of training and test and indicates 88% F state detection rate, 75% G state detection rate, 5.3% unknown states (samples cannot be classified) and one false alarm in 116 H samples.

The responses of the three output units of the 4-4-3-3 ANN to the training and test samples are presented in Figure 4.8 and 4.9. For a successfully trained ANN, it is expected that an input pattern will activate only the unit representing the class to which the pattern belongs. In Figure 4.8, the 9 F samples have activated the epicyclic unit (desired output of about 0.9). Meanwhile, the responses of the healthy and mast fault units have corresponded to the minimum desired response which is 0.1. Figure 4.9 indicates that all healthy samples have activated the expected unit. Referring to

Figure 4.8, the second and third G samples have incorrectly activated the H unit and four samples produced correct activation values; activation values above 0.5 for the G unit and below 0.5 for the other units. Sample 5 is considered to be indicative of the mast fault since the G unit has the highest activation value.

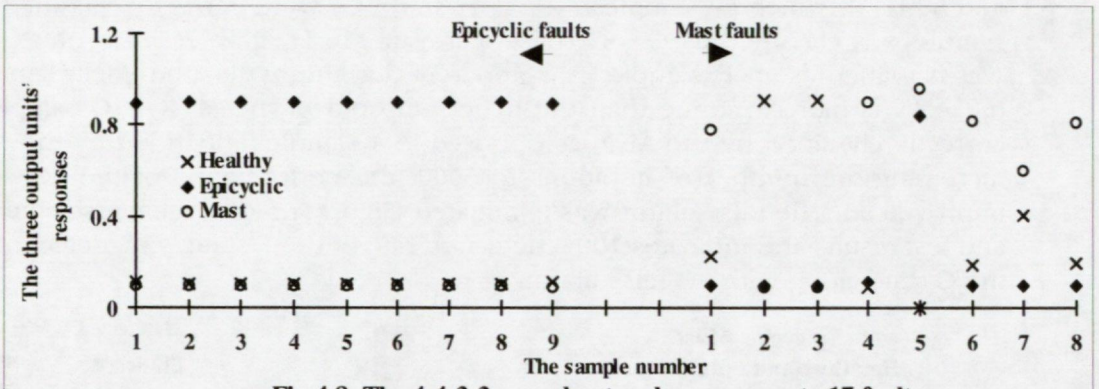


Fig.4.8 The 4-4-3-3 neural network responses to 17 fault cases

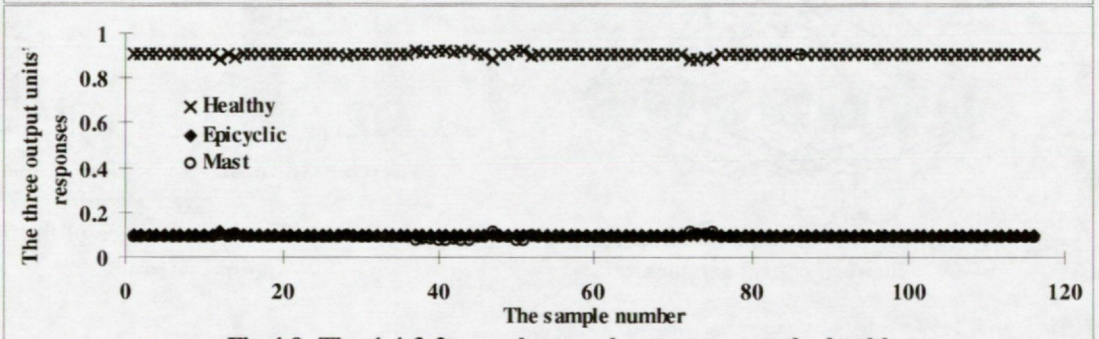


Fig.4.9 The 4-4-3-3 neural network responses to the healthy cases

In Sections 2 and 3, the benefits of model-based data pre-processing have been demonstrated. In order to further demonstrate the importance of data pre-processing, a 4-4-3-3 network was trained and tested using the same data sets. However, the four inputs used in this case were raw concentration levels of Fe, Mg, Al and Ag. After 5000 iterations only 2 G cases were correctly classified. All other training and test samples were classified as H.

4.2.3 *Diagnostic Reports*

The remaining important task of the supervised process is to establish the practical interpretation of the process output which can offer a robust diagnostic. The criteria that define a robust diagnostic are as follows:

- A fault must be identified before it can cause a catastrophic failure or excessive maintenance costs.
- The number of false alarms must be minimum. In a practical situation, removing all false alarms can result in reducing fault visibility.

The demonstration of the previous section suggests two warning messages which a line engineer can readily understand:

- The first message will be triggered if the output of a fault unit is well above the minimum value (0.1) but below a predefined maximum value (0.8), e.g. the first sample of the mast bearing fault. The message will indicate a possible

component fault but costly maintenance actions will not be recommended. The message is stored and the unit activation value which can reflect the fault intensity is closely monitored.

- The second warning message will be triggered if the output of a fault unit reaches the maximum value (e.g. the fourth and fifth samples of the mast bearing fault). The end results will be a work card, cross referenced with the appropriate section in the maintenance manual.

4.3 Conclusions

The function of the supervised process is to draw more precise boundaries around each mechanical state and offer robust diagnosis. The process captures cause-effect relationships, even when the precise relationships cannot be explicitly expressed. This has been demonstrated in this section by training a Multi-Variate Regression system and an Artificial Neural Network as a means to discriminate between the three mechanical states. The training data has contained the known mechanical states (causes) and the associated pre-processed measurements (effects). Testing the systems by presenting pre-processed measurements which have not been used for training has indicated that robust diagnosis is possible. For example, the neural network has classified correctly 100% of the epicyclic bearing spalling related samples, 75% of the mast bearing spalling related samples and 100% of the normal bearing related samples. Only two mast bearing related samples have been mis-classified as normal samples. These two samples had been taken from the M186 gearbox about 239 hours before a magnetic chip detector indicated this fault. Nevertheless, the network output has indicated that early detection of the mast bearing fault would have been possible 170 hours prior to the magnetic plug diagnosis. By analysing the MVR and ANN results, it has been demonstrated that the output of the supervised process could be presented in a work card form.

5 THE INTELLIGENT DATA MANAGEMENT SYSTEM

5.1 General

Section 2 has demonstrated the benefits of unsupervised learning techniques, Section 3 has indicated the importance of data pre-processing, and the benefits of supervised learning techniques have been presented in Section 4. This section concentrates on the **5th objective** of the CAA contract, namely demonstrating the performance of an IDM system by identifying its ability to recognise previously reported and unreported features in HUMS data.

5.2 The Intelligent Data Management Process

A practical IDM system needs to offer a framework that is capable of organising intelligent interactions between data pre-processing, unsupervised learning and supervised learning.

5.2.1 *The IDM Process*

The IDM process of this section starts by identifying initial data pre-processing methods that condition the data before it is operated upon. The unsupervised learning core algorithm is applied on the pre-processed data in order to identify

atypical data patterns even when the cause is unknown. If component samples persistently fall into atypical data clusters, inspection is triggered in order to identify the cause. Initially, the inspection is not expected to identify component faults. However, the inspection results can be used as a means to refine the data pre-processing methods in order to enhance the fault visibility and reduce the number of false alarms. In Section 2, it has been demonstrated that the unsupervised process can achieve its main objective and identify atypical, significant clusters.

As mentioned in Section 1.1.3, the initial choice of a set of features (attributes) defines the frame of reference within which the clusters relevant to these features can only be identified. Therefore, the selected features must be relevant to the pre-defined purpose of the unsupervised process. For example, if the purpose of analysis is to identify product patterns which appeal to the consumer, the colour and geometric shape of the product can be significant features. For the purpose of this demonstration, the IDM process is targeted at the identification of non-adjustable faults in the main rotor system with particular reference to frequency adaptor faults. The relevant features are considered to be harmonics extracted from vibration measurements combined with blade positional information at various operational conditions.

5.2.2 *Frequency Adaptor Faults*

The stiffness and damping values of a faulty frequency adaptor are considerably dissimilar to those of normal units. The drag-wise lag deflection is proportional to both the stiffness and the drag-wise loads. Therefore, the tip lag displacement can be used in order to identify a faulty frequency adaptor. If a frequency adaptor of a blade is stiffer than those of other blades, this blade will lead the other blades in climb and hover (relatively high loading conditions). In descent and at MPOG (Minimum Pitch angle On Ground), the blade will lag the other blades (or lead by values less than the climb-hover values). It is also expected that the value of the 1R lateral vibration phase angle in climb (or hover) will not be the same as the value in descent (or MPOG). These signatures have been proposed to identify a faulty damper. The stiffness of a frequency adaptor S_i can be expressed as follows:

$$1/S_i = D_i/F_{cd} + E|1/S_i| \ , \ i = 1, b \qquad (5.1)$$

where:
- S_i is the stiffness of the i^{th} frequency adaptor.
- D_i is the difference between the lag values in climb and descent of the i^{th} blade.
- F_{cd} is the difference between the loads in climb and descent at 180 degrees azimuth.
- E| | is an operator that evaluates the expected value of the operand between square brackets.
- b is the number of blades.

The load evaluation is beyond the scope of the current investigation. Therefore, it is only possible to assume that the stiffness of the first adaptor and the value of F_{cd} are known. In this case, the stiffness of the other adaptors can be determined from the equation:

$$F_{cd}/S_i = F_{cd}/S_1 - (D_1 - D_i) \ , \ i = 2, b \qquad (5.2)$$

The value of F_{cd}/S_1 for this report was estimated from flight and test measurements. This value must be greater than the maximum value of 'D_1-D_i', otherwise, the solution indicates a frequency adaptor of negative stiffness or an infinitely rigid adaptor.

A comprehensive mathematical model [Reference 6] has been implemented to investigate the track-lag disparities which can be induced by a faulty frequency adaptor, an adjustable pitch link error and various blade irregularities. The model simulates a non-uniform induced velocity wake, detailed non-linear aerodynamic data and a non-linear flap-lag-torsion motion. The model combines the dynamics of the rotor, the fuselage, and other components. The model has highlighted the practical difficulties that can arise because of the shared symptoms between the frequency adaptor fault, the pitch link error and blade irregularities; the symptoms of one of these faults can be similar to those of another. Therefore, it has been concluded that not only the frequency adaptor fault manifests itself through features extracted from low and high loading conditions, but also other main rotor non-adjustable faults.

Essential to the successful estimation of the frequency adaptor stiffness is the removal of the effects of main rotor adjustable faults which can be mass unbalance, tab error and pitch link error. In this report, these effects are estimated from vibration and track measurements in cruise and at MPOG using measured sensitivities. Nevertheless, it has been found that the values of the measured sensitivities need revising and hence, the estimated values of adjustable faults are not expected to be accurate.

5.3 **The HUMS Data**

A data set from the HUMS database was used to establish and refine the unsupervised process. The data was pre-processed before the clustering analysis. Another data set was used to test the significance of the analysis and demonstrate the performance of the IDM process by establishing its ability to recognise previously reported and unreported features in HUMS data.

5.3.1 *The Training Data*

A HUMS data set covering the period from 1/1/1994 to 31/8/1994 was used to refine the IDM process as described in Section 1. The data set was extracted from 5585 data downloads; most of the downloads were associated with revenue earning flights. Vertical and lateral vibration harmonics along with track and lag measurements from the following 23 Super Puma MK I helicopters were used for the purpose of this demonstration: G-BLPM, G-BLRY, G-BLXR, G-BLXS, G-BMCX, G-BTCT, G-TIGB, G-TIGC, G-TIGE, G-TIGF, G-TIGG, G-TIGI, G-TIGK, G-TIGL, G-TIGM, G-TIGO, G-TIGP, G-TIGR, G-TIGS, G-TIGT, G-TIGU, G-TIGV and G-TIGW. The vibration harmonics were extracted at multiples of the main rotor frequency. The number of maintenance actions which corresponded to the 5585 downloads was 2193. The maintenance actions included the following 138 main rotor non-adjustable maintenance actions:

Table 5.1

No. of faults	fault description
11	Main rotor frequency adapter pair replaced.
45	Main rotor replacement of blades.
42	Main rotor assembly refitted/replaced.
17	Main rotor pitch rod replaced on a blade.
10	Main rotor support struts refitted/replaced.
13	Main rotor spindle replaced on a blade.

5.3.2 *Data Pre-processing*

Practical experience and mathematical models suggest that the faults of interest can manifest themselves through features extracted at high and low loading conditions (see section 5.2.2). Therefore, only climb, descent, cruise and MPOG measurements were selected and stored in a database. The database consisted of 17062 records of which 1295 records were associated with maintenance actions. Each record contained information about a test point such as climb and each download contained a number of test points. Vibration measurements and track and lag displacements were filtered using the following linear filter:

$$V_{FL} = \left(\sum_{j=L}^{L-N+1} V_j W_j \right) / \left(\sum_{j=L}^{L-N+1} W_j \right) \qquad (5.3)$$

V_{FL} is the L^{th} filtered value.
V_j is the j^{th} raw value .
W_j is the weight of the j^{th} raw value.
L indicates the present sample.
N is the window width .

For each helicopter, the width of the window was increased from a value of 1 up to a value of 7, and then kept constant at a value of 7. After each maintenance action, the width of the window was re-set to a value of 1. In this way, the noise in the first record is not influenced at all by the averaging process (window width of 1), and the noise in the second record is slightly reduced by averaging the values of the first and the second samples. Therefore, the first five records following a maintenance action were removed to ensure that the random noise was adequately attenuated. Each flight that contained information about climb, descent, cruise and MPOG was considered, and the information was stored as a record in another database table. The database table consisted of 1358 records. It was assumed that maintenance actions associated with vibration alerts could have been carried out up to 14 days after the alert. There were 1637 such maintenance actions. These included the following 104 main rotor non-adjustable faults:

Table 5.2

No. of faults	fault description
10	Main rotor frequency adapter pair replaced
35	Main rotor replacement of blades.
31	Main rotor assembly refitted/replaced
13	Main rotor pitch rod replaced on a blade
8	Main rotor support struts refitted/replaced
7	Main rotor spindle replaced on a blade

Adjustable main rotor faults, namely mass unbalance, tab errors and pitch link errors which could have been associated with each record, were estimated by an RTB algorithm from the filtered measurements in cruise and at MPOG. The influence of these faults on vibration and blade displacements were removed by the RTB algorithm. Diagnostic features were extracted from the residual vibration and blade displacements. The features were ratios between vibration amplitudes, differences between phase angles (cruise-MPOG and climb-descent) and predicted stiffness values of blades' frequency adaptors. The stiffness values have not previously been reported by the HUMS system. The IDM analysis of this report was based on these features.

5.3.3 *The Test Data*

A data set covering the period from 14/8/1994 to 6/11/1994 was used to test the significance of the cluster analysis. The data represented 1812 downloads of which 1412 downloads did not overlap with the training data. The overlapped data was only included to ensure smooth filtered results and, was not used in testing. The data was gathered from the following 20 helicopters: G-BLXR, G-BLXS, G-BMCX, G-BTCT, G-TIGB, G-TIGC, G-TIGE, G-TIGF, G-TIGG, G-TIGI, G-TIGK, G-TIGL, G-TIGM, G-TIGO, G-TIGR, G-TIGS, G-TIGT, G-TIGU, G-TIGV, G-TIGW. The number of maintenance actions which corresponded to the 1412 downloads was 459. The maintenance actions included the following 30 main rotor non-adjustable maintenance actions:

Table 5.3

No. of faults	fault description
4	Main rotor frequency adapter pair replaced
11	Main rotor replacement of blades.
12	Main rotor assembly refitted/replaced
3	Main rotor pitch rod replaced on a blade
0	Main rotor support struts refitted/replaced
0	Main rotor spindle replaced on a blade

The test data was pre-processed in the same way as the training data. The measurements which had been acquired in climb, descent, cruise or MPOG were filtered and the first five records succeeding a maintenance action were removed to ensure that the random noise was adequately attenuated in the remaining samples. Each flight that contained information about climb, descent, cruise and MPOG were

considered and represented by a database record. This resulted in a database that consisted of 446 records. The number of distinct maintenance actions which were carried out within a 14 day period and associated with these records was 336. Removing the overlap with the training data, the numbers of records and maintenance actions covering the period from 1/9/1994 to 6/11/1994 were 385 and 277 respectively. The maintenance actions during this period included the following 19 main rotor non-adjustable maintenance actions:

Table 5.4

No. of faults	fault description
1	Main rotor frequency adapter pair replaced.
8	Main rotor replacement of blades.
7	Main rotor assembly refitted/replaced
3	Main rotor pitch rod replaced on a blade
0	Main rotor support struts refitted/replaced
0	Main rotor spindle replaced on a blade

5.4 Demonstration of the Performance of an IDM System

5.4.1 *Pre-processing and Features Selection*

The following list summarises the data pre-processes applied and describes the features selected:

- Vibration harmonics along with relative main rotor blade track and lag displacements were extracted from HUMS data downloads.
- A linear filter was applied to attenuate the random noise (the non-correlated noise).
- The influences of adjustable main rotor faults on measurements were removed. These influences can be considered as correlated noise effects. An RTB algorithm using the Singular Value Decomposition (SVD) method was developed for this purpose.
- The following four features were selected: the differences between the stiffness values of opposite blades (D1 and D2), the ratio between the amplitudes of the 1R lateral vibration in climb and descent (R1LAC) and the difference between the phase angles of the 1R lateral vibration in climb and descent (R1LPC). This selection was based on practical experience and substantiated by mathematical models.

5.4.2 *Unsupervised Learning*

Each feature was scaled (divided by its standard deviation) and an optimised – iterative – self-organising data clustering algorithm was implemented. The clustering analysis of the pre-processed training data (Section 5.3.2) produced 6 clusters. The first and second clusters contained two flight records (samples) having very high R1LAC values indicative of faulty sensors (118.97 ips and 54.952 ips) and, hence, can be considered as outlying clusters. The third and sixth clusters (51 and 592 samples) are characterised by small R1LPC values. The fourth and fifth clusters (189 and 524 samples) can be identified by relatively high values of D1 or D2 along with high

values of R1LPC. Considering the number of samples in each cluster, and ignoring the outlying samples, two atypical candidates were nominated based on a relatively small number of samples. These were cluster 3 and cluster 4. Whilst the unsupervised analysis of the training data was not used to trigger alarms, the maintenance actions which could be associated with the clusters were scrutinised and listed in Table 5.5.

Table 5.5 Unsupervised cluster analysis of the training data

Cluster serial number	1	2	3	4	5	6	1	2	3	4	5	6
	Number						%					
Number of helicopters in the cluster	1	1	5	12	21	23	1.6	1.6	7.9	19.0	33.3	36.5
Number of records (samples)=1358	1	1	51	189	524	592	0.1	0.1	3.8	13.9	38.6	43.6
Number of associated maintenance actions =1961	3	0	112	329	828	689	0.2	0.0	5.7	16.8	42.2	35.1
Main rotor frequency adapter pair replaced	0	0	0	5	5	2	0.0	0.0	0.0	41.7	41.7	16.7
Main rotor replacement of blades.	0	0	3	7	15	16	0.0	0.0	7.3	17.1	36.6	39.0
Main rotor assembly refitted/replaced	0	0	2	3	19	14	0.0	0.0	5.3	7.9	50.0	36.8
Main rotor pitch rod replaced on a blade	0	0	0	0	11	3	0.0	0.0	0.0	0.0	78.6	21.4
Main rotor support struts refitted/replaced	0	0	1	2	6	2	0.0	0.0	9.1	18.2	54.5	18.2
Main rotor spindle replaced on a blade	0	0	0	2	2	3	0.0	0.0	0.0	28.6	28.6	42.9

As can be seen, cluster 3 did not portray main rotor frequency adaptor faults. Cluster 4 contained 50% of the frequency adaptor faults and a number of non-adjustable main rotor faults. Cluster 5 (38.5% of the samples) which was considered typical, contained the other 50% of the frequency adaptors' faults. The relative size of cluster 4 (13.9%) and the number of the faults of interest within the cluster (19 faults; 18.2% of the non-adjustable rotor faults) suggested that the clustering results were insignificant and could produce a large number of false alarms.

5.4.3 *The Refinement Process*

The unsupervised process can be refined by:

- **mathematical models:** a better understanding of the underlying laws of physics can bring to light new features and enhanced strategies that allow discrimination between typical and atypical samples and split the atypical samples in a robust way.

- **selected features:** a larger set of features or another set can provide better discrimination capabilities.

- **another cluster algorithm:** it is likely that one cluster algorithm can reveal aspects of a problem which can not be seen by another algorithm.

A frequency adapter fault can manifest itself through features emerging from direct comparison between high and low loading conditions; this is the fundamental hypothesis of the model upon which the analysis was based. Whilst a comprehensive mathematical model was used to validate this practical observation, it was not used to construct a better model or enhance the existing one. The favourable impact of the use of the laws of physics have been demonstrated in Sections 2 and 3 and therefore, this demonstration has concentrated on the other two methods of refinements.

5.4.3.1 Features' Selection

Three additional vibration amplitude ratios were introduced to characterise the 1R lateral and vertical vibration ratios between cruise and MPOG (R1LAF and R1VAF) and the 3R vertical vibration ratios between climb and descent (R3VAC). These ratios could be, to a great extent, insensitive to fault intensity but sensitive to fault types and, therefore, would entail better discrimination. The optimised – iterative – self-organising clustering analysis and the subsequent examination of the associated maintenance actions produced the following results:

Table 5.6 Unsupervised cluster analysis of the training data

Cluster number	No. of Samples	Number of associated maintenance actions	Frequency adaptors' faults	Main rotor non-adjustable faults	Main characteristics of standardised features (cluster centres)
1	1	0	0	0	Very high R1LAC
2	1	3	0	0	Very high R1LAC
3	1	29	1	13	Very high R1VAF
4	2	48	0	2	Very high R1VAF
5	6	0	0	0	Very high R3VAC
6	20	129	3	19	High R1VAF, R1LPC, D1, D2
7	26	69	1	17	Very high R1LAF
8	53	121	0	10	Moderate R1VAF and D1
9	80	169	2	17	High R3VAC
10	218	333	3	17	Very high D2 and high D1
11	385	526	3	28	Very high D1
12	565	742	3	43	Small values of D1 and D2

By considering the size of clusters and the degree of persistence of helicopter samples in a cluster, the significance of the unsupervised process can be assessed. Bearing this in mind, clusters 1 to 4 could be considered atypical outlying clusters because of their small number of samples that had very high vibration indicative of faulty sensors. Clusters 5 to 9 would have been regarded as atypical clusters that could relate to the faults under consideration in some sense. It turned out that cluster 5 did not relate to the main rotor non-adjustable faults and, cluster 8 did not indicate frequency adaptor faults. The total number of the frequency adaptor faults which were associated with clusters 5 to 9 (13.6% of the samples) was found to be 6. For clusters 10 and 11, the total number was found to be 6. The results of Table 5.6

suggested that further unsupervised analysis of the majority clusters will be required especially if the number of features is relatively large.

5.4.3.2 Supervised classification

Most of the cluster algorithms are sensitive to the initial statistics of the data and hence, the boundary of the clusters can be unstable. In other words, the addition of a small number of samples can significantly change the position of these boundaries if the statistics of the new samples are different to those of the initial samples. A technique such as the Adaptive Resonance Theory (ART) of Reference 7 which possesses relatively stable boundaries and produces new clusters, is expected to benefit the IDM system. Nevertheless, each cluster technique can only be optimised to highlight some aspects of the problem under consideration. Therefore, the inclusion of more than one algorithm is recommended. This will require the introduction of an intelligent interpreter that can manipulate the results of different algorithms and produce collective conclusions.

Whilst the resources of the current project did not allow for further investigation into the above issues, another supervised clustering algorithm was introduced for the purpose of the demonstration. The algorithm was driven by fuzzy rules. Two sets of rules were chosen based on experience. The first set implemented the following features, D1, D2, R1LAC, R1LPC, R2LPF, R3LPC and R3VPC (Note that in R1LAC, the second letter indicates the harmonic number, the third indicates Lateral or Vertical vibration, the fourth indicates Amplitude or Phase angle, the fifth indicates cruise-MPOG (F) or Climb-descent features). The second set implemented the following features: D1, D2, R1LAC, R1LPC, R2LPC, R2LPF , R3LPC and R3VPC. The statistics of the data were used to determine the feature values that could enable the rules to split the data into minority and majority clusters. The split criterion constructed a minority class having the least possible number of samples and the maximum possible number of frequency adapter maintenance actions. In this way, each supervised clustering analysis produced a minority cluster and a majority cluster. The statistics of the minority cluster obtained along with the associated maintenance actions were as follows:

Table 5.7 Supervised classification of the training data; minority cluster details

Atypical cluster produced by :	Set 1		Set 2	
		%		%
Number of helicopters in the cluster	11	48	8	35
Number of records (samples)	95	7	49	3.6
Number of associated maintenance actions	259	16	251	15
Main rotor frequency adapter pair replaced	5	50	5	50
Main rotor replacement of blades.	9	26	8	23
Main rotor assembly refitted/replaced	5	16	4	13
Main rotor pitch rod replaced on a blade	6	46	5	38
Main rotor support struts refitted/replaced	4	50	3	38
Main rotor spindle replaced on a blade	2	29	2	29

It is clear that the above results are more significant than those produced previously, the size of the atypical cluster is small (7% and 3.6%) and the concentration of fault cases is high (50% of the frequency adaptor faults and, 30% and 26% of the main rotor non-adjustable faults). Care must be taken, however, not to over-model and produce a very small cluster which is only significant for the training data. Over-modelling can be avoided by careful training and thorough understanding of the laws of physics.

5.4.4 *Analysis of the Test Data*

Having used the training data to refine the IDM process, the first set of rules was used to cluster the test data which covered the period from 1/9/94 to 6/11/1994 (the pre-processed test data of Section 5.3.3). The clustering analysis produced two clusters; a minority cluster and a majority cluster. The minority cluster was considered to be an atypical cluster. This cluster contained 34 samples (8.8%). The number of associated maintenance actions was found to be 75 (27%). The following table shows the atypical samples' detail.

Table 5.8 Classification of the test data, minority cluster details

Record	Tail-No	date	comments
1	GBLXR	4/10/1994	(1) There were no associated maintenance actions during this period, (2) There were four persistent samples (15 to 18).. (3) The last flight is contained in the atypical cluster Actions: **alarms** were triggered to the line engineer. Response on the 26/11/94: The helicopter is undergoing heavy maintenance.
2	GBLXR	4/10/1994	
13	GBLXR	29/10/1994	
15	GBLXR	1/11/1994	
16	GBLXR	2/11/1994	
17	GBLXR	4/11/1994	
18	GBLXR	5/11/1994	
133	GTIGE	2/ 9/1994	The maintenance actions included main rotor blade replacement and main rotor assembly maintenance. Actions: actions were not required
134	GTIGE	8/ 9/1994	
143	GTIGE	14/10/1994	
144	GTIGE	14/10/1994	
212	GTIGK	19/ 8/1994	
238	GTIGK	29/ 9/1994	The maintenance actions included main rotor blade replacement on 28/10/94
239	GTIGK	4/10/1994	
241	GTIGK	4/10/1994	
242	GTIGK	5/10/1994	Actions: actions were not required
247	GTIGK	17/10/1994	
248	GTIGK	18/10/1994	
353	GTIGT	25/ 9/1994	There had been three persistent samples (374, 375 and 376) after which the helicopter migrated to a majority cluster (two flights). Tail rotor maintenance was carried out on 4/11/94. Actions: alarms were not triggered. A query however was raised to whether other actions were carried out. Response on 26/11/94: The DAPU was changed on 3/11/94
355	GTIGT	26/ 9/1994	
356	GTIGT	26/ 9/1994	
371	GTIGT	17/10/1994	
372	GTIGT	17/10/1994	
374	GTIGT	28/10/1994	
375	GTIGT	1/11/1994	
376	GTIGT	2/11/1994	
399	GTIGU	30/ 9/1994	There were three persistent samples after which the helicopter migrated to the majority group (3 flights). Actions: actions were not required
417	GTIGU	27/10/1994	
418	GTIGU	28/10/1994	
419	GTIGU	29/10/1994	
443	GTIGW	28/ 9/1994	(1) There were four persistent samples. (2) The helicopter did not fly between 30/9/94 and 6/11/94 Maintenance actions were carried out on 30/9/1994 which included main rotor blade replacement and main rotor assembly maintenance. Action: alarms were not triggered.
444	GTIGW	28/ 9/1994	
445	GTIGW	29/ 9/1994	
446	GTIGW	30/ 9/1994	

As a consequence of the clustering analysis, an alarm was triggered regarding G-BLXR and a query was raised regarding G-TIGT (see the above table). The analysis suggested that G-BLXR had a main rotor non-adjustable fault, which was highly likely to be a frequency adapter fault. The analysis also reported a signature which had not been reported before. The signature indicated that the effect of the fault is equivalent to a difference in the flap-wise stiffness between the blue and red blade, which are opposite to each other, of about 200 DAN (DAN is the unit used by the aircraft manufacturer, ECF). The observations of the line engineer regarding the two helicopters substantiated, to some extent, the clustering analysis. G-BLXR was under heavy maintenance and the Data Acquisition and Processing Unit (DAPU) of G-TIGT was replaced.

The line engineer also indicated his concern regarding G-TIGO. The clustering analysis did not identify this as an atypical helicopter. In the period from 1/9/94 to 6/11/94, 36 maintenance actions were carried out on this helicopter, perhaps, to rectify the cause of the concern. These actions included the replacement of frequency adapters on 16/9/94 and 29/10/1994 along with main rotor head assembly maintenance. Nevertheless, the last valid flight record for this helicopter was on 24/9/1994; absence of a flight condition of interest (e.g., cruise or climb) invalidates the flight record. It seems that the search for the cause of the problem is still going on, but the lack of valid flight records after the 24/9/1994 prohibited MJAD from participating in this search.

It is worth mentioning that some of the maintenance actions could have been miss-targeted actions or routine maintenance. In other words, the maintenance actions can fail to rectify the symptoms that triggered the alarms or are carried out as scheduled maintenance. Therefore, 100% concentration of a maintenance action in a cluster is not expected. To this end, it is recommended that line engineers indicate the reason for the maintenance action. In the case of more than one maintenance action, it will be useful to indicate which maintenance action is considered to be targeted.

It is also recommended that features other than the vibration harmonics which have been extracted from signal averages be considered. This recommendation is based on an initial assessment involving another data set of 233 downloads of raw vibration measurements from 20 helicopters in cruise. Whilst the practical significance of analysing this small data set is low, the signatures which were seen in this data set are expected to enhance the visibility of the main rotor faults.

The above results demonstrate the performance of the IDM system by identifying its ability to recognise previously reported and unreported features in HUMS data, which may be used in order to diagnose mechanical defects. Nevertheless, it is important to emphasise that the above process is developed only for demonstration purposes. The realisation of a practical system for the main rotor non-adjustable faults requires a dedicated programme. In this programme, it is recommended that a large data set be used and to concentrate on the refinement issues, which include mathematical modelling, features' selection and the interpretation of various cluster algorithms.

5.5 Conclusions

Central to the intelligent management of a large data set are data pre-processing, unsupervised learning and supervised learning. A HUMS Intelligent Data

Management system (IDM) is therefore a framework that organises effective interactions between these three intelligent processes and, at each interactive stage, reports intermediate results.

Pre-processing is the mechanism that allows the use of generic tools to extract features from HUMS measurements so that the problem of interest is adequately portrayed. The implementation of the underlying laws of physics is an application dependent pre-processing task which can be carried out during an initial refinement stage. During the life time of the IDM system, pre-processing may be revisited at reasonable time intervals (one or two year period) and hence, the system must be flexible enough to be efficiently upgraded and re-certified. The upgrading process can be guided by practical observations, data-related knowledge and mathematical simulations.

Data-related knowledge is extracted from features by statistical pre-processors and a number of unsupervised algorithms. Each algorithm is targeted at a specific cluster type or a specific aspect of the problem. Unsupervised learning is the mechanism that can report pattern abnormalities even when the underlying cause is unknown. If samples persistently fall in atypical data clusters, inspection is triggered in order to identify the cause. Initially, the inspection is not expected to identify component faults. However, the inspection results can be used as a means to refine the data pre-processing methods in order to enhance the fault visibility and reduce the number of false alarms.

As a consequence of the interaction between the operator and the IDM system, a reasonable number of cause-effect examples can be generated. Each example consists of a detected fault along with the symptoms of the fault. By using these examples, a supervised learning process can be trained to identify faults given symptoms. The learning ability of the supervised learning process is dependent on the range covered by the examples and, the features that describe the symptoms. Again pre-processing may be revisited to re-establish the features that can offer sufficient fault discrimination.

Whilst the IDM framework must facilitate interactions between pre-processing, unsupervised learning and supervised learning, it is required to provide an initial refinement mechanism and a long term upgrading mechanism which aid the improvement of mathematical models, and seek for a representative set of features and/or enhanced core algorithms (supervised or unsupervised). Section 2 has concentrated on refinements using mathematical models. In this section, the impact of the clustering algorithms and feature's selection has been investigated. A refined process has been used to demonstrate the performance of the IDM system by identifying its ability to recognise previously reported and unreported features in HUMS data. The process has reported stiffness disparities which can be associated with main rotor non-adjustable faults; the main target has been frequency adapter faults. The process has also triggered an alarm concerning a helicopter and raised a query regarding another. The feedback observations from the line engineer have substantiated the analysis results. The first helicopter is currently undergoing heavy maintenance and the DAPU of the second helicopter was replaced.

In order to realise the full potential of the IDM system it is recommended that more than one cluster algorithm be included and, consider features other than the vibration harmonics which have been extracted from signal averages. Revising the measured sensitivities of the main rotor adjustable faults can have a favourable

impact on the IDM analysis. It is also recommended to state in the HUMS database the reasons that triggered a maintenance action and, in the case of more than one action, indicate which maintenance action is considered to be targeted.

6 CONCLUSIONS

There is much still to be gained from the patterns concealed in helicopter HUMS measurements. This report has demonstrated the feasibility and performance of a computer based Intelligent Data Management (IDM) process. The IDM process detected, by using unsupervised learning methods, abnormal patterns in the large volume of HUMS data even when the underlying cause was unknown. The process also assimilated the relationships between mechanical faults and abnormalities by using supervised learning methods. Not only mechanical faults induce abnormal data patterns but also other factors such as atypical operational conditions and equipment noise. These factors can trigger false alarms. The IDM process was able to discriminate between the vibration signatures induced by such factors and those induced by faults. In this way, the report has re-emphasised that central to the intelligent management of a large data set are data pre-processing, unsupervised learning and supervised learning. A HUMS IDM system should organise effective interactions between these three intelligent processes and, at each interactive stage, report intermediate results.

Associated with the success of detecting mechanical faults is the selection of appropriate measurements and how they are conditioned. Pre-processing is the mechanism that extracts features from HUMS measurements and conditions the features so that high visibility of abnormal patterns is achieved. Successful pre-processing should employ model-based processes that use the underlying laws of physics as portrayed by mathematical models. The demonstration presented in this report deduced non-linear functions of Flight Data Recorder (FDR) parameters and used them to mitigate operational influences on measured airframe vibration. Mathematical representation of the effects of metal wear on the composition of oil samples also led to pre-processing mechanisms that reduced noise effects across features and between oil samples. In general, it is unlikely that the initial data pre-processing functions would be optimum. In this work therefore, quantifiable checks were proposed in the form of quality curves that could indicate the effectiveness of the chosen pre-processing mechanisms. During the life time of an IDM system, pre-processing should be revisited and, if required, refined at reasonable time intervals (one or two year period). The IDM system should be therefore flexible enough to be efficiently upgraded and re-certificated. The upgrading process should be guided by practical observations, data-related knowledge and mathematical simulations.

Unsupervised learning searches for data clusters where features extracted from measurement samples within each cluster are similar to each other. Unsupervised learning is the mechanism that can report pattern abnormalities even when the underlying cause is unknown. This report has demonstrated that atypical clusters having a small number of measurement samples are diagnostically significant in that such measurements can be related to the development of mechanical faults. This work also suggested that if samples persistently fell into atypical clusters, inspection should be triggered in order to identify the cause. Initially, the inspection would not necessarily identify component faults. However, the inspection results should be used as a means to refine the data pre-processing methods in order to enhance the fault visibility and reduce the number of false alarms.

As a consequence of the interaction between the operator and an IDM system, a reasonable number of detected faults (causes) and associated measurements (effects) will be generated. By using these cause-effect examples, a supervised learning process can be trained to identify faults given symptoms. The supervised learning process draws more precise boundaries around each mechanical fault and can offer robust diagnosis. The process captures cause-effect relationships, even when the precise relationships cannot be explicitly expressed. The process has been demonstrated in this report by training systems to discriminate between three reported mechanical states comprising two faults and a health state. Testing the systems indicated robust diagnostic that produced very high fault detection rates and very low false alarm rates.

This report demonstrated the importance of system refinements, which should be based on realistic mathematical models, representative extracted features, enhanced cluster algorithms and engineering knowledge from in-service experience and HUMS data. Therefore, the IDM framework should facilitate interactions between pre-processing, unsupervised learning and supervised learning, and should include refinement and upgrading mechanisms.

It is important to appreciate that the IDM process of this report has been developed for demonstration purposes. A practical IDM system for main rotor non-adjustable faults should be based on the framework described in this report. The system should not only possess the benefits of the IDM processes, but the software should also be capable of accommodating the refinements without the need for re-designing or re-certifying the system. In order to maximise the benefits of the system, a dedicated programme should consider using a large data set and concentrate on the refinement issues through mathematical modelling, features selection and alternative supervised and unsupervised algorithms.

ACKNOWLEDGEMENTS

Acknowledgements are expressed to Mr David Howson and Mr Steve James for their interest in this investigation and their discussions and valuable comments. The author would like also to acknowledge the co-operation of Bristow Helicopters Limited.

REFERENCES

(1) Azzam, H and Andrew, M, A modular intelligent data administration approach for helicopter health and usage monitoring systems, Journal of Aerospace Engineering, Proc Instn Mech Engrs Vol. 209, IMechE 1995.

(2) Azzam, H and Callan, R E, Vibration and oil debris analysis for advanced helicopter health and usage monitoring systems. Journal of Condition Monitoring and Diagnostic Engineering Management, Volume 3 No 1 August 1992.

(3) Azzam, H, Investigation of single and twin rotor behaviour, Ph D thesis, University of Southampton, 1986.

(4) Rumelhart, D E, Maclelland, J L, Parallel Distributed Processing, The MIT Press, Vol.1 (1986)

(5) Azzam, H, A review of neural network applications and limitations with special reference to fault diagnosis, MJAD/R/120/93, January 1993.

(6) Azzam, H, The use of mathematical models and artificial intelligence techniques to improve HUMS prediction capabilities, Innovation in Rotorcraft Technology Proceedings, The Royal Aeronautical Society, 24-25 June 1997.

(7) Carpenter, G A and Grossberg, S, The ART of Adaptive pattern recognition by a self-organizing neural network, IEEE Computer, Vol.21, No.3, pp77-88, March 1988.

Appendix A: A math Model For SOAP Analysis

In this appendix, a math model is developed to highlight limitations of current analysis methods and establish reliable diagnostic techniques that can exploit the full potential of oil debris monitoring.

A.1 Elements' Concentration Levels

As mentioned previously, SOAP quantifies elements present in oil by spectrum analysis. Elements' concentration levels are measured in parts per million (ppm). One part per million is equivalent to one gram in a million millilitres (one milligram/litre). The features which are often used by the current monitoring systems are the measured concentration levels C_i ($i = 1, N_e$, where N_e is the number of elements present in oil samples). The diagnostic effectiveness of these raw features can be evaluated by analysing the following equation:

$$C_i(t) = C_{ci}(t) + C_{ni}(t) - \sum_j \left\{ C_i(t_j - \varepsilon)\ \theta(t - t_j) \right\} V_{ad)j} / V_{T)j} \qquad \text{(A.1)}$$

where:

- t — is the effective time at which the concentration level is measured. The word effective is used to emphasis that this time is equivalent to the actual wear time.
- $C_i(t)$ — is the concentration level of the i^{th} element measured at the time t.
- $C_{ci}(t)$ — is the continuous, deterministic function which can represent $C_i(t)$ fully if and only if there is no oil leakage, oil addition, contamination, equipment noise or measurement errors. Component wear results in elements deposited continuously in oil from a number of alloys and, hence, the concentration $C_{ci}(t)$ has in general an upward trend with time.
- $C_{ni}(t)$ — is the non-deterministic error (stochastic noise) of $C_i(t)$ which can be induced by factors such as contamination, equipment noise and measurement errors.
- $\theta(t - t_j)$ — is the Heaviside distribution;
- $\theta(t - t_j) = 0$ if $t < t_j$
- $\theta(t - t_j) = 1$ if $t \geq t_j$
- $t_j - \varepsilon$ — is the time just before oil addition.
- $V_{ad)j}$ — is the oil volume added at the time t_j.
- $V_{T)j}$ — is the oil total volume at the time t_j.

It is clear that the actual wear time relates to instantaneous rotational speeds of components as well as component loads, and does not necessarily match the flight hours. For example, a component running at load levels which are higher than those of similar components, is expected to wear faster than the other components. The evaluation of the time 't' requires extensive knowledge about operational conditions and is therefore, very difficult. Perhaps, SOAP analysis will be a more practical tool only if the wear time is assumed to be related directly to the flying (operational) hours. The assumption implies the introduction of an additional error (noise). This noise as well as the noise associated with equipment or concentration measurement errors is often characterised by a symmetric probability distribution; the error term fluctuates with time above and below a zero mean value such

that the positive and negative amplitudes are equally likely. Also, small amplitude values are more likely to take place, and the probability distribution is often independent of flying hours or sampling times. The average value of the error term associated with contamination depends on factors such as environment, age of components and maintenance procedures. For example, the silicon concentration level can be above the expected average for a vehicle operating in the desert. An additional noise of a Gaussian type may be caused by possible errors in evaluating volumes of added oil or concentration levels just before oil addition. In order to simplify the analysis, the error term $C_{ni}(t)$ will be considered to represent the combined effect of all noise sources described above.

A feature (e.g., a concentration level) will be regarded as an acceptable feature (input) for diagnostic systems if the following two conditions are satisfied:

- The feature can describe component wear modes.
- The noise associated with the feature is relatively insignificant.

It is evident that measured concentration levels do not necessarily satisfy these two conditions. The error term $C_{ni}(t)$ is not small and the concentration levels associated with normal wear modes can be higher than those associated with failure modes. This is attributed to the fact that the measured values of $C_i(t)$ depend on oil addition, oil leakage and accumulation time as well as the wear mode. Customarily, available SOAP databases, do not contain information that allows the evaluation of the third term of Equation (A.1) accurately. It is therefore concluded that systems based on raw concentration levels can not provide robust diagnosis and will be characterised by a high degree of false alarms and/or low detection rates.

Pre-processing of measured concentration levels is essential to realise a robust diagnostic system. The major purpose of pre-processing is the reduction of noise and the production of features that relate closely to wear modes.

A.2 Elements' Instantaneous Weights

Consider the weight of the ith element W_{pi} present in oil at the time t:

$$W_{pi}(t) = C_i(t)\, V_e(t)$$

$$V_e(t) = V_{ec}(t) + V_{en}(t) + \sum_j \left\{ V_{ad)\,j}(t-\varepsilon)\ \theta(t-t_j) \right\} \qquad \text{(A.2)}$$

where:
- $V_e(t)$ is the existing volume of oil at the time t.
- $V_{ec}(t)$ is the continuous, deterministic function which can represent $V_e(t)$ fully if and only if there is no oil addition or volume measurement errors. For example, if the oil is consumed (or leaking) at constant rate, then $V_{ec} = - m_{ve}\, t + k_{ve}$ where m_{ve} is the constant rate and k_{ve} is a constant.
- $V_{en}(t)$ is the non-deterministic error (stochastic noise) of $V_e(t)$ which can be induced by volume evaluation errors.

By combining the noise terms, Equations (A.1) and (A.2) give:

$$W_{pi}(t)=\left[C_{ci}-\sum_j\left\{C_{i)j}\ \theta_{tj}\ V_{ad)j}/V_{T)j}\right\}\right]\left[V_{ec}+\sum_j\left\{V_{ad)j}\ \theta_{tj}\right\}\right]+W_{pni}(t) \qquad \text{(A.3)}$$

W_{pni} is the combined noise terms. The weight W_{pi} present in oil at the time t is not an acceptable feature for diagnosis since the error term W_{pni} is not small and W_{pi} does not represent the total weight $W_{Ti}(t)$ shed from the oil wetted components.

A.3 Elements' Total Weights

The total weight $W_{Ti}(t)$ of the i^{th} element at the time t is:

$$W_{Ti}(t) = \int w_i\ dt = W_{pi}(t) + W_{si}(t) \qquad \text{(A.4)}$$

$$W_{si}(t) = \int C_i(t)\ (dV_s/dt)\ dt \qquad \text{(A.5)}$$

where:
- w_i is the rate at which the i^{th} element is shed from the oil wetted components. Generally, this rate is a function of time; it may change with age. However, from practical point of view, it can be assumed to be a function of the wear mode only.
- $W_{si}(t)$ is the instantaneous weight of the i^{th} element which is shed and lost with oil leakage.
- $V_s(t)$ is the instantaneous volume of the leaked oil.

The volume of the leaked oil can be expressed as follows:

$$V_s(t) = V_{sc}(t) + V_{sn}(t) = V_A - V_{ec}(t) - V_{en}(t) \qquad \text{(A.6)}$$

where:
- V_A is the total initial volume, $V_A = V_T$ at t=0.
- $V_{sc}(t)$ is the continuous, deterministic function which can represent $V_s(t)$ fully if and only if there is no volume measurement errors; $V_{sc} = V_A - V_{ec}$.
- $V_{sn}(t)$ is the error of $V_s(t)$; $V_{sn} = -V_{en}$.

Naturally occurring noise is usually characterised by bell shaped symmetric probability distributions which are known as the Normal distributions. Differentiation, subtraction and addition of signals can magnify the noise. Integration however reduces the random noise considerably. By combining the noise terms of Equations (A.3) to (A.5), the total weight $W_{Ti}(t)$ of the ith element becomes:

$$W_{Ti}(t)=[C_{ci}-\sum_j\{C_{i)j}\ \theta_{tj}\ V_{ad)j}/V_{T)j}\}][V_{ec}+\sum_j\left\{V_{ad)j}\ \theta_{tj}\right\}-$$
$$\int\left\{C_{ci}-\sum_j\{C_{i)j}\ \theta_{tj}\ V_{ad)j}/V_{T)j}\}][dV_{ec)j}/dt]\right\}dt+W_{Tni}(t) \qquad \text{(A.7)}$$

The standard deviation of the combined noise term $W_{Tni}(t)$ is expected to be high.

The values of $W_{Ti}(t)$ relate to the total wear of components and, if they exceed acceptable thresholds, can indicate that the serviceable lives of components are expired. However, an accurate evaluation of total weight values is not a straight forward task and further analysis is essential to attenuate the expected high level of noise. It is worth mentioning that in-service simplified formulae for total wear evaluation have been used and indicated partial success.

A.4 Fundamental Pre-processing Equations

A fundamental pre-processing equation is simply an equation that contains ***measured*** concentration levels and a parameter (or parameters) that indicates the wear mode of interest. The wear rates w_i are effected by the wear modes. Therefore, a fundamental pre-processing equation can be derived from Equation (A.7) by differentiation:

$$W_i=[C_{ci}-\sum_j \{C_{i)j}\,\theta_{tj}\,V_{ad)j}/V_{T)j}\}]'[V_{ec}+\sum_j \{V_{ad)j}\,\theta_{tj}\}]+W_{dni}$$
$$+[C_{ci}-\sum_j \{C_{i)j}\,\theta_{tj}\,V_{ad)j}/V_{T)j}\}][\sum_j \{V_{ad)j}\delta_{tj}\}] \qquad \text{(A.8)}$$

where: δ_{tj} is the Dirac Delta at the time t_j; $=\delta(t-tj)$.
W_{dni} is the noise contribution.

The first term of Equation (A.8) can be differentiated if the concentration levels are expressed in a rigorous form:

$$C_i(t) = C_{ci}(t) + C_{ni}(t) - \sum_j \{\int C_i(t-\varepsilon)\delta(t-t_j)V_{ad}/V_T\,dt\} \qquad \text{(A.9)}$$

In this case, it can be shown that the wear rate between oil addition is:

$$W_i = C_{ci}'\,[Vec+\sum_j \{V_{ad)j}\,\theta_{tj}\}+W_{dni}\,; \qquad t \text{ is not equal to } t_j \qquad \text{(A.10)}$$

where: W_{dni} is the noise term.

For an ordinary gearbox, the normal wear results in elements deposited continuously in oil from a number of alloys at low rates. The ratios between the elements of each alloy are constant and hence, the ratios between the wear rates are expected to be constant. As a gearbox defect grows, another number of components start to deposit elements in the oil and hence, the wear rates as well as the ratio between the wear rates are expected to change significantly.

Equation (A.8) can be further simplified by assuming that the wear rates during a particular wear mode are constants and that the oil leaks at a constant rate; $V_{ec} = V_T - m_{ve}\,t$. If this is the case, the concentration will increase with time in an exponential manner between oil addition:

$$C_i = C_{i)initial} - [w_i / m_{ve}] \ \ln[\ 1 - m_{ve}\, t/(V_T + \sum_j \{V_{ad})\theta tj\})\ ] \qquad \text{(A.11)}$$

At t=0, the concentration level is equal to an initial value. As the oil volume reduces, the concentration rapidly increases with time and approaches infinity at very small volume values. Typically, the volume of oil varies with time. In this case, Equation (A.7) rules out the measured concentration levels as acceptable diagnostic features even if the noise is absent. As m_{ve} approaches zero, C_i becomes:

$$C_i = C_{i)initial} - [w_i /(V_T + \sum_j \{V_{ad})_j \theta tj\})\ ] \qquad \text{(A.12)}$$

Equation (A.12) is only applicable between successive samples where t is relatively small and the amount of consumed (or lost) oil can be neglected (otherwise, the inclusion of V_{ad} in the above equation is physically meaningless).

Appendix B: Helicopter Math Model

The main objective of the mathematical model presented in this appendix is to establish effective vibration data pre-processing methods. The nbR airframe vibration originates from blade aerodynamics and inertia loads modified by structural dynamics of blades and fuselage (n is an integer, R refers to a frequency of one main rotor revolution per second and b is the number of blades). The blade loads at the radial position r are as follows:

$$M_f(r) = \int_r^R \left\{ \left(F_k - ma_k\right)_\rho (\rho - r) + \left(ma_i - F_i\right)\left(z_\rho - z_r\right) \right\} d\rho \tag{B.1}$$

$$M_d(r) = \int_r^R \left\{ \left(F_j - ma_j\right)_\rho (\rho - r) + \left(ma_i - F_i\right)\left(y_\rho - y_r\right) \right\} d\rho \tag{B.2}$$

$$M_t(r) = \int_r^R \left(M_{air} - M_{il}\right) d\rho \tag{B.3}$$

$$S_f(r) = \int_r^R \left\{ \left(F_k - ma_k\right)_\rho \right\} d\rho \tag{B.4}$$

$$S_d(r) = \int_r^R \left\{ \left(F_j - ma_j\right)_\rho \right\} d\rho \tag{B.5}$$

$$a_i = a_{bi} + a_{hi}, \quad a_j = a_{bj} + a_{hj}, \quad a_k = a_{bk} + a_{hk} + g \tag{B.6}$$

a_b	Acceleration of a blade particle, m/s^2.
a_h	Hub acceleration, m/s^2.
a	Net acceleration of a blade particle, m/s^2.
i , j and k	Hub rotating axes: i and j lie in the hub plane and k points upwards.
F_i	Aerodynamic load per unit span along the i axis, N/m.
F_j	Aerodynamic load per unit span along the j axis, N/m.
F_k	Aerodynamic load per unit span along the k axis, N/m.
g	Gravity acceleration, m/s^2.
m	Blade mass per unit length, k/m.
M_{air}	Pitching moment about the elastic axis, N-m.
M_d	Dragwise moment at r, N-m.
M_f	Flapwise moment at r, N-m.
M_{i1}	Inertial pitching moment, N-m.
M_t	Torsion moment at r, N-m.
S_d	Dragwise shear force, N.
S_f	Flapwise shear force, N.
ρ	Radial co-ordinate, m.
	Subscripts refer to either direction or a value of a function variable.

The blade element theory is implemented to evaluate the aerodynamic loads from information about local angles of attack along the blade, aerodynamic coefficients and the local velocity distribution. The angle of attack α is approximated by the following equation:

$$\alpha = \theta_o + x\theta_T - A_{1c}\cos\psi - B_{1c}\sin\psi - U_{lk1}/U_{lj1} \qquad \text{(B.7)}$$

θ_o	Collective pitch angle, radians.
θ_T	Blade twist angle at the tip, radians.
x	Radial co-ordinate/blade radius.
A_{1c}	Lateral cyclic pitch angle, radians.
B_{1c}	Longitudinal cyclic pitch angle, radians.
U_{lk1}	Normal velocity component, m/s.
U_{lj1}	Chordwise velocity component, m/s.
ψ	Blade azimuth angle, radians.

The aerodynamic coefficients are functions of the Mach number and the Reynolds number. In Reference 3, a simple model and wind tunnel measurements have been used to deduce an analytical expression for the lift curve slope a:

$$a = a_{1o} + V_T\left(a_{10} + a_{11}\mu^2 + a_{12}\mu\sin\psi\right)/a_s \qquad \text{(B.8)}$$

a_{lo}, a_{l1} and a_{l2}	are constants.
a_s	The speed of sound, m/s.
μ	= V/(ΩR); the advance ratio.
R	The blade radius, m.
V	The helicopter forward speed, m/s.
V_T	The blade tip speed = ΩR, m/s.
Ω	is the rotor rotational speed, radians.

For the purpose of this investigation, the lift curve slope will be assumed to consist of a constant term, first harmonic terms (sinψ and cosψ terms) as well as second harmonic terms (sin2ψ and cos2ψ). The drag coefficient will be represented by the following expression:

$$C_D = C_{Do} + \alpha^2 C_{D\alpha} \qquad \text{(B.9)}$$

α	is the angle of attack.
C_{Do} and $C_{D\alpha}$	are drag coefficients.

The motion expressed in terms of local axes fixed in the blade can be related to axes fixed in the rotating hub:

$$\begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix} = \begin{bmatrix} \cos\beta\ \cos\eta & \sin\eta & \sin\beta\ \cos\eta \\ -\cos\beta\ \sin\eta & \cos\eta & -\sin\beta\ \sin\eta \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} i \\ j \\ k \end{bmatrix} \qquad \text{(B.10)}$$

$$\begin{bmatrix} i \\ j \\ k \end{bmatrix} = \begin{bmatrix} \cos\beta\ \cos\eta & -\cos\beta\ \sin\eta & -\sin\beta \\ \sin\eta & \cos\eta & 0 \\ \sin\beta\ \cos\eta & -\sin\beta\ \sin\eta & \cos\beta \end{bmatrix} \begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix}$$

β is a rotation angle about the flapping axis.
η is a rotation angle about the drag axis.

Assuming the helicopter moves with velocity components U_x, U_y and U_z, rotates about its centre of gravity with angular velocity components p, q and r, the angular velocity **A** of the hub axes i, j and k becomes:

$$\mathbf{A} = \omega_x \mathbf{i} + \omega_y \mathbf{j} + \Omega \mathbf{k}$$
$$\omega_x = q\sin\psi - p\cos\psi \qquad \omega_y = q\cos\psi + p\sin\psi \tag{B.11}$$

The velocity of a point '**d**' (having x_1, y and z co-ordinates, $x_1 = x + eR$, e is the flapping hinge offset) relative to the hub axes i, j, and k becomes:

$$\mathbf{V_d} = \left\{\dot{x} + \left(\omega_y z - \Omega y\right)\right\}\mathbf{i} + \left\{\dot{y} + \left(\Omega x_1 - \omega_x z\right)\right\}\mathbf{j} + \left\{\dot{z} + \left(\omega_x y - \omega_y x_1\right)\right\}\mathbf{k} \tag{B.12}$$

The local axes i_1, j_1 and k_1 are defined by the local angles of the deflected blade β_l and η_l. Therefore, the velocity of 'd' relative to the local axes becomes:

$$\begin{aligned}
\mathbf{V_d} = \{ & (\dot{x} + \omega_y z - \Omega y) \quad \cos\beta_\ell \cos\eta_\ell + (\dot{y} + \Omega x_1 - \omega_x z)\sin\eta_\ell \\
& + (\dot{z} + \omega_x y - \omega_y x_1)\sin\beta_\ell \cos\eta_\ell\} \qquad \mathbf{i} \\
+ \{ & -(\dot{x} + \omega_y z - \Omega y) \quad \cos\beta_\ell \sin\eta_\ell + (\dot{y} + \Omega x_1 - \omega_x z)\cos\eta_\ell \\
& - (\dot{z} + \omega_x y - \omega_y x_1)\sin\beta_\ell \sin\eta_\ell\} \qquad \mathbf{j} \\
+ \{ & -(\dot{x} + \omega_y z - \Omega y) \quad \sin\beta_\ell + (\dot{z} + \omega_x y - \omega_y x_1)\cos\beta_\ell\} \qquad \mathbf{k}
\end{aligned} \tag{B.13}$$

The net components of the local velocity of the air at the point 'd' becomes:

$$\begin{aligned}
U_{\ell i1} = & \quad [V_{xt}\cos\psi - V_{yt}\sin\psi - \dot{x} - \omega_y z + \Omega y]\cos\beta_\ell \cos\eta_\ell \\
& - [V_{xt}\sin\psi + V_{yt}\cos\psi + \dot{y} + \Omega x_1 - \omega_x z]\sin\eta_\ell \\
& + [V_{zt} - \dot{z} - \omega_x y + \omega_y x_1]\sin\beta_\ell \cos\eta_\ell + v_{i1} \\
U_{\ell j1} = & \quad [V_{xt}\cos\psi - V_{yt}\sin\psi - \dot{x} - \omega_y z + \Omega y]\cos\beta_\ell \sin\eta_\ell \\
& + [V_{xt}\sin\psi + V_{yt}\cos\psi + \dot{y} + \Omega x_1 - \omega_x z]\cos\eta_\ell + \\
& + [V_{zt} - \dot{z} - \omega_x y + \omega_y x_1]\sin\beta_\ell \sin\eta_\ell + v_{i2}
\end{aligned}$$

$$U_{\ell k1} = -[V_{xt}\cos\psi - V_{yt}\sin\psi - \dot{x} - \omega_y z + \Omega\, y]\ \sin\beta_\ell + [V_{zt} - \dot{z} - \omega_x y + \omega_y x_1]\ \cos\beta_\ell - v_{i3} \quad \text{(B.14)}$$

U_x, U_y and U_z	are the helicopter velocity components relative to x, y and z.
p , q and r	are the helicopter angular velocities about its centre of gravity.
V_{xt}, V_{yt} and V_{zt}	are the net velocities that contain the effect of p, q and r.
v_{i1}, v_{i2}, v_{i3}	are induced velocities at the point 'd'.

Equation (B.14) is a general form which can be applied to steady and unsteady cases. Regarding the Super Puma MK I HUMS FDR parameters, p and q are not measured and the steady forward speed is only considered and hence Equation (B.13) can be simplified by considering wind axes x_n and y_n as follows:

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} k_1 & k_2 \\ -k_2 & k_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k_1 & -k_2 \\ k_2 & k_1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix}$$

$$k_1 = V_{xt}/V_h \qquad k_2 = V_{yt}/V_h \qquad V_h = \sqrt{V_{xt}^2 + V_{yt}^2} \quad \text{(B.15)}$$

$$-U_{\ell j1} = V_h \sin\psi + \dot{y} + \Omega x_1$$

$$U_{\ell k1} = -V_h \cos\beta_\ell + V_{zt} - \dot{z} - v_{i3}$$

The blade elastic deflections are used to evaluate the net acceleration of the blade particle. These deflections are obtained through the use of modal analysis. For example, the flapwise degrees of freedom can be estimated from the following equation:

$$\left(\int_0^R m g_n^2 dr\right)\left(\ddot{\beta}_n + \omega_f^2 \beta_n\right) = \int_0^R \left\{F_k - m a_k + m\ddot{z} + \left(m a_{fi} - F_i\right)\beta_1\right\} g_n dr$$

$$a_{fi} = a_i + \rho\Omega^2$$

g_i is the ith flapwise mode shape.

Throughout this investigation, the first flapwise mode shape is only considered. The Flapping angles are related to the control angles by simplifying and solving the analytical equations of Reference 3. The induced velocity in the vertical k direction is evaluated from the well known Glauert formula:

$$\lambda = \lambda_o\left(1 + kx\cos\psi\right)$$

λo	is assumed to be a function of torque.
k	is a function of the velocity state.
x	is a radial non-dimensional co-ordinate.

Considering the above equations and along with a fixed frame of reference, the 4R vibration can be approximated by the following equation:

$$
\begin{aligned}
V_{nR}=\rho(1-w^2)\,\Omega^2\,\{ & V_b^{\,2}(\phi_1+\phi_2\theta_o+\phi_3 A_{1c}+\phi_4 B_{1c}+\phi_5\mu_z\ + \\
& +\phi_6 A_{1c}^{\,2}+\phi_7 B_{1c}^{\,2}+\phi_8\theta_o A_{1c}+\phi_9\theta_o B_{1c}+\phi_{10}A_{1c}B_{1c} \\
& +\phi_{11}\mu_z A_{1c}+\phi_{12}\mu_z z6\ B_{1c})+(1-w)V_b(\psi_1+\psi_2\theta_o+\psi_3 A_{1c} \\
& +\psi_4 B_{1c}+\psi_5\mu_z+\psi_6 A_{1c}^{\,2}+\psi_7 B_{1c}^{\,2}+\psi_8 A_{1c}B_{1c}) \\
& +(1-w)^2(\vartheta_1+\vartheta_2\theta_o+\vartheta_3 A_{1c}+\vartheta_4 B_{1c}) \\
& +(1-w)(\chi_1+\chi_2 B_{1c}+\chi_3\mu_z) \\
& +V_b^{\,2}(\xi_1+\xi_2\theta_o+\xi_3 A_{1c}+\xi_4 B_{1c}+\xi_5 A_{1c}^{\,2}+\xi_6 B_{1c}^{\,2}+\xi_7 A_{1c}B_{1c}) \\
& +V_b^{\,4}(\eta_1+\eta_2\theta_o+\eta_3 A_{1c}+\eta_4 B_{1c}+\eta_5 A_{1c}^{\,2}+\eta_6 B_{1c}^{\,2}+\eta_7 A_{1c}B_{1c})\}
\end{aligned}
\qquad \text{(B.16)}
$$

μ_z	$= V_{zl}/(\Omega R) - \lambda o$
φ_i, ψ_i, ν_i,	are functions of helicopter configuration parameters, structural properties,
χ_i, η_i, ξ_i	Mach number and Reynolds number.

The parameters of the model can be reduced through Multi-Variate Regression analysis (e.g., Principle Component Analysis). The tail rotor parameters and the attitude angles can be used to assess the helicopter trim state. As more data becomes available, the above analysis can be revisited in order to establish the optimum pre-processing model.

Appendix C: The Neural Network Basic Building Block

The human nervous system contains hundreds of different types of neurons. The brain itself has a large number of neurons and communication paths of the order of 10^9 neurons and 10^{11} synapses. ANNs are made of a relatively small number of neurons and synapses. The ANN is a mathematical tool with a potential parallel hardware implementation. The tool can perform challenging computational tasks such as function approximation, classification, filtering, modelling and control.

The different ANN architectures can be constructed from the basic building block shown in Figure C.1.

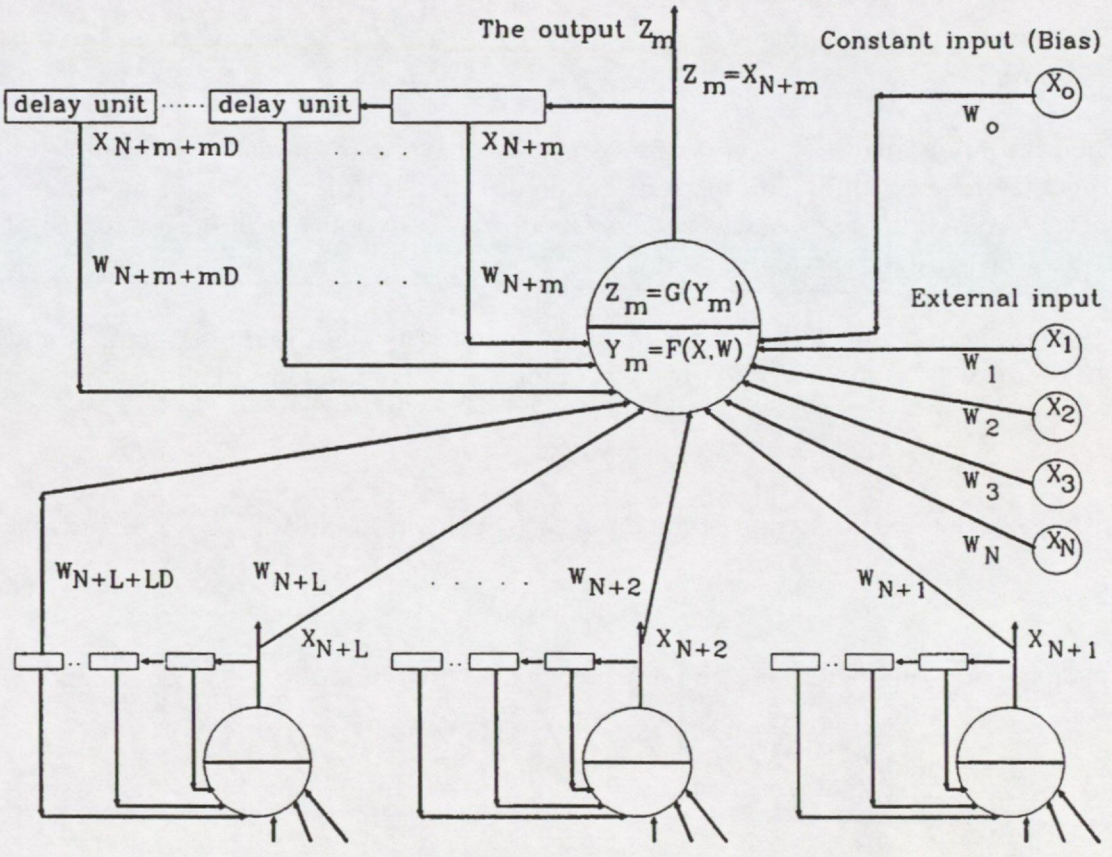


Fig.C.1 A basic building block for ANNs

Any network consists of 'L' neurons. Each neuron receives input values X_O, X_1, X_2, .., X_{N+L+DL} through synapses. The real numbers W_O, W_1, W_2, .., W_{N+L+DL} (the weights) simulate the synapses' strengths. Each neuron combines its input values and the associated weights using the function F to produce a real number Y. The neuron applies a non-linear function G to Y and produces the neuron output Z. As shown in Figure C.1, the m^{th} neuron represented by the central circle receives four types of inputs:

- ***N inputs (X_1, X_2, ..., X_N) from N external sources.***

For example, the inputs may be N numbers that represent concentration of elements in SOAP samples or amplitudes and phases of vibration signals.

- ***A constant internal input X_o (X_o=1) referred to as the bias.***

- ***The feedback from the m^{th} unit to itself.***
- Assume that the output of the m^{th} unit is a function of time and that the current output is a function of itself as well as D values of previous outputs. The feedback values to the m^{th} unit will be:

$X_{N+m} = Z_m(t)$,
$X_{N+m+(m-1)D+1} = Z_m(t-1)$, ...,
$X_{N+m+mD} = Z_m(t-D)$
The delay units shown in Figure C.1 store the previous values of the output Z_m.

- ***Input from other neurons.***

The output X_{N+i} of the neuron j (j = 1, 2, ..., L and j ≠ m) is passed to the m^{th} neuron. Also, the D delayed outputs $X_{N+j+(j-1)D+1}$, $X_{N+j+(j-1)D+2}$, ..., X_{N+j+jD} may be passed to the m^{th} neuron.

The neuron inputs may not necessarily comprise all types of inputs. For example, some neurons may receive inputs from external sources only. The networks that have feedback loops are referred to as *the dynamic or recurrent* networks as opposed to *the static* networks which are those without feedback.

Quite often, the function F determines the correlation between the input vector **X** and the weight vector **W** by combining the products of the vector components in a linear manner.

$$Y_m = X_O W_O + X_1 W_1 + ... + X_{N+L+K} W_{N+L+K} = \mathbf{W}^T \mathbf{X} \quad (C.1)$$

F may be chosen to be a function of the difference between **X** and **W**.

$$Ym = F\{(X_O - W_O),(X_1 - W_1), ... (X_{N+L+K} - W_{N+L+K})\} = F\{\mathbf{X}-\mathbf{W}\} \quad (C.2)$$

For example Y_m can determine the Euclidian similarity measure between **X** and **W** as follows:

$$Y_m = d_m^2 = (X_O - W_O)^2 + (X_1 - W_1)^2 + ... (X_{N+L+K} - W_{N+L+K})^2 \quad (C.3)$$

Thus, a small value of the distance d_m indicates that the input vector **X** is close and similar to the weight vector **W**.

Generally, the weights are evaluated during training sessions. The weights can be regarded as the knowledge that has been acquired and distributed over the network synapses.

Usually, the non-linear function G referred to as the activation function takes one of the following forms:

$$Z_m = 1 \quad \text{for } Y_m > 0.0$$
$$= 0 \quad \text{for } Y_m \leq 0.0 \quad (C.4)$$

$$Z_m = 1 / (1 + e^{-cY_m}) \quad (C.5)$$

$$Z_m = \tanh (Y_m) \tag{C.6}$$

$$Z_m = e^{- Y_m / (2\sigma^2)} \tag{C.7}$$

$$Z_m = Y_m \tag{C.8}$$

Equations (C.4) and (C.5) are referred to as the hard limiting non-linearity (or signum function) and the sigmoid function respectively. The sigmoid function is an S-shaped continuous, differentiable function that varies monotonically from 0 to 1 as its input Y_m varies from $-\infty$ to $+\infty$. The constant c determines the steepness of the transition region and is often chosen to be equal to unity. The tanh function is an S-shaped continuous, differentiable function that varies monotonically from -1 to 1 as its input Y_m varies from $-\infty$ to $+\infty$.

Equation (C.7) represents a Gausian differentiable function; if equation (C.3) is used to evaluate Y_m and the weights are regarded as a representative pattern, the closer an input $\mathbf{X}$ is to the representative pattern $\mathbf{W}$, the larger the response of the function will be. The Gausian function belongs to a class of functions called radial basis functions. In this case, the weight $\mathbf{W}$ and σ are referred to as the centre and the width of the function.

A number of ANNs use equation (C.8) for output neurons. A special case of neuron type is the input neuron which passes the input unchanged ($X_m = Y_m = Z_m$). Another special case is the high order neuron that combines its input in a non-linear manner, for example:

$$Y_m = F(\mathbf{X}) = X_1 X_2 \tag{C.9}$$

$$Z_m = Y_m \tag{C.10}$$

In this case, the neuron produces a new input by using a priori knowledge. This input is passed to non-linear neurons along with the rest of the inputs to yield significant results.

ANNs are made of layers of neurons. Often, the neurons of each layer can perform computational tasks independent of each other (i.e. in parallel to each other).

Study II

Intelligent Management of HUMS Data – The Use of Artificial Intelligence Techniques to Detect Main Rotor Gearbox Faults

Summary

This report details the results of using advanced Artificial Intelligence (AI) fault detection techniques to analyse two sets of seeded fault data from the Civil Aviation Authority (CAA) / GKN–Westland Helicopters (GKN–WHL) S61 Main Rotor GearBox (MRGB) seeded defect test programme. Vibration and Spectrometric Oil Analysis (SOA) data were analysed for both seeded defects, to enable consideration of data fusion aspects of HUMS data management.

The analysis primarily makes use of supervised and unsupervised AI analysis techniques. These were initially configured for 'blind' analysis of the first fault, optimised in the light of feedback, and then essentially 'frozen' for the analysis of the second seeded defect. In this way, the ability of these techniques to identify the existence and location of the new seeded defect without any prior knowledge, and despite having been optimised for the first defect, is demonstrated.

Data integrity checks, data modelling and trend detection methods were developed and refined during the analysis of the seeded defects, and resultant improvements in detection efficiency demonstrated.

The results from the analysis of the second fault build on the success achieved in the detection of the first fault, and show good potential for implementation as part of in-service helicopter Health and Usage Monitoring Systems (HUMS). The report makes recommendations to achieve this successful implementation.

Contents

List of Figures

List of tables

Glossary

AI	Artificial Intelligence
CAA	Civil Aviation Authority
FDR	Flight Data Recorder
GI	Gear Indices
GKN-WHL	GKN – Westland Helicopters Limited
HUMS	Health and Usage Monitoring System
IDM	Intelligent Data Management
IFE	Iterative Feature Extraction
MJAD	MJA Dynamics Limited
MRGB	Main Rotor Gearbox
PLATO	Pattern Learning Algorithm Toolkit
s.d.	standard deviation
SGDS	Smart Gear Diagnostic System
SOA	Spectrometric Oil Analysis
SQL	Structured Query Language
VIS	The SGDS Visibility GI (see Appendix A for details of other GI parameters)
1R	The amplitude of the vibration component at a gearshaft's rotational frequency
2R	The amplitude of the vibration component at twice the gearshaft's rotational frequency.

1 INTRODUCTION

1.1 Background

Helicopter Health and Usage Monitoring Systems (HUMS) produce huge quantities of data. The current generation of in-service diagnostics are generally efficient at recognising faults for which the HUMS has already been manually configured. However current systems are limited in their ability to detect previously unseen faults and lack the ability to automatically learn the characteristics of faults once they have occurred.

In addition, HUMS operators have to make decisions based upon interpretation of more than one type of data. For example, to diagnose a gearbox fault, consideration of both vibration and oil debris data may be appropriate. The problem of 'fusing' this data to reach a decision is made more difficult by differing rates of sampling, and fault-related patterns within the data.

MJA Dynamics have developed a range of Artificial Intelligence (AI) techniques aimed at removing the current HUMS limitations. This report details the application of these AI techniques to data from two seeded fault tests on an S61 Main Rotor GearBox (MRGB) to demonstrate their potential use in future HUMS.

MJA Dynamics' AI techniques employ two strategies for fault detection, namely supervised and unsupervised machine learning.

Unsupervised machine learning is a process which automatically identifies atypical behaviour, without any prior knowledge of the reason for this behaviour.

Supervised machine learning can establish relationships between causes (mechanical faults) and effects (measurements). Once a fault has been identified the data which are characteristic of that particular fault are used to define a fault area. The fault area can then be used to classify new data to establish if that particular fault exists.

Both of these learning strategies are incorporated into MJA Dynamics' Pattern Learning Algorithm Toolkit (PLATO), which was used to produce the majority of results contained in this report.

1.2 The S61 MRGB Test Rig

The seeded fault tests used WHL's closed loop back-to-back test rig. In this configuration the test S61 gearbox is connected to an identical slave gearbox. The output of the slave gearbox is then routed back to the input of the test box. By introducing a known quantity of wind up in this closed loop arrangement it is only necessary to supply enough power to overcome frictional losses, whilst operating the test gearbox under full load conditions.

The S61 MRGB contains two input pinions, a number of speed reducing/combiner gears and a single epicyclic stage. The signals produced by 11 accelerometers placed at various locations around the gearbox were recorded on magnetic tape. This data was analysed by WHL using MJA Dynamics' SMART gear diagnostic system (SGDS). This performs synchronous signal averaging of vibration data and computes a range of gear indices, each related to a specific aspect of the signal average. Groups of accelerometers were used to monitor individual gears. Table 1 shows which accelerometers were used to monitor which gears.

MJA Gear Reference	Monitoring Accelerometers
1	C, D, K
2	F, G, H, I
3	A, B, C, I, K
4	A, B, C, I, K
5	A, C, I, K
6	B, C, I, K
7	B, C, I
8	A, C, I
9	B, C, I
10	A, C, I
11	G, H, I
12	D, E, F, J
13	D, E, F, J
14	D, E, F, J

Table 1 Accelerometer groups for all gears

A defect was deliberately seeded in the gearbox before each test began. Its nature and position were not revealed to MJAD. The growth of this defect was monitored by WHL over the complete duration of each test.

1.3 Data Supply and Analysis Protocol

The data was supplied as a number of discrete samples, where each sample contained gear indices (GI) derived from the vibration signal using MJA Dynamics' Smart Gearbox Diagnostic System (SGDS), and elemental concentrations derived from Spectrometric Oil Analysis (SOA).

The GI and SOA data were supplied to MJAD for analysis. The time interval between the acquisition of each set was not revealed. MJAD were advised that the first batch of samples for each fault could be considered as healthy since the defect had not become significant during the first stages of the test. WHL believed the remaining samples contained detectable fault characteristics.

The test data were supplied in the following order:

1 Vibration and SOA data for the pre-initiation phase. This was used to characterise the gearbox's normal behaviour using unsupervised learning techniques.

2 Vibration and SOA data for the remainder of the test (the 'post-initiation' samples). This was analysed 'blind' using the analysis configuration resulting from the unsupervised analysis of the normal data.

3 Details of the fault and its development. This was used to permit 'supervised' analysis of the data, and to refine the pre-processing once the 'blind' analysis had been performed.

4 Additional 'unseen' vibration and SOA samples to permit further blind testing of the refined processing.

1.4 Aim

The aim of the work reported is to demonstrate the ability of AI based analyses to detect S61 MRGB gear faults from vibration and SOA data.

2 OVERVIEW OF THE AI TECHNIQUES EMPLOYED

2.1 Unsupervised Machine Learning

Unsupervised learning involves the identification of distinct patterns in the data without labelling information. A meaningful characteristic can often be attributed to a pattern after it has been discovered by unsupervised learning. For example, if the members of a pattern are found to be exclusively fault related, it is likely that the pattern itself is fault related. Patterns based on high occurrence rate and similarity of parameter values are called clusters.

In order to detect clusters in the data, a multi-axis graph is effectively generated with selected SGDS Gear Indices (GI) and/or SOA data forming each of the axes.

The first step in the analysis is to plot the pre-initiation samples onto the graph. The area of the graph can be termed 'vector space'. PLATO analyses the distribution of the samples on the graph looking for densely populated areas (referred to as '*clusters*' or '*groups*' in this report). Often clusters will be developed that comprise one particular gear type because of its unique behavioural characteristics. There may also be clusters containing a single point due to statistical spread, but the clusters which are of interest are those which contain a large number of samples from a particular gear. These clusters can be said to define the 'normal' or healthy behaviour of that gear. Sometimes clusters contain more than one gear type. The centre point of each cluster is calculated together with Standard Deviation (s.d.) along each axis.

The second step is to effectively plot the post-initiation samples onto the same graph on which the pre-initiation clusters have been defined. The distance from each data point to the centre of each pre-initiation cluster is calculated and an assessment made of the likelihood of it being part of the cluster based on the cluster's statistics. This process is termed '**Group Assignment**'.

2.1.1 *Fault Recognition Criteria*

The final step of the analysis is the interpretation of the distance values. If a gear develops a fault then this will affect one or more of the GI values. As the GI values change, the samples plotted on the graph will progressively move away from the pre-defined clusters. Typically one cluster is selected to represent 'normal' behaviour and movement considered relative to this.

Progressive faults initially manifest themselves as small changes in the parameter values of a healthy sample. As a fault develops, the difference between its parameter values and those of a typical healthy case will gradually increase. This will be seen in vector space as a progressive movement of the fault sample away from the healthy cluster(s).

In cases where there is considerable variability between data from different gears, selecting a single cluster to represent the behaviour of all gears in healthy state, may result in reduced sensitivity to faults. In general, where there is considerable variability in the data, the alternative approach of defining healthy clusters for individual gears (or groups of gears exhibiting similar behaviour) will provide increased sensitivity. The process of recognising clusters in the healthy data which were predominantly composed of data from individual gears could be automated. This process would be facilitated using larger quantities of data than were available for the current investigations. The individual data points would then need to be group assigned to the relevant healthy cluster(s).

There are a range of different possible group assignment strategies. Each strategy returns a set of values which correspond to the likelihood of a sample belonging to each of the clusters. The likelihood of a sample belonging to a cluster depends on two things:

- The Euclidean distance between the sample and the cluster centre.
- The standard deviation of the cluster (i.e. its spread).

Each strategy aims to combine these factors to give the most reliable estimation. Issues of noise and outlying data mean that the best strategy to use will depend on the nature of the data.

2.1.2 *Strategy for Analysis of the S61 MRGB Seeded Fault Data*

The pre-initiation data for the first seeded fault was analysed to establish the normal cluster(s). A range of different group assignment strategies was then investigated for categorisation of the post-initiation data. From feedback on the unsupervised blind analysis' success in finding the fault, an optimum analysis strategy was determined.

Once the optimum strategy was determined for the first fault. The same analysis strategy was applied to the second fault in order to assess its general applicability. This assessment relates to both the performance of the strategy, and the degree to which the normal cluster(s) apply between builds of the same gearbox.

2.2 Supervised Machine Learning

Supervised learning techniques work by detecting features which distinguish one data set from another. In the case of fault detection, the goal is to distinguish a data set relating to a gear fault condition from the healthy data set. For this trial, all post-initiation samples were considered to be faulty while the pre-initiation samples were all considered to be healthy.

Various techniques exist for detecting features which distinguish one data set from another. Most rely on the principle of 'plotting' samples in multi-dimensional vector space. Each dimension corresponds to one of the parameters which is associated with a data sample (i.e. GI values, or SOA element concentrations).

All data samples are plotted in vector space and labelled as either healthy or faulty. The goal of the learning algorithm is to identify areas of the vector space which exclusively and repeatedly contain all the samples from one data set. Subsequent data falling into this area are likely to belong to the same data set. If an area of vector space can be found which exclusively and repeatedly contains faulty data, new samples falling into it are likely to be faulty.

2.2.1 *Iterative Feature Extraction (IFE) Technique*

Techniques vary in the way vector areas are identified and bounded. Iterative feature extraction (IFE) identifies fault areas iteratively and encloses them with linear boundaries.

Identification is achieved by iterative outlier exclusion. At the first iteration, linear boundaries are chosen for each dimension such that all fault points are just enclosed within an area of vector space. Two boundary values are chosen for each dimension, effectively an upper and lower limit. Considering one dimension (or parameter) of all the fault samples, the minimum value is used as the lower limit while the maximum value is used as the upper limit. This process is repeated for each dimension in turn until all have an upper and lower limit set which just encloses the fault data.

At this stage it is possible that some healthy data may have been included in the selected area which, ideally, should not contain any. This is likely if a fault data sample has an unusually high or low value compared with the rest of the fault data. By moving one of the linear boundaries *inwards* it is possible to eliminate one or more of the mis-classified healthy samples and any outlying faults. In a similar fashion to a computer chess program, the algorithm tries all the possible boundary moves, and then chooses the best one. The best move is defined as the one which minimises the number of healthy cases contained in the area while maximising the number of faulty cases remaining.

After each move is made, the number of healthy samples still remaining in the reduced area is counted. If this is greater than zero, that is to say there are still healthy cases which have been mis-classified as faulty cases, the algorithm chooses and makes another boundary move. This makes the area enclosing the fault cases progressively smaller and smaller until no healthy cases are mis-classified. At this point the algorithm stops adjusting the boundaries and reports the number of fault cases still remaining in the derived fault area.

If this number is high compared with the total number of fault samples in the whole data set, then the fault-related data set is highly separable and the boundaries identified should be capable of correctly classifying new samples presented to the system. If however this number is low, then the chosen parameters are not capable of discriminating healthy and fault-related data and are therefore poor fault indicators.

Before the first iteration the space which just encloses the fault data always gives 100% fault detection. All fault cases in the training data are contained within this enclosure and will therefore always be classified as fault related. At this stage however, there are likely to be healthy samples which are also contained within the enclosure. These have been mis-classified and will constitute false alarms.

For example, consider a data set which has 12 samples, 7 of which are faulty. Before the first iteration all of the faulty samples are guaranteed to be in the fault enclosure. Unfortunately 2 healthy samples have also been caught in the enclosure. This means that all 7 fault samples will be spotted but there will also be 2 false alarms.

After the last iteration the fault area will not contain any healthy samples. This means there will be no false alarms whatsoever. The level of fault detection will, however, have fallen. At this stage there will be faults which lie outside the enclosure. These will be classified as healthy although they are in fact faulty.

Continuing the example, after the last iteration there might be 4 samples left in the enclosure all of which are guaranteed to be faulty. This means that no false alarms will be generated. Unfortunately, the 3 faults which had to be removed in order to eliminate the false alarms will now be classified as healthy.

As the algorithm iterates, the rate of false alarms gradually drops as the rate of missed faults increases. Since the algorithm records all of its moves, a compromise can be chosen. The level at which the compromise is set is determined by the operator, and depends only on the relative cost of false alarms verses the cost of missed faults.

In comparison to other approaches such as neural networks, the IFE technique is ideally suited to cases where comparatively small amounts of training data are available. The IFE training times are also considerably faster, and it provides a clear identification of how it has learnt to recognise the fault-related patterns.

3 UNSUPERVISED BLIND ANALYSIS OF FAULT 1 DATA

3.1 Analysis of Pre-Initiation Vibration Samples

The first step in the unsupervised blind analysis was to analyse the pre-initiation data from the first seeded defect to identify the normal clusters.

3.1.1 *Data Integrity Checking of Vibration Data*

The initial pre-initiation samples were considered to be representative of a healthy gearbox. None of the clusters discovered in this data would therefore be expected to be fault related. The VIS parameter was not calculated for certain gears due to the fact that phase effects often cause cancellation of mesh frequency components for these components. It was therefore necessary to analyse these gears separately to those for which VIS was calculated.

Initial analysis of all the pre-initiation data (i.e. data from both 'VIS' and 'non-VIS' gears) for the first seeded fault, produced unexpected results which contained several outlying data points. Inspection of the vibration parameters for these samples suggested that an erroneous acquisition had occurred. Another cluster analysis was performed on a subset of the data for those gears for which VIS was calculated, with these outliers removed. This produced the results which can be seen in the composition table, Table 2. For each gear-sensor combination (e.g. where 'GEA4_A' indicates gear 4 monitored from sensor location A), this shows how many of the 5 normal data samples for each gear-sensor combination fell into each of the groups. The composition table shows 13 distinct groups in the data. An ideal result would show each gear consistently appearing in the same group. This would indicate that the parameters had a consistent pattern which PLATO was able to identify as 'normal'. At this stage however, there were no distinct groupings for some gears, particularly gears 10 & 11.

Closer inspection of the signal averages for gears 10 and 11 revealed that the tachometer signal had been acquired incorrectly for these gears. This explained the unexpected results for these gears shown by the unsupervised gear analysis. Aware that greater confidence would be achieved using the complete data set without having to selectively remove outlying datapoints, a request was made for re-acquisition of selected data using a modified acquisition set-up. The remaining outlying points were eliminated by customising the SGDS analysis configuration for the S61 MRGB, only 6 out of the total 52 gear sensor combinations required changes. Such modification of the analysis configuration to suit a particular gearbox application is fairly common once an initial dataset has been analysed. It does not represent customisation to suit the particular seeded fault.

Following customisation of the SGDS configuration and appropriate data reacquisition, the pre-initiation data for the first seeded defect (i.e. samples 1 to 5) was split into three categories and analysed appropriately. The three data categories were as follows :

- Vibration data from gears exhibiting characteristics which allowed visibility ('VIS') to be measured.
- Vibration data from gears for which 'VIS' was not appropriate.
- Spectrometric Oil Analysis ('SOA') data produced by analysis of gearbox oil samples.

Group:	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13
Gear/Sensor:													
GEA14_D	3										1		1
GEA14_E	1										4		
GEA14_F	5												
GEA14_J	5												
GEA01_C	4												
GEA01_D	2									2		1	
GEA01_K	4	1											
GEA02_F									1				4
GEA02_G	2					1							1
GEA02_H	1								1			1	2
GEA02_I									1				4
GEA03_A	2					1				1			1
GEA03_B	5												
GEA03_C	3									2			
GEA03_I	4									1			
GEA03_K	2					1				2			
GEA04_A										1			4
GEA04_B												4	1
GEA04_C									1			1	3
GEA04_I									1			2	2
GEA04_K					1					1		1	2
GEA05_B												2	3
GEA05_C		1										2	2
GEA05_I	2						1					2	
GEA05_K				1						1		2	1
GEA06_A				2			1					2	
GEA06_C									2			1	2
GEA06_I	1									1		3	
GEA06_K										1		3	1
GEA07_B												1	4
GEA07_C	1									1		3	
GEA07_I	1					1						2	1
GEA08_A	4					1							
GEA08_C												5	
GEA08_I	2					1				1		1	
GEA09_B												1	4
GEA09_C		1	1							3			
GEA09_I									1	1		1	2
GEA10_A								1	1	1			1
GEA10_C										2			3
GEA10_I		1								2			1
GEA11_F										3		1	1
GEA11_G	5												
GEA11_H	1					3				1			1

Table 2 Composition table of initial cluster analysis

Group:	G1	G2	G3	G4	G5	G6	G7
Gear/Sensor:							
GEA14_D					1	1	3
GEA14_E					3		2
GEA14_F							5
GEA14_J							5
GEA01_C					1		4
GEA01_D					1		4
GEA01_K					2		3
GEA02_F		1		1		3	
GEA02_G					1	1	3
GEA02_H			1	1		2	1
GEA02_I		1		1		3	
GEA03_A						1	4
GEA03_B							5
GEA03_C						1	4
GEA03_I							5
GEA03_K					1		4
GEA04_A				1		4	
GEA04_B						5	
GEA04_C		1				4	
GEA04_I				1		4	
GEA04_K				1		4	
GEA05_B						5	
GEA05_C	1			1		3	
GEA05_I	1					2	2
GEA05_K	2					3	
GEA06_A	4				1		
GEA06_C	1			2		2	
GEA06_I						4	1
GEA06_K	2			2		1	
GEA07_B						5	
GEA07_C	2			1		1	1
GEA07_I	1					3	1
GEA08_A	1						4
GEA08_C	1					4	
GEA08_I	1				1	1	2
GEA09_B				1		4	
GEA09_C						1	4
GEA09_I						2	3
GEA10_A							5
GEA10_C							5
GEA10_I						1	4
GEA11_F	2					3	
GEA11_G							5
GEA11_H	1			2	1	1	

Table 3 Composition table of final cluster analysis

The following sections describe the analysis of the vibration data, the SOA data is considered separately (Section 3.4).

3.1.2 *Gears with 'VIS'*

The distinction between clusters appeared to be related to gear position and the monitoring sensors. Seven clusters were observed. A composition table is shown in Table 3.

This table represents all the data which was available, given the removal of outliers and repeat acquisitions.

Gears 3, 4 and 10 exhibited the most consistent behaviour, regardless of sensor location. Gears 3 and 10 fell into cluster 7, while gear 4 fell into cluster 6. Other gears had samples in several clusters. The majority of the data fell into cluster 7. Table 4 shows the cluster populations and their characteristics.

Cluster Number	Population	Characteristics
1	20	Low – none High – AMIK, IMP, AMIE
2	3	Low – VIS High – EIIV, EIIE
3	1	Low – IMP, STB, VIS High – WEA, EIIV, EIIE
4	15	Low – STB High – none
5	13	Low – WEA, AMIK High – STB
6	79	Low – none High – AMIE
7	89	Low – AMIK High – VIS

Table 4 Cluster Populations and Characteristics

3.1.3 *Gears without 'VIS'*

Initially the same cluster analysis was performed for those gears for which VIS is not computed as was used with the other gears. A large number of small clusters with little significance were produced. This suggested that the iterative process used by PLATO had formed potentially significant clusters, but then sub-divided these into smaller and less meaningful ones. This phenomenon tends to occur when a non-representative sample of a statistical population is analysed in isolation. Generally this situation may be rectified by adding data to the sample in order to make it more representative of the whole population. Often this can be achieved by the analysis of increased quantities of data. In a production system, it would be possible to automate the process of iteratively using different selections of additional data, until adequately stable and meaningful clusters were identified.

For the current analysis, there was no more data available from gears without VIS which related to the pre-fault initiation state. For this reason, vibration samples from three gears with 'VIS' (gears 2, 6 and 7) were included in the data. These gears were chosen since they approximately doubled the number of samples to be analysed, and also provided a reasonable spread of data. This increased the diversity of the data presented to PLATO without affecting the distribution of the original data. This

had the required effect of encouraging the clustering algorithm to form clusters of a useful size. This approach produced 12 clusters.

Clusters 3, 10, 11 and 12 picked out the original non-VIS gear data. The remaining clusters contained the data which had been deliberately added. Table 5 shows the cluster populations.

Cluster Number	Population	Characteristics
1	4	Low – none High – IMP, AMIE, AMIK
2	1	Low – none High – AMIK
3	1	Low – Imp High – AMIK
4	5	Low – STB High – EIIV, EIIE
5	2	Low – WEA High – IMP, EIIV, EIIE
6	2	Low – WEA, AMIE High – IMP, EIIV, EIIE
7	3	Low – AMIE, WEA High – none
8	1	Low – IMP, STB High – EIIE, EIIV, WEA
9	1	Low – AMIE High – EIIE, EIIV
10	32	Low – none High – none
11	11	Low – none High – none
12	18	Low – none High – none

Table 5 Cluster populations and characteristics for 'non-VIS' gears

3.2 Analysis of Fault Related Samples

The post-initiation fault-related data samples 6 to 12 were compared with the distribution of healthy vibration data derived from samples 1 to 5. It was anticipated that faults in the gearbox would cause atypical patterns in the relative values of the signal average parameters. If this happened, fault cases would be immediately apparent since they would not fall into any of the healthy clusters.

3.2.1 *Standard Deviation Assignment Technique*

In the composition tables shown in Tables 2 and 3, a sample was allocated to the cluster which had the nearest centre. This assignment technique is termed 'nearest neighbour' assignment. Standard deviation assignment is a slight variation on this approach. Instead of returning the raw Euclidean distance of each cluster, this strategy returns the membership likelihood of each data point. The normalisation takes account of the differences in the standard deviation of each cluster.

Consider the situation in which a tightly packed cluster called 'A' occurs next to a sparse cluster called 'B'. The standard deviation of the clusters will be very small and very large respectively. Because of this, the membership radius (or size), of 'A' will be much smaller than that of 'B'. Data points occurring just inside the edge of the larger cluster 'B' could in fact be closer to 'A' while still belonging to 'B'.

The normalisation process involves dividing the Euclidean distance between the data point and the cluster by the standard deviation of the cluster. This is then mapped on to an approximation of a Gaussian distribution to give a membership likelihood value. Figure 1 shows the behaviour of the output of this assignment technique.

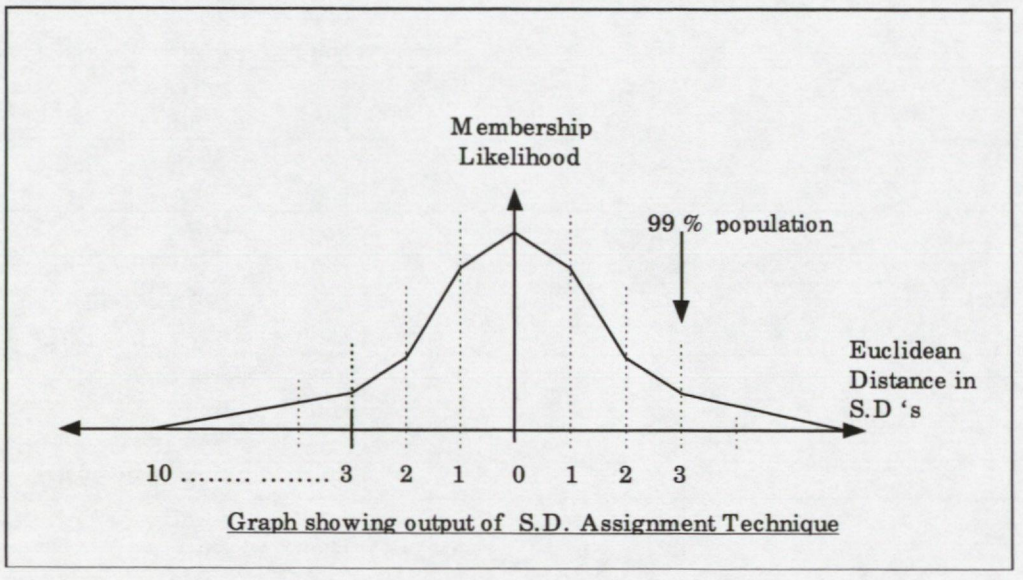


Figure 1 Output of Standard Deviation Assignment technique

The advantage of this strategy is that it takes account of the standard deviation of clusters. This provides a more accurate indicator of the cluster to which a data point belongs.

The disadvantage of this strategy is that it cannot cope with clusters having only one data point. This is due to the unavailability of a standard deviation value for the cluster. Furthermore, this strategy does not work well with data points which are a long way from a cluster. It returns zero for all data points greater than 10 standard deviations from the centre of the cluster. As a result it is impossible to differentiate between a data point which is 10 standard deviations from a cluster and one which is 100 standard deviations away. Figure 2a/b shows the assignment values produced for one gear-sensor combination to the seven cluster centres (C1 to C7) identified in the unsupervised analysis results (presented in Table 3). Figure 2 a/b illustrates the difficulty in assessing trends away from the cluster centres.

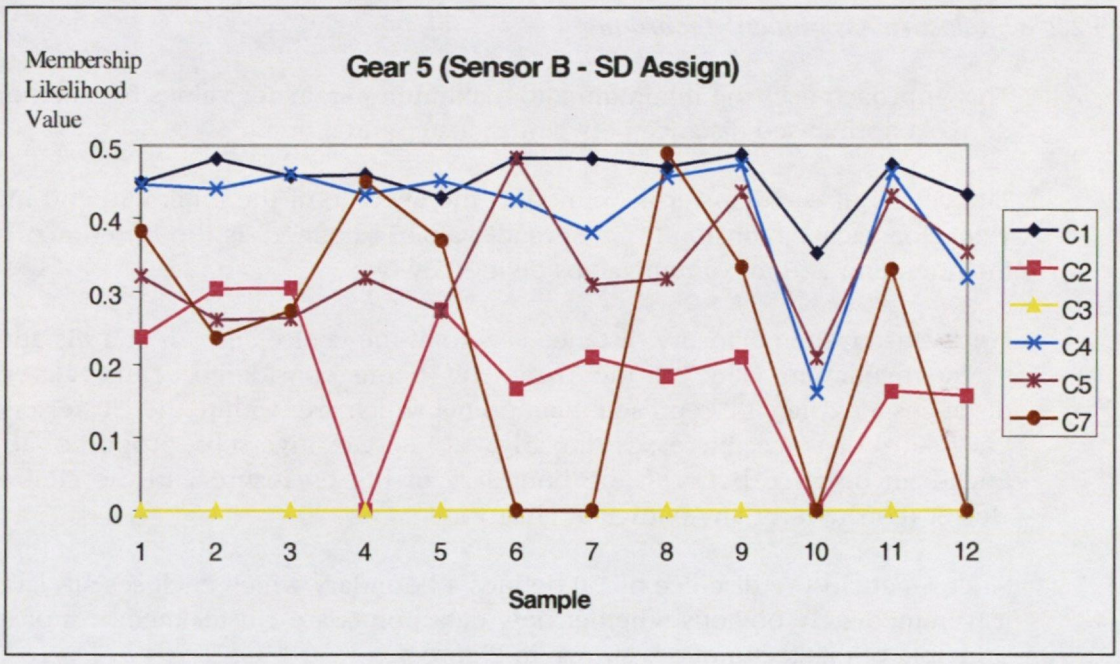


Figure 2 Results of the Standard Deviation Assignment technique

(a) standard scaling. (C1 to C7 are clusters identified from unsupervised analysis – see Table 3)

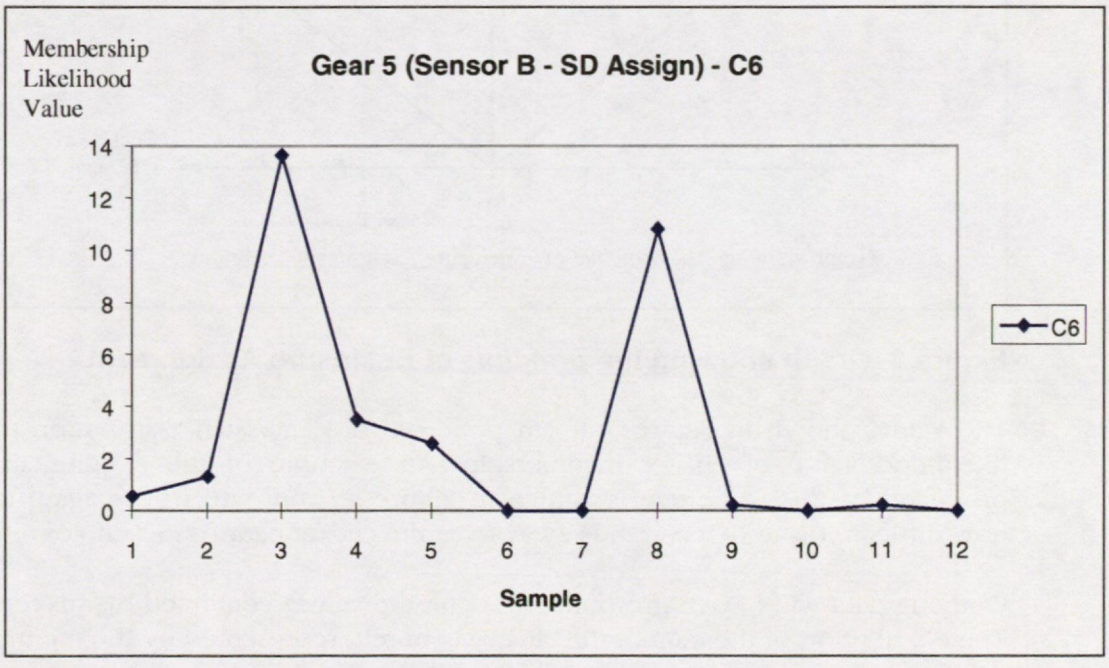


Figure 2 Results of the Standard Deviation Assignment technique

(b) Re-scaled for C6 (C6 is the cluster identified from unsupervised analysis – see Table 3)

3.2.2 *Enclosure Assignment Technique*

This approach uses the minimum and maximum parameter values for each cluster to derive a normalised distance between a cluster and a point.

The centre of each cluster is defined as the average of the minimum and maximum value for each dimension. The 'spread' value is defined as the difference between the minimum and maximum values divided by two.

For a cluster, the enclosure distance is zero at the centre and 1.0 at a distance equal to the maximum value of the parameter being considered. Thus, all enclosure distances less than 1.0 represent data points which are 'within' the cluster enclosure. Enclosure distance increases linearly, at a rate inversely proportional to the Euclidean distance between the boundary of the enclosure and the centre of the cluster (constant for any particular cluster).

Since an enclosure distance of 1.0 defines a boundary which encloses all data points, it is immediately obvious whether new data points are cluster members or not. The principle of this technique is shown in Figure 3.

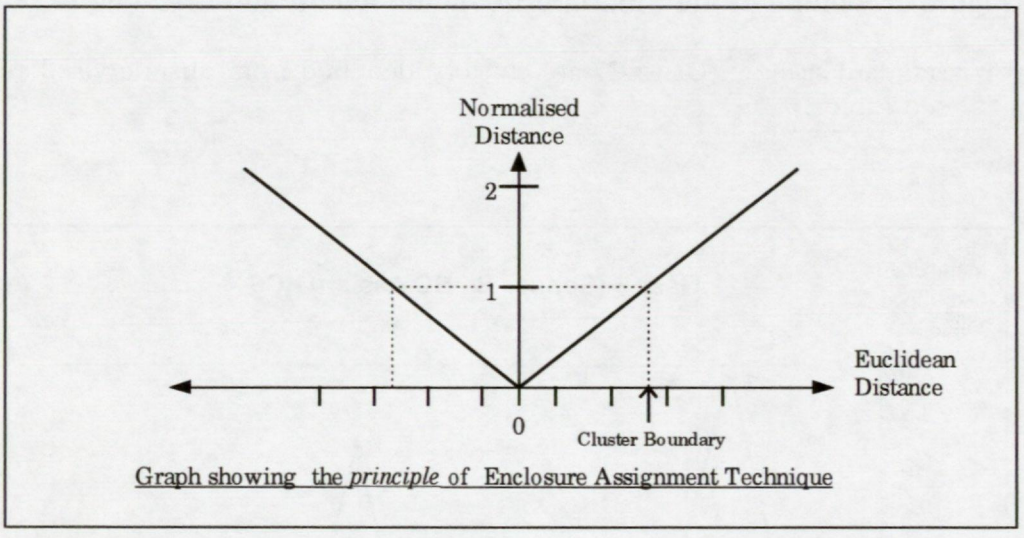


Figure 3 Graph showing the *principle* of Enclosure Assignment

The values shown in Figure 3 are mapped on to a Gaussian distribution to give a likelihood value of cluster membership. An example of the assignment results produced by this technique is shown in Figure 4, from which it is again apparent how difficult it is to detect trends away from the cluster centres in samples 6 to 12.

Outlying data will have a profound effect on the values generated by this technique. The distribution of the data in the cluster is purely represented by the minimum and maximum values. This can be mis-leading if the data in the cluster is tightly packed around the centre with one or two values at the boundary.

A cluster containing a singleton point will cause a division by zero error since the difference between the maximum and minimum is zero. (This value is used to normalise the distance of a point from the centre of the cluster).

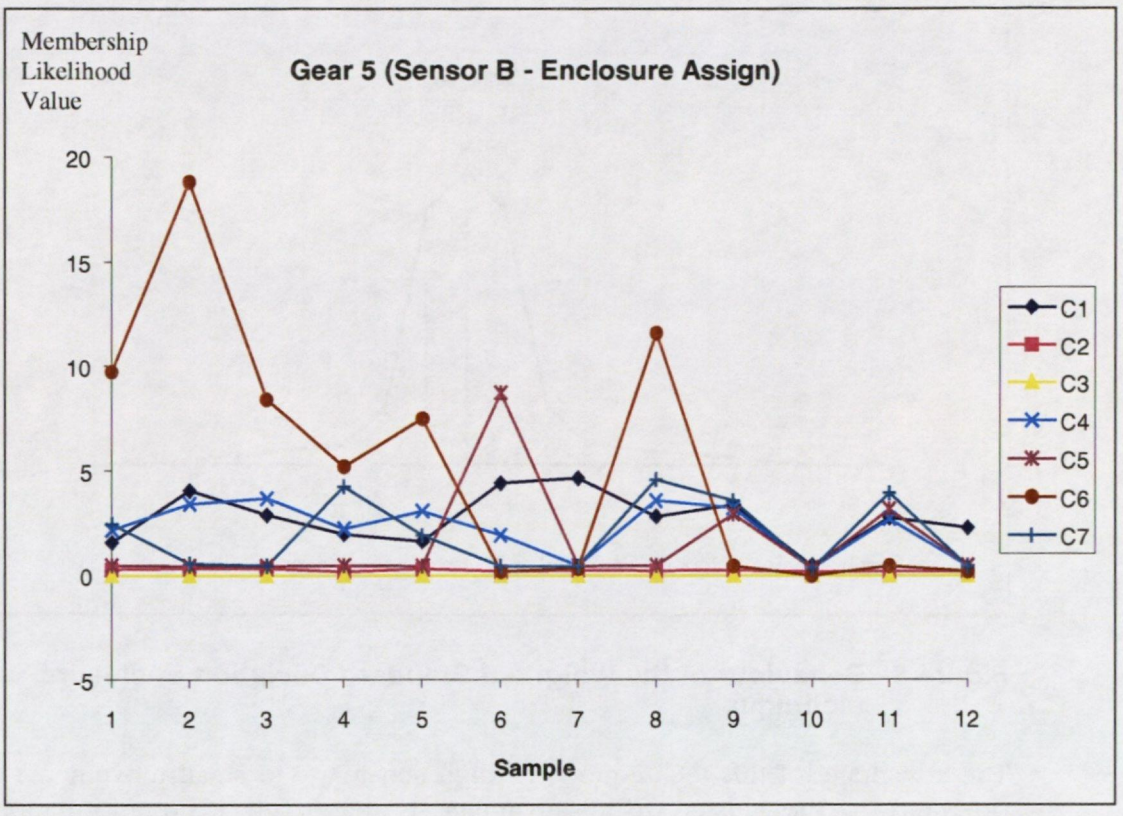


Figure 4 Graph showing results produced by the Enclosure Assignment technique

(C1 to C7 are the clusters identified in the unsupervised analysis results – see Table 3)

3.2.3 *Weighted StDev. Enclosure Assignment Technique*

This assignment technique is an improved version of the enclosure assignment technique. It is weighted by the value the standard deviation technique would produce given the same data point.

The weighting is done by multiplying the values from both techniques together and converting the result into a likelihood value. The affect of this is to increase the rate at which data moves away from a cluster. The behaviour graph is therefore sharper than the other two techniques. This is shown in Figure 5.

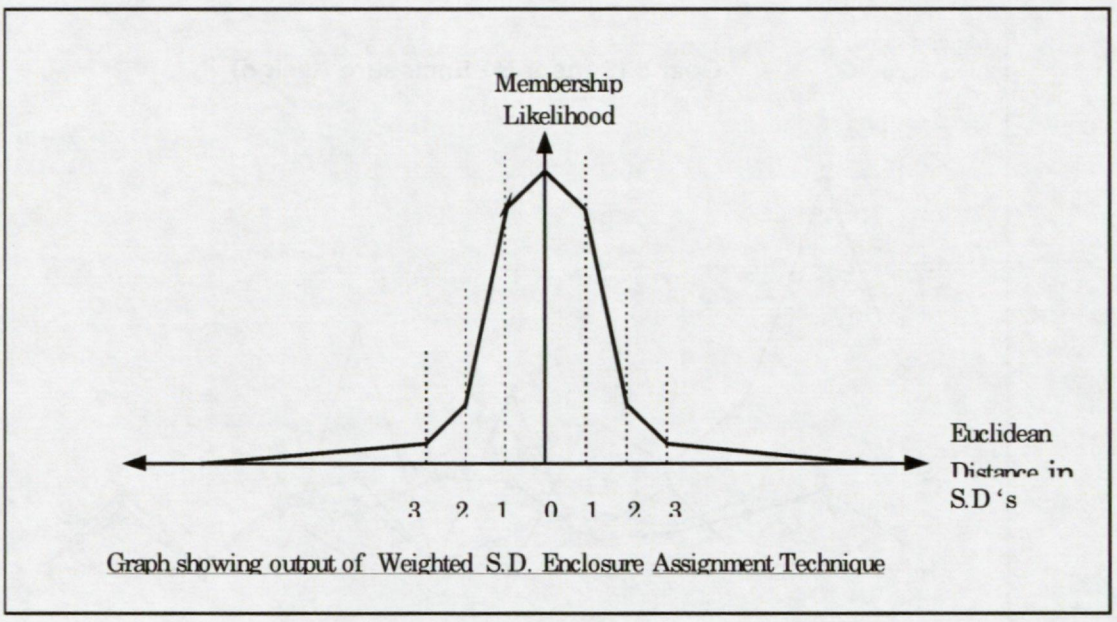


Figure 5 Behaviour of the Weighted Standard Deviation Enclosure technique

The advantage of this technique is its high sensitivity to small movements in close proximity to the cluster. Its disadvantage is that it gives no information about movements just outside the cluster boundaries. The results of this technique are shown in Figure 6 a/b.

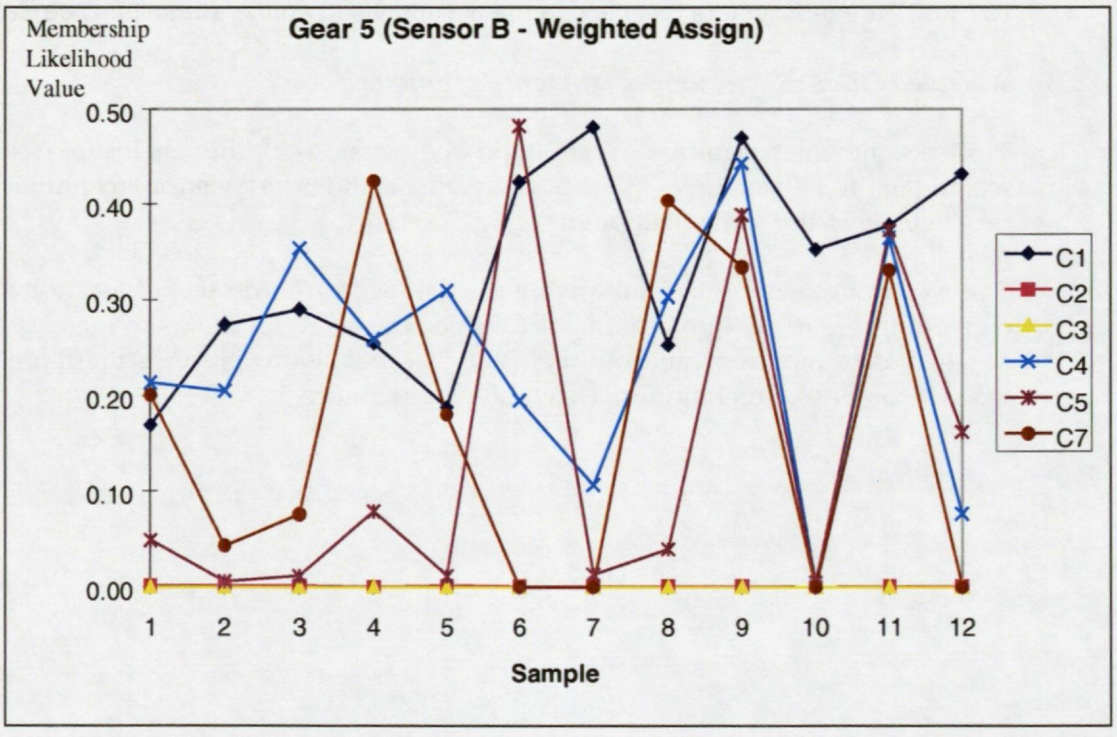


Figure 6 Results of the Weighted Standard Deviation Enclosure technique

(a) Normal scaling

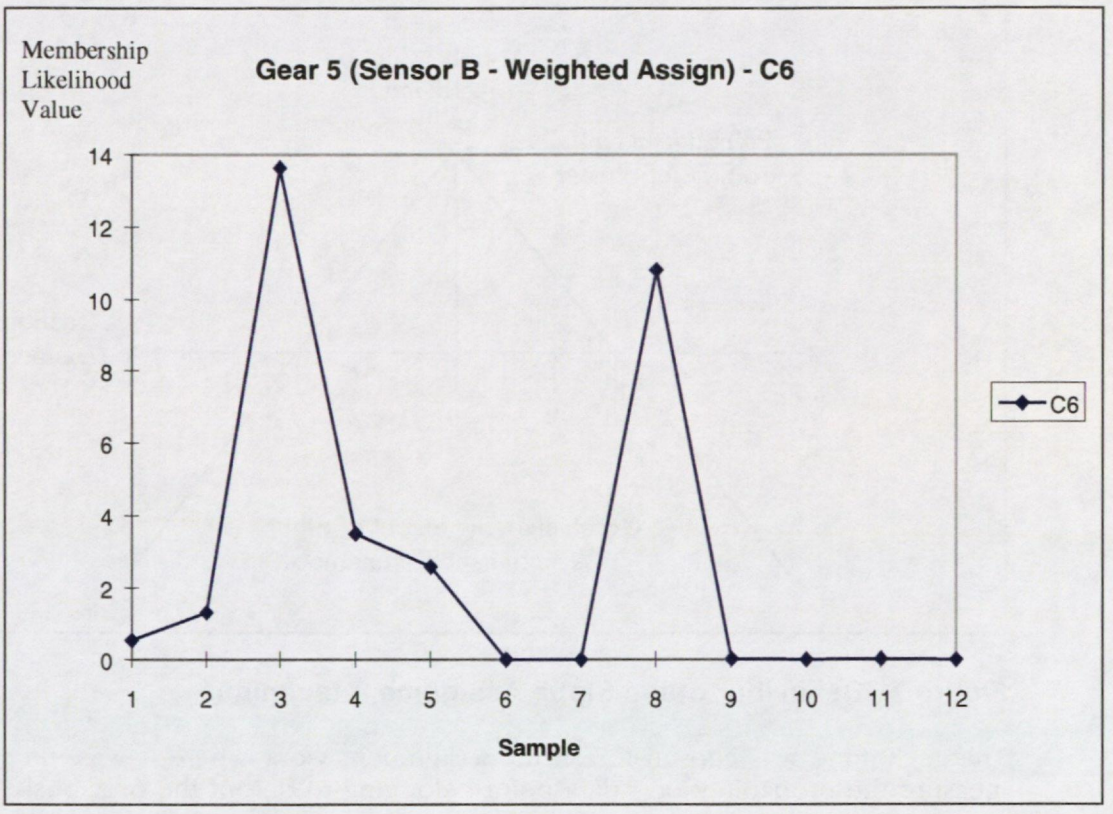


Figure 6 Results of the Weighted Standard Deviation Enclosure technique.

(b) Re-scaled for C6.

(C1 to C7 are the clusters identified in) the unsupervised analysis results – see Table 3)

3.2.4 *Slope Assignment Technique*

This technique is a linear approximation of the standard deviation technique. The Euclidean distance between the cluster and the data point is simply multiplied by a constant based on the standard deviation of the cluster.

This technique effectively fits a straight line to each side of a normal distribution. The line cuts the X-axis at plus and minus 3 standard deviations and meets at the Y-axis to form a point. Beyond 3 standard deviations, the values produced by this technique become increasingly more negative. This is illustrated in Figure 7.

The advantage of this technique is its linearity. Distances greater than 3 standard deviations from a cluster are represented by negative values rather than values tending towards zero. This allows the distance between the data points and the cluster to be monitored even when it is very large.

The results of the slope assignment technique are shown in Figure 8. This graphical representation did not convey the information adequately so an inverted format was established. An example of the inverted format is shown in Figure 9. When compared with the alternative assignment techniques, this presentation format provides some evidence of trends associated with samples 6 to 12.

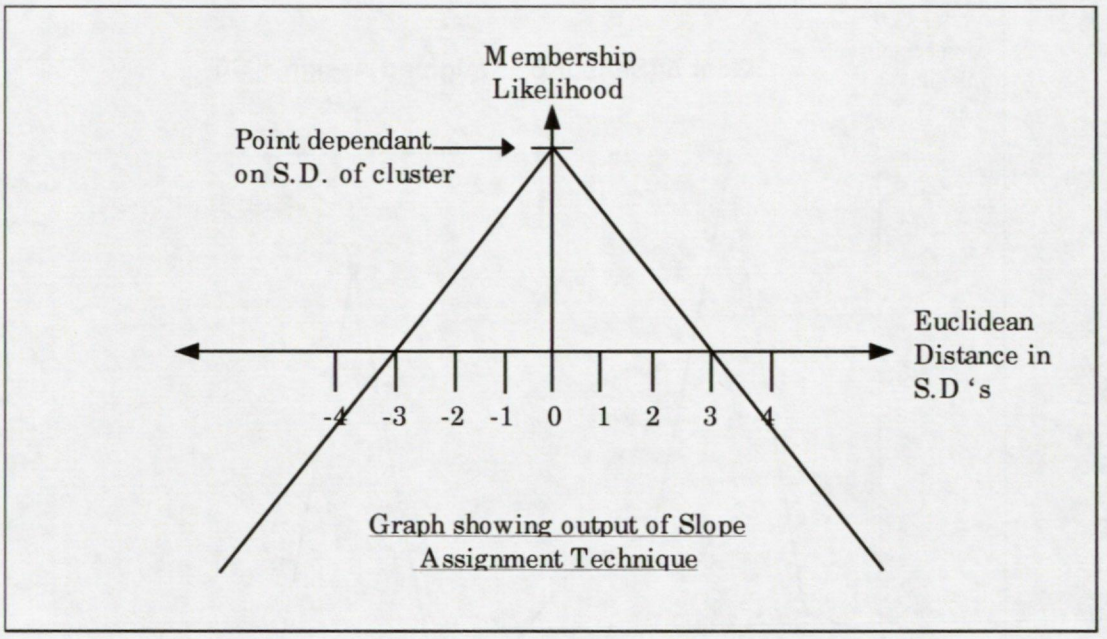


Figure 7 Behaviour of the Slope Assignment technique

This technique was chosen for all the assignment work where it was necessary to measure the probability of a data point belonging to one of the previously derived clusters.

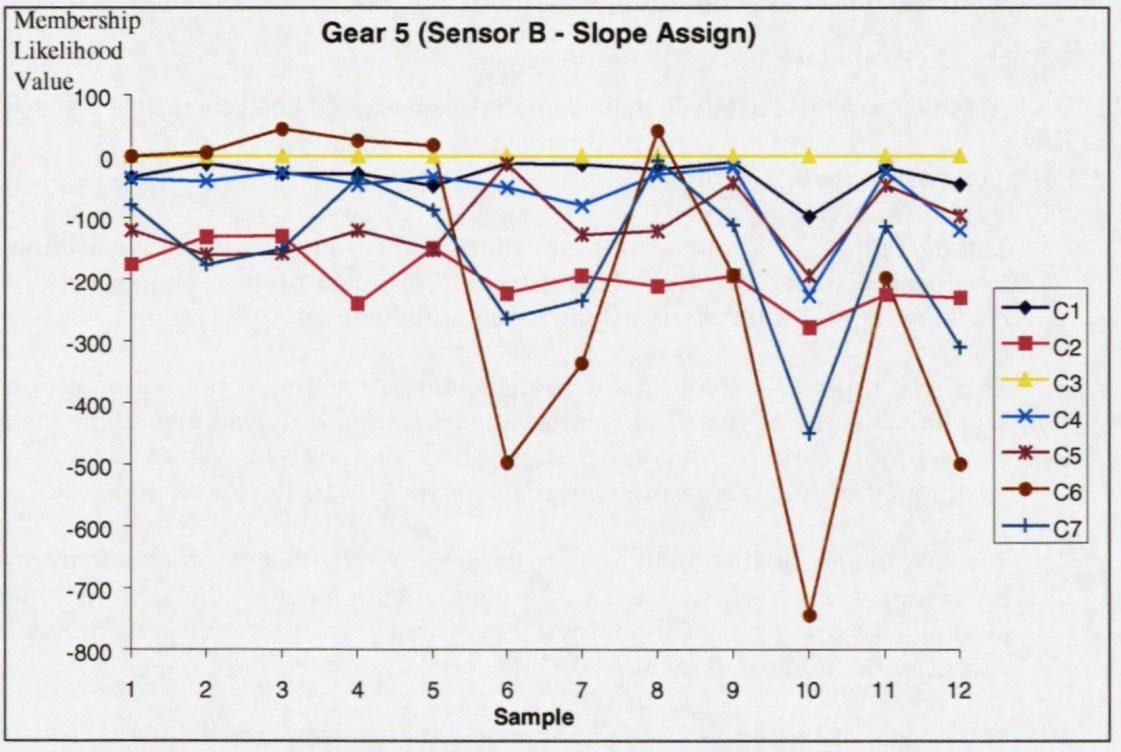


Figure 8 Results from Slope Assign technique

(C1 to C7 are the clusters identified in the unsupervised analysis results – see Table 3)

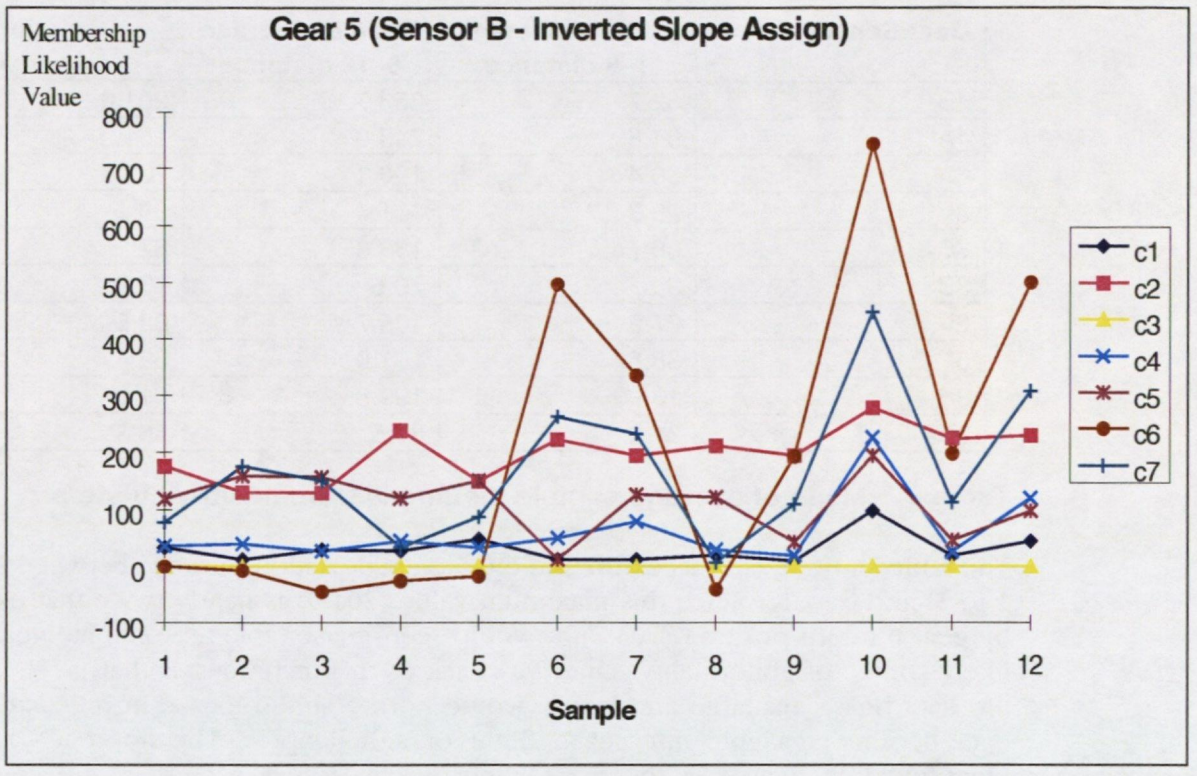


Figure 9 Inverted presentation of results from Slope Assignment technique

(C1 to C7 are the clusters identified in the unsupervised analysis results – see Table 3)

3.3 Fault Recognition Criteria

Since cluster analysis of the vibration data was only permitted on the healthy data, it was anticipated that fault-related data points would be recognised by their deviation from normal clusters. In practice, automating the detection of this deviation was far from straight forward.

3.3.1 *Fault Detection by Comparison of Averages*

The samples produced by a faulty gear were expected to be further away from the healthy clusters than samples from healthy gears. Calculating the average distance of a sample from all the healthy clusters gave an indication of fault probability. A small average indicates a low fault likelihood, whereas a large average indicates a high fault likelihood. For each gear, the mean of the average values for samples 1 to 5, and 6 to 12 were calculated. A large difference between these two values signified a change in the behaviour of a gear, potentially a fault.

Table 6 gives the gear-sensor combinations with the 10 largest deviation distances.

Gear/Sensor	Mean of Average 1–5 distance	Mean of Average 6–12 distance	Deviation distance
4B	52.7	253.7	201.0
4A	91.8	243.0	151.2
6A	86.4	190.1	103.7
3I	28.3	111.4	83.1
7C	50.7	133.3	82.6
3B	3.5	76.5	73.0
6I	37.2	105.7	68.5
5I	38.6	101.0	62.4
9B	138.0	199.8	61.8
8A	31.2	87.5	56.3

Table 6 Mean of average sample distances from healthy clusters

The table shows gear 4 as exhibiting the largest deviation distance between samples 1 to 5, and 6 to 12. Since this maximum value (201.0) is nearly twice that exhibited by gear 6 in 3rd position (103.7), it would seem reasonable to nominate gear 4 as a likely source of abnormality. One drawback to this technique is that it relies on all the data being available and hence would not be suited to use in practice as new data became available after each flight of a helicopter. The approach has also identified that gear 4 is the most likely gear to have a fault in the context of knowledge that a fault is present. This, again, is not the case for in-service 'real time' fault identification.

3.3.2 *Fault Detection by Maximum Deviation*

A similar technique involved scoring each gear-sensor combination according to the maximum distance it moved away from any healthy cluster. The average distance of samples 1 to 5 was calculated for each cluster and recorded. This was also done for samples 6 to 12. The first set of results were then subtracted from the second. For each gear/sensor combination, this gave the average movement away from each normal cluster. There were 7 of these movement values for each gear-sensor combination. The largest was selected from each gear/sensor combination and used to rank the results.

Table 7 shows the top ten gear-sensor combinations according to this technique.

Gear/Sensor	Max movement away from cluster	Movement value
7b	6	397.8
5b	6	365.5
5c	6	336.1
6a	6	316.6
4b	7	293.8
4a	7	234.3
3c	6	228.3
14d	6	202.9
3b	6	168.7
3i	7	156.5

Table 7 Top ten maximum movement values for gear/sensor combinations

Clusters 6 and 7 both contained a high percentage of the total data. From this technique it would be inferred that gear 5 or gear 7 was fault related. It is worth noting that both these gears are on the same shaft.

The disadvantage of this technique is that it does not recognise the significance of consistent *trends* in the data. Inspection of the data revealed that gears 5 and 7 do not show any such trends. Figure 10 shows two maxima for samples 6 and 7 in cluster 6 which are highly uncharacteristic of the samples produced by this gear-sensor combination. The maximum deviation technique, however, is sufficiently influenced by these outliers to conclude that gear 7 is the most likely source of abnormality. As with fault detection by comparison of averages, this technique also relies on all the data being available and would therefore not be viable in practice.

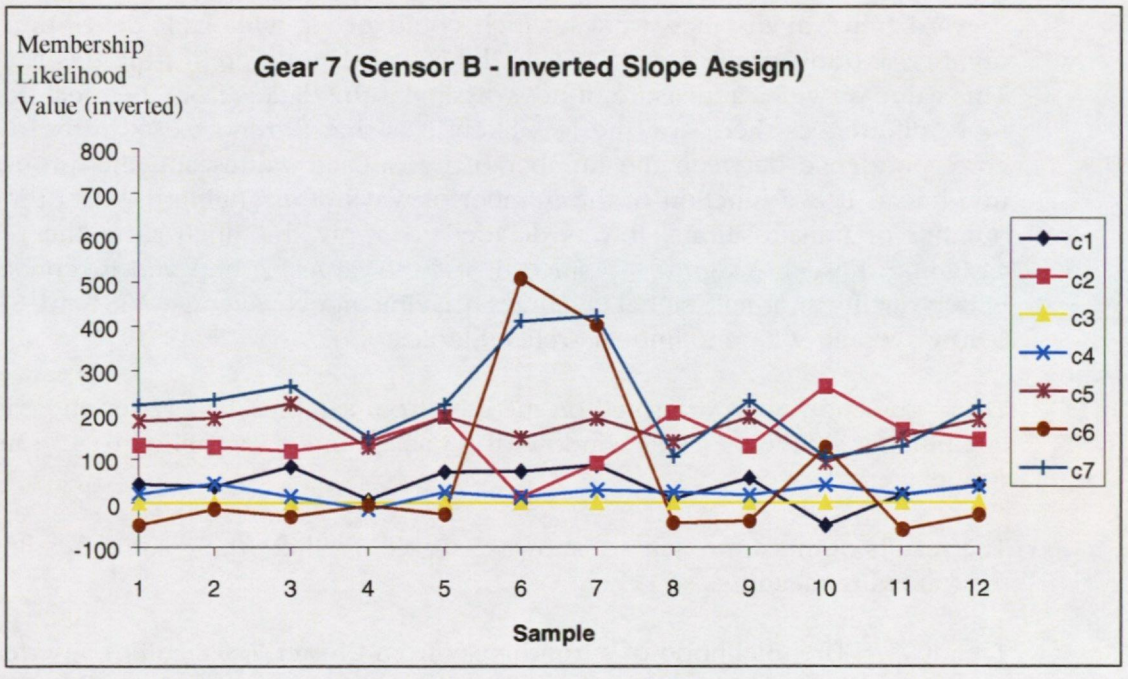


Figure 10 Distance of gear 7 sensor B samples from normal clusters

3.3.3 *Fault Detection by Trend Analysis*

Following an interim review of the project, the CAA advised MJAD that more emphasis should be placed on trend detection. This recommendation was fully endorsed by MJAD and the subsequent work is discussed in the following sections.

3.3.3.1 Gears with 'VIS'

This technique utilised a trend detection algorithm to identify samples moving in a consistent manner. It was anticipated that the distance of abnormal samples from normal clusters would ideally show an increasing trend between samples 6 and 12, as the defect propagated.

The trend detection algorithm considers a window of five samples. The current sample together with four previous samples are considered by the algorithm. In practise this has the effect of reducing trend noise due to one-off samples which do not conform to the trend. For example, data which shows an increase on 4 of 5 occasions is likely to be trending upwards.

A linear regression algorithm was not used for the task of trend detection for two reasons:

- The data contains outliers which would adversely affect the results of a linear regression algorithm.

- The time interval between each sample was variable and unknown. A constant interval approximation could have generated misleading results.

The chosen algorithm analyses the distance of the samples from a normal cluster centre, and reports the statistical likelihood of the distance values increasing with sample number. It does this by comparing the actual data with a sorted version of the data. The sorted version is in ascending order and therefore represents the ideal upward trend in distance values which could occur with fault development. The number of transpositions which make the actual data different from this is counted. This value provides a measure of how dissimilar the data is from the ideal trend. The more differences there are, the less likely it is that a trend exists in the data. The correspondence between the number of transposed values and this probability is non-linear. It is a function of the number of ways of arranging the data for a given number of transpositions. It was decided to apply this analysis to the results of assigning the data to group 7, since this is the largest group and it represents the better quality 'normal' signal averages (having above average VIS and STB, and below average WEA and impulse-related indicators).

Trend detection was performed on the data from samples 1 to 12 for all gear sensor combinations. Since a trend window of 5 data points was applied to 12 samples, 8 results were derived.

The results of this approach are shown in Table 8, where the column headings have the following meanings:

TrendC7 – The likelihood of a trend away from Cluster 7: 0 signifies a perfect trend away, 0.5 signifies no trend and 1.0 signifies a perfect trend towards.
MeanC7 – This number represents the average distance from the centre of group 7 in the moving window of 5 data points.
StDevC7 – The standard deviation of distances from the centre of group 7 in the moving window of 5 data points.
Sample – The sample number (i.e. a value of 6 represents the results of using a moving window from samples 2 to 6 inclusive).

Table 8 displays those records where TrendC7 < 0.5, MeanC7 >233 (i.e. 10 s.d. from the cluster centre), and StDevC7 >100.

From the results in Table 8, it is clear that there are a large number (>50%) of occurrences of gear 4 from different sensors. In particular, sensor C shows consistent behaviour for samples 9 through 12, and sensor I is included at sample 6. The mean and standard deviation for the whole data set can be seen in the bottom right-hand corner of Table 8.

The average trend value was also calculated for each gear, by averaging the TrendC7 values across all the sensors monitoring each gear. This indicated how much the data for each gear was moving away from, or towards a particular cluster when averaged

across all the sensors monitoring the gear. Values were derived for all clusters and all gears. Table 9 shows the results.

Gear-Sensor Combination	Sample	TrendC7	MeanC7	StDevC7
GEA04_A	12	0.2	280	133
GEA04_C	9	0.3	251	111
GEA04_C	10	0.4	272	103
GEA04_C	11	0.4	296	112
GEA04_C	12	0.4	401	116
GEA04_I	6	0.1	240	152
GEA04_I	7	0.4	290	110
GEA04_I	12	0.3	382	167
GEA04_K	6	0.3	285	103
GEA04_K	12	0.3	434	154
GEA05_I	10	0.3	261	208
GEA05_I	11	0.4	305	191
GEA05_I	12	0.3	288	191
GEA06_I	10	0.3	321	242
GEA06_I	11	0.2	331	247
GEA06_I	12	0.4	343	236
GEA07_B	7	0.2	295	108
GEA07_B	8	0.4	263	132
GEA09_B	12	0.4	312	126
		Mean:	106	77
		StDev:	103	56

Table 8 Trend analysis of slope assign data

Gear	Clust 1	Clust 2	Clust 3	Clust 4	Clust 5	Clust 6	Clust 7	Mean
4	4.38	4.73	n/a	4.45	4.33	4.70	4.30	4.48
6	4.31	5.06	n/a	4.22	4.82	4.69	4.09	4.53
10	4.21	5.25	n/a	5.25	4.63	4.46	3.79	4.60
3	4.43	5.38	n/a	4.18	4.68	4.63	4.35	4.60
7	4.38	4.42	n/a	4.38	6.00	4.58	5.13	4.81
1	5.29	4.00	n/a	4.92	4.96	5.00	4.96	4.85
9	4.21	5.38	n/a	4.96	4.75	5.21	4.83	4.89
8	5.08	5.00	n/a	4.71	5.04	5.04	4.50	4.90
14	5.53	4.41	n/a	5.22	4.53	5.28	4.91	4.98
5	4.94	5.28	n/a	4.97	6.09	4.72	4.56	5.09
2	5.25	5.35	n/a	5.69	4.88	5.31	4.66	5.19
11	5.79	5.75	n/a	5.88	5.08	5.46	4.79	5.46

Table 9 Average transpositions for each gear cluster combination

Note: It was not possible to obtain values for cluster 3 since it contains only one data sample and therefore does not have a valid standard deviation.

A value less than 5.0 in Table 9 indicates the samples from a gear are trending away from a normal cluster. A value of 5.0 indicates no trend, and a value greater than 5.0 indicates a trend towards a cluster. The scale is a non-linear function of the number of values which are mis-positioned when compared to an ordered version of the sequence. The function returns a value which is related to the number of ways the given sequence can be re-ordered to improve its trend. For example, there will be numerous ways of improving a sequence which shows very little trend, but very few for one which is already well ordered. The function, therefore, has a low sensitivity to changes in well ordered data, but a high sensitivity to badly ordered data. Thus, small changes either side of 5.0 are equally as significant as large changes near 0.0 and 10.0.

Table 9 shows that ***gear 4 is the only gear which shows a trend away from all groups***. Every other gear shows a trend *towards* at least one of the healthy clusters. This suggests that while healthy gears are moving between normal clusters, gear 4 is the only gear moving away from all of them. This analysis therefore clearly identifies that ***gear 4 is the only gear exhibiting consistent fault-related behaviour.***

3.3.3.2 'Non-VIS' gears

The slope assignment results for this data did not show any obvious trends.

3.4 SOA Data

Oil samples were taken from the gearbox after each run and analysed for trace element composition. This provided a completely separate data source which was independent of the vibration analysis. The SOA data reported the concentration levels of 8 chemical elements (Fe, Mg, Al, Ti, Cu, Ag, Zn and Cd) found in the gearbox oil. The accuracy of these samples was one part in a million. Concentrations less than 1 ppm did not show up in the analysis. As a result, the values for all elements apart from iron and zinc were zero. Data for 11 oil samples was provided which related to the same time period as the first 5 vibration samples.

In a similar way to the vibration data, cluster analysis was applied to the SOA samples. Since only two elements were detected, the analysis was two dimensional and can therefore be visualised as shown in Figure 11.

In Figure 11, the individual SOA samples are denoted as either a number or a number followed by a letter. The number indicates the vibration sample to which they relate, and the letter the number of SOA samples since the last vibration sample. For instance SOA sample 3b is the second repeat SOA sample since the third vibration sample.

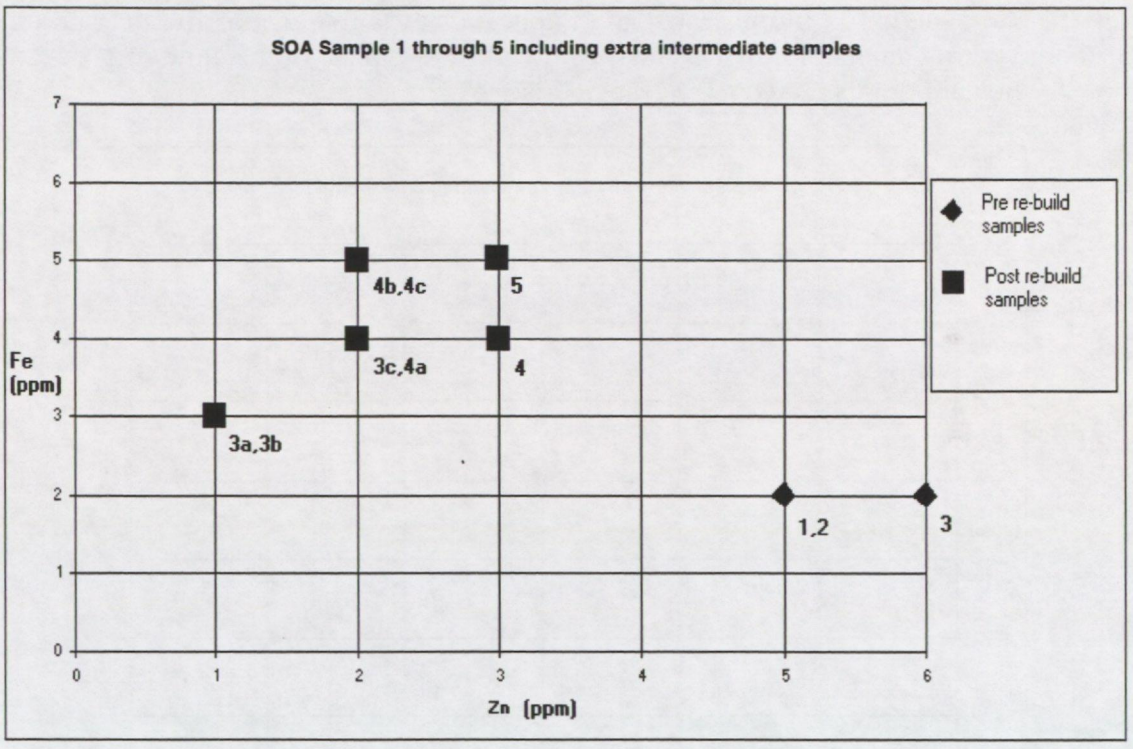


Figure 11 Location of first 11 SOA samples and PLATO clusters

Following sample 3, there was a major rebuild of the gearbox. This was picked up by the cluster analysis and is shown by the discrimination of samples 1 to 3 and 3a to 5.

Samples taken after the rebuild show an approximately linear trend of wear before the fault is visible. Approximately equal amounts of zinc and iron appear to be worn away in each time interval. This implies that the wear rate is equal for both metals.

The small number of monitored elements together with the 1 ppm accuracy prevented any form of alloy matching. Without the absolute elapsed time between samples, even wear rate monitoring could not be used as a reliable fault indicator. If the sampling rate was increased as the fault developed, a fixed approximation of sample intervals would produce misleading results since any increase in the wear rate would be cancelled out.

3.4.1 *SOA Data Unsupervised Analysis*

SOA data from a further 7 oil samples covering the post-initiation phase was supplied to MJA Dynamics. This was of the same accuracy as the first 11 samples and only showed results for iron and zinc.

PLATO was used to assign the new data to the existing clusters derived from the first set of data. The results can be seen in Figure 12.

As can be seen from Figure 12 the movement of the data points after sample 5 appears to be a linear continuation of the behaviour before this sample. This suggested that the fault was not showing in the SOA analysis. The approximately equal increase in the levels of both iron and zinc during the trial was more consistent with normal wear than an abrasive fault. On the assumption that a fault had in fact

been seeded in the gearbox, the implication was that it must be of a non-abrasive type not amenable to SOA analysis. Cracks are typical of this kind of gear fault since they are not expected to release debris.

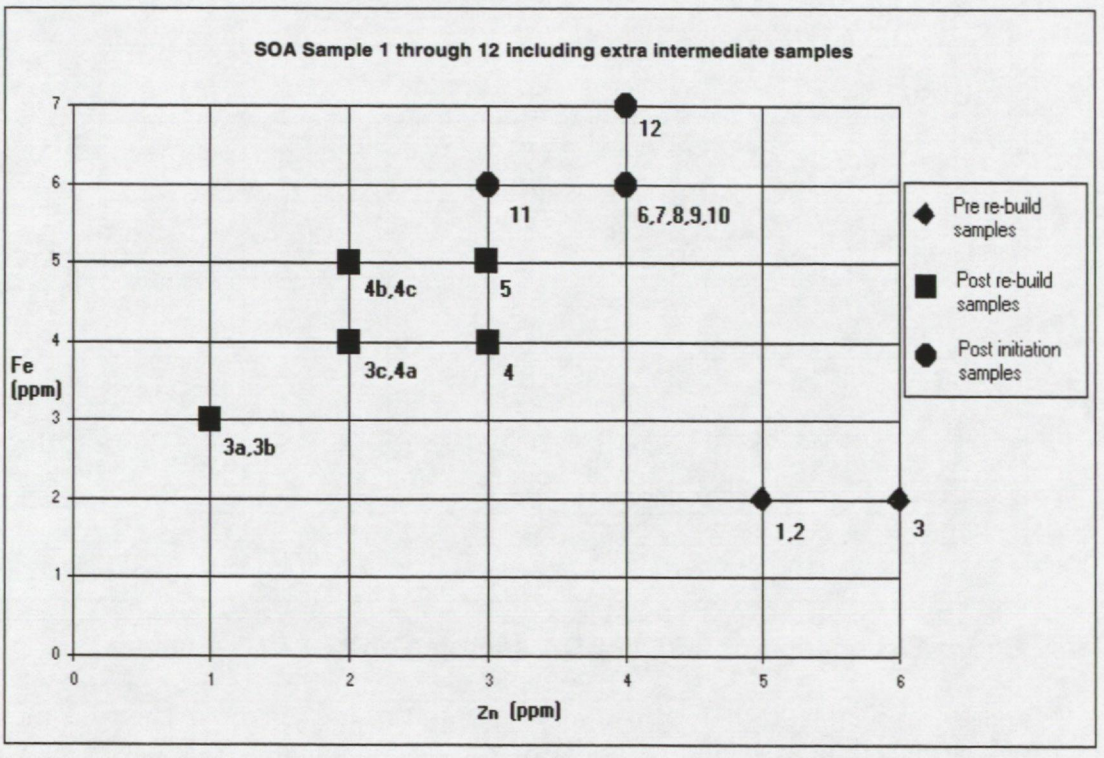


Figure 12 A complete plot of the SOA data vector space

3.5 Conclusions

Consideration of the vibration data from gears with and without measurable 'VIS' gave an indication of the location and nature of the fault. The SOA data provided supplementary evidence concerning the nature of the fault. The following conclusions were drawn.

- The unsupervised analysis of the vibration data pointed to a fault in gear 4.
- The normal behaviour of the SOA data indicated the fault was of a non abrasive type.

Since cracks are a common type of non-wearing fault, and the vibration analysis pointed to gear 4, the most likely fault was a crack in gear 4.

Following advice on the nature of the seeded fault, the above deduction was proved to be accurate. The seeded fault turned out to be a progressive crack in the webbing of gear 4.

A defect was spark eroded at the root of a geartooth. This developed into a small crack during samples 1 to 5, but was not deemed to be significant at this stage. From sample 5 onwards the crack extended rapidly, eventually entering the web during sample 12. The test was stopped at this point.

The results in Table 8 demonstrate that gear 4 was consistently recognised as faulty from sensor C from sample 9 onwards. This finding was corroborated from sensors A, I, and K at sample 12. The other gears identified in Table 8 prior to sample 12 are all on shafts which mesh with gear 4's shaft. These results indicate that the change in the health of gear 4 has affected the meshing of these gears, since the GI values are sensitive to the characteristics of the vibration synchronous with a gear's rotation (i.e. affected by its meshing characteristics).

4 SUPERVISED ANALYSIS OF FAULT 1 DATA

Supervised learning techniques work by detecting features which distinguish one data set from another. In the case of fault detection, the goal is to distinguish the post-initiation (i.e. fault present) data set from the pre-initiation (i.e. healthy) data set. For the first seeded defect trial, gear 4 samples 6 to 12 were considered to be faulty while all other samples from all other gears were considered to be healthy.

4.1 Initial Iterative Feature Extraction Results

The Iterative Feature Extraction technique (see Section 2.2.1) was applied to the vibration data. The VIS, IMP, STB, WEA, AMIK, AMIE, EIIV and EIIE gear indicator values were used (see Annex A for further details of these gear indicators). The separation results are shown for each of the sensors monitoring gear 4 in Table 10 below. The 'Cross over point' column indicates the detection rate at the point where the number of 'false alarms' (i.e. pre-initiation data points falling within the fault-related area) equals the number of 'missed faults' (i.e. post-initiation samples classified as 'healthy').

Gear Sensor	No. of iterations to achieve no false alarms	Fault Capture Rate for zero false alarms.	Cross over point Fault Capture Rate
4a	0	100%	100%
4b	4	43%	81%
4c	2	71%	93%
4i	4	43%	86%
4k	2	71%	86%

Table 10 Number of iterations to find solution giving 0% false alarms

The following conclusions can be drawn from the above table:

- The iterative feature extraction technique is quick to learn and does not require much training data. Neural networks are renowned for requiring large amounts of training data and taking a long time to learn.
- The fault data from gear 4 sensor A is highly separable from the normal data.
- The fault data from gear 4 sensor C is reasonably separable from the normal data.
- Iterative feature extraction returns the dimensions of the enclosure space it has learned to recognise. These values could be used to set HUMS limits from known fault data and assign reliability probabilities to the diagnosis. This would not be possible using a neural network.

4.2 Enhanced Pre-processing

Following the success of Iterative Feature Extraction (IFE), as described above, it was decided to apply additional statistical pre-processing to the SGDS parameters before the IFE analysis. The results obtained using the IFE technique on the raw parameters for all the gear-sensor combinations are shown in Table 11. All the IFE results' tables in this section show those samples which fell within the boundaries identified by the IFE technique, at the point at which processing was stopped.

The iterative process was set to suspend calculations when the false alarm rate dropped below 30%. This value was chosen arbitrarily as a benchmark for comparing the performance of different pre-processing techniques. For a practical system it may be necessary to achieve a lower false alarm rate by allowing the algorithm to continue iterating beyond this point. It should be noted that the definition of false alarm rate chosen expresses the number of 'healthy' samples falling within the enclosure as a percentage of all the samples falling within the enclosure. If the number of false alarms are considered as a percentage of the total number of samples, much lower values result.

SENSOR_MEL	SAMPLE	VIS	IMP	STB	WEA	AMIK	AMIE	Fault
GEA14_D.ME	1	0.03	3.35	0.43	1.00	2.51	0.52	FALSE
GEA04_A.ME	10	0.04	2.60	0.47	0.69	2.26	0.43	TRUE
GEA04_A.ME	12	0.08	2.89	0.28	0.90	3.07	0.50	TRUE
GEA04_B.ME	12	0.02	2.50	0.18	0.99	4.46	0.53	TRUE
GEA04_C.ME	8	0.15	2.35	0.34	0.92	2.94	0.44	TRUE
GEA04_C.ME	10	0.15	2.79	0.52	0.93	2.77	0.38	TRUE
GEA04_C.ME	11	0.21	2.66	0.35	0.90	2.87	0.45	TRUE
GEA04_C.ME	12	0.01	2.73	0.28	0.53	4.84	0.60	TRUE
GEA04_I.ME	5	0.12	2.78	0.46	0.87	2.66	0.44	FALSE
GEA04_I.ME	6	0.11	2.41	0.31	0.93	3.19	0.46	TRUE
GEA04_I.ME	9	0.08	2.95	0.35	0.88	2.72	0.43	TRUE
GEA04_I.ME	10	0.06	3.04	0.42	0.91	3.02	0.45	TRUE
GEA04_K.ME	1	0.05	3.06	0.31	0.35	2.70	0.44	FALSE
GEA04_K.ME	6	0.08	2.84	0.41	0.58	3.04	0.42	TRUE
GEA04_K.ME	9	0.04	3.06	0.25	0.28	2.18	0.51	TRUE
GEA04_K.ME	10	0.07	2.57	0.30	0.39	2.28	0.41	TRUE
GEA04_K.ME	11	0.16	2.80	0.39	0.43	2.87	0.44	TRUE
GEA04_K.ME	12	0.05	4.27	0.44	0.49	2.28	0.44	TRUE
GEA05_C.ME	1	0.09	2.98	0.44	0.59	3.83	0.53	FALSE
GEA06_C.ME	1	0.10	2.75	0.49	0.58	3.26	0.46	FALSE
GEA09_B.ME	5	0.14	2.73	0.46	0.98	3.20	0.45	FALSE

Table 11 Contents of Fault Enclosure formed by IFE.
(Note : The SENSOR_MEL column identifies specific gear-sensor combinations)

(Overall Fault Capture Rate : 42.85% for False Alarm Rate : 28.5%)

The large number of correctly classified fault samples, shown as 'TRUE' in the right-hand column of Table 11, demonstrates that even the raw parameters are exhibiting a degree of fault related behaviour which allows them to be largely isolated from the healthy data. It was anticipated that pre-processing the parameters would reduce the number of incorrectly classified samples. The results in Table 11 indicated that there was no consecutive misclassification of data from any one gear-sensor combination. This suggested that it was uncharacteristic behaviour associated with single samples which caused the misclassifications. It was decided to investigate the use of averaging the data to reduce such effects.

4.2.1 *The effect of applying a moving average window*

The results of using a moving average of two samples to reduce the effect of uncharacteristic spikes in the data are shown in Table 12. (Note that wherever two point moving average resutls are presented, the sample number will refer to the second of the two samples used).

SENSOR_MEL	SAMPLE	VIS	IMP	STB	WEA	AMIK	AMIE	Fault
GEA04_A.ME	10	0.08	2.71	0.52	0.59	2.66	0.45	TRUE
GEA04_A.ME	11	0.06	2.69	0.56	0.66	2.48	0.43	TRUE
GEA04_A.ME	12	0.08	2.83	0.47	0.76	2.89	0.46	TRUE
GEA04_C.ME	6	0.49	2.81	0.57	0.75	2.30	0.43	TRUE
GEA04_C.ME	8	0.35	2.60	0.55	0.82	2.57	0.43	TRUE
GEA04_C.ME	9	0.13	2.47	0.38	0.84	2.55	0.45	TRUE
GEA04_C.ME	10	0.13	2.69	0.47	0.84	2.47	0.43	TRUE
GEA04_C.ME	11	0.18	2.72	0.44	0.91	2.82	0.41	TRUE
GEA04_I.ME	5	0.12	2.54	0.55	0.90	2.42	0.47	FALSE
GEA04_I.ME	6	0.12	2.59	0.39	0.90	2.93	0.45	TRUE
GEA04_I.ME	8	0.27	2.44	0.54	0.87	2.27	0.46	TRUE
GEA04_I.ME	9	0.17	2.76	0.46	0.88	2.54	0.44	TRUE
GEA04_I.ME	10	0.07	2.99	0.39	0.89	2.87	0.44	TRUE
GEA04_K.ME	2	0.19	2.91	0.46	0.53	2.52	0.45	FALSE
GEA04_K.ME	4	0.18	2.68	0.44	0.59	2.37	0.46	FALSE
GEA04_K.ME	5	0.08	2.66	0.40	0.49	2.45	0.43	FALSE
GEA04_K.ME	6	0.10	2.77	0.49	0.57	2.91	0.43	TRUE
GEA04_K.ME	7	0.10	2.87	0.49	0.61	2.95	0.41	TRUE
GEA04_K.ME	8	0.10	2.85	0.56	0.62	2.81	0.40	TRUE
GEA04_K.ME	9	0.06	2.93	0.40	0.45	2.46	0.46	TRUE
GEA04_K.ME	10	0.05	2.82	0.28	0.33	2.23	0.46	TRUE
GEA04_K.ME	11	0.11	2.69	0.34	0.41	2.58	0.43	TRUE
GEA06_C.ME	3	0.20	2.70	0.56	0.73	2.76	0.45	FALSE
GEA07_B.ME	6	0.35	2.71	0.54	0.80	2.72	0.45	FALSE
GEA09_B.ME	11	0.32	2.76	0.57	0.92	2.58	0.46	FALSE

Table 12 Fault enclosure formed after applying moving average to the data
(Note : The SENSOR_MEL column identifies specific gear-sensor combinations)

(Overall Fault Capture Rate : 60% for False Alarm Rate : 28%)

The moving average improved the separability of the data from sensors A and C but reduced it for K and I.

A possible explanation involved the proximity of the sensors to the faulty gear. Sensors A and C may have a better pre-disposition for detecting abnormalities in gear 4. Analysis of vibration paths was not possible due to a lack of detailed technical drawings showing the exact location of the sensors. Even with drawings, such assessment is non-trivial, and would have required confirmation by practical measurements.

There was also a concern that gear 4 was exhibiting fundamentally different behaviour from all other gears even before the fault started to develop. If this was the case, the IFE technique may be separating the data on these differences rather than on those related to the fault.

This raised considerable concern, so a strategy was devised to test it.

4.2.2 *The effect of standardisation on the moving averaged data*

By standardising samples 1 to 12, by samples 1 to 5, for each gear-sensor combination, it was possible to ensure any natural differences between the gears were completely eliminated. This process involved calculating a moving average for each sample between 1 and 12, and then dividing each of these values by the standard deviation of samples 1 to 5. The standardised values effectively measured variations away from the normal on a uniform scale for all gears. This new data was processed using the IFE technique. The results are shown in Table 13.

SENSOR_MEL	SAMPLE	VIS	IMP	STB	WEA	AMIK	AMIE	Fault
GEA13E.ME	6	0.00	-0.47	-2.15	-0.30	-0.23	0.28	FALSE
GEA13E.ME	7	0.00	-0.11	-1.68	0.93	-0.25	-0.22	FALSE
GEA02_F.ME	10	1.29	1.39	-2.40	1.67	0.15	1.13	FALSE
GEA03_A.ME	12	-0.61	3.05	-1.42	0.34	-0.31	0.79	FALSE
GEA04_A.ME	10	-0.64	0.90	-1.56	-2.74	-0.40	0.47	TRUE
GEA04_A.ME	11	-0.77	0.79	-1.32	-2.11	-0.91	-0.26	TRUE
GEA04_A.ME	12	-0.63	1.37	-1.88	-1.09	0.29	1.08	TRUE
GEA04_C.ME	6	1.46	0.28	-4.96	-0.43	-1.22	0.04	TRUE
GEA04_C.ME	7	1.48	0.33	-2.10	-0.45	-1.51	-0.05	TRUE
GEA04_C.ME	8	0.54	-0.49	-6.05	0.69	-0.75	-0.12	TRUE
GEA04_C.ME	9	-0.83	-0.96	-14.16	0.97	-0.79	1.09	TRUE
GEA04_C.ME	10	-0.82	-0.15	-9.78	1.00	-0.93	-0.06	TRUE
GEA04_C.ME	11	-0.50	-0.03	-11.53	2.19	-0.33	-0.59	TRUE
GEA04_I.ME	9	-0.01	0.58	-1.42	0.96	-0.29	-0.22	TRUE
GEA04_I.ME	11	-0.48	1.96	-1.27	0.98	0.01	-0.58	TRUE
GEA04_K.ME	10	-0.82	0.25	-1.08	-1.27	-1.02	0.22	TRUE
GEA08_C.ME	10	-0.61	-0.65	-2.12	0.82	-0.11	0.60	FALSE

Table 13 Fault enclosure formed after applying moving average and standardisation to the data

(Note : The SENSOR_MEL column identifies specific gear-sensor combinations)

(Overall Fault Capture Rate : 40% for False Alarm Rate : 29.4%)

Table 13 shows that the standardisation process has not reduced the separability of gear 4 from sensors A and C. This indicates that the IFE analysis has identified fault characteristics in the data and is able to use them to distinguish faulty samples from healthy ones from these sensor locations. The longer unbroken sequence of faulty samples detected by sensor C indicates that the standardisation process has actually improved the visibility of the fault from this sensor. This is further evidence to suggest that the standardisation process, combined with a moving average, provides a powerful tool for ensuring that fault data is not separated out based upon characteristics of gear vibration which were present prior to fault initiation. The fact that the final fault sample (i.e. sample 12) was only recognised from sensor A, indicates that the indicator values returned from this sample are significantly different from those returned from the earlier samples. This suggests that it may ultimately be better to train the system twice, once to recognise the early stages of failure, and then a second time to recognise the final stages of the fault's development. This would result in improved fault capture rates, and provide an indication of fault development.

4.3 Test On Unseen Data

Following the successful classification of fault data from sensor C, it was decided to test the set of limits derived by the IFE process on the previously unseen data samples (13–19). Since the limits were derived exclusively from data samples 1 to 12 with absolutely no 'knowledge' of the unseen data, this was considered to be a challenging test, and representative of the manner in which data would be analysed in service (i.e. by applying new data as they were acquired to boundaries which had previously been identified from analysis of data known to relate to the existence of a fault).

Table 14 shows the results obtained, again applying the same standardisation and two point moving average pre-processing to the new data as was used to define the original boundaries. (Note: Only the data from sensors monitoring gears 3,4,5 and 6 were considered. These gears mesh with, or are on the same shaft as, the fault gear. In a real system it would be both practical and beneficial to monitor gears in localised groups, since such gear-specific analysis should provide increased sensitivity when compared to developing fault analysis criteria which have to apply to the whole gearbox).

Sensor_MEL	Sampl	VIS	IMP	STB	WEA	AMIK	AMIE	Fault
GEA04_A.ME	14	–0.33	2.03	–1.64	–0.74	–0.25	0.86	TRUE
GEA04_A.ME	16	–0.02	0.68	–1.46	–0.68	–0.97	–0.35	TRUE
GEA04_C.ME	14	0.30	–0.22	–9.12	1.16	–0.44	–0.45	TRUE
GEA04_C.ME	16	0.63	–0.62	–4.52	–0.63	–0.85	–0.07	TRUE
GEA04_C.ME	17	0.49	–0.58	–6.71	0.78	–0.47	0.21	TRUE
GEA04_C.ME	18	–0.12	–0.72	11.09	1.64	–0.22	0.66	TRUE
GEA06_I.ME	19	–0.38	0.96	–2.07	–0.05	–0.29	1.24	FALSE

Table 14 Classification of unseen data using exactly the same fault enclosure.

(Note : The SENSOR_MEL column identifies specific gear-sensor combinations)

(Overall Fault Capture Rate : 20% for False Alarm Rate : 14%)

The results shown in Table 14 are extremely encouraging. Samples 14,16,17 and 18 from sensor C are all correctly classified. Of the unseen fault data, only sample 15 from sensor C was missed. In the same way as sample 1, sample 13 cannot be processed since it has no predecessor with which to form a moving average. It is assumed that samples 12 and 19 were taken just before the rig was stopped, they exhibit extreme behaviour and are likely to have unique characteristics which prevents them from being classified with the other fault related data.

One hundred and eight unseen samples were classified during the test (samples 14–19 for all sensors monitoring gears 3,4,5,6). Out of these, only 1 false alarm was generated (gear 6, sensor I, sample 19). Six of the 30 fault related samples recorded by the sensors monitoring gear 4, were successfully classified. Interestingly, the false alarm was generated at the very end of the test when extreme vibration from the fault gear may have affected other gears. In practice it is highly unlikely that a fault would go undetected before it reached the catastrophic level recorded during sample 19. If the last samples from all of the sensors are discounted because of their extreme nature, the technique can be considered as having detected 6 (samples 14 & 16 from sensor A, and samples 14,16,17,18 from sensor C) in 25 (samples 14 to 18

from sensors A, B, C, I,and K) of the unseen fault samples with no false alarms. Considering the unseen data from sensor C alone, this figure rises to 80% (4 in 5) with no false alarms.

It is common current HUMS practice to use one or more 'preferred' sensor locations to monitor particular gears, based upon the performance exhibited (e.g. the higher percentage fault capture rates). In this case, for the processing and analysis used, sensor locations C and A appear best suited to detection of the first seeded defect.

4.4 Graphical Analysis

Monitoring the IFE technique progress as it processed the data, indicated that the fault data was being classified predominantly by adjusting the boundaries related to the STB and WEA parameters. A graphical representation of the pre-processed data is shown in Figure 13 from which the variation between healthy and faulty samples can be seen.

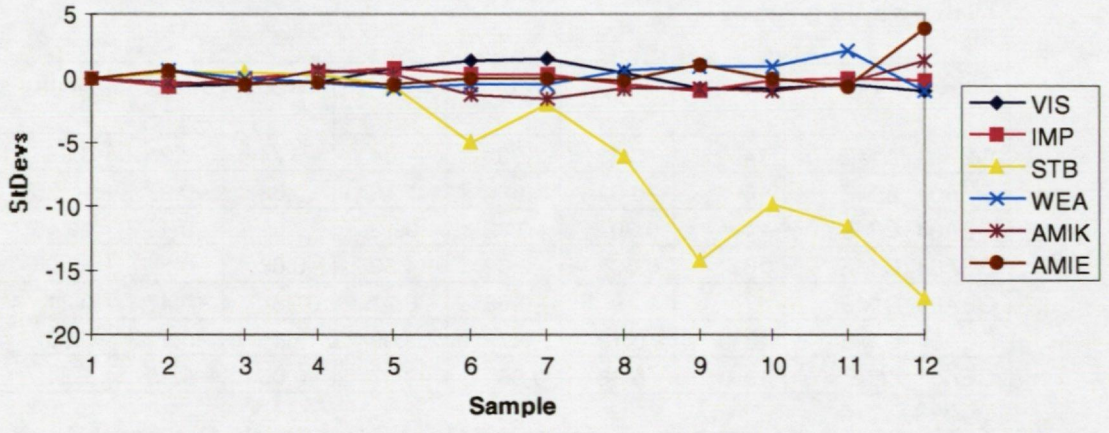


Figure 13 The value of pre-processed gear parameters monitored by sensor C

Clearly, the pre-processed value of STB from sensor C is a good indicator of gear health in this particular case. Figure 14 shows the pre-processed value of the STB parameter for several gears. STB reflects the degree to which the fundamental gear mesh vibration component dominates the lower frequency ranges in the signal average, and would not necessarily have been predicted as being sensitive to the particular fault. The result obtained confirms the advantage of the approach used, which considers *all* of the parameters, rather than restricting the analysis to individual parameters which are anticipated to be sensitive to the fault.

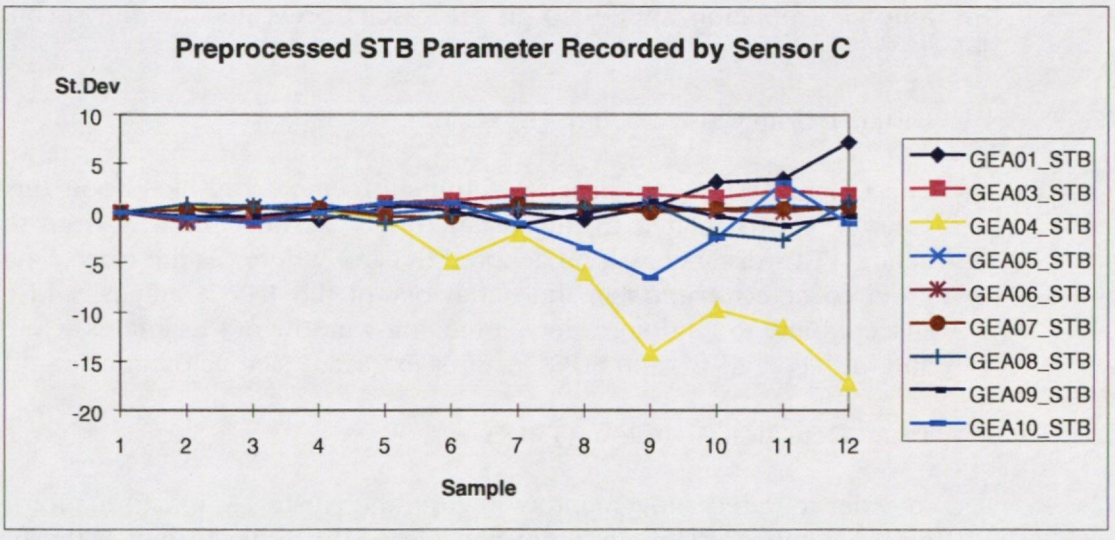


Figure 14 Pre-processed STB parameter for gears monitored by sensor C

5 ANALYSIS OF FAULT 2 DATA

5.1 Introduction

The primary objective of the analysis of the data for the second seeded defect was to demonstrate the performance of the AI techniques developed and configured for the first seeded defect when applied to a different defect. There are a number of steps in this process and success will be judged based on the satisfaction of each of them.

1 To demonstrate that the fault 1 supervised machine learning set-up was capable of classifying the pre-initiation samples from gearbox build 2 as healthy.

2 To demonstrate that the unsupervised machine learning developed during the analysis of seeded fault 1, was capable of identifying data from the post-initiation samples of gearbox build 2 as being potentially faulty.

3 To improve the ability of the unsupervised machine learning to detect a fault, by modifying the data modelling, reconfiguring the system if necessary, and/or modifying the decision criteria.

4 To demonstrate a method of analysing SOA data which can be fused into an Artificial Intelligence system.

5 To modify the set-up of the supervised machine learning developed for the first fault so that it is capable of identifying characteristics of the second seeded fault.

Six pre- and eight post-initiation samples were initially supplied for the second seeded defect (as opposed to 5 and 7 for the first defect).

The following notation applies to all the results presented for the second seeded defect :

1 Distance from Cluster Centre **(INVC*).**

The Group Assignment process outputs a membership likelihood (probability) value which is related to the distance that the new point is from the cluster centre. This distance will be denoted INVC*. Where * is the cluster number. In order to detect ***trends*** in the behaviour of the **INVC*** values, and to reduce susceptibility to random events, a moving window of 5 samples was examined and the TPC*, AVC*, and SDC* features extracted (see below for definitions).

2 Trend Detection Algorithm **(TPC*)**

In order to detect movement away from the cluster, a transformation algorithm was developed. The algorithm considered the order in which the probability values appeared against the size ranked order. A '**trend value**' between 0 and 1 was returned. A perfect movement away (e.g. 35, 71, 107, 150, 2000) returned 0. A perfect trend towards returned 1. No significant trend returned the mid value.

3 Average Distance **(AVC*)**

The average value of the windowed INVC* values, gives an indication of the general movement and attenuates any inconsistent changes.

4 Standard Deviation **(SDC*)**

The Standard Deviation of the windowed INVC* values, gives an indication of the significance of the movement away and hence how rapidly the fault is propagating.

Where * is the cluster number in all the above descriptions.

5.2 Data Integrity Checks

The SGDS scheduler software was used by WHL to acquire the vibration data and compute and store the gear indicator values for the second seeded defect. The scheduler software stores a greater range of results-related information to a relational database. This provides more information than the SM2ACQ/SM2GEA software used by WHL to acquire and analyse the data for the first seeded defect.

At an early stage in the analysis of the second seeded defect, data integrity checks carried out on the status information contained in the SGDS Scheduler results database showed that there was a possible problem with some of the measurements.

The problem related to the convergence of the signal average which was poor in a number of cases. Poor convergence particularly influences the energy parameters (i.e. 1R & 2R, etc.). The affected signal averages were removed from the analysis. This means that there may be differing numbers of samples for each gear/sensor combination used in the analysis of the second seeded defect. This is a situation that may also apply to real HUMS data.

5.3 Results

The results have been presented below, with each subsection relating to the analysis steps contained in the fault 2 introduction (Section 5.1). **MJAD were not told of the nature or location of the fault until after analysis step 3, the initial analysis was carried out 'Blind'.**

5.3.1 *Classification Of The Fault 2 Pre-Initiation Data To The Fault 1 Supervised Analysis*

The first part of the analysis of the second seeded fault data was to classify the 'pre-initiation' data to the fault boundaries defined by the supervised analysis technique as configured to detect fault 1. The analysis simply identifies whether or not the data falls within the bounded region (i.e. no numerical output values result from the analysis).

The result was that none of the samples fell within the region previously identified for fault 1 and therefore **no false alarms** were generated.

5.3.2 *Unsupervised Analysis Using The Fault 1 Set-Up And Criteria*

5.3.2.1 Gears with VIS

The approach adopted was to take the unsupervised analysis as set up for successful detection of fault 1, and use it to analyse the fault 2 data. This permits the criteria developed for detection of fault 1 to be tested using the new data.

As mentioned above the fault 1 set-up identified one cluster (group 7) which was considered normal for the gears with VIS. The fault 2 post-initiation samples were **'group assigned'** to this cluster and the same fault recognition criteria applied to the features derived from the distance or probability information. This analysis was ***identical*** to that reported for the first seeded defect.

The results are shown in Table 15 in ranked order of increasing INVC7.

Decision Criteria

Distance From Cluster **(InvC*)**	*Descending*
Trend Detection Algorithm **(TPC*)**	<0.5
Average Distance **(AVC*)**	*see Note 2*
Standard Deviation **(SDC*)**	>100

MEAS_N *(see Note 1)*	Sample	InvC7	TPC7	AVC7	SDC7
GEA05_B	9	475.1	0.117	186.82	145.7925
GEA07_C	9	398.3	0.408	169.84	152.2157
GEA02_D	13	372.2	0.242	146.76	119.2092
GEA10_C	10	370.7	0.242	42.26	164.5408
GEA11_F	10	341.1	0.042	149.5	101.3947
GEA04_A	12	332.5	0.042	185.22	161.125
GEA05_K	12	323.1	0.242	113.76	116.1944
GEA04_A	11	306.7	0.242	154.96	143.9116
GEA14_D	13	301	0.408	71.04	115.0517
GEA01_D	13	284	0.242	13.44	135.4443
GEA01_I	10	266.5	0.042	40.92	118.9702
GEA03_I	13	205.7	0.242	69.58	121.4646
GEA04_A	7	181.2	0.242	211.78	108.9478
GEA05_K	13	175.4	0.117	141.16	111.2405
GEA02_D	14	172.4	0.117	159.72	117.761
GEA01_D	14	159.5	0.042	54.44	142.2536
GEA04_A	13	95.3	0.242	194.72	153.9847
GEA01_I	12	32.9	0.408	62.7	106.6042
GEA07_C	10	26.3	0.408	116.34	146.1332
GEA03_I	14	25.9	0.242	73.86	119.4565
GEA03_K	13	16.9	0.408	64.32	130.1155
GEA03_K	11	0	0.242	59.36	132.2862
GEA10_C	11	–4.1	0.117	49.76	161.3715
GEA01_I	11	–10.9	0.117	46.1	116.0002

Table 15 Details of the second seeded fault post-initiation data which were furthest away from Fault 1, Group 7

Notes:

1. Each measurement (MEAS_N) analyses one gear from one sensor. The notation given to each measurement takes the form:–

   GEA04_A – This is Gear 4 monitored from sensor A.

   GEA14_D – This is Gear 14 from sensor D.

   and so on.

2. The unsupervised system as set-up after fault 1 had a filter of 233 applied to AVC7. With this filter in place none of the samples in Table 15 would be returned and therefore **none of the gears with VIS, based on this set-up would be considered as faulty.**

5.3.2.2 Gears without VIS :

The post-initiation data from the second seeded defect non-VIS gear data was 'group assigned' to the three normal groups previously identified by the fault 1 set-up (Groups 10,11,12), and the results analysed. The results are shown in Table 16.

Decision Criteria

Distance From Cluster **(InvC*)'**		*Descending*			
Trend Detection Algorithm **(TPC*)**			< 0.5		
Average Distance **(AVC*)**					
Standard Deviation **(SDC*)**					>100
MEAS_N	Sample	InvC10	TPC10	AVC10	SDC10
GEA13E	14	4196	0.242	907	1645
GEA13D	14	2354	0.242	541	915
GEA13F	14	2186	0.117	529	837
GEA13J	14	1901	0.242	436	735
GEA12F	11	340	0.408	88	129
GEA13F	12	337	0.242	116	118
GEA13D	10	312	0.242	89	127
GEA12F	12	111	0.408	108	123
GEA12F	13	95	0.408	114	122
GEA13D	11	–39	0.242	55	133

Table 16a Second seeded fault post-initiation 'non-VIS' gear data which were furthest away from Group 10

(AVC11 > 233 highlighted)

Decision Criteria

Distance From Cluster **(InvC*)**		*Descending*			
Trend Detection Algorithm **(TPC*)**			<0.5		
Average Distance **(AVC*)**					
Standard Deviation **(SDC*)**					>100
MEAS_N	Sample	InvC11	TPC11	AVC11	SDC11
GEA13E	14	5305.8	0.242	1209.54	2051.68
GEA13D	14	3463.2	0.408	748.98	1358.69
GEA13F	14	3053.4	0.117	814.1	1159.89
GEA13J	14	2909.8	0.242	717.66	1103.93
GEA13F	12	837.8	0.408	251.04	300.29
GEA12F	11	822.1	0.242	192.36	318.69
GEA13D	6	447.3	0.408	107.66	171.96
GEA13E	11	412.8	0.408	151.14	141.00
GEA12F	12	172.3	0.408	231.64	301.22
GEA13E	12	155.6	0.408	163.84	137.94
GEA12F	13	149.9	0.408	240.6	297.97
GEA13E	13	106.4	0.242	150.4	139.60
GEA13D	7	38.5	0.117	103.52	173.32
GEA13D	8	17.9	0.408	103.64	173.26

Table 16b Second seeded fault post-initiation 'non-VIS' gear data which were furthest away from Group 11

(AVC11 > 233 highlighted).

Decision Criteria

Distance From Cluster **(InvC*)**	*Descending*
Trend Detection Algorithm **(TPC*)**	< 0.5
Average Distance **(AVC*)**	
Standard Deviation **(SDC*)**	>100

MEAS_N	Sample	InvC12	TPC12	AVC12	SDC12
GEA13E	14	3854	0.242	851	1505
GEA13J	14	2307	0.408	543	889
GEA13F	14	2217	0.117	590	851
GEA13F	12	668	0.408	169	252
GEA12F	11	649	0.242	135	257
GEA13J	8	380	0.117	115	133
GEA13E	12	56	0.408	78	115
GEA13E	13	55	0.408	79	114
GEA13D	7	41	0.242	100	134
GEA12F	12	39	0.408	145	252
GEA12F	13	32	0.408	146	252

Table 16c Details of the second seeded fault post-initiation 'non-VIS' gear data which were furthest away from Group 12

(AVC12 > 233 highlighted).

The criteria previously used for the identification of significant trends in the fault 1 data (TPC*<0.5, AVC*>233, and SDC*>100) were applied to this data. (Note that the AVC >233 criterion was not applied to the results shown in Table 16). The AVC>233 criteria is equally applicable to this analysis, since it is a value which has been normalised by the spread of the cluster (i.e. the standard deviation of the parameter values within the cluster). This value of AVC results from application of the slope assignment process (described in Section 3.2.4), and represents a distance of 10 times the standard deviation of the parameter values within the cluster.

Application of the above criteria clearly identifies **a problem with gear 13 at sample 14 from all 4 sensors (D, E, F, and J) used to monitor the gear.** The problem was highlighted when the analysis was performed with any of the 'normal' groups (10,11,12). In the case of group 11, sample 12 for the gear 13 from sensor F also satisfied the criteria, although the AVC11 value of 251 is close to the 233 threshold applied.

Inspection of the results revealed that the system had detected increased levels of IMP, EIIE, EIIV and to a lesser extent AMIK. All of these indicate the presence of an impulsive event.

The system has at this stage positively identified Gear 13 to be displaying fault behaviour at sample 14 with a slight indication at sample 12, using the detection criteria set-up based upon the fault 1 data. Aware that early detection of the fault was one of the underlying goals further analysis was carried out.

5.3.3 *The Effect Of Improved Pre-Processing on the Unsupervised Machine Learning*

Drawing on the conclusions of the analysis of the first seeded defect, it was decided to pre-process the data in an attempt to detect the fault at an earlier stage.

Improved pre-processing techniques were developed during the supervised analysis of the data from the first seeded defect. They require computation of the change in the post-initiation GI values as a proportion of the statistical spread in the pre-initiation values, followed by application of a moving average filter to the data (as described in Section 4.2).

The effect of the above method is to highlight changes in the GI's performance by comparing them to the pre-initiation results on an individual gear / sensor basis. The previous method compared the GI's performance with the spread of values contained within each particular cluster. The objective of using the pre-processing was to improve the sensitivity before the clustering stage.

Because the GI values have been pre-processed differently, the previously defined clusters are no longer valid. The mean centre standardisation process has the effect of displacing the data from a number of places on the multi-axis graph onto the origin. As before, the pre-initiation samples are used for defining the clusters.

The results of applying the modified pre-processing techniques to the unsupervised analysis of data from the second seeded defect are detailed in the subsections below.

5.3.3.1 Gears with VIS

The clustering analysis resulted in the definition of just one cluster.

The previous trend interpretation criteria were applied to the data from gears for which VIS was computed. The results are shown in Table 17.

Inspection of Table 17, reveals that from sample 9 onwards gears 1, 2, and 3 (and to a lesser extent gear 4) consistently exhibit the greatest deviations from the pre-initiation results. It should be noted that gears 1 and 2 are on the same shaft, as are gears 3 and 4. Gears 1 and 3 mesh.

Decision Criteria

Distance From Cluster **(InvC*)**
Trend Detection Algorithm **(TPC*)** <0.5
Average Distance **(AVC*)** >233
Standard Deviation **(SDC*)** >100

Sample 7 MEAS_N	Sample 8 MEAS_N	Sample 9 MEAS_N	Sample 10 MEAS_N	Sample 11 MEAS_N	Sample 12 MEAS_N	Sample 13 MEAS_N	Sample 14 MEAS_N
	GEA02_E	GEA01_E	GEA01_E	GEA01_E	GEA01_E	GEA14_D	GEA14_D
		GEA02_E	GEA01_I	GEA01_I	GEA01_I	GEA01_E	GEA01_E
		GEA02_E	GEA02_E	GEA02_E	GEA02_E	GEA01_I	GEA01_F
		GEA03_I	GEA02_D	GEA02_E	GEA02_E	GEA02_D	GEA01_I
		GEA04_K	GEA02_E	GEA02_G	GEA02_G	GEA02_E	GEA02_D
			GEA03_I	GEA03_I	GEA03_I	GEA02_D	GEA02_E
			GEA04_A	GEA03_K	GEA04_I	GEA02_E	GEA02_F
			GEA04_K	GEA04_A	GEA04_K	GEA02_F	GEA02_D
			GEA05_I	GEA04_K	GEA05_I	GEA02_G	GEA02_E
			GEA10_C	GEA05_I	GEA06_A	GEA03_I	GEA02_F
			GEA10_I	GEA07_I	GEA08_A	GEA03_K	GEA02_G
				GEA10_C	GEA11_G	GEA04_C	GEA03_I
						GEA04_I	GEA03_K

Table 17 Results of applying the alternative pre-processing technique

(The relevant INVC, TPC, AVC, and SDC parameter values are tabulated in Annex B)

5.3.3.2 Gears without VIS

Decision Criteria

Distance From Cluster **(InvC*)**
Trend Detection Algorithm **(TPC*)** <0.5
Average Distance **(AVC*)** >233
Standard Deviation **(SDC*)** >100

MEAS_N	Sample	InvC1	TPC1	AVC1	SDC1
GEA13J	14	3711	0.242	952	1395.363
GEA13D	14	2439	0.242	683	889.51
GEA13F	14	2365	0.042	627	873.824
GEA12F	12	974	0.042	402	460.874
GEA13E	11	742	0.042	298	270.601
GEA13J	9	654	0.242	247	302.162
GEA13E	12	472	0.117	375	253.996
GEA13J	11	460	0.408	347	273.292
GEA12F	13	114	0.242	425	443.087
GEA13J	10	45	0.242	255	295.812

Table 18 Results of applying the improved pre-processing technique to gears without VIS

With reference to Table 18, the results were generally in line with those previously found, (i.e. the gear 13 data at sample 14 is clearly atypical). The result from gear 13 sensor J has consistently high values from sample 9 to sample 14. The only sample for sensor J in the range 9 to 14 which has not satisfied the selection criteria is sample 12, since sample 13 was discarded due to poor convergence of the signal average.

5.3.3.3 Analysis Of Complete Gearbox Using One Unsupervised System

At this stage it was difficult to compare the overall performance of the gearbox as the analysis was split into two groups depending upon whether VIS was computed for that gear. So it was decided to investigate combining the data from the gears with VIS and without VIS into one system by the exclusion of VIS from the processing. The idea being that the interpretation of all the gears' behaviour could be implemented in a single analysis in which the trend analysis criteria would depend on the characteristics of a single cluster.

The results showed that the later samples from gear 13 displayed a significant deviation from their normal behaviour. Gears 1, 2, 3 and 4 were also showing significant changes and this was consistent with the previous results.

Single occurrences of particular gears appeared in the abnormal behaviour listing but only single samples and not from different sensors. Investigation of this revealed that the gears' presence was not due to the fact that there was a large increase in a GI value but because the pre-initiation samples were extremely stable, the resultant small standard deviation (s.d.) values falsely amplifying the normalised results.

5.3.3.4 Pre-Initiation Stable Parameter Compensation

Some parameter values varied only very slightly during the pre-initiation phase, resulting in a very small s.d. value for that particular GI value.

Similarly, some parameter values were constant during the pre-initiation phase, resulting in a zero value for the relevant s.d.. Initially the s.d. values for these parameters were set to unity. However this will tend to diminish the importance of these parameters compared to those which exhibit a fraction more variability during the pre-initiation phase. The result being that some gears might be incorrectly highlighted and some distinct changes overlooked.

For example, compare the following *(hypothetical)* stability (STB) values for two different gears

Gear (STB)	GEAR 04	GEAR 01
Sample Number		
1	1.0	1.0
2	1.0	1.0
3	1.0	0.982
4	1.0	1.0
5	1.0	1.0
Average	1.0	0.9964
S.D.	0.0	0.012
6	0.98	0.98
7	0.95	0.95
8	0.90	0.90
9	0.88	0.88
10	0.85	0.85
11	0.80	0.80

Table 19 STB values for Gear 04 and Gear 01 (Hypothetical Data)

As can be seen, the trends in values for samples 6 to 11 are the same for both gears but GEAR 04 has a s.d. of 0 and therefore the trend is not highlighted since the post initiation values will be normalised by unity. Whereas GEAR 01 with a very small s.d. will produce a very large value when standardised, this example illustrates the situation which had to be overcome.

Ideally it would be better to obtain more pre-initiation data so that an accurate s.d. could be calculated. However the problem was compensated for by constraining the minimum s.d. value. The minimum s.d. value was based upon the mean value of the pre-initiation samples. By comparison of the non-zero pre-initiation s.d.s with the equivalent mean values, a judgement was made to set the s.d. for any data with an s.d. less than 2% of the average value, to 2% of the average value.

The results obtained using stable value compensation are shown in Table 20. (Note that the previously used AVC*>233 criterion was not applied to the results shown in Table 20. However, the effect of applying this criterion can be inferred from examination of the numerical results in Table B2 of Annex B).

Decision Criteria

Distance From Cluster **(InvC*)**
Trend Detection Algorithm **(TPC*)** <0.5
Average Distance **(AVC*)**
Standard Deviation **(SDC*)** >100

Sample 7 MEAS_N	Sample 8 MEAS_N	Sample 9 MEAS_N	Sample 10 MEAS_N	Sample 11 MEAS_N	Sample 12 MEAS_N	Sample 13 MEAS_N	Sample 14 MEAS_N
	GEA13_F	GEA14_D	GEA14_D	GEA14_D	GEA14_E	GEA14_E	GEA14_D
	GEA13_J	GEA13_J	GEA13_E	GEA14_J	GEA12_C	GEA14_J	GEA14_E
	GEA01_E	GEA12_K	GEA13_F	GEA13_D	GEA12_F	GEA13_D	GEA14_J
	GEA10_I	GEA01_E	GEA13_J	GEA13_E	GEA01_D	GEA13_E	GEA13_D
		GEA01_E	GEA12_K	GEA13_J	GEA01_I	GEA12_C	GEA13_E
		GEA02_A	GEA01_D	GEA12_C	GEA02_F	GEA12_F	GEA13_F
		GEA03_B	GEA01_E	GEA12_F	GEA02_E	GEA01_D	GEA13_J
		GEA03_I	GEA01_I	GEA01_E	GEA02_G	GEA02_E	GEA12_C
		GEA03_B	GEA02_D	GEA01_I	GEA03_A	GEA02_G	GEA12_F
		GEA04_B	GEA02_E	GEA02_D	GEA03_I	GEA02_H	GEA01_D
		GEA05_B	GEA02_G	GEA02_G	GEA06_A	GEA03_A	GEA02_D
		GEA06_A	GEA03_A	GEA02_H	GEA08_I	GEA03_I	GEA02_E
		GEA08_I	GEA03_G	GEA03_K	GEA10_C	GEA04_K	GEA02_F
		GEA11_G	GEA03_B	GEA04_K	GEA11_G	GEA06_A	GEA02_G
		GEA11_H	GEA03_I	GEA05_B		GEA11_G	GEA03_I
			GEA04_K	GEA06_A			GEA04_K
			GEA05_B	GEA07_I			GEA06_A
			GEA06_A	GEA08_I			GEA07_C
			GEA07_I	GEA10_C			GEA11_F
			GEA10_C	GEA11_F			GEA11_G
			GEA10_H	GEA11_H			
			GEA11_G				
			GEA11_H				

Table 20 Results of applying stable parameter compensation

(The relevant INVC, TPC, AVC, and SDC parameter values are tabulated in Annex B)

These results clearly show that gear 13, and gears 1 and 2 are trending away (TPC<0.5) from their normal behaviour at a significant rate (SDC>100), with the results from particular sensors returning consistently abnormal behaviour.

At this stage, the analysis was clearly showing the capability of unsupervised analysis to identify abnormal behaviour and categorise the severity of changes using the trending characteristics. The SGDS scheduler also calculates energy-related parameters, such as 1R and 2R information, and it was decided to include these in the analysis to see if an improved result could be obtained.

5.3.3.5 Addition of the energy parameters 1R and 2R

The final clustering analysis was carried out using IMP, STB, WEA, AMIK, AMIE, EIIE, EIIV, 1R & 2R. (As noted in Section 5.2, the data was pre-filtered to ensure that cases where signal average convergence was poor were excluded to ensure that stable 1R and 2R values were used). As a result of the different analysis configuration, there was a greater spread of trend analysis values. Consequently the absolute values used for the criteria were adjusted accordingly. In addition, the results seemed less variable than previously, and it was possible to set selection criteria based upon the INVC1 value of the individual sample, rather than using AVC1 (the average value of INVC1 in a window of 5 samples) as was previously necessary for fault 1. The results from this analysis are shown in Table 21.

INVC1	>433
TPC1	<0.5
AVC1	
SDC1	>433

Sample 7 MEAS_N	Sample 8 MEAS_N	Sample 9 MEAS_N	Sample 10 MEAS_N	Sample 11 MEAS_N	Sample 12 MEAS_N	Sample 13 MEAS_N	Sample 14 MEAS_N
	GEA13_F	GEA13_J	GEA13_E	GEA13_D	GEA12_E	GEA14_F	GEA14_D
		GEA01_E	GEA13_F	GEA13_E	GEA02_E	GEA14_J	GEA14_E
		GEA02_E	GEA13_J	GEA13_J	GEA03_K	GEA13_D	GEA14_F
		GEA02_F	GEA12_E	GEA12_C		GEA13_E	GEA13_D
			GEA01_E	GEA01_E		GEA01_D	GEA13_E
			GEA02_E	GEA02_F		GEA01_E	GEA13_J
			GEA02_F	GEA03_K		GEA01_K	GEA12_E
			GEA10_C	GEA10_C		GEA02_E	GEA01_D
						GEA03_C	GEA01_E
							GEA01_K
							GEA02_E
							GEA03_C

Table 21 A combined analysis of all the gears including the 1R and 2R parameters

(The relevant INVC, TPC, AVC, and SDC parameter values are tabulated in Annex B)

The result shows that the unsupervised system has detected significant changes in the behaviour of various gears at a very early stage. With generally increased numbers of 'alerts' generated as the fault developed. The alerts tend to be from gears 12,13, 15, 1 and 2, with gear 13 exhibiting the greatest number of alerts. From sample 8 it can be seen that gear 13 could therefore be considered as a suspect gear. This suspicion is reinforced as each additional sample is analysed.

At this stage MJAD were told that the faulty gear was gear 13.

A review of the previous results will quickly identify that gear 13 was always the most likely fault gear. The AI system's ability to detect the fault was clearly proven.

The presence of gear 1, gear 2 and gears 12 and 14 in the above table, does not diminish the result in any way. Indeed their presence is useful to confirm the result because these gears are in the same area of the gearbox and it may not be

unreasonable to expect the disruption of one gear to affect the behaviour of one meshing with it or close by.

This result confirms the recommendation in the earlier report that gears should be considered in localised groups.

Closer examination of the normalised GI data revealed that the initial indications of fault behaviour were detected at sample 8 by the Enhanced Impact Indicator (EIIE and EIIV), the 1R and 2R also recorded increases. As the fault progressed the Amplitude Modulation Indicator increases together with the Wear (WEA) indicator. Towards the final stages of failure the impact indicator (IMP) returned significantly high values. The Stability Indicator (STB) showed a consistently decreasing value with increasing gearbox usage.

It is important to note that all four sensors used to monitor the gear 13 provided useful information giving an indication of the gear's deteriorating condition. This meant that even though some samples had been lost due to data integrity checks, the overall system was able to deal with the situation and still return at least one fault instance for every sample from 8 onwards except for sample 12.

The behaviour of the standardised GI values and the 1R and 2R energy parameters are detailed in Annex C.

5.3.4 *Spectrometric Oil Analysis*

As for the first seeded defect, the pre-fault initiation SOA samples for the second defect only contained non-zero values of iron and zinc (Fe and Zn). The raw SOA values for the second defect are shown in Figure 14, from which it is apparent that the pre-initiation values (samples 1 to 6) were again indicative of an approximately linear trend of wear.

5.3.4.1 Additional SOA Data Pre-Processing

Information on oil additions was available for the fault 2 SOA data, which permitted compensation of the results for oil consumption. The compensated results are shown in Figure 15. The results were compensated further for the normal wear rates occurring during the pre-initiation samples (Figure 16). Since there was no information regarding the intervals at which the samples were taken, the wear rate compensation assumed equi-spaced intervals. It was considered useful to apply this compensation even though the sampling intervals were unknown, since it is illustrative of the processing which would be necessary in an in-service system. In addition, the resultant increasingly negative compensated values indicate that the time interval between samples may well decrease with increasing sample number.

5.3.4.2 SOA Data Unsupervised Analysis

The fault 1 and fault 2 pre-initiation SOA data only contained values greater than 1ppm for Fe, and Zn. Unsupervised analysis of the fault 2 post-initiation SOA data immediately revealed the presence of **abnormal amounts of Cu from sample 8 onwards**.

Cu can be generated from wear of copper, brass or bronze materials. Typical % elemental composition of brass and bronze are shown in Table 22.

Alloy Material	Copper (Cu)	Zinc (Zn)	Tin (Sn)	Lead (Pb)
Brass	66 %	33 %		
Bronze	89 %	0.5 %	10 %	0.5 %

Table 22 Elemental Composition of Brass and Bronze

The lack of any significant change in the compensated Zn values indicated that the wear was not attributable to a brass part. Further differentiation between Copper (Cu) and Bronze parts was not possible since tin (Sn) and lead (Pb) values were not available in the SOA data. In addition, since the SOA values provided have a resolution of 1 ppm, there would have to be 170 ppm Cu generated before Pb could be detected. The maximum Cu value measured for the second seeded defect was 16 ppm. It is believed that oil analysis can be performed with an order of magnitude greater sensitivity, and that Sn and Pb levels can be monitored. Such improvements would have considerably increased the scope for analysis of the data.

Figure 15 SOA data for second seeded fault (raw values)

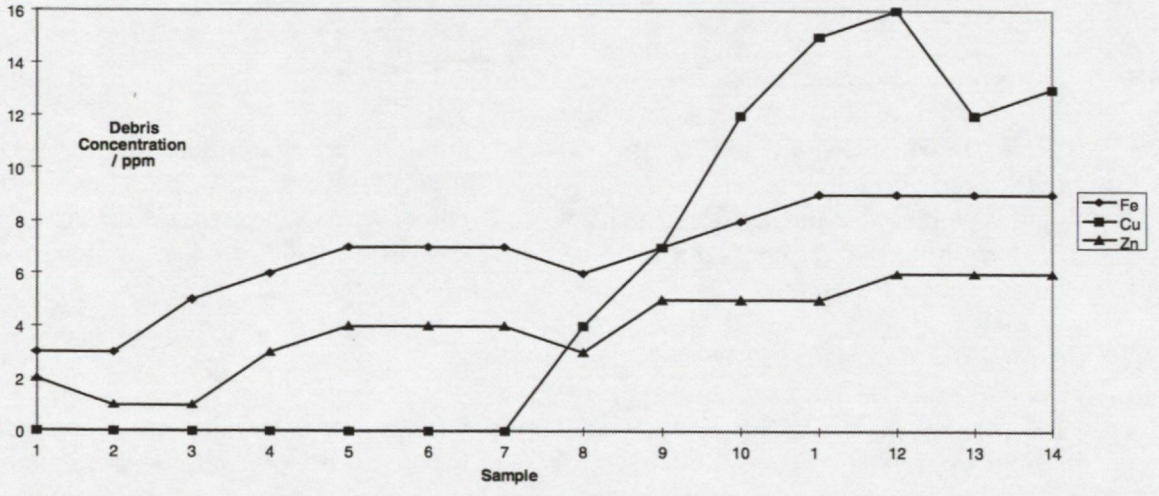


Figure 16 SOA data for second seeded fault (compensated for oil loss)

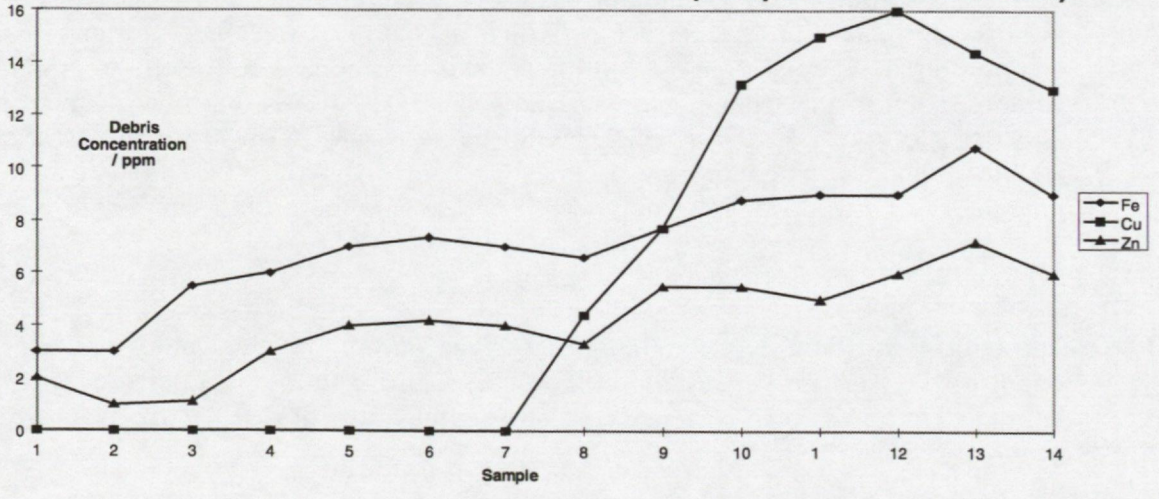
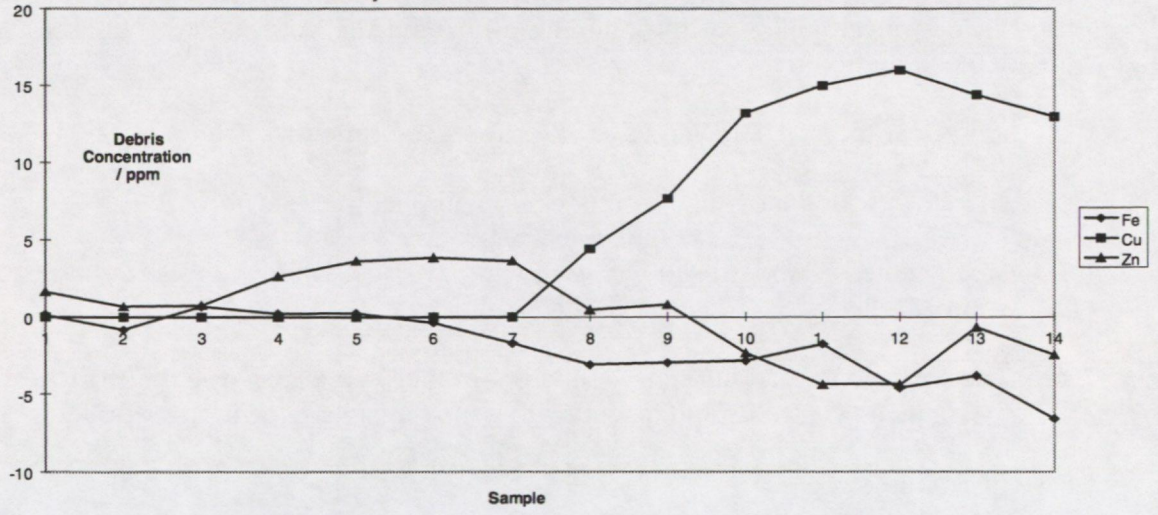


Figure 17 SOA data for second seeded fault (compensated for oil loss and normal wear)



There are a range of copper, brass and bronze parts within the S61 MRGB. Without increased resolution, or a greater range of elements, it was impossible to unambiguously identify the source of the debris generated.

Although it was not possible to unambiguously identify the source of the Cu debris from SOA analysis alone, it was possible to 'fuse' the information manually with the results of the vibration data analysis. Combination of the results of the two independent analyses allowed the possible sources of Cu to be correlated with the fact that the AI analysis of the vibration data independently identified a problem with gear 13. There is a copper-based component associated with this gear. The high levels of Cu identified from analysis of the SOA data therefore provide further substantiation that there was indeed a problem with gear 13. When the gearbox was stripped, the copper-based component associated with gear 13 (i.e. the gear with the seeded defect) was found to have experienced severe wear.

In this way, on the assumption that the vibration and SOA results are correlated, the vibration data can be used to pin-point the source of the oil debris more accurately than if the existing SOA analysis was used in isolation. In addition, the fact that oil debris is generated suggests that the gear fault detected by the vibration analysis is either causing, or exacerbating, wear of the copper-based component.

The SOA data clearly provided valuable information which has been manually fused at a higher level in the decision process. Unfortunately it can not currently be fused effectively at a low level in the analysis because the SOA data relates to the gearbox health as a whole and not individual gears, and would require detailed alloy composition information for the detection of individual component wear.

5.3.5 *Supervised analysis of the second seeded fault*

The second defect was seeded on a gear for which VIS cannot be calculated due to phase cancellation affects. The IFE algorithm was implemented using IMP, STB, WEA, AMIK, AMIE, EIIV and EIIE (i.e. the same indicators as were used for the first seeded defect, with the exclusion of VIS).

In order to demonstrate the importance of data pre-processing, IFE was performed after each of four processing stages. Stages 1 to 3 replicate processing previously used in the development of the fault 1 set-up (and the equivalent sections for the first defect are referenced).

1 Raw Gear Indices (GI) Values (c.f. Section 4.1 for the first defect).

2 GI values with a moving average window applied (c.f. Section 4.2.1).

3 Mean centred standardised with a moving average window applied (c.f. Section 4.2.2).

4 The inclusion of additional parameters (which were not available for the first fault), mean centred standardised with a moving average window applied.

5.3.5.1 System Training

As for the unsupervised stage, data that was identified as being potentially corrupt due to poor convergence of the signal average was removed from the analysis.

The faulty gear in this case was Gear 13. Samples 7-14 from this gear were considered faulty (the number of fault samples varied for each sensor because some had been removed because of poor convergence).

The sensors were first analysed collectively and then individually. The measurements which are being isolated are shown in the 'sensors' column in Table 23. (Note that because the boundaries which are determined for each sensor individually are different to those for the collective analysis, there is no guarantee that the sum of the faults bounded, or total samples bounded for each sensor will equal the value from analysis of all the sensors collectively).

The % fault capture rate is the measure of how successful the IFE process has been in isolating all the faulty samples. The % false alarm rate is a measure of how many healthy samples have been bounded incorrectly by IFE (as a proportion of the total number of samples bounded).

1 Raw GI values

Sensors	Number of Samples	Number of fault samples	Faults bounded	Total bounded	% Fault Capture rate	% False Alarm rate
GEA13 All Sensors	703	25	7	10	28	30
Sensor D	703	5	1	1	20	0
Sensor E	703	7	1	1	14	0
Sensor F	703	6	2	2	33	0
Sensor J	703	7	1	1	14	0

Table 23 Raw GI values

This result was similar to that obtained for fault 1, where an initial overall fault capture rate of 42.8% was obtained, compared with a value of 28% in Table 23 above.

2 GI values with a moving average window applied (two data points).

Sensors	Number of Samples	Number of fault samples	Faults bounded	Total bounded	% Fault Capture rate	% False Alarm rate
GEA13 All Sensors	643	25	3	3	12	0
Sensor D	643	5	1	1	20	0
Sensor E	643	7	2	2	29	0
Sensor F	643	6	5	6	83	17
Sensor J	643	7	5	6	71	17

Table 24 GI values with a moving average window applied

The results in Table 24 show an increased fault capture rate for individual sensors, particularly sensors F and J.

3 Mean centred standardised with a moving average window applied

The data has been modelled by determining the change in each of the parameter values as a proportion of the spread of the results from a healthy gearbox build. The data was presented in terms of the deviation from the mean, with a moving average window of two applied.

Sensors	Number of Samples	Number of fault samples	Faults bounded	Total bounded	% Fault Capture rate	% False Alarm rate
GEA13 All Sensors	643	25	4	5	16	20
Sensor D	643	5	4	4	80	0
Sensor E	643	7	3	3	43	0
Sensor F	643	6	5	5	83	0
Sensor J	643	7	4	4	57	0

Table 25 Mean centred standardised with a moving average window applied

This technique returns a marked improvement for sensors D and E.

4 The inclusion of additional parameters, mean centred standardised with a moving average window applied.

Two additional parameters, 1R and 2R, have been included. (These parameters were not available in the database for the first fault).

Sensors	Number of Samples	Number of fault samples	Faults bounded	Total bounded	% Fault Capture rate	% False Alarm rate
GEA13 All Sensors	643	25	18	22	72	18
Sensor D	643	5	4	4	80	0
Sensor E	643	7	6	6	86	0
Sensor F	643	6	5	5	83	0
Sensor J	643	7	5	6	71	16

Table 26 The inclusion of additional parameters, mean centred standardised with a moving average window applied

An extremely well defined fault area has been defined with 3 sensors showing good fault capture results and no false alarms.

5.3.5.2 System Testing (Classification Of Unseen Data)

Additional data was made available (designated samples 15 to 24) this was used to test the IFE technique. The test consisted of classifying the new data points against the IFE boundaries defined for the second seeded defect (i.e. gear 13). As most success had been achieved with the pre-processed data, (sub paragraph 4 above), these boundaries were chosen and the 'unseen data' processed accordingly. Samples 15 to 24 have been acquired from the test rig in sequence but no indication was given as to whether the samples were healthy or not.

The sensors were first analysed collectively i.e. the post initiation samples for gear 13 from every sensor were used to define the fault boundaries and then each sensor was considered individually.

In order to measure the success, the samples available for classification have been shown in the table column 'available samples'.

Subsequent to performing the analysis MJAD were informed that samples 17 to 24 were post-initiation. This enabled the fault capture rate to be computed.

1 Classification of unseen data to pre-defined boundaries

Collective Analysis (all sensors)

Gear Measurement	Sensor Name	Available Sample Numbers	Sample Numbers Bounded	% Fault Capture rate
GEA13	Sensor D	15,17,20,22, 23,24	20	20
	Sensor E	15,16,17,18,19, 20,21,22,24	16,19,20, 21,22	57
	Sensor F	20,21,24	–	0
	Sensor J	16,17,18,20,24	20	25

Table 27 Classification Of All Measurements To The Collective Fault Boundaries

Table 27 shows the result of classifying the 'unseen' gear 13 data to the boundaries defined by all the sensors (72 % of the original data was used to define the boundaries). As can be seen Sensor E has shown a fair detection efficiency. (Note that the two point moving average process results in one less sample being available. This is taken into account when computing % Fault Capture Rate values in this Section of the report).

At this stage the measurements for all the gears and from every sensor have been classified and there were only 7 'false alarms', as shown in Table 28.

Gear Measurement	Sample Number
GEA14_F	21
GEA12_D	18
GEA12_D	20
GEA12_D	21
GEA02_D	24
GEA04_H	21
GEA13_E	16

Table 28 False alarm gear samples

It can be seen that sensor D appears to suffer from an increased number of false alarms, although since there were in actual fact approximately 100 measurements classified, the false alarm rate was quite small.

Tables 27 and 28 highlight the ability of the IFE algorithm to isolate the fault related samples. Particularly consistent results are obtained for sensor E, but even the results from sensors D and J are encouraging. What has to be borne in mind when examining the data is that the IFE fault enclosure area is tightly defined. It may be that there is slight variability in the GI values and the remaining samples are just outside the boundaries. Nevertheless in a practical system a fault alarm could have been triggered at sample 20 on the basis of a consecutive alarm from one sensor or as an alarm from two separate sensors. The false alarms are from gears in the same area as gear 13 and as there is obviously disruption of the general performance of the gearbox it is open to interpretation as to whether or not these would be considered as false alarms.

To ascertain whether improved performance could be obtained, the unseen data was then analysed using the enclosure areas defined for the individual sensors. The unseen data did not classify into any of the individual sensors' fault enclosure areas. Further investigation showed that for a given 'unseen' sample one or more of the GI values were just outside the boundary of the tightly defined enclosure area even though the actual behaviour of the GI and energy parameters was very similar. The Normalised GI values can be seen in Annex C.

2 Optimised Boundaries

The result of analysing the unseen data (i.e. samples 15 –24) indicated that the boundaries may have been over-tightly defined from a restricted data-set. The Supervised Learning technique draws boundaries around the faulty samples. The objective however is to detect faults with no false alarms, so it is a valid argument to move the boundaries to just before the next 'healthy' sample. So each of the boundaries was examined to see if it could be moved without capturing any of the healthy training data. The new enclosed area may be thought of as a 'non-healthy' area rather than a 'fault' area. At this stage any measurement which was not from gear 13 (sample 7–14) was considered to be healthy.

In order to optimise the boundaries as described above each of the sensors was considered separately. Table 29 shows the results. The available samples are shown in a column in order to enable the success of the classification to be judged.

The optimised boundaries have improved the classification of the unseen data. Sensor E has returned the most consistent result, with the correct classification of 3 consecutive samples, although a further improved result was anticipated by taking the optimisation one step further. As stated above, the boundaries were optimised by moving them to the nearest 'non-fault' sample and this included samples related to other gears.

Gear Measurement	Available Sample Numbers	Sample Numbers Bounded	% Fault Capture rate	Mis-classified results Gear I Sensor I Sample
GEA13 Sensor D only	15,17,20,22, 23,24	24	20	Gear 1 I Sensor D I 18
GEA13 Sensor E only	15,16,17,18, 19,20,21,22, 24	19,20,21	43	NONE
GEA13 Sensor F only	20,21,24	NONE	0	NONE
GEA13 Sensor J only	16,17,18,20, 24	16	0	GEA12I Sensor J I 20, 21 GEA13ISensor JI16

Table 29 Classification Of The Gear 13 Data On An Individual Sensor Basis

5.3.5.3 Optimisation of boundaries using fault gear data only

The data supplied relates to one gear/fault type. The IFE analysis so far has attempted to separate out the post-fault initiation data for this gear, from data which includes data for a range of other gears.

In practice, other gears could exhibit similar behaviour without any fault being present on those gears. In addition a given gear/fault type could manifest itself differently at each sensor location. It is therefore appropriate to repeat the IFE analysis using only the gear 13 data since any real HUMS system could be configured to only use the enclosure identified to detect a gear 13 fault from a given sensor.

Gear Measurement	Available Sample Numbers	Sample Numbers Bounded	% Fault Capture rate
GEA13 Sensor D only	15,17,20,22, 23,24	20,22,23, 24	80
GEA13 Sensor E only	15,16,17,18,19, 20,21,22,24	15,19,20, 21,24	57
GEA13 Sensor F only	20,21,24	NONE	0
GEA13 Sensor J only	16,17,18,20, 24	16,17,18, 20,24	100

Table 30 Classification Of The Unseen Data To Optimised Boundaries For Fault Gear Data

The results of this final optimisation process appear to be very successful, with very high capture rates for sensors D and J, and reasonably good results for sensor E. Whilst there were a couple of false alarms, it is likely that these could be removed at the expense of a reduced fault detection rate if required.

The reason for the poor behaviour of sensor F can be inferred from the graphical results presented in Figures 7 and 8 of Annex C. Inspection of this data, and comparison with the equivalent results for the other sensors, reveals a difference in the sensor F's parameter values between the training data, and the additional unseen data. The behaviour of the STB parameter in particular is noticeably different. Given the fact that sensor F is in a similar physical position relative to the fault gear compared to the other sensors, the physical reasons for this difference in behaviour are not clear. It is however noticeable that sensor J, which returned the best fault capture rate, was also the sensor which produced the best quality signal averages from the pre-initiation data.

6 OVERALL DISCUSSION OF THE RESULTS

6.1 Unsupervised Analysis

The unsupervised analysis has successfully identified the presence and location of both seeded faults. This has been achieved without knowledge of the faults' existence, and/or tailoring of the approach to suit the particular fault type and location. Furthermore, the approach adopted should not be specific to the S61 MRGB application, but could reasonably be expected to work for any other helicopter type's MRGB.

The only pre-requisite for the approach adopted is the existence of a set of data which is known not to contain faults (i.e. represents a 'healthy' gearbox). Whilst this might be more difficult to guarantee in an in-service application of the technique, the fact that there would be considerably greater quantities of data available, should offset this limitation.

In the work reported, a knowledge of the nature and location of the seeded fault was then successfully used to optimise the processing applied to the data. This would also be possible in a real application, as the results of stripping suspect gearboxes became available. This approach ensures that the system continues to build on experience, and so the performance should continue to improve with time, and would be able to reflect failure modes which might only become apparent once a gearbox has been in service for a period of time.

6.2 Supervised Analysis

It has been possible to train the iterative feature extraction technique used to recognise both the seeded defects. Furthermore, application of the data from the second defect to the boundaries identified for the first defect did not produce any false alarms.

The best results were obtained by making use of the maximum amount of information known about the fault. This included only applying the analysis to those sensors which might reasonably have been expected to have visibility of each fault, based upon their locations. This is an entirely reasonable constraint to apply when

configuring the system for practical in-service application. It also ensures that the fault can be localised to an individual gear, and that the same fault may be recognised in slightly different ways for different gears. Using this approach, fault capture rates of 80 to 100% were achieved with very low false alarm rates. In some cases zero false alarm rates were achieved during training (see, for example, the results in Tables 10 and 26).

6.3 Data Fusion

The restricted information content of the SOA data has precluded detailed investigation of automating the process of fusing vibration and SOA data to arrive at more positive conclusions.

The work performed has, however, illustrated how the two data sources can be used to corroborate and/or refine the fault diagnosis. For instance, in the first seeded defect, there were no trends visible in the SOA data. This would be consistent with the fault which was detected by the vibration being a crack, rather than the result of a wear mechanism. Conversely, for the second fault the SOA trends observed were consistent with wear, not for the gear itself but an associated copper-based component.

In this way the SOA data has produced a verdict (such as 'no wear present' or 'wear present in a Copper or Bronze part'), which can then be combined with the output from the vibration analysis (such as 'localised defect to gear 4'), to produce an overall diagnosis. The fusion of data would therefore be achieved by the process of combining the outputs from the individual subsystems. This would be entirely feasible to achieve in service, and would have the added advantage that it probably reflects the way in which the human expert currently arrives at an overall decision. This has advantages when such systems are put into service, since the human operator can relate to the way in which the system has arrived at a decision, and is therefore more likely to trust it.

6.4 Considerations for Practical Implementation

6.4.1 *Database Contents and Structure*

The degree of success which can be achieved using the Intelligent Data Management (IDM) techniques described for in-service routine analysis of HUMS data, is fundamentally constrained by the contents and structure of the database used to contain the HUMS information. The major requirements are outlined below :

- **Querying Facilities :** It must be possible for the system to automatically generate and refine requests for data. This might be by the use of a database engine which supports industry standard Structured Query Language (SQL).

- **Sampling Rates :** The system must be able to correlate data which may originate from different sources at different sampling rates (e.g. vibration measured routinely every flight, SOA data periodically sampled). Typically this would require different tables within the database schema, linked by a common field (such as airframe hours).

- **Maintenance information :** The system requires access to maintenance information which can be related to the primary data (i.e. SOA or vibration data). This can be used for three main purposes :

  - To screen inappropriate data (e.g. if a gearbox has been replaced, then data prior to the replacement should not be included in trends following the replacement).

  - To pre-process data (e.g. use the quantity of oil added to compensate SOA values).

  - To automate supervised learning (since a particular maintenance action might be used to represent a particular fault, then data immediately prior to the action could be assumed representative of that fault type and used for supervised training purposes).

Maintenance information should be recorded in a coded or numeric fashion to avoid any ambiguities that might otherwise arise through the use of free text, which would prevent an automated system from making efficient use of the information.

- **Quality of data :** The recorded values should maintain the original measurement accuracy (e.g. if the SOA values are measured to greater than 1ppm, then they should be recorded to that accuracy). Additionally, any parameters which indicate the quality of the measurement (such as the convergence of the signal average) should also be recorded, since they will permit the system to learn to reject particular measurements as being of inadequate quality.

- **FDR Data –** Another source of variability in HUMS data will be the flight conditions under which the data was acquired. Most HUMS systems acquire the data in specific 'flight regimes'. Hence an in-service system would be able to analyse data from each flight regime individually, and then 'fuse' the results to gain further corroboration of a particular fault. Rather than just record the flight regime which has been identified however, it is preferable to store the actual flight data recorder (FDR) parameters in the database. This is because the flight regimes are recognised from combinations of particular flight parameters falling within given ranges of values. Since there may be variability in the measured vibration data within a particular flight regime, access to the stored FDR parameters would enable identification of the cause of any such variability.

- **Component material information :** SOA data relate to particles suspended in oil samples. These particles could derive from any part which could be subject to wear within the gearbox. Without information on the materials used to make every part within the gearbox, it is impossible to accurately corroborate the results derived from independent analysis of vibration data.

6.4.2 *Unsupervised Analysis*

Unsupervised analysis works best with larger quantities of data. For this reason, the unsupervised analysis presented in this report considered all the data from the gearbox as a whole. For the routine analysis of in-service HUMS data, much greater quantities of data would be available. It is likely that greater sensitivity could be achieved by analysing the vibration data for each gear (or even each gear-

sensor/flight regime combination) individually. Given the greater quantities of data available, it is likely that there would be sufficient statistical spread of data to prevent any clusters associated with normal behaviour becoming fragmented. If fragmentation did occur, this could be identified automatically by the system. Additional data could then be added to the analysis sample (e.g. data from other gears, sensors, or flight regimes), or the criteria used to define clusters adjusted iteratively. These processes could be automated.

In the work reported, the major requirement for the initial identification of 'normal' clusters is the existence of a set of data which is known not to contain faults (i.e. represents a 'healthy' gearbox). This may be more difficult to guarantee in real HUMS data. However, provided that the majority of the data related to non-fault cases the approach should still be successful. Indeed, it may even be possible to use the initial analysis of the data to identify problem gearboxes, since these might form clusters which were distant from the main 'normal' cluster.

For the analysis of real data, the success of the approach will be critically dependent upon the quality of the data used. Prior to automated in-service unsupervised analysis of HUMS data, it will be essential to screen out any data which is known to be of poor, or questionable quality (e.g. where the signal average has not converged, or the data is near the border of a flight regime). However, since the approach adopted has made use of trends in the data, the analysis of more than one data sample will make the analysis less sensitive to spurious indications, than an alerting system based upon analysis of individual samples. If the system detects trends which are subsequently attributable to factors other than gearbox health, it should then be possible to use the supervised learning facilities to detect future occurrences of the same event. Additional work would be required to confirm that the number of samples included in the trend analysis window length was adequate for in service application.

6.4.3 *Supervised Analysis*

The success of supervised analysis of in-service HUMS data will be dependent upon a knowledge of particular faults which have occurred.

Whilst it may be possible to automate the training process in the longer term, initially it is likely that it would have to be manually initiated. The training process would require that the system enabled the user to identify particular samples as relating to a fault. The data might derive from a number of instances of the same fault type. It is likely that the fault enclosures defined would be specific to particular gears as monitored by one or more sensors.

Once the fault enclosures were defined, they could routinely be used to analyse new data downloaded from the aircraft.

With the greater quantities of data available, it might also be worthwhile investigating more complex neural net configurations since they could effectively accommodate more complex fault enclosures than the IFE technique reported. However, the IFE technique can be applied relatively simply to existing HUMS systems, and can provide the user with an indication of how the system has learnt to recognise the fault.

7 CONCLUSIONS

The following conclusions are drawn from the work reported in this document :

Both unsupervised and supervised analysis of the vibration data was able to detect and locate both seeded faults.

The unsupervised techniques which were developed for the first seeded defect were found to be equally applicable to the second seeded defect.

The supervised analysis technique as configured for the detection of the fault 1 characteristics classified the pre-initiation samples from fault 2 as healthy. The configuration did not therefore generate any false alarms even though the fault 2 data had been acquired after the gearbox had been completely rebuilt.

Some gears have variability in the gear parameters they produced. A range of 'noise' rejection/trend analysis techniques were developed for efficient detection of the fault related abnormal behaviour.

SOA data produced no indication of the first defect, although a rig rebuild was detectable. This indicated that the fault was of a type which did not produce fine debris. This was subsequently confirmed.

An effective pre-processing set-up was used on the SOA data for the second defect which highlighted the fact that there was a gearbox problem from sample 8 onwards. The results from this analysis were judged to be most effectively fused manually at a higher level rather than treating the SOA data in the same way as each GI value. The inclusion of additional elements and/or increased resolution would have enabled identification of the source of the wear debris. SOA data may be of more use with increased accuracy since this might permit alloy matching at lower levels of debris concentration.

8 RECOMMENDATIONS

The use of AI techniques to analyse data from both the first and second seeded defects on the S61 MRGB has been very successful. The potential advantages from the inclusion of such techniques in HUMS to address the problems of recognising faults which the system has not previously seen, and automating the fault detection process, are evident.

To achieve successful implementation within in-service HUMS, further development and refinements are required as outlined below. These refinements reflect the limitations of the investigation possible with the data to date, namely that it is seeded fault rig data for one gearbox type gathered under well controlled conditions.

The refinements required could be progressively addressed (not necessarily in the order below), so that success at each stage could be clearly demonstrated before progressing to another stage.

1 **Restricted range of faults :** The techniques to date have only been demonstrated on two gear faults. Direct in-service implementation of the techniques without exposure to additional fault data could be premature.

Additional seeded fault data from the S61 rig should be used to test the developed AI techniques further. The MJAD SGDS scheduler software could be used to cost-effectively gather increased amounts of data. This would be particularly useful during the pre-initiation stages which would enable more accurate pre-processing of the data.

2 **Restricted baseline data :** The data to date has been derived from 2 gearbox builds. The variability in results between different gearboxes of the same type, and/or different builds of the same gearbox is therefore not known. In a real system it would be preferable not to have to establish each aircraft's normal vibration pattern before pre-processing could be carried out. Analysis of the pre-initiation samples from a number of gearbox builds would help establish if a healthy baseline dataset for a particular gearbox type was achievable, or whether fault detection sensitivity was compromised.

3 **Single gearbox type :** To date the techniques have been successfully demonstrated on the S61 MRGB. This is only one of many gearbox types which HUMS systems monitor. The applicability of the techniques to other gearbox types should be investigated prior to implementation within HUMS. Data from an alternative seeded fault programme could be used for this purpose, if such data were available.

4 **Restrictions of seeded fault rig data :** The data to date has been derived from seeded faults in a gearbox rig. An artificial fault generation mechanism has therefore been used, and the rig operated under well controlled (known) conditions. Naturally occurring 'real world' faults may therefore exhibit different characteristics, and the conditions under which data is gathered will almost certainly be less well controlled, resulting in a greater degree of variability in the data. The use of vibration data recorded by an in-service HUMS to test the effect of such variability on the performance of the system as set up, would certainly be a significant step forward towards removing these limitations. Obviously any data containing a fault will be very scarce. If an undetected fault was discovered during routine overhaul the preceding HUMS data could be examined for evidence. An alternative approach would be the examination of HUMS data where a number of false alarms have been generated. In this case, the AI techniques could be used to determine their susceptibility to similar false alarms.

3 **SOA data limitations :** Analysis of the SOA data clearly showed its ability to be processed into a very useful health indicator. However, the following additional information would greatly increase its ability to discriminate the source of wear debris :

- resolution greater than 1ppm
- increased range of elements
- details of parts' material alloy elemental composition
- details of operating hours

The improvements in the SOA analysis techniques developed with the above information could be investigated using either rig and/or in-service data. In-service data could also be used to ascertain the degree to which the rig data was representative of real data, and to provide greater quantities of data.

Premature implementation of the techniques into in-service HUMS without additional investigation as outlined above could lead to poor initial performance and end user perception. However, given the correct additional development results, the techniques are capable of alleviating many of the problems experienced with existing HUMS systems, and greatly increasing the systems' performance.

Annex A SGDS Gear Indicators

Table A1 : SGDS Gear Indicators

GI	Long Name	Function
VIS	Visibility Indicator	To indicate how visible a gear is in the signal average.
IMP	Simple Impact Indicator	To detect the presence of localised gear defects which produce low level impulses in the data.
STB	Stability Indicator	To detect the presence of strong submesh frequency components.
WEA	Wear Indicator	To detect the presence of distributed faults on gears.
AMIK	Amplitude Modulation Indicator (Kurtosis)	To detect the localised gear defects from their effect on the gearmesh component.
AMIE	Amplitude Modulation Indicator (Energy)	The detection of modulation effects associated with whole gear faults (e.g. misalignment, eccentricity).
EIIV	Enhanced Impact Indicator Value	To detect the presence of localised gear defects which produce extremely low level impulses in the data.
EIIE	Enhanced Impact Indicator Event count	To detect the presence of localised gear defects which produce extremely low level impulses in the data

Annex B Tabulated Unsupervised Analysis Results

Table B1 Results of applying the alternative pre-processing technique for unsupervised analysis of the data from gears without VIS.

(Note : This table details the parameter values for the results presented in Table 17 in the main body of the report)

Decision Criteria

Distance From Cluster **(InvC*)**	*Descending*
Trend Detection Algorithm **(TPC*)**	< 0.5
Average Distance **(AVC*)**	>233
Standard Deviation **(SDC*)**	>100

MEAS_N	Sample	InvC1	TPC1	AVC1	SDC1
GEA02_E	14	7126	0.242	6676	1426
GEA01_E	14	3698	0.117	2956	956
GEA02_F	14	3392	0.242	931	1234
GEA14_D	14	1844	0.042	758	870
GEA03_I	14	1209	0.042	878	250
GEA04_K	14	1177	0.117	891	477
GEA02_G	14	1032	0.042	705	247
GEA02_D	14	992	0.242	698	221
GEA01_F	14	940	0.042	377	284
GEA06_A	14	931	0.242	838	323
GEA11_G	14	775	0.117	540	262
GEA09_B	14	752	0.042	422	230
GEA08_A	14	434	0.042	286	114
GEA02_E	13	8624	0.242	6578	1409
GEA01_E	13	4068	0.242	2637	921
GEA14_D	13	1801	0.242	403	700
GEA04_K	13	1596	0.408	922	499
GEA06_A	13	1343	0.117	746	348
GEA03_I	13	1044	0.242	769	194
GEA02_D	13	914	0.408	597	174
GEA02_G	13	888	0.242	607	188
GEA11_G	13	843	0.242	470	236
GEA09_B	13	629	0.242	331	162
GEA07_I	13	477	0.117	352	103
GEA01_I	13	437	0.117	538	195
GEA02_F	13	327	0.242	276	116
GEA02_E	12	7525	0.242	5405	1638
GEA01_E	12	3382	0.242	1993	815
GEA03_I	12	946	0.242	626	201
GEA06_A	12	942	0.117	511	247
GEA02_G	12	760	0.242	459	201
GEA11_G	12	598	0.242	329	173
GEA01_I	12	373	0.042	481	250
GEA10_C	12	37	0.408	665	765

Table B1 (continued) : Results of applying the alternative pre-processing technique for unsupervised analysis of the data from gears without VIS.

(Note : This table details the parameter values for the results presented in Table 17 in the main body of the report)

MEAS_N ....cont.	Sample	InvC1	TPC1	AVC1	SDC1
GEA02_E	11	4755	0.242	3899	2319
GEA01_E	11	1802	0.242	1321	775
GEA10_C	11	1532	0.117	664	766
GEA01_I	11	758	0.042	413	310
GEA03_K	11	598	0.042	238	284
GEA02_D	11	576	0.042	453	155
GEA03_I	11	527	0.242	474	188
GEA07_I	11	469	0.008	242	153
GEA06_A	11	462	0.117	327	196
GEA05_I	11	385	0.117	354	283
GEA02_G	11	384	0.242	305	206
GEA02_E	10	5347	0.042	2945	2717
GEA01_E	10	1828	0.042	955	886
GEA10_C	10	1669	0.042	356	657
GEA01_I	10	789	0.008	260	290
GEA05_I	10	757	0.042	281	311
GEA03_I	10	662	0.042	368	263
GEA02_D	10	608	0.008	344	212
GEA04_K	10	401	0.117	508	508
GEA10_I	10	273	0.408	242	128
GEA02_E	9	6637	0.117	1884	2605
GEA01_E	9	2106	0.042	589	826
GEA04_K	9	1331	0.117	431	547
GEA03_I	9	664	0.042	235	249
GEA02_E	8	2762	0.408	557	1103

Table B2 Results of applying pre-initiation stable parameter compensation for unsupervised analysis of the data from all the gears

(Note : This table details the parameter values for the results presented in Table 20 in the main body of the report)

Decision Criteria

Distance From Cluster **(InvC*)**	*Descending*
Trend Detection Algorithm **(TPC*)**	< 0.5
Average Distance **(AVC*)**	
Standard Deviation **(SDC*)**	>100

MEAS_N	Sample	InvC1	TPC1	AVC1	SDC1
GEA13_E	14	6441	0.117	1675	2392
GEA13_D	14	3876	0.008	1787	1251
GEA13_J	14	3584	0.242	1057	1274
GEA13_F	14	2778	0.408	720	1034
GEA01_D	14	2605	0.008	1249	918
GEA02_E	14	1650	0.042	1386	374
GEA14_D	14	1559	0.242	877	573
GEA03_I	14	964	0.242	861	218
GEA02_F	14	928	0.008	327	301
GEA06_A	14	897	0.242	810	315
GEA02_G	14	873	0.042	594	209
GEA04_K	14	821	0.117	524	328
GEA11_G	14	575	0.117	398	205
GEA14_J	14	572	0.042	325	226
GEA12_C	14	417	0.408	268	161
GEA14_E	14	408	0.042	257	177
GEA11_F	14	312	0.242	187	104
GEA07_C	14	274	0.042	113	114
GEA02_D	14	242	0.117	315	172
GEA12_F	14	170	0.408	355	306
GEA13_D	13	2234	0.008	992	877
GEA01_D	13	2055	0.008	723	723
GEA02_E	13	1926	0.008	1049	645
GEA06_A	13	1302	0.117	721	339
GEA03_I	13	1156	0.042	716	320
GEA04_K	13	1015	0.042	366	337
GEA02_G	13	741	0.042	455	209
GEA11_G	13	648	0.242	347	186
GEA14_J	13	582	0.008	199	229
GEA14_E	13	448	0.042	182	177
GEA13_E	13	417	0.117	401	268
GEA03_A	13	381	0.408	294	134
GEA02_H	13	251	0.042	135	127
GEA12_C	13	128	0.242	179	178
GEA12_F	13	95	0.117	321	333

Table B2 (continued) : Results of applying pre-initiation stable parameter compensation for unsupervised analysis of the data from all the gears

(Note : This table details the parameter values for the results presented in Table 20 in the main body of the report)

MEAS_N....cont.	Sample	InvC1	TPC1	AVC1	SDC1
GEA02_E	12	1281	0.008	657	585
GEA03_I	12	969	0.042	482	338
GEA06_A	12	914	0.117	494	241
GEA12_F	12	735	0.008	298	352
GEA01_D	12	720	0.042	314	320
GEA02_G	12	646	0.042	299	226
GEA01_I	12	472	0.408	512	222
GEA11_G	12	429	0.408	242	124
GEA14_E	12	334	0.042	78	139
GEA12_C	12	299	0.042	145	198
GEA08_I	12	230	0.242	157	110
GEA03_A	12	224	0.242	215	171
GEA10_C	12	31	0.408	628	734
GEA13_D	11	1500	0.117	660	621
GEA10_C	11	1461	0.117	626	736
GEA13_E	11	761	0.042	336	294
GEA12_F	11	720	0.117	160	281
GEA01_E	11	696	0.242	539	312
GEA01_I	11	685	0.042	409	316
GEA03_K	11	568	0.117	215	279
GEA02_D	11	520	0.008	266	215
GEA12_C	11	457	0.408	802	1403
GEA13_J	11	451	0.408	349	240
GEA06_A	11	444	0.117	314	191
GEA02_G	11	325	0.042	162	177
GEA02_H	11	319	0.008	82	122
GEA14_D	11	307	0.242	289	200
GEA14_J	11	305	0.042	74	138
GEA07_I	11	268	0.008	174	107
GEA11_F	11	256	0.008	121	107
GEA04_K	11	254	0.042	163	124
GEA08_I	11	250	0.042	106	123
GEA05_B	11	44	0.242	114	114
GEA11_H	11	7	0.242	211	271

Table B2 (continued) : Results of applying pre-initiation stable parameter compensation for unsupervised analysis of the data from all the gears

(Note : This table details the parameter values for the results presented in Table 21 in the main body of the report)

MEAS_N....cont.	Sample	InvC1	TPC1	AVC1	SDC1
GEA10_C	10	1590	0.042	331	630
GEA02_E	10	1238	0.117	396	537
GEA01_I	10	774	0.042	271	316
GEA01_E	10	727	0.042	393	370
GEA01_D	10	670	0.042	161	268
GEA13_E	10	622	0.042	178	229
GEA03_I	10	596	0.042	283	281
GEA03_K	10	541	0.408	103	221
GEA11_H	10	530	0.042	200	280
GEA02_D	10	519	0.008	160	194
GEA06_A	10	492	0.008	218	219
GEA14_D	10	483	0.042	211	247
GEA02_G	10	384	0.117	93	167
GEA03_A	10	368	0.117	170	191
GEA04_K	10	303	0.042	113	128
GEA07_I	10	258	0.008	116	118
GEA11_G	10	244	0.042	146	123
GEA10_I	10	198	0.242	141	110
GEA05_B	10	174	0.042	102	123
GEA12_K	10	142	0.117	136	103
GEA13_J	10	134	0.117	259	268
GEA13_F	10	122	0.117	152	167
GEA03_B	10	–29	0.408	56	116
GEA01_E	9	863	0.042	242	357
GEA02_E	9	833	0.117	141	346
GEA13_J	9	622	0.008	232	286
GEA03_I	9	621	0.117	167	246
GEA11_H	9	555	0.242	103	228
GEA14_D	9	500	0.117	107	219
GEA01_I	9	498	0.117	112	202
GEA06_A	9	455	0.117	128	176
GEA03_A	9	434	0.242	103	168
GEA11_G	9	319	0.042	100	121
GEA05_B	9	310	0.008	61	126
GEA12_K	9	277	0.042	105	119
GEA08_I	9	255	0.117	55	104
GEA03_B	9	197	0.408	62	112
GEA13_J	8	538	0.008	98	221
GEA01_E	8	426	0.042	61	183
GEA13_F	8	379	0.117	124	181
GEA10_I	8	282	0.042	100	119

Table B3 : Results of unsupervised analysis of data from all the gears, including the 1R and 2R parameters

(Note : This table details the parameter values for the results presented in Table 21 in the main body of the report)

Decision Criteria

Distance From Cluster **(InvC*)**	>433 *Descending*
Trend Detection Algorithm **(TPC*)**	< 0.5
Average Distance **(AVC*)**	
Standard Deviation **(SDC*)**	>433

MEAS_N	Sample	InvC1	TPC1	AVC1	SDC1
GEA14_F	14	20873	0.117	6031	7786
GEA13_E	14	9186	0.042	2926	3154
GEA13_D	14	7063	0.242	3774	2364
GEA01_E	14	3961	0.242	2358	884
GEA13_J	14	3930	0.408	1848	1043
GEA12_E	14	3197	0.042	1548	1080
GEA01_D	14	2732	0.042	1377	889
GEA02_E	14	2322	0.242	3694	2079
GEA14_E	14	2315	0.042	730	810
GEA03_C	14	1815	0.242	841	744
GEA14_D	14	1664	0.117	905	596
GEA01_K	14	1293	0.042	1082	437
GEA14_F	13	6609	0.408	1957	2464
GEA02_E	13	5949	0.042	3239	2529
GEA13_D	13	2243	0.242	2342	2092
GEA01_D	13	2083	0.042	831	710
GEA01_E	13	2039	0.242	1653	715
GEA01_K	13	1810	0.008	826	588
GEA03_C	13	1674	0.408	515	586
GEA13_E	13	1615	0.117	1219	483
GEA14_J	13	651	0.408	891	727
GEA02_E	12	6457	0.008	2048	2371
GEA12_E	12	1812	0.042	903	840
GEA03_K	12	1357	0.008	837	447
GEA13_D	11	4535	0.242	2017	2205
GEA13_E	11	1751	0.242	1035	472
GEA10_C	11	1462	0.242	661	710
GEA01_E	11	1396	0.117	1257	912
GEA03_K	11	1268	0.008	566	460
GEA13_J	11	1238	0.117	991	487
GEA02_F	11	1072	0.242	1302	912
GEA12_C	11	554	0.408	915	1522

Table B3 (continued) : Results of unsupervised analysis of data from all the gears, including the 1R and 2R parameters

(Note : This table details the parameter values for the results presented in Table 21 in the main body of the report)

MEAS_N ....cont.	Sample	InvC1	TPC1	AVC1	SDC1
GEA01_E	10	2557	0.008	971	1038
GEA02_E	10	2296	0.008	753	952
GEA02_F	10	2122	0.117	1087	1057
GEA12_E	10	1965	0.042	543	754
GEA10_C	10	1595	0.117	369	614
GEA13_E	10	1449	0.242	679	468
GEA13_J	10	1444	0.008	744	599
GEA13_F	10	1026	0.042	519	494
GEA02_F	9	2506	0.242	661	980
GEA01_E	9	1837	0.042	461	708
GEA02_E	9	1447	0.117	292	578
GEA13_J	9	1365	0.042	456	536
GEA13_F	8	1143	0.117	312	452

Annex C Normalised GI values from the second gear fault

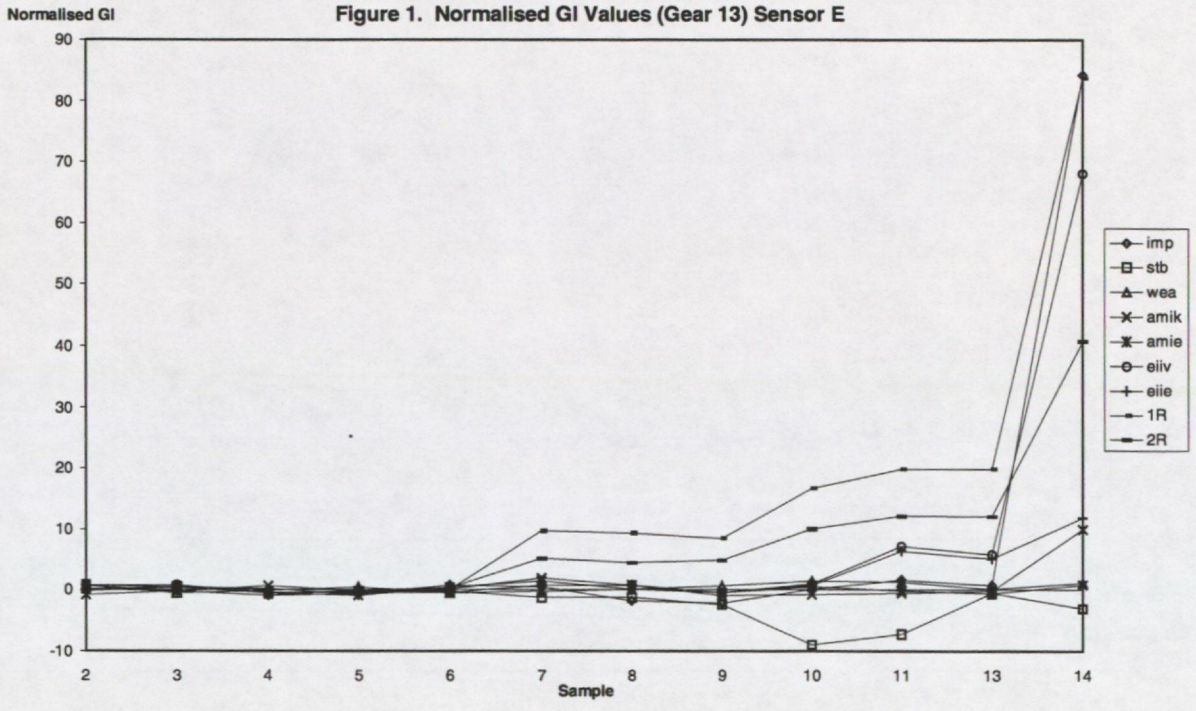


Figure 1. Normalised GI Values (Gear 13) Sensor E

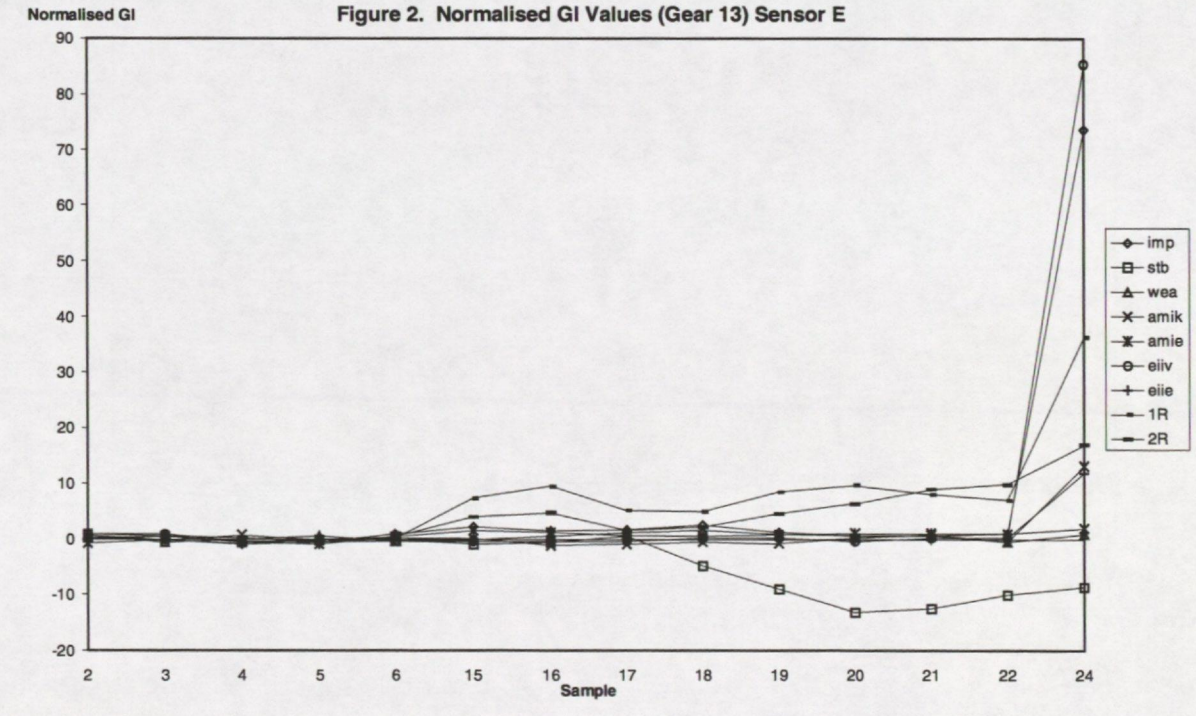


Figure 2. Normalised GI Values (Gear 13) Sensor E

Note :

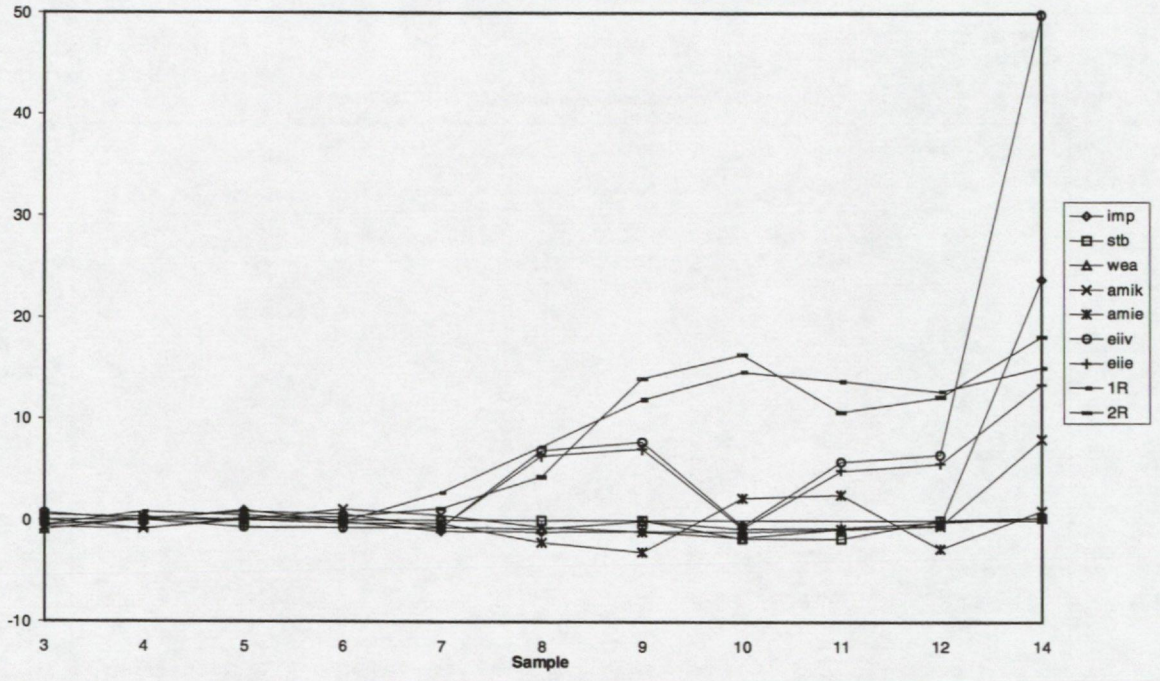
1 Figure 1 includes samples 7–14, which were used as training data for the supervised learning.

2 Figure 2 includes samples 15–24 (i.e. the additional 'unseen' data) which were used to test the supervised machine learning.

3 Only good quality data is plotted (i.e. poor signal average convergence data has been removed).

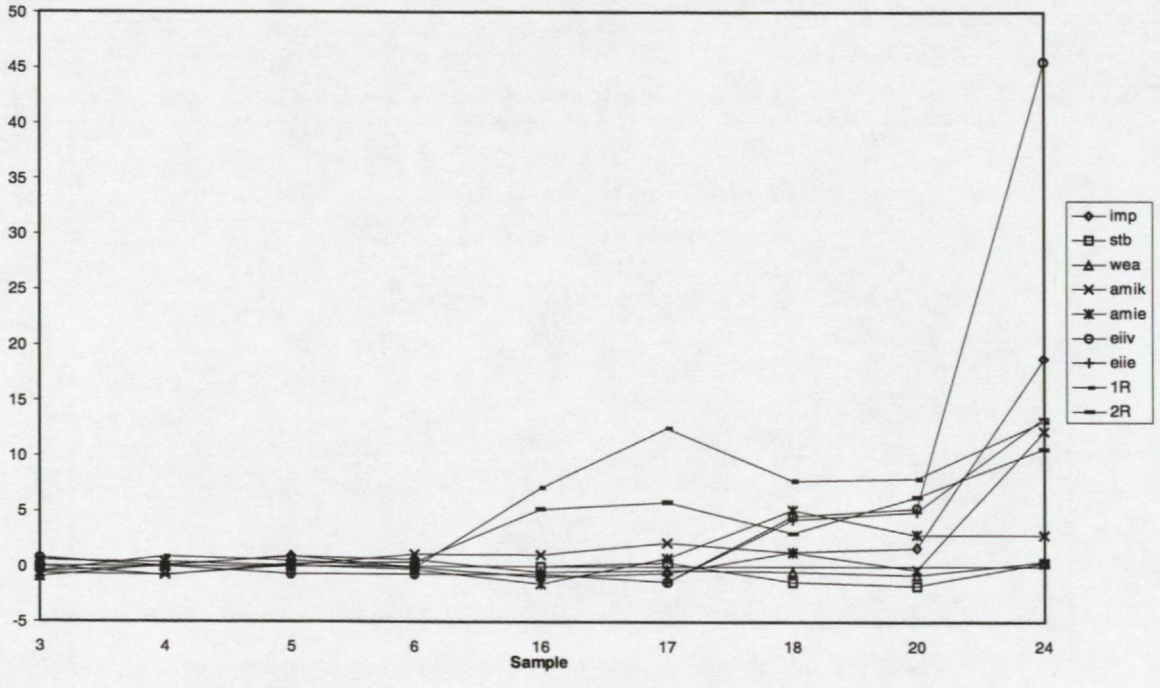
Normalised GI

Figure 3. Normalised GI Values (Gear 13) Sensor J



Normalised GI

Figure 4. Normalised GI Values (Gear 13) Sensor J



Note :

1 Figure 3 includes samples 7–14, which were used as training data for the supervised learning.

2 Figure 4 includes samples 15–24 (i.e. the additional 'unseen' data) which were used to test the supervised machine learning.

3 Only good quality data is plotted (i.e. poor signal average convergence data has been removed).

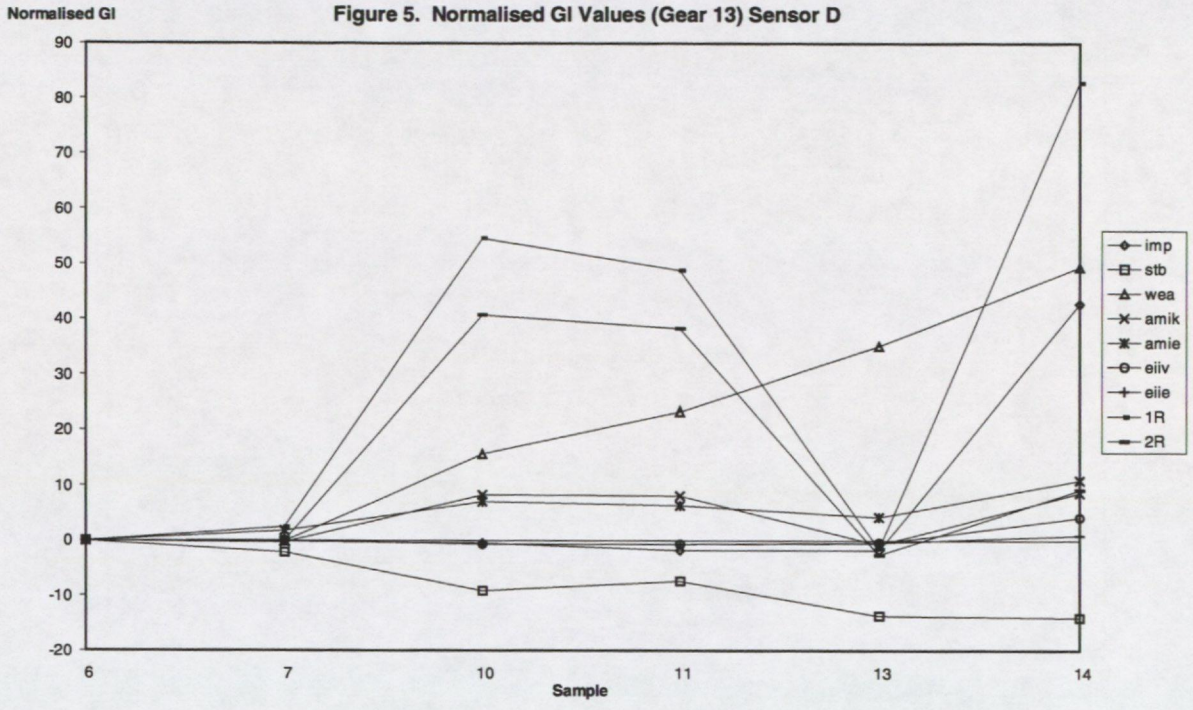


Figure 5. Normalised GI Values (Gear 13) Sensor D

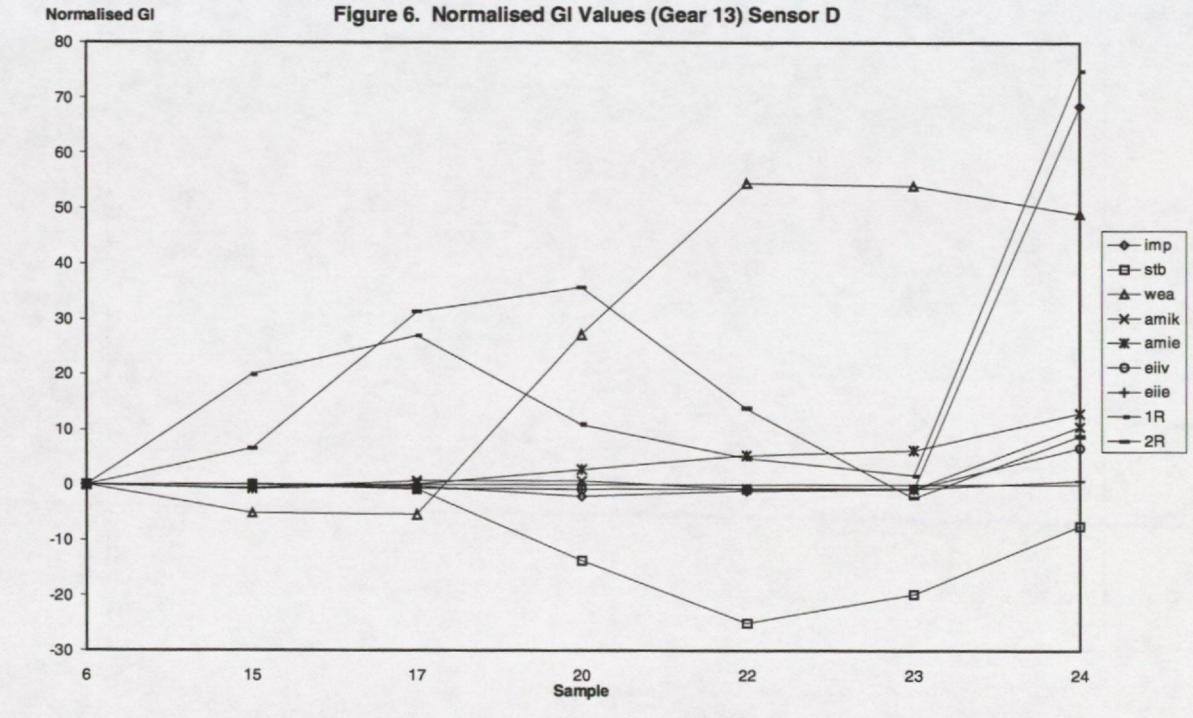


Figure 6. Normalised GI Values (Gear 13) Sensor D

Note :

1 Figure 5 includes samples 7–14, which were used as training data for the supervised learning.

2 Figure 6 includes samples 15–24 (i.e. the additional 'unseen' data) which were used to test the supervised machine learning.

3 Only good quality data is plotted (i.e. poor signal average convergence data has been removed).

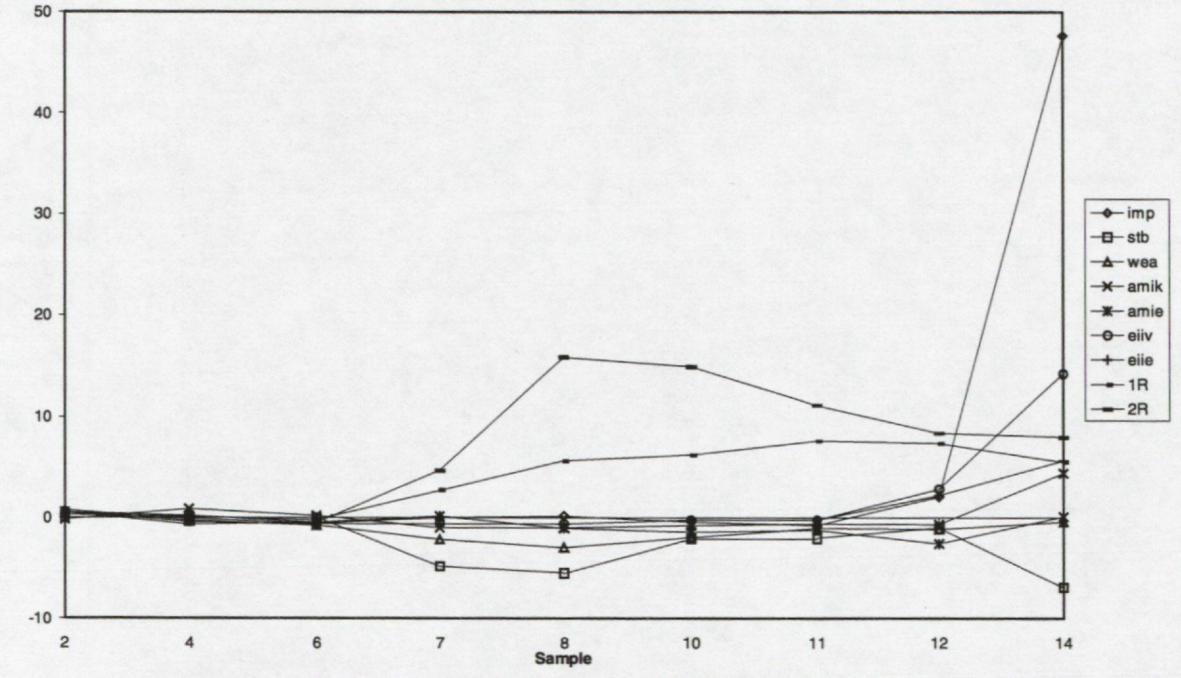


Figure 7. Normalised GI Values (Gear 13) Sensor F

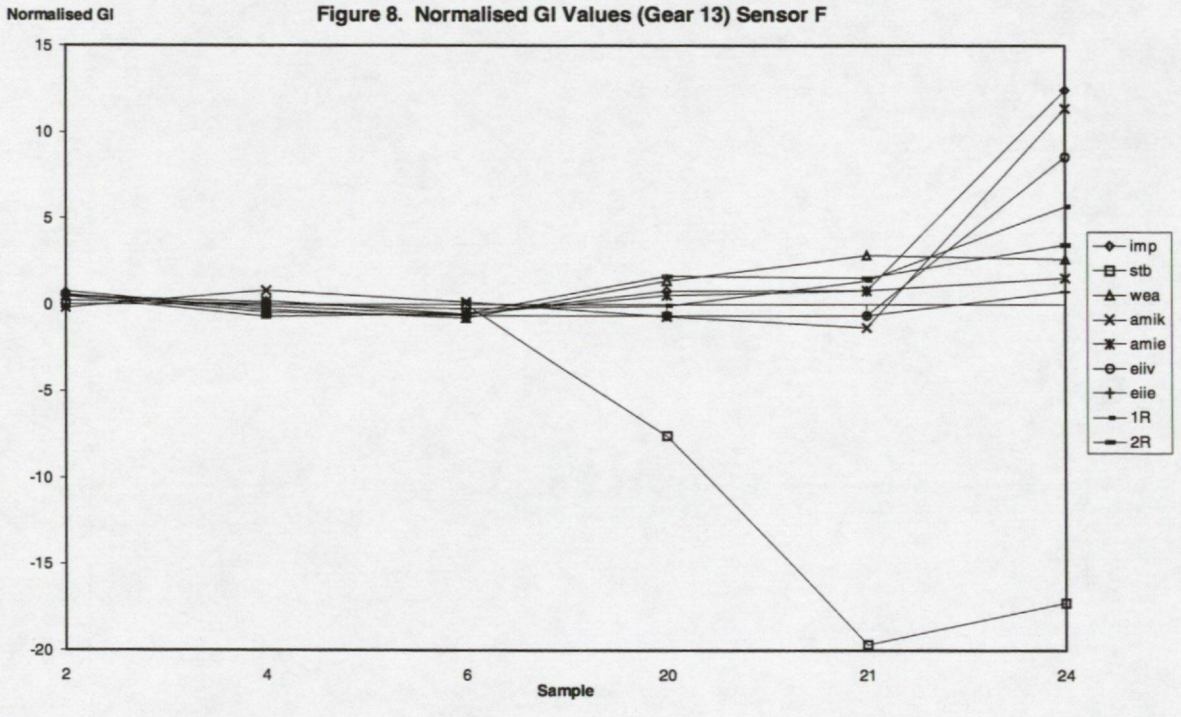


Figure 8. Normalised GI Values (Gear 13) Sensor F

Note :

1 Figure 7 includes samples 7–14, which were used as training data for the supervised learning.

2 Figure 8 includes samples 15–24 (i.e. the additional 'unseen' data) which were used to test the supervised machine learning.

3 Only good quality data is plotted (i.e. poor signal average convergence data has been removed).