# Probabilistic approach to the condition monitoring of aerospace engines

**S King**[1]*, **P R Bannister**[2], **D A Clifton**[2], and **L Tarassenko**[2]

[1]Rolls-Royce plc, Derby, UK

[2]Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford, UK

**Abstract:** The provision of TotalCare® styled service offerings by original equipment manufacture (OEM) suppliers of high-integrity assets is intended to provide improved levels of system availability to the operator. A key element of such service offerings is the ability to minimize unplanned equipment downtime, and the utilization of advanced diagnostic and prognostic monitoring tools is a significant component in achieving this. Monitoring methods, founded on novelty detection technologies, are now a well-established condition monitoring technique. This approach is particularly appropriate for monitoring high-integrity plant where fault conditions arise with extremely low levels of probability. The approach described in this article is to establish empirically based models of normality that are guided by engineering knowledge and utilize key features normally used by expert engineers. However, rather than consider generic modelling approaches, it is proposed that application of models that adapt their sensitivity to the operation of individual assets offer greater prognostic efficiency. This article demonstrates how this can be achieved by considering asset-specific models that adapt the threshold of alerting in accordance with the observed normal running of the plant.

**Keywords:** probabilistic, novelty detection, condition monitoring, health monitoring, aero engines

## 1 INTRODUCTION

### 1.1 Overview of condition monitoring

The provision of services, such as TotalCare® and power by the hour arrangements, are now regarded as an essential element of delivering asset operation. The primary aim of these service contracts is to increase availability of the product for all operational requirements. Clearly, the ability to increase availability requires a robust method of health monitoring. The emphasis of this service is not to accurately diagnose events as they occur, but rather to detect incipient signs of problems long enough in advance such that serious outages are avoided, with a minimal cost of disruption (ideally zero). Modern aero engines are designed to be extremely reliable, typically operating for many thousands of hours before requiring a major overhaul. This poses a significant challenge in the implementation of dependable health monitoring systems where design assumptions are made in the context of an abundance of normal data. This is not unique to the aviation industry. In many industrial sectors, wide-scale usage of high-valued assets has led to the development of various maintenance and condition monitoring strategies. Condition-based maintenance is now an established form of proactive maintenance adopted by many OEM suppliers and operators of high-integrity plant. It is considered to be an alternative to corrective maintenance, usually initiated when a failure occurs, and conventional maintenance based on a fixed operational time, which is triggered when some measure of operational time at a given condition has reached a pre-defined level. A key aspect of this type of monitoring is therefore its prognostic ability. If the detection horizon is so short that there is insufficient time to plan relevant work scope within a scheduled maintenance activity, then the likelihood is that monitoring will consist of measuring operational

*Corresponding author: Rolls-Royce plc, PO Box 31, Derby, DE24 8BJ, UK.

email: steve.king@rolls-royce.com

time against a fixed limit and then invoking a pre-defined set of maintenance actions. Hence, equipment health monitoring is only applicable to a certain percentage of failure modes and an appropriate time lag between inception and final failure must be present. Moubray [1, 2] defined the concept of the P–F interval as the time between the point at which a potential failure can first be detected (P), and that of actual failure (F). Different monitoring approaches will provide varying levels of P–F interval. The example in Fig. 1 indicates how such intervals could vary for different methods used for monitoring bearing faults.

Three main factors influence the time at which a potential failure can first be detected.

1. Ability of the sensor output to show a characteristic change in response to the incipient event – this should also extend to the conditioned signal that is observed by the monitoring algorithm.
2. Sensitivity of the monitoring algorithm to detect the change at the sensor output.
3. Any time lag in the entire monitoring system between the signal change occurring at the sensor element and the time at which a report to the operator can be issued. This must also account for any significant time involved in signal/data transmission.

The objective of using advanced health monitoring techniques is therefore to provide as much prognostic capability as possible. Although this motivates the entire design of a complete monitoring system, this article will focus on different types of appropriate monitoring algorithm. The method adopted must be extremely robust in terms of its ability to accurately detect incipient failures, but it is also important that the false-positive detection rate is well understood and kept to a minimum, since any reported abnormality (genuine or otherwise) will generate additional work scope and therefore cost. Trust in the detection ability is therefore dependent on avoiding no-fault-found



**Fig. 1** Variation of *P–F* interval with various example monitoring methods

outcomes (i.e. the result of a maintenance action following a false-positive detection).

## 1.2 Overview of existing condition monitoring approaches

During the design stage of any high-integrity complex machinery, emphasis is given to the understanding of potential failure mechanisms. Depending on assessed levels of the probability of failure and the estimated severity of impact, the considered failure mode will either be designed out of the physical product, or the risk mitigated by in-service monitoring. This is a common approach adopted by many industrial sectors and is known as failure mode effect and criticality analysis (FMECA) [3, 4]. In the case of the aero gas-turbine engine, the most significant modes to consider will be possible bearing failure mechanisms, fatigue-induced cracks in mechanical systems including combustors, compressor/turbine blades and rotating assemblies, potential secondary effects arising from foreign object damage (for example, bird strike), and aerodynamic instabilities such as fan flutter. As explained above, for those areas where a short P–F interval is expected, hard-life monitoring will be used as a measure of the asset's condition. However, the inherent reliability and robustness of high-integrity plant (which can typically operate for many thousands of hours before requiring a major overhaul) makes it difficult to design algorithms for detecting failure types with longer P–F intervals. This is mainly due to examples of abnormal behaviour being relatively few in comparison to the quantity of normal examples. Consequently, conventional fault-specific detection schemes are usually limited to identifying a small subset of known, well-understood modes of failure.

Various attempts have been made to apply methods for intelligent monitoring and fault diagnosis using model-based and knowledge-based techniques [5, 6]. In addition, expert systems (consisting of *if–then*-type rules) have been successfully used for the analysis of blade vibration [7]. However, successful implementation of these methods can only be achieved when domain-specific engineering knowledge is available that can be expressed in a concise format for suitable representation in a computer program [8]. In situations where physics-based models are either too complex or not available, alternative methods need to be considered.

Other statistically based techniques have been used to monitor data to identify trends, usually focussing on monitoring an observed parameter to determine if it exceeds a predefined limit. Although based on engineering judgement, such limits are often arbitrarily set and are made to apply across all similar asset platforms, despite potential inter-system variability.

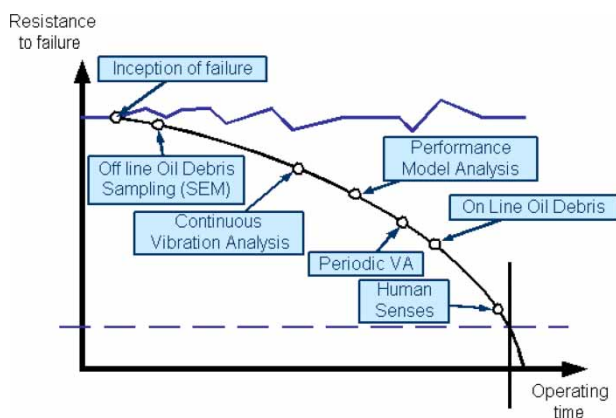The novelty detection paradigm for identification of rare and unexpected features is now a well-established

method and has been successfully demonstrated for the detection of abnormalities in vibration-based monitoring systems. Previous examples include automatic analysis of vibration signatures during engine pass-off testing [**9**, **10**] and continuous monitoring of shaft tracked-order vibration for signs of imbalance or foreign object damage [**11**]. In both of these examples, the model was constructed from normal data collected from a number of engine examples, which therefore was a generic model of engine behaviour. Examples of this approach are now used for on-line monitoring of the latest Trent family of engines.

Equivalent models, using summary data collected from overhaul pass-off testing, have been extended to offer insight into fleet-wide behaviour. This is made possible by the integration of visualization techniques that provide a two-dimensional (2D) view of complex vibration data [**12**]. Various approaches have been investigated for establishing models of normality. Use of neural networks for classification of novelty was evaluated by Markou and Singh [**13**] by simulating abnormal conditions using simulated artificial data. This approach clearly makes bold assumptions about the distribution of abnormal conditions and is likely to prove extremely difficult to validate on high-integrity applications. Statistically based methods involving clustering and various implementations of principal component analysis (PCA) have also been considered. Xue and Yan [**14**] employed techniques such as use of the Mahalanobis distance in combination with Gaussian mixture models to define a distribution of expected residuals from training data, and hence provide estimates of novelty. Other approaches for novelty detection include the use of dynamic modelling techniques such as Kalman filters and Hidden Markov Models. While these techniques are suitable for modelling dynamical systems, the framework adopted by the authors (of which a detailed technical account is given in reference [**15**]) allows us to explicitly model abnormal regions of feature space near the boundary of normality. In addition, our adopted framework provides a mechanism for establishing such boundaries of normality in a principled, automated manner based on the observed data collected during the early stages of asset operation. Later in this article, the concept of adaptive models is discussed, which demonstrates how asset-specific models provide an assessment of novelty in an individual asset's operation, rather than an expected fleet-wide norm. This unique approach is particularly advantageous, since experience shows that from a group of engines of a similar design type, a range of conditions can be observed for the same operating criteria. For example, two independent engines can both operate within design limits, but exhibit different vibration levels for a given speed condition. If the lower-running engine were then to experience a 10 per cent increase in vibration, and still fall within the predefined fleet limit, it would

be considered an uncharacteristic condition relative to an engine-specific model. Hence, utilizing threshold levels specific to the individual asset provides a mechanism for early detection of novel conditions (i.e. increased algorithmic sensitivity).

### 1.3 Article overview

Having justified the use of novelty detection for health monitoring, this article provides examples based on robust modelling techniques that offer reliable novelty alerts. The approach underpinning the use of these techniques is based on the novelty detection framework proposed in reference [**15**].

Section 2 provides an introduction to the proposed methodology of novelty detection in the context of equipment health monitoring of high-integrity plant. The initial step in this framework addresses various examples of feature extraction techniques and is explained in section 2.2.

Data visualization is the next major step in the process of deriving models of normality, and is essential in gaining an understanding of the underlying structure of the data. Methods such as PCA, Sammon's mapping and NeuroScale are all viable techniques, each of which can be used in combination with other statistical and signal processing methods to gain insight into the data. In this article, discussion is restricted to model creation, and the reader is referred to other works on visualization techniques for further reading [**16**, **17**].

Methods of characterizing normality within an empirically derived model are then discussed in section 2.3. The final phase of the novelty detection framework is described in section 2.4, which addresses the critical area of defining novelty thresholds in a principled manner using an extension of extreme value statistics. A key benefit of this method is the ability to automatically derive thresholds that provide engine-specific limits. It is therefore the view of the authors that encapsulating this extension of extreme value statistics within our novelty detection framework offers a unique and robust mechanism for developing such capability within engine-health monitoring applications.

Case study examples are described in section 3, with a discussion of key results in section 4. Finally, conclusions and an outline of future work are presented in section 5.

## 2 METHODOLOGY

### 2.1 Introduction to novelty detection

The concept of novelty detection is based on the premise that a bounded model of normality can be constructed. It is usual for the model to be derived

from data-driven methods (for example, neural networks, statistical clustering, etc.) using data collected from observations representing normal system operation. Newly observed data can then be tested against this model in order to determine their relative novelty.

Traditional approaches typically assess the distance of each new observation from the model (in a Euclidean sense), classifying data as abnormal if they exceed some threshold defined on that distance. The more a point exceeds the defined boundary of normality, the greater its measure of novelty. A more principled and less heuristic approach is to assume a probabilistic, generative model of normality, which can be established from density estimation techniques to obtain an estimate of the underlying data distribution [**16**]. A threshold can then be defined on the resultant probability distribution to define a region of normality within areas of data space. This approach allows novelty to be expressed in terms of the probability of observing data that do not belong to the derived distribution.

The following factors are therefore key considerations for the construction of robust models of normality and are discussed in subsequent sections of this article.

1. Availability of representative data that accurately characterize normal operation. This will not consist of the raw signals measured by sensors, for which the same assumption must apply, but will instead relate to key diagnostic indicators (*features*) that can be extracted from the raw data.
2. Appropriate model selection. As already mentioned above, many techniques are available, such as those described in reference [**15**].
3. Derivation of a robust novelty threshold. This reflects the expected boundary of normality (in probabilistic terms) and will be used in the generation of alerts if new data exceed its value. It is therefore important that this threshold is well defined and understood, since if it is too sensitive then there will be a risk of false-positives alerts. Conversely, if the threshold is too insensitive then there is a risk of a false-negative condition in which abnormal data remain undetected.

## 2.2   Feature extraction

As already indicated, a key consideration in the detection of novelty lies with the ability to detect subtle changes in the conditioned signal that relate to incipient events. Clearly, it is important that appropriate signal conditioning is utilized so that relevant features can be extracted from the sensor data.

Often, prior knowledge of the system being monitored provides a useful guide for feature selection. Where systems involve rotating machinery, vibration is often a key discriminator between normal and abnormal behaviour. Civil gas-turbine engines are comprised of two or more rotating shafts. This gives rise to vibration excitation, which is transmitted through the mechanical structure of the engine and hence can be monitored as carcass vibration using standard case-mounted accelerometers. Additional aerodynamic and combustion effects will also be reflected in the measured vibration signal. Additional sensors (for example microphones, strain gauges, etc.) can also provide an indication of vibration content.

Key features, extracted from the frequency domain, can then be used for identifying signs of engine abnormality. For example, out-of-balance shaft conditions can be detected by observing changes in the amplitude profile extracted from the fundamental rotational frequency (the first *tracked-order* component) of the shaft as it is accelerated and decelerated. Other shaft-order components are also of interest (for example, $0.5\times$ fundamental frequency, $1.5\times$ fundamental frequency, and other multiples) as these can reveal events relating to blade rubs, foreign object impact, blade cracks, and even certain aerodynamic instabilities on a given blade row. Problems relating to bearing assemblies can also be detected by the observation of side-band components around harmonic tracked-order components.

Monitoring signals from within the gas path (for example, pressures, temperatures, etc.) can also yield features that provide useful measures of deterioration and incipient signs of failure. The output from the model is then used as a *novelty score*, which increases with increasing system abnormality. In section 2.4, it is shown that this is expressed in terms of a meaningful probability.

## 2.3   Density estimation

We have assumed a fixed, underlying generative model of normality, which, in the investigation described by this article, is provided using Parzen windows [**18**] to estimate the unconditional probability density $p(x)$ of normal training data. This method places an identical Gaussian kernel $K(x)$ on each of the $N$ training data, where each kernel has width $\sigma$, and where the resultant data density is defined to be

$$p(x) = \frac{1}{N\sigma^{D}} \sum_{i=1}^{N} K\left(\frac{x - \mu_i}{\sigma}\right)$$

for kernel

$$K(x) = \frac{1}{(2\pi)^{D/2}} \exp\left(-\frac{1}{2}x^2\right)$$

where the data are $D$-dimensional, and where the $i$th kernel is centred at $\mu_i$.

Typically, the value of $\sigma$ is set to ensure a compromise between underfitting and overfitting the training data.

### 2.4 Principled setting of novelty thresholds

In order to classify data $x_i$ as either normal or abnormal, we define a decision boundary $H$ on $P(x)$, the integrated form of $p(x)$, such that $x_i$ is classified normal if $P(x_i) < H$, else $x_i$ is classified abnormal, where $P$ is defined to be the probability mass enclosed by $p(x)$. We term this decision boundary the *novelty threshold.*

Previous work [15] has shown that extreme value theory (EVT) [19] can be used to set novelty thresholds on Gaussian kernels, which characterizes the expected distribution of extreme values generated from an underlying normal distribution – effectively modelling the tails of that distribution. The novelty threshold corresponding to $P(x) = H$ occurs at some radius $r$ from the Gaussian kernel

$$r = \frac{\sigma}{\sqrt{2\ln m}}\left[2\ln m - \ln(-\ln H) - \frac{\ln\ln m + \ln(4\pi)}{2}\right]$$

which corresponds to an equivalent threshold, $p(r)$. Here, EVT provides a probability distribution describing where the maximum of $m$ values drawn from the model of normality are expected to lie. The above equation integrates this probability distribution, providing a threshold such that the maximum of those $m$ values will lie within the novelty threshold with probability $H$.

### 3 DATA

### 3.1 Introduction to data

The probabilistic modelling approach for novelty detection is illustrated using aero gas-turbine vibration measurements. Transducers mounted on the engine case measure the broadband vibration signal during flight, which is processed by an engine health monitoring (EHM) system, transforming it into amplitude versus frequency spectra five times per second.

Software modules (known as feature detectors) process this data in order to monitor individual engine parameters. It is the resulting outputs from these modules, or feature detector scores, which are used as input features in our modelling approach. It is then possible to compare new scores (obtained from further system operation) to the model to determine if the condition of the engine is abnormal.

Two datasets are described in this work. Example A is an engine that suffered fan liner damage and example B is an engine where a compressor blade failed. In both cases there were no significant precursors to the fault detected using existing EHM systems. For each example, there are 7 days of flight test data available leading up to the time of failure.

The objective of the investigation described by this article is to show that it is possible to construct a model using data from early in the life of the engine, in order to demonstrate increased novelty in tests immediately leading up to each failure.

### 3.2 Construction of models

For the purpose of this investigation, we focus on the following data types: the fundamental Tracked Order (TO), which is a measure of vibration amplitude at the frequency of engine shaft rotation; the fractional and multiple TOs associated with a set of common shaft harmonics; the residual energy (i.e. the energy in the signal that is not already attributed to one of the TOs); and the broadband energy in the range 0–1 kHz.

The score produced by each feature detector corresponding to one of these parameters is median filtered, and truncated so that only data acquired above the idle speed of the engine are retained. Data are partitioned such that the operation relating to the three main subsystems of the engine are grouped separately, leading to three distinct sets of models.

For each of these engine subsystem models, the preprocessed feature detector scores are combined to form a feature vector for each sample. A typical feature vector is of the form:

> [Fundamental TO; 2nd harmonic TO; 3rd harmonic TO; 4th harmonic TO; 5th harmonic TO; 1.5th harmonic TO; residual energy; broadband]

Once a set of feature vectors has been constructed, each element is component-wise normalized [17] in
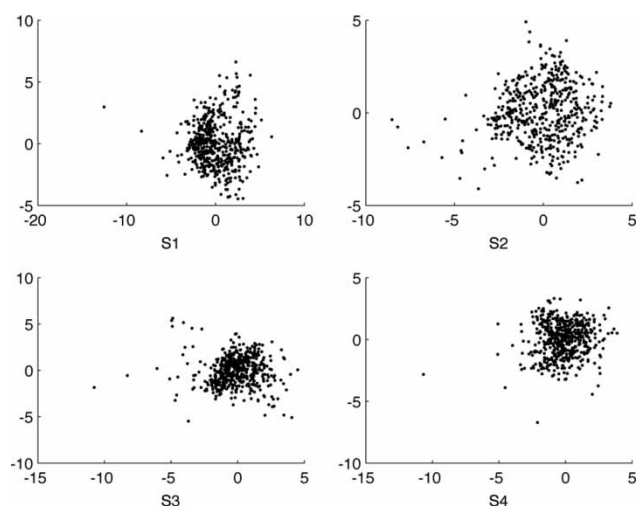


**Fig. 2** NeuroScale projection of model training centres for the Fan liner loss example. The training data are partitioned into four, corresponding to four speed subranges, each represented by 500 prototype feature vectors generated from three flights of normal operation. Speed subranges are labelled S1–S4 in ascending order

order to approximately equalize the dynamic ranges of the outputs of the various feature detectors.

Practical experience shows that the vibration response of an aero engine will vary according to the regime of the flight. For example, scores generated during cruise (i.e. steady state) will differ from those obtained during a high-power manoeuvre such as takeoff. In order to provide discrimination between various modes of aircraft operation, the training data for each engine subsystem are partitioned into four subsets according to a partitioning of the speed range into quartile subranges. Each of the four subsets

of data is used to construct a separate model of normality.

Another pertinent observation regarding the data is that a disproportionate amount of time is spent during cruise conditions, compared to other engine speeds. When constructing a model, the selected data are filtered to ensure that feature vectors insufficiently different in speed to the previous feature vector are not retained. This step prevents the model being over-fit to cruise conditions.

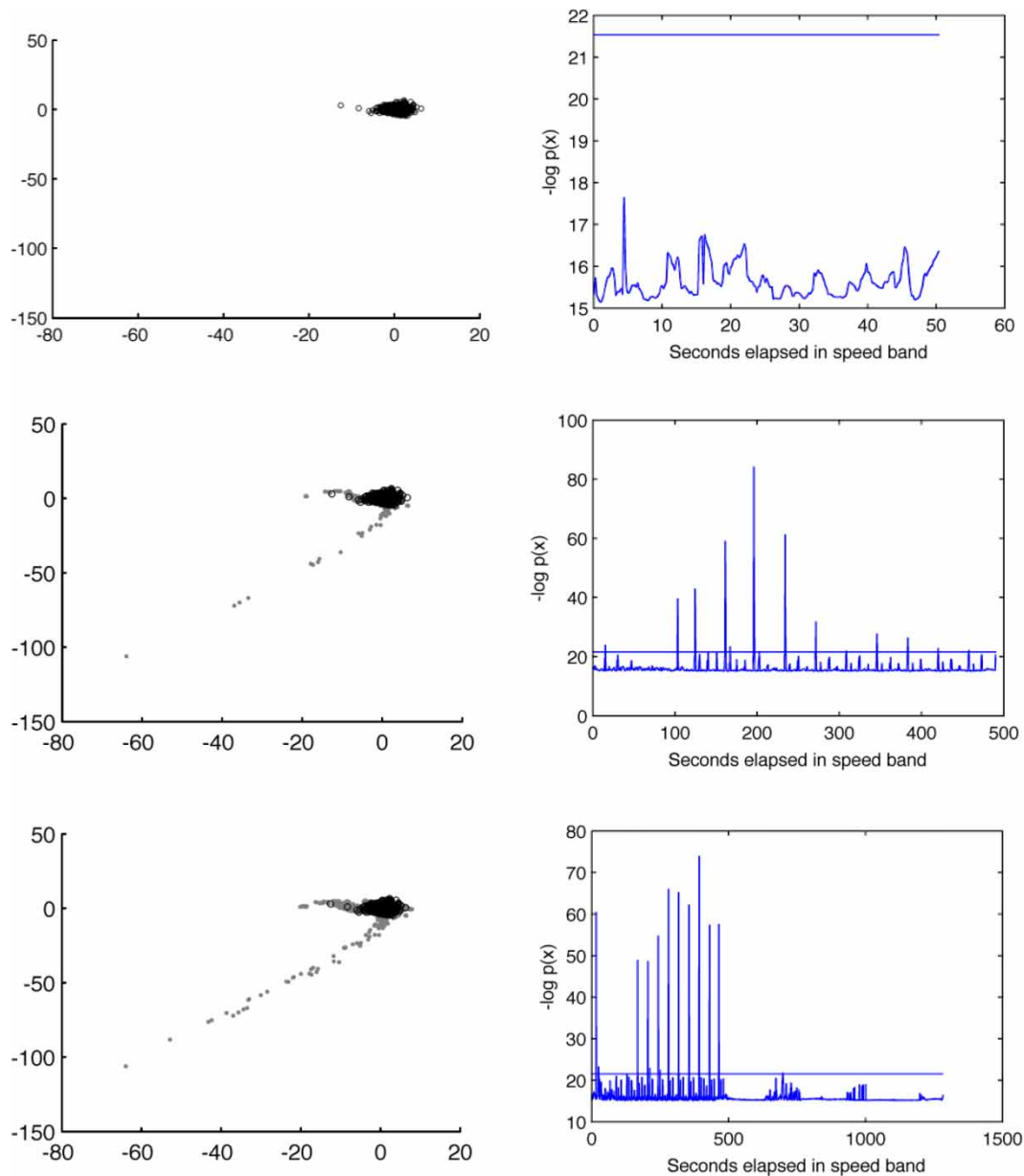Data used to construct each model of normality are also used to train a projection function (used to



**Fig. 3**   NeuroScale plot of new data against model (left) and associated novelty score against novelty threshold (right) for flights occurring after the training period for the fan liner loss event. In the NeuroScale plots, the model prototypes are plotted in black, while new test data are plotted in grey. The earliest flight is shown in the top row and the latest flight on the bottom row

display the high-dimensional feature vectors in 2D). This mapping from high-dimensional feature vector space into 2D is achieved through the use of a NeuroScale mapping [20]. In this approach, a radial basis function neural network is trained using normal data to give a mapping into 2D, which seeks to make the Euclidean distances between pairs of image patterns in the 2-D visualization space as close as possible to the Euclidean distances between the corresponding pair of patterns from the original higher-dimensional space. Further description of the NeuroScale method for reducing high-dimensional data into 2D for visualization is deferred to [20].

The NeuroScale algorithm is trained using a set of prototype feature vectors that are generated by applying the $k$-means clustering algorithm on the full set of feature vectors gathered to train each model, generating a subset of cluster centres in the high-dimensional space. The same centres are used to train a Parzen window estimator as described in section 2.3.

## 4 RESULTS

In this section we present NeuroScale visualizations of training and test data, and compute novelty scores for test data in order to determine if the abnormal behaviour can be detected before the point of failure. This analysis is carried out for both of the example datasets introduced in section 3.1.

### 4.1 Example a – fan liner loss

Seven flight tests were available culminating in the loss of fan liner material from this engine. The first three tests were grouped together to form a training set, and the visualizations in Fig. 2 show that the four sets of $k = 500$ prototypes (each corresponding to a different subrange of engine speeds) cluster together within their respective speed subranges. Axes for NeuroScale plots are unitless, but we have found that data representing

a period of prolonged normal operation will cluster within $\pm 10$ in both $x$ and $y$ directions for this particular application.

This initial visualization is vital in order to confirm that the training data are indeed representative of the same engine condition.

We will focus on the first of the four models with the techniques, discussion being directly applicable to the other three models.

The visualizations in the left-hand column of Fig. 3 show a clear drift away from the model centres (plotted in black) over the course of three flights (each plotted in grey). A novelty score is computed for each feature vector using the method described in section 2.4. Note that the novelty score for feature vector $x$ is defined to be $-\log_e p(x)$, which will take high values for data occurring with a low probability. The novelty threshold is shown in each novelty plot, and while the novelty scores for the first test lie below this value, the subsequent flights exceed the threshold by a significant margin.

This demonstrates the prognostic capabilities of the scheme with detection of abnormal engine condition achieved 2 days prior to the failure event.

### 4.2 Example b – blade failure

Again for this second example, only seven tests are available. Data from the first five tests were pooled in order to populate the training set across the full speed range of the engine. In this example, more tests are required for training purposes as the tests themselves are much shorter in length and contain less vibration data.

Figure 4 shows the NeuroScale visualization and novelty scores associated with the highest speed subrange for the day prior to the failure.

It is clear that the operation of the engine is significantly different from that of the learned model and the threshold exceedance would suggest that the engine
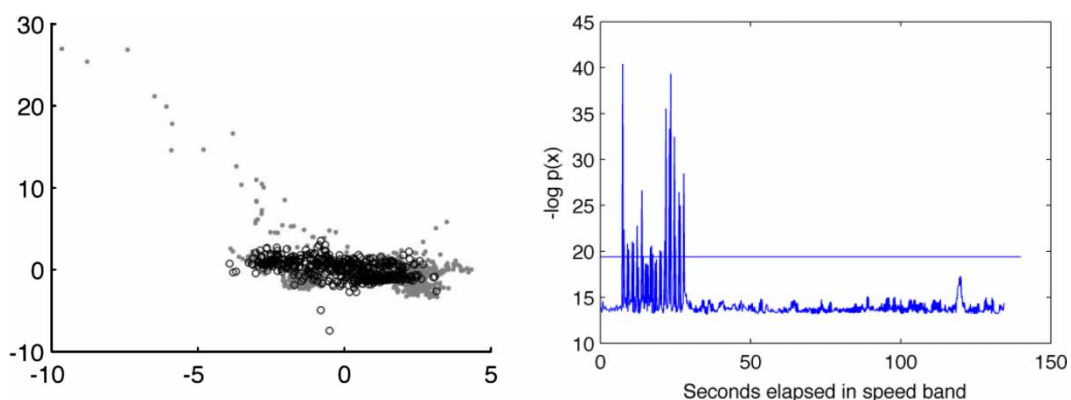


**Fig. 4** NeuroScale plot of new data against model (left) and associated novelty score against novelty threshold (right) for the flight prior to the blade failure event. In the NeuroScale plot, the model prototypes are plotted in black, while the new test data are plotted in grey

should be examined for damage at this point. Without such prognostic tools available at the time, the engine was left on the aircraft and suffered a blade failure during the next flight.

## 5 CONCLUSIONS

This article has demonstrated the effectiveness of a principled approach to developing modelling techniques for visualization of complex engine health data and deriving asset-specific thresholds for early detection of abnormalities. Two case studies have been presented (that is, fan liner loss and blade failure), which highlight how these techniques can be used to provide early signs of incipient failure; usually in situations where prognostic capability was previously unavailable. Retrospective analysis of the Fan Liner Loss data demonstrated 2-day prior warning of the impending failure. Similar prognostic behaviour was demonstrated in the analysis from the blade failure example, where indications of problems were highlighted the day prior to actual failure. These results demonstrate a significant increase in the ability to determine abnormal engine operation in comparison with conventional techniques, which detected only the final engine event itself.

Additionally, experience has also shown that these methods offer a generic capability for identifying abnormality in a range of other fault scenarios. For example, in the detection of combustor injector blockage, foreign object damage, crack propagation in rotating assemblies, and aerodynamic instabilities [8].

The authors consider that the successful application of these methods lies with the use of relevant features extracted from observed data, founded on domain expert knowledge, in combination with appropriate model selection and the principled, automatic definition of model boundaries.

This offers the advantage of requiring minimal expert intervention both in initial model construction, and, more importantly, with subsequent analysis of data where model exceedances directly relate problems to specific sensor/engine fault locations. However, their disadvantage lies with the initial arbitrary partitioning of training data over different engine speeds and flight conditions. Although to date, this has had negligible impact on results, there is scope for utilizing an approach in the selection of training data that more closely corresponds to actual engine operational regimes.

Although this article has concentrated on actual modelling techniques, the introduction identified some important business-related issues. These were concerned primarily with reliability of detection and effective prognostic capability, both of which directly influence costs associated with subsequent work-scope activity in the event of a fault being identified. Therefore, prior to adopting any new health monitoring technique, it is important that any costs associated with its implementation can be quantified (such as, 'What is the likely cost of false alerts versus likely savings generated as a result of improved prognostic ability?'). Hence, the main benefit of applying the described modelling methods within the framework of extreme value statistics is that information relating to sensitivity and specificity of the detector [21] can be defined in support of the above business concerns where expected false-positive rates can be quantified in advance of deployment.

### 5.1 Future work

The detection techniques discussed above have concentrated on vibration features as prime indicators of engine condition. Many aspects of what has been presented in this article are now being incorporated within the service support of the latest Trent family of engines. However, it is recognized that other indicators of engine condition are also available. Engine performance signals are currently trend monitored using conventional off-line analysis methods. Further work will therefore investigate potential benefits of applying the above methods to these data both for on- and off-line analysis. This is likely to involve extending the idea of using extreme value statistics to multi-parametric data. As already indicated in the previous subsection, further work will also consider more robust heuristic methods for automatic partitioning of engine speed and flight phase data for training and modelling purposes.

So far, in the application of the modelling methods described by this article, it has been possible to use either the entire dataset obtained from engine testing, or invoke some simple sampling regime. However, as the complexity of monitoring systems grows, and with the volume of generated engine health data continuing to increase, there will be a point at which certain techniques will not scale and more efficient sampling strategies will be required. For example, Sammon's mapping technique requires approximately $4n^2$ bytes of memory when projecting $n$ feature vectors, irrespective of the dimensionality of those vectors. Hence, even a modest 100 000 observations will require 40 Gbytes of working memory. Therefore, techniques similar to those described in reference [22] will require development so that existing modelling techniques can make effective use of very large datasets.

© Rolls-Royce plc 2009

## REFERENCES

1 **Moubray, J.** *Reliability-centered maintenance*, 1997 (Industrial Press, Inc., New York City, NY).
2 **Rausand, M.** and **Hoyland, A.** *System reliability theory: models, statistical methods, and applications*, 2nd edition, 2004 (John Wiley & Sons, Hoboken, New Jersey).
3 Analysis techniques for system reliability – procedure for failure mode and effects analysis (FMEA), International Standard IEC 812, December 1997.
4 Procedure for performing a Failure mode, Effects and Criticality Analysis, Military Standard 1629A, August 1998.
5 **Merrington, G. L.** Fault diagnosis in gas turbines using a model-based technique. *Trans. ASME*, 1994, **116**, 374–380.
6 **Piers, M. A.** and **Donker, J. C.** A knowledge-based assistant for diagnosis in aircraft maintenance. In Proceedings of the AGARD Avionics Panel Symposium on Machine Intelligence for Aerospace Electronic Systems, Lisbon, May 1991.
7 **Allwood, R., King, S.,** and **Pitts, N.** The automatic interpretation of vibration data from gas turbines. *Aeronaut. J. R. Aeronaut. Soc.*, 1996, **100**(993), 99–107.
8 **King, S., Anuzis, P., King, D., Tarassenko, L., Utete, S.,** and **McGrogan, N.** A review of applications for advanced engine health monitoring in civil aircraft engines. In Proceedings of the 13th International Congress of Sound and Vibration, Vienna, 2006.
9 **Cowley, P. H.** and **Carr, H. R.** Synopsis of application of neural networks to aero engine vibration monitoring. In Proceedings of the SAE32 Symposium on Advanced Vibration Monitoring Techniques, Aix-en-Provence, October 1997.
10 **Nairac, A., Townsend, N., Carr, H. R., King, S., Cowley, P. H.,** and **Tarassenko, L.** A system for the analysis of jet engine vibration data. *Integr. Comput. Aided Eng.*, 1999, **6**, 53–65.
11 **King, S., Anuzis, P, Astley, K., Tarassenko, L., Hayton, P.,** and **Utete, S.** The use of novelty detection techniques for monitoring high-integrity plant. In Proceedings of the IEEE International Conference on Control Applications, Glasgow, UK, September 2002, pp. 221–226.
12 **Eciolaza, L., King, D., Fairbrother, C., Walters, M., Ault, D., Sever, I., King, S.,** and **Kadirkamanathan, V.** A decision support system for efficient data management, visualisation and analysis in jet engine health monitoring. In Proceedings of the Fifth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Edinburgh, UK, July 2008.
13 **Markou, M.** and **Singh, S.** A neural network-based novelty detector for image sequence analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2006, **28**(10), 1664–1677.
14 **Xue, F.** and **Yan, W.** Parametric model-based anomaly detection for locomotive subsystems. In Proceedings of the International Joint Conference on Neural Networks, Florida, USA, August 2007.
15 **Tarassenko, L., Clifton, D. A., Bannister, P. R., King, S.,** and **King, D.** Novelty detection. In *Encyclopaedia of structural health monitoring*, 2008 (John Wiley and Sons, New York).
16 **Bishop, C.** *Pattern recognition and machine learning*, 2006 (Springer, New York).
17 **Tarassenko, L., Nairac, A., Townsend, N., Buxton, I.,** and **Cowley, P. H.** Novelty detection for the identification of abnormalities. *Int. J. Syst. Sci.*, 2000, **31**(11), 1427–1439.
18 **Silverman, B. W.** *Density estimation for statistics and data analysis*, 1986 (Chapman and Hall, London).
19 **Fisher, R. A.** and **Tippett, L. H. C.** Limiting forms of the frequency distributions of the largest or smallest members of a sample. In Proceedings of the Cambridge Philosophical Society, 1928, p. 24.
20 **Nabney, I.** *NETLAB: algorithms for pattern recognition*, 2002 (Springer-Verlag, Berlin).
21 **Clifton, D., McGrogan, N., Tarassenko, L. King, D., King, S.,** and **Anuzis, P.** Bayesian extreme value statistics for novelty detection in gas-turbine engines. In Proceedings of the IEEE Aerospace, Montana, USA, 2008.
22 **Sundaram, S., Strachan, I., Clifton, D., King, S.,** and **Palmer, J.** A data mining approach to reveal hidden patterns in aircraft engine and operational data. In Proceedings of the Fifth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Edinburgh, UK, July 2008.

JAERO414

Proc. IMechE Vol. 223 Part G: J. Aerospace Engineering