Proceedings of the 2002 IEEE Aerospace Conference, Big Sky MT, March 2002

A Testbed for Data Fusion for Engine Diagnostics and Prognostics[1]

Tom Brotherton, Paul Grabill, and Dariusz Wroblewski
The Intelligent Automation Corp
13029 Danielson St., Poway, CA 92064
(858) 679-4140
tom.brotherton@iac-online.com
paul.grabill@iac-online.com
dariusz.wroblewski@iac-online.com

Richard Friend and Bill Sotomayer
Aeronautical Systems
Wright-Patterson AFB, OH 43558
(937) 255-2734
richard.friend@wpafb.af.mil
william.sotomayer@wpafb.af.mil

John Berry
US Army Aviation and Missile Command
Redstone Arsenal, AL 35898
(256) 313-4815
john.berry@redstone.army.mil

Abstract— A key to producing reliable engine diagnostics and prognostics resides in fusion of multisensor data. It is believed that faults will manifest effects in a variety of sensors. By 'integration' (fusion) of information across sensors detections can be made of faults that are undetectable on just a single sensor. Data to support development of prognostic techniques is very rare. The development requires continuous collection of significant amounts of data to capture not only "normal" data but also capture potential fault event data well before the fault is detected by existing techniques, as well as capture data related to rare events. The collected data can be analyzed to develop processing tailored to new events and to continuously update algorithms so as to improve detection and classification performance and reduce false alarms. IAC in collaboration with the Air Force and the Army is developing a testbed to perform data collection and to develop fusion techniques for gas turbine engine health monitoring. The testbed and examples of its operation are presented here.

TABLE OF CONTENTS

1. INTRODUCTION

The key to producing reliable engine diagnostics and predictive diagnostics / prognostics resides in the fusion of multisensor data. It is believed that faults that occur will manifest effects in a variety of sensors. By 'integration' (or fusion) of information across sensors detections can potentially be made of faults that are otherwise undetectable on just a single sensor. Fusion saves cost and weight; no new sensors are required. Fusion reduces false alarm rates; faults are seen across multiple sensors. Diagnostic performance is improved by allowing detection of unique fault patterns seen on sets of sensors instead of a single sensor. Fusion enables prognostics; low signal-level information is integrated across a variety of sensors so potential faults can be detected earlier.

However the successful development of such a system requires real data that represent nominal operation, data with known faults, and most importantly for prognostics, data that has been collected well ahead of the time that a fault becomes obvious. Currently good data sets to support prognostic algorithm development and validation are rare or do not exist. Typically data is saved when a fault is detected; too late to be useful for prognostics development.

Table 1. Table of acronyms

ACRONYM	MEANING
AD	Anomaly detector
AEDC	Arnold Engineering Development Center
BN	Bayesian network
CART	Classification and regression tree
CBM	Condition based monitoring
DHMS	Distributed health management system
DAS	Data and analysis server
F-GBS	Facility ground based station
iMDS	Intelligent machinery diagnostics software
GUI	Graphical User Interface
NN	Neural network
PCA	Principal components analysis
P-GBS	Portable ground based station
VMU	Vibration management unit

Intelligent Automation Corporation (IAC) in collaboration with the Air Force and the Army is developing a *distributed health management system* (DHMS) to perform data collection and monitoring of real aircraft collected data. The system is also contains components for the development and application of data fusion techniques for health monitoring

[1] This effort is being funded by both the U.S. Air Force and the U.S. Army

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 MAR 2002**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **A Testbed for Data Fusion for Engine Diagnostics and Prognostics1**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **The Intelligent Automation Corp 13029 Danielson St., Poway, CA 92064**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**14**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

of gas turbine engines as well as vibrations from other aircraft components. The system uses a combination of signal and information processing algorithms to perform data fusion for engine and aircraft fault diagnostics and prognostics to support individual aircraft facility maintenance, fleet maintenance, as well as the development of new diagnostics and prognostics algorithms using real data.

Signal processing algorithm development is centered around a Matlab toolbox. Other commercially available software is used to perform data mining and Bayesian network development. Figure 1 shows a top level architecture for the system being developed to achieve these goals.

The next section describes the overall DHMS architecture and presents the operating philosophy of the system. Following this, tools being developed for analysis and solution of diagnostic and prognostic problems are presented. Next, approaches to data fusion are discussed. Finally examples of using the system for detection of events on a test cell F100 engine using collected vibration data are presented.

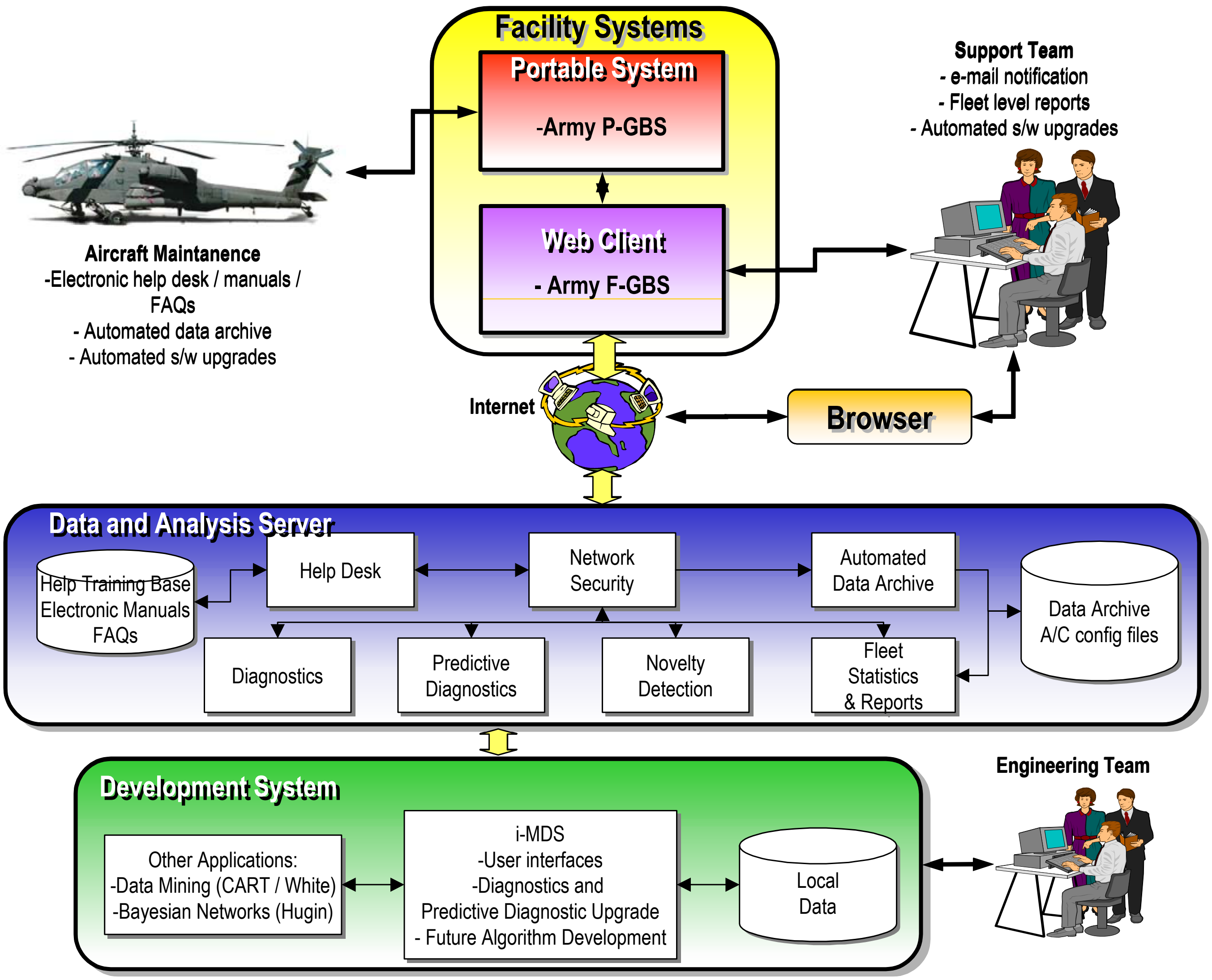

Figure 1. Overall system architecture

2. THE DISTRIBUTED HEALTH MANAGEMENT SYSTEM: DHMS

The overall architecture of the DHMS is shown in Figure 1. The four main components of DHMS for the Army application are the on-aircraft *Vibration Management Unit* (VMU), the *Ground-Based Systems* (GBS), the *Data and Analysis Server* (DAS), and the *Development System*. These four systems are briefly described below.

Overall system objectives are:

- Transfer of aircraft collected data from the aircraft to the Facility Systems followed by automated transfer to the DAS;
- Archiving of data on the DAS;
- Processing on the DAS to detect *novel* events;
- Development and improvement of detection, diagnostics, and prognostic algorithms for all system components;
- Automated transfer of s/w upgrades, new algorithms, and new parameter settings from the Development System to the DAS, to the Facility Systems, and eventually to the VMU.

DHMS recognizes three types of system users: maintenance personnel, maintenance supervisors and engineers. These users have different requirements that must be considered when designing portions of the systems used by a class of users

On-Board Systems

Data is collected from individual Army aircraft using an on-aircraft *vibration management unit* (VMU) developed by IAC [1]. The data includes not only engine related information, but vibration information in the form of *condition indicators* (CIs) [1] collected from various locations on the aircraft as well as rotor tracking information. The system described in [1] has been upgraded to include a 1553 aircraft bus interface that collects engine related data as well as several analog inputs for collection of environmental information such as outside air temperature and pressure altitude. All the data collected by the VMU is stored in flash memory for transfer to a *portable ground based station* (P-GBS).

Facility Systems

At each facility a ground-based station (GBS) is required for download and analysis of aircraft data. The GBS can be installed in two different modes:

Portable GBS (*P-GBS*)-The P-GBS is the system used by line maintenance personnel initially for setup of an aircraft VMU and for subsequent transfer of VMU vibration and engine performance data from that aircraft. The software runs on a ruggedized laptop computer.

The P-GBS is intended for use by line maintenance personnel involved in routine daily aircraft maintenance operations. These users are mainly interested in receiving status information on the relative health of the aircraft and providing suggested corrective actions to perform maintenance based on the data processed on the VMU and P-GBS. The system also will transfer data collected on the aircraft to the GBS system and transfer of OBS software and configuration updates from the GBS

Facility GBS (*F-GBS*)-The facility GBS serves as a central repository of data for a single Army facility. The F-GBS will typically be hosted on a powerful PC that has Internet access and may also be located on a LAN. Data collected by the facility's P-GBS systems is imported into the F- GBS. The F-GBS will serve as the distribution site for VMU upgrades; these will be passed down to the portable GBS systems and then on to the on-aircraft systems. Because the F-GBS will contain information about all of the facilities aircraft, it can provide facility-wide summaries and reports. The F-GBS will also be able to perform more extensive analysis and visualization of the data than a P-GBS. The F-GBS will interface with the Data and Analysis Server (DAS) to periodically transfer new data from the F-GBS to the DAS and also check for any upgrades that need to be disseminated to the facility's F-GBSs, P-GBSs, or VMUs.

The F-GBS is intended for use by facility maintenance supervisors interested in maintenance operations for all the aircraft at their facility. The maintenance supervisor will be expected to use the system to monitor the overall state of the fleet of aircraft he or she supervises. They will also perform in-depth analysis of the state of individual aircraft.

Data and Analysis Server

The Data and Analysis Server (DAS) is a single, Internet accessible system that will combine the data and information from participating facilities and provide a centralized repository for that data. Facility GBSs will periodically transfer new data they have collected and stored to the DAS. Facility GBSs will also be able to transfer to the F-GBS, P-GBS, or VMU upgrades received from the DAS as needed. The DAS will also contain novelty detection processing to automatically screen incoming data to detect new or never seen before events [2,3,4] and bring that data to the engineering team's attention. The DAS will also provide an interface that will allow users to be able to access fleet data, statistics, trending, and summary reports from any computer connected to the web and running a standard web browser.

The DAS, F-GBS, and browser connections are intended for use by fleet-wide (more then one facility) maintenance supervisors. Fleet maintenance supervisors require visibility to fleet-wide system symptoms for projecting depot and supply resources. Comparisons of unit averages, statistical values, and trending tools are required.

Development System

The Development System is a "behind the scenes" toolbox used by development engineers. It contains software tools for performing advanced engineering analysis on data stored on the DAS and elsewhere. The toolkit will allow engineers to prototype algorithms that can later be incorporated into upgrades for the DAS, GBS and VMU systems. The development system will be standalone from the other DHMS system components; however, it will have the ability to download and process data from the DAS or in the future from other data sources such as the Air Force CETADS system. Major components of the development system are based on COTS software packages: MathWork's *Matlab*, Hugin's *Expert*, and Salford System's *CART*. Matlab is used for signal processing development. A Matlab machinery diagnostic toolbox is being created that may be used in the Matlab Simulink programming environment. Hugin Expert software is used for probabilistic network development and Salford System's CART software is used for data mining. The development system should be flexible enough that additional third party components can be added. The Development System will <u>not</u> be accessible from the GBS and DAS Server systems.

Matlab intelligent Machinery Diagnostic System (iMDS) Toolbox–The system has as its core for signal processing development a Matlab intelligent Machinery Diagnostic System (iMDS) development suite. iMDS was initially developed and used at IAC for Army helicopter vibration monitoring processing as well as anomaly detection processing for the JSF. It is currently being upgraded on the project described here to include fusion processing for engine diagnostics and prognostics.

The iMDS toolbox is built on a foundation of MATLAB. The iMDS library of diagnostic algorithms are supported by both MATLAB M-file scripts and C code compiled to run in the Simulink environment. The algorithms are developed and tested in Simulink before being compiled for use in real-time applications. The Matlab iMDS model-based tools involve the use of Condition Indicator (CI) algorithms for processing of vibration data. The model-based tools use *a priori* knowledge of the mechanical system as a basis for the fault diagnosis. This *a priori* knowledge includes information about rotational speed, mechanical construction (such as gear ratios and inner and outer race data on bearings), and information on structural vibration or acoustic resonance of the system to be diagnosed. A condition indicator uses some form of measured data as input and produces a single real number as output. This single number can be thresholded, trended, fused or otherwise analyzed to provide an indication of the location and type of fault condition. There is a large body of literature on mechanical signature analysis, which is used to develop the knowledge base for the diagnostic toolbox [5,6].

Figure 2 shows the iMDS toolbox in the Simulink™ environment. Figure 3 shows a sample Simulink script for processing of gear related vibration and tachometer signals.

CIs–There are a variety of CI's included in the iMDS Matlab toolbox. Most of the CI's have been developed for extraction of features relevant to helicopter and engine gear, bearing, and acoustic event vibrations.

Measurement CI–Algorithms that are designed as pre-processors for vibration measurements. These algorithms include the basics such as filtering, averaging, and re-sampling.

Neural Network–Tools that allow the fusing and evaluation of non-model based tools. The neural network is trained with good and known fault conditions so that it can recognize normal, novelty and faults. The training allows the neural network to use one or many CIs to determine the machinery fault condition. The data fusion characteristics of the neural network allows for higher probability of detection of mechanical faults with lower false alarms. The neural net tools have been used extensively for performing *novelty detection* of aircraft data. A novel event is a never seen before nor anticipated event.

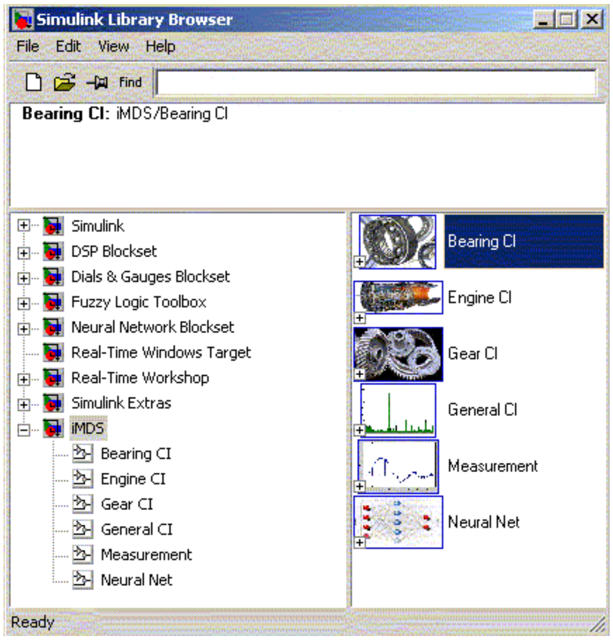

Figure 2 intelligent Machinery Diagnostics System Toolbox Simulink™ interface

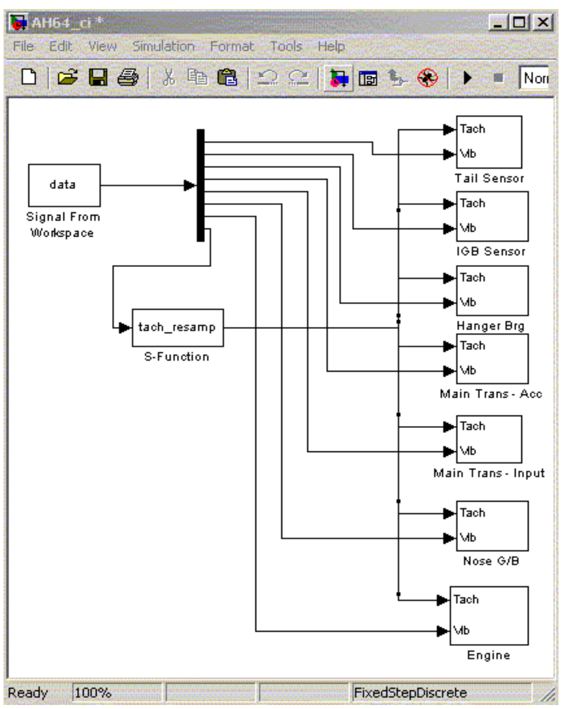

Figure 3 Sample iMDS script

GUI and Visualization Tools – Since iMDS has Matlab as its core, the development of custom user interfaces is straightforward. Figure 4 shows an example of a display developed for engineers to monitor an engine in a test cell for detection of low bypass turbofan augmentor events. The example display can show raw data, processed data, and include virtual 'switches' with which to select different data sets for input and for display of different sensor output spectra.

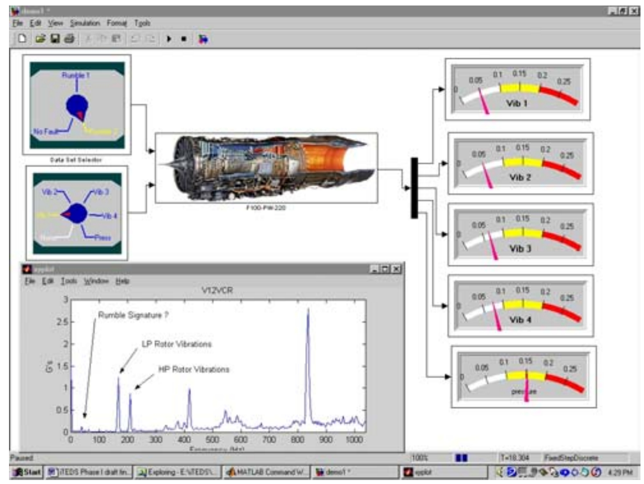

Figure 4 iMDS GUI / Visualization

3. The Prognostics Problem

Figure 5 illustrates the different aspects of the gas turbine engine health-monitoring problem that need to be addressed. The figure shows the trajectory of a particular engine component's health as a function of time. When the engine component is new, its health is considered 100 percent. As time goes on and the component begins to wear out, it's health, defined here somewhat arbitrarily, drops. This figure assumes the component is following a known fault life degradation path. In the following, an *anomaly* is any off nominal operating condition. Anomalies come in two types. The first is a *fault*. A fault is a known off nominal condition. It is assumed that fault-specific algorithms have been developed to detect a fault. The second anomaly is a *novel* event. A novel event is an unknown off-nominal condition. That is, the novel event is not nominal nor is it classified in any of the known fault conditions. *It's something completely new.* We do not know if the novel event is an active failure, an incipient failure, or an "I don't care". Prognostic algorithms are designed to respond to "known faults" that correspond to known failure modes (and not novel events). This is because an important part of the prognostics is the modeling for prediction of the engine component health trajectory shown in Figure 5. In order to develop that model, something about the trajectory of a component from nominal to a known fault condition is required.

Novelty detection is an important component in the operation of DHMS. All incoming data is screened to detect, set aside, and flag for engineering analysis anomaly events. Engineers will not have to continuously process "normal" events.

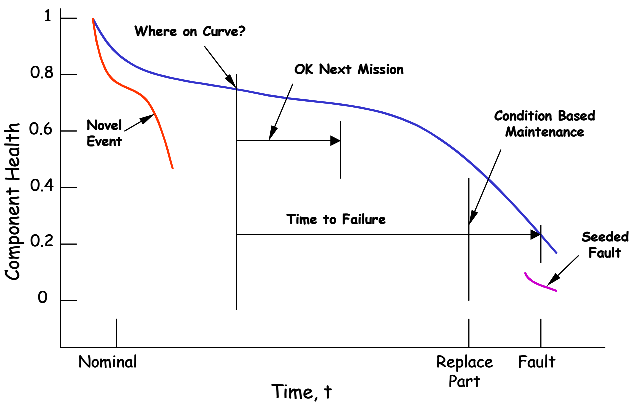

Figure 5 The health-monitoring problem

Component health monitoring determines where the engine is on the curve shown in Figure 5. Is the engine "nominal"? Does some "anomaly" condition exist? Or, is it somewhere between those two extremes? Note that a normal engine health curve may encompass a variety of behaviors and thus this curve represents a single region or single fault trajectory rather than a series of strictly defined points. Determining where we are on the engine health curve is the first step in prognostics.

Fault detection / diagnostic reasoning as discussed above, determines if an engine component has moved away (degraded) from 100% along a known path, as indicated in Figure 5, to a point where engine performance may be compromised. Novelty detection determines if the engine component has moved away from what is considered acceptable nominal operations *and* away from all known fault health (diagnostics as defined above) propagation paths.

Prognosis is the assessment of the engine's current health and a prediction of the engine's future health. There are two variations of the prediction problem. The first prediction type may have just a short horizon time—is the engine good to fly the next mission? The second type is to predict how much time we have before a particular fault will occur and, by extension, how much time we have before we should replace it. Or it may be longer term—tell me when to schedule removal of an engine for overhaul. Accurate prognosis is a requirement for implementing condition based monitoring (CBM).

The creation of a prognostic algorithm is a challenging problem. There are several areas that need to be addressed in order to develop a prognostic that achieves a given level of statistical performance.

What Curve are we on? & Where are we on the Curve?

The first step in prognosis is determining "where" on the overall health curve the component resides. Along with "where" is "what" fault curve we are on. This is similar to the "fault detection" problem. However the equivalent signal-to-noise ratio (SNR) of the related fault signatures that we are looking for to determine component health will be much lower than for the fully developed fault.

This will have two effects. First, because the health component signatures SNR are low, we are always operating in the "gray" area between nominal and a fully developed fault. Because we are in the gray area, even knowing what fault trajectory we are operating on is a challenge. Likely several different fault hypotheses will need to be carried along by the system until a clear-cut condition becomes apparent. Likely a large number of the hypotheses are false so that ultimately no maintenance operation will be required.

Second when we are on the "flat" part of the overall health curve of the component as shown in Figure 6, it is hard to resolve in time where we are on the curve. Suppose that the best we can do in resolving the "health" of a component is to determine that it is in a range of 60-80% of perfect. The component is still quite acceptable. However as indicated in by the green band in Figure 6, we cannot resolve where we are on the curve. Predictions for short time horizons will be reliable (i.e. in determining "good-to-go" for the next mission decisions), but determining remaining life is not possible. The conservative approach would be to assume the worst; that we are at the end of the green part of the curve. Or, we can couple the prognostic with life usage models. The life usage model (assuming one exists) will form the basic estimate of the component health and the prognostic is just used to perturb that basic result.

Prediction uncertainty

Once we determine what the current health of the component is, we need to predict what the health of the component will be sometime in the future. As discussed this prediction can be for a short time horizon or an estimate of the time till the part needs to be replaced or a failure will occur. There are a variety of issues that need to be considered.

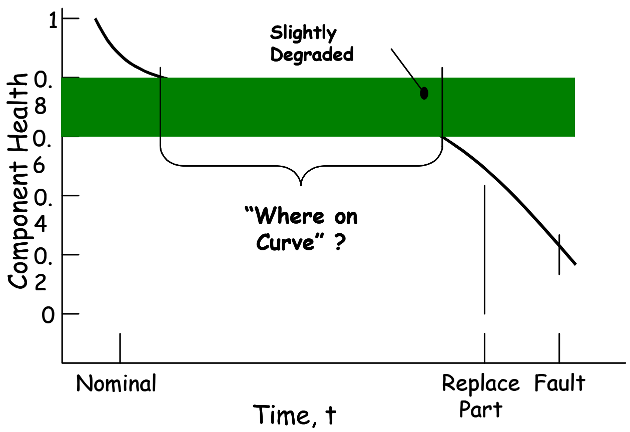

Figure 6 Where are we on the fault curve?

The model will need to accurately predict into the future. Those predictions will be required to be unbiased and to have a small variance in order to be useful. Figure 7 illustrates these problems. In this figure the red line is the prediction of the health of the component from the current state. It is a biased estimate of the true trajectory. However, the model does accurately predict the health / time to replace the component. Is this sufficient?

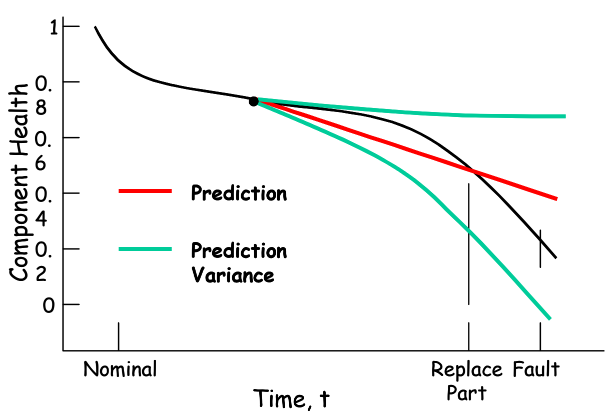

Figure 7 Prediction uncertainty

The green lines represent the error bars for the prediction. The true value of component health curve should fall inside of these error bars as is does. Thus the model is sufficient since it always includes "truth". How useful is it?

The spreading of the error bars defines the time horizon and resolution that can be achieved with this model for performing prognostics. If the error bars spread rapidly then predictions are reliable for only a short time horizon. If they are narrow and follow the true trajectory accurately, then the information from the predictions is useful for longer time horizons.

Data collection = Improved prognostics performance

One of the goals of the DHMS system is the collection of data to not only develop diagnostic and prognostic algorithms but to also improve existing algorithms with knowledge gained. The problem with the time resolution of "where are we on the curve" shown in Figure 6, is the uncertainty of the component's health. As data is collected for a particular fault it may be that this uncertainty will be reduced. This will result in an improvement of determining where we are on the curve as illustrated in Figure 8.

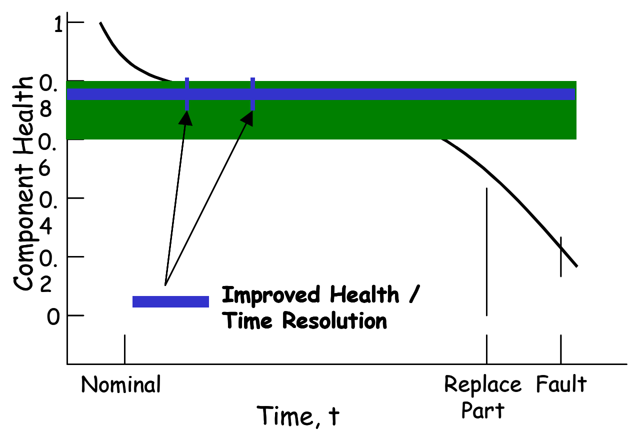

Figure 8 Decreased uncertainty = improved time resolution

Similarly the variation on the prediction of the engine's health into the future will also be reduced. This results in improvement of the prediction uncertainty and thus more reliable estimates of the state of the engine in the future as illustrated in Figure 9.

4. DATA FUSION APPROACHES FOR PROGNOSTICS

There are many different approaches for the development of prognostic algorithms. For practical purposes, these approaches can be generalized into three basic forms. The first are physical models. These are models that have been developed by experts in the component field and validated on large sets of data to show that they are indeed accurate. The second are systems that embody rules of thumb that have been developed and refined by human engineering and maintenance experts. Examples of these systems are rule-based expert systems and fuzzy logic systems. The third are statistical models that 'learn' from examination of real data that contain nominal and known fault conditions. Examples of these are neural net and data mining systems.

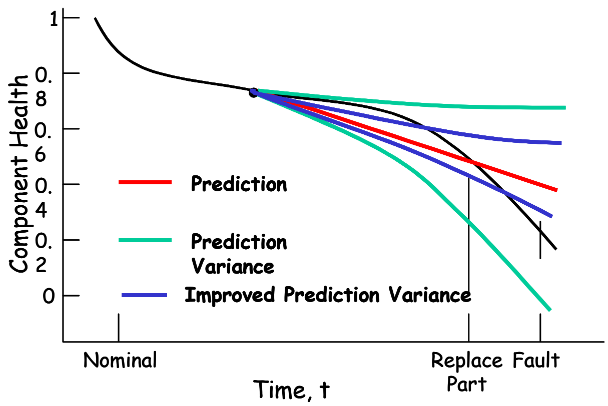

Figure 9 Reduced prediction variance = more reliable prognostics

Physical models and rule-based systems contain information for anticipated fault events that have yet to occur on the component that is being monitored. On the other hand 'learning' systems are good because they can process a wide variety of data types and potentially have performance superior to rule-based systems because they exploit the nuances in the data that are not covered by general rules. This is particularly true for new sources of data for which expert analysis, physical models, and rules have not been developed. Physical models and rule-based systems are only as good as the design engineer can anticipate the variety and nature of faults. Learning systems are only as good as the data from which they have been trained. Obviously with the fusion of these systems the best of all worlds can be achieved.

There are a variety of modeling techniques that are being investigated and included into the development system. These include:

Physical models–-an engine manufacturer is supplying a physical model for a particular Air Force engine. The model is of the gas path of the engine operating under nominal steady state conditions. That model can be used for performing anomaly detection and fault isolation of gas path related faults. It can also be fused with empirical models developed from analysis of real data samples to consider secondary effects that may occur to other engine components such as bearing or oil system failures.

Features–-There are a wide variety of features that may be considered. Features are essentially a compressed representation of the data being analyzed. For spectral data these features include the *condition indicators* (CIs) mentioned previously [1]. There are several 'extensions' of standard spectral estimation that may be useful. Higher order spectra such as the 3rd and 4th order spectra may be useful for non-Gaussian processes; 'instantaneous' time / frequency representations such as the AOK transform, and wavelet representations [7]. Features derived from linear models of the data such as autoregressive (AR), moving average (MA), and autoregressive – moving average (ARMA) models and their poles and zeroes that are associated with the data spectra have proven useful in the past [2]. There are a variety of statistical measures that may be useful such as the skew and kurtosis of data.

Prediction–-Prediction of the trajectory of components over time is being developed for scalar and multivariate data. The algorithms are used to perform prediction of the trajectories of measured features into the future. Prediction can be performed by simple trend analysis such as fitting straight lines or polynomials to scalar data; by extension of fitting polynomials to multivariate data; by regression analysis with models such as autoregressive (AR), moving average (MA), and autoregressive – moving average (ARMA) models, fit to multivariate data; neural networks and fuzzy logic techniques.

Classification–-Once a prediction of a feature or the state of a component is made, classification of the health of the component is required. Empirical classification is primarily a pattern recognition problem. Techniques included are correlation and coherence processing of raw and/or model residual data, neural networks, fuzzy logic, and data mining approaches. Classification also includes novelty detection; the ability to determine that the pattern under test does not match any previously considered pattern. For data mining we are using COTS software developed by Salford Systems. That software allows the user to output a C / C++ version of the decision tree that can then be compiled for a release version of the code.

Reasoning–-For sorting out the causal relationships between observed effects and likely failure modes, we are using a Bayesian Network approach. The development environment COTS software developed by Hugins. That software allows the user to output a C / C++ version of the network that can then be compiled for a release version of the code. A major problem in developing Bayesian Networks is determination of the initial state probabilities and the various transition probabilities associated with different internal states of the network as 'evidence' is included. To do this we are combining the network developed with processing to use the real data collected to determine these probabilities adaptively as new data is collected.

5. EXAMPLE

An example of vibration data fusion for enhanced fault detection is demonstrated with a process used to detect the onset of turbine engine augmentor instabilities known as "rumble" and "screech". These combustion instabilities can be potentially harmful to the engine and are known to cause premature failures in gas turbine components. The rumble and screech events are not true faults. Rather they are the result of current engine operating conditions and environment.

Turbine engines for high-speed applications typically utilize an afterburner or augmentor to increase the thrust and speed of the aircraft for a relatively short period of time. The augmentor portion of the engine is located immediately behind the turbine sections and forward of the exhaust nozzle. It operates similarly to a ramjet, where the core engine bypass air is mixed with fuel, ignited and burned. The augmentor operation often leads to combustion instabilities that can be potentially detrimental to the engine. These instabilities are usually classified as screech and/or rumble.

Augmentor screech is a high-energy tone generated by acoustic feedback loops [8]. An instability wave of the jet is generated by acoustic disturbances near the nozzle exit where the mixing layer is thin and most receptive to excitation. The instability wave grows as it propagates downstream by extracting energy from the flow of the jet. At a distance of about four to five shock cells downstream, the instability wave, having acquired large amplitude, interacts strongly with the shock cell structure inside the jet plume. This interaction results in the emission of intense acoustic waves, some of which propagate upstream outside the jet. Upon reaching the nozzle exit the acoustic disturbances excite the shear layer of the jet, thus generating a new instability wave. In this way the feedback loop is closed. Screech tones from jets operating at low supersonic Mach number are axisymmetric with respect to the jet axis. However, at higher jet Mach numbers they acquire a helical or flapping configuration through the effect known as "mode switching".

Low frequency afterburner instability is known as rumble. Rumble mainly occurs at high fuel-air ratios and at flight Mach numbers and altitudes where low duct inlet air temperatures and pressures exist. Augmentor rumble is generally associated with longitudinal combustion instabilities with acoustic frequencies between 50-100 Hz

In the near field the intensity of these tones can be as high as 160 dB [8]. At this sound pressure level the tones can cause structural fatigue and other undesirable vibratory problems. Design engineers are aware of the predicted screech frequencies when a prototype engine is under development. Care is taken to insure all the augmentor accessories and structural components will not resonate at the predicted screech frequency. Despite these design considerations, an engine that spends a significant amount of time in a screech or rumble condition will experience rapid deterioration of the augmentor components such as flame holders and variable nozzle actuator linkage. It has been documented that sonic fatigue damage can occur to aircraft structures. If combustion instability is detected, reducing the fuel flow to the augmentor can control it. A test cell system able to detect the onset of the screech or rumble condition and warn the test cell engine operator can help prevent premature engine wear.

The data fusion method developed to detect the fault condition involved the use of Principal Component Analysis (PCA). The PCA approach was tested on data from F-100 engine tests at Arnold Engineering Development Center (AEDC) with known rumble conditions and was found to be superior to the standard spectral analysis in detecting the fault condition.

Principal Component Model

Any complex system (mechanical or other) may be characterized by a set of 'features' - results of diagnostic measurements. The feature vector consists of all the available features. In the usual machine monitoring application, large number of measurements of the feature vector may be performed for the nominal state (or, states) of the system. These nominal data are than used to construct an empirical model of the normal state of the system. Principal component analysis is one of many ways to model a system characterized by a large number of features.

Principal components are a set of orthogonal vectors ($\mathbf{v_1}$ $\mathbf{v_2}$ $\mathbf{v_3}$.... $\mathbf{v_N}$) obtained as linear combinations of the original feature vectors. Also, each original feature vector, **F,** may be represented by a linear combination of PCs:

$$\mathbf{F} = a_i\,\mathbf{v_i} + a_2\,\mathbf{v_2} + a_3\,\mathbf{v_3} + \ldots + a_N\,\mathbf{v_N} \qquad (1)$$

Where a_1.... a_N are the principal component coefficients. The principal components represent data features and feature combinations with most of the variance. The number of PCs that can accurately reproduce the data is usually much smaller than the size of the original data vector (N<<length(**F**)), which enables a significant reduction in the number of features that can adequately represent the system. The diagnostic data is then concisely represented by a small number of coefficients a_i. An important feature of the PCA is that correlated features and features with small variance are effectively eliminated. Note that the PCA model encompasses not only the data statistics (mean and variance of each data feature) but also correlations between the features.

Anomaly Detection

As described above the ability of the monitoring system to detect an anomaly is especially important for knowledge-based systems, i.e., systems that in one way or another encode the previous knowledge about the system. The novel data encountered by the system needs to be examined to establish the nature of novelty. After proper labeling, these data can be included in the knowledge base of the system. Anomaly detection is of special importance in problems for which there is very little data in one class, relative to others. This situation is commonly encountered in machinery diagnostics where most of the collected data is normal and anomalies (novelties) may be usually associated with emerging faults.

In the PCA anomaly detection, nominal state of the system is characterized by the principal components determined for the nominal data set. Thus, the nominal data is modeled with a high accuracy. The test data (new data that was not used for PC derivation) are modeled (approximated) with the same set of nominal principal components. A large error in the PC approximation (i.e., inability of the nominal PC set to accurately approximate the actual data) indicates novelty. Thus, the novelty indicator, A is constructed as a measure of the difference between the measured and PC approximated data:

$$A = | \boldsymbol{F} - \boldsymbol{F_0}|, \qquad (2)$$

where $\boldsymbol{F}$ and $\boldsymbol{F_0}$ are the current and nominal feature vectors, respectively. (Other definitions of the novelty may use the root-mean-square (RMS) of the difference).

Moreover, through a comparison of the actual and PC approximated measurements one can identify the features that produced the novelty, which leads to an important function of ‘feature discovery’. The features identified through the feature discovery can be subsequently used to construct fault-specific detection algorithms.

The PCA novelty detection approach has several desirable properties:

- It is a data driven learning system, easily updated to include new diagnostics and measurements.
- Both statistical distributions of the data and feature correlations are modeled.
- The method is computationally efficient.

The PCA anomaly detection approach is particularly useful for spectral data, which are generally characterized by isolated peaks and relatively large number of features (spectral bins) that do not change significantly for measurements performed at different operating regimes.

The PCA anomaly detection was applied for feature discovery and for construction of rumble detector in a gas turbine. The preliminary results shown below were obtained for a limited dataset but show the usefulness of this approach.

Experimental Set-Up

The data used in this study were obtained from F100-PW-220 engine test at AEDC. Three data sets with varying magnitude of the rumble instability (no rumble, small and large rumble) were examined. Each data set included signals from four vibration sensors and one pressure sensor sampled at 5000 Hz. The vibration sensors were standard case mounted accelerometers that were located on the engine as shown in Table 2 and Figure 10. The pressure sensor was installed in the augmenter at station 6. For each data set, the engine was operated with a Snap acceleration from Idle to Max.

Spectral Analysis

Figure 11 shows the results of spectral analysis of the vibration data for the run with large rumble run. The spectral data were obtained with asynchronous frequency domain processing of the time-domain vibration data. The processing included filtering, decimation and fast Fourier transform. It was performed for moving time window with 90% overlap between consecutive windows. The frequency range was from 2 to 100 Hz. The pressure sensor shows a clear signature of the rumble in the frequency range of about 50 to 76 Hz (time 50 and later in Figure 11). This is in agreement with the expected frequency range of 50-100 Hz. There seems to be a small signature of rumble in the vibration spectra for sensor VKFLV.

Table 2. *F100-PW-220 Vibration Sensors*

Sensor	Name	Description
Vib 1	V12VCR	12^{th} Vane and Case Radial
Vib 2	VAP6	Access Port 6 Low Pressure Turbine
Vib 3	VKFLV	K-Flange Vertical
Vib 4	VMGBV	Accessory Gearbox Vertical
Pressure	PAB68	Afterburner Duct Static Pressure

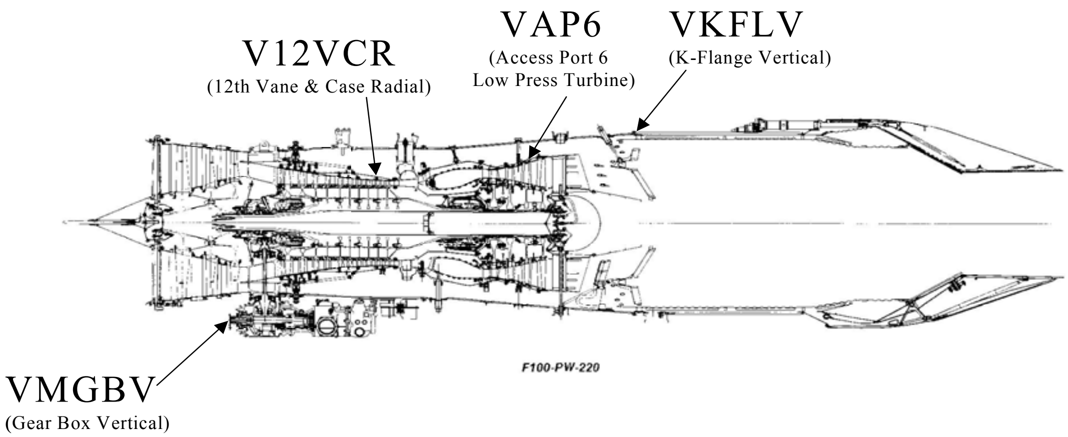

Figure 10. Locations of vibration sensors for diagnostic of F100-PW-220 engine.

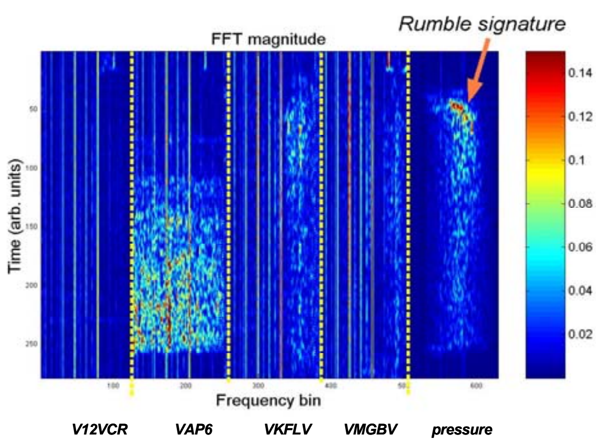

Figure 11. Spectral analysis of vibration and pressure data for the run with large rumble instability. Data for all available sensors are shown side-by-side. The frequency range is 2-100 Hz. The rumble signature is seen clearly in the pressure sensor spectrum

Feature Discovery with PCA

Rumble is an acoustical phenomenon and thus is readily seen in the pressure signal. Our aim was to identify its signature in the vibration measurements and to construct a rumble detector based solely on these measurements.

The PCA anomaly detection technique was applied to the composite frequency spectrum consisting of spectra obtained for four vibration sensors in the frequency range from 50 to 100 Hz. As outlined above, the principal components were determined for the normal spectral data, i.e., the data obtained for the run that did not show the rumble signature in the pressure sensor signal. Other processing details were as described above for the spectral data.

Figures 12 and 13 show the plots of PCA anomaly indicator defined as the absolute value of the difference between the actual spectra and their PC approximation (cf. Equation 2), obtained with the use of 30 most significant principal components. These 30 components account for about 97% of the variance in the original (normal) data and, as expected, provide an excellent approximation for the whole normal data set. For small rumble (Figure 12) and large rumble (Figure 13) cases, the anomaly is detected for VKFLV sensor in the frequency range from about 68 to 80 Hz (time ~ 50-100), and for VAP6 sensor (time > 100 for the small rumble set, and time > 150 for the large rumble set). The novelty detected for VKFLV is well correlated with the rumble signature in the pressure signal and thus, can be used as the rumble indicator.

As seen, though four channels of vibration measurements were used in the processing, only a single channel is really required in order to perform the detection. However the absence of a signal in the other three channels maybe used to isolate the rumble event.

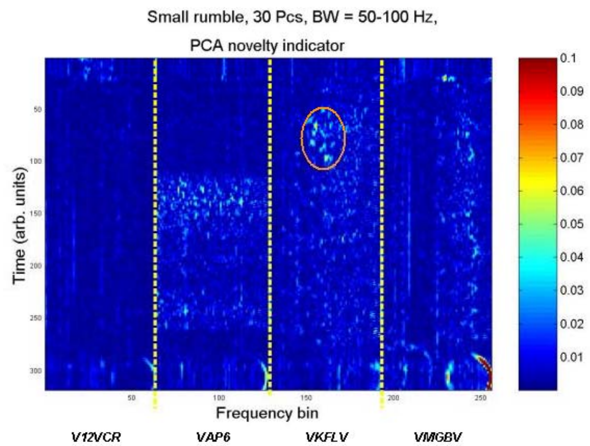

Figure 12. PCA novelty for the test run with low level rumble instability. Data for all available (four) vibration sensors are shown side-by-side. The frequency range is 50-100 Hz. Novelty associated with rumble is seen in the third sensor (VKFLV) signal in the range 68.5 – 79 Hz (circled area).

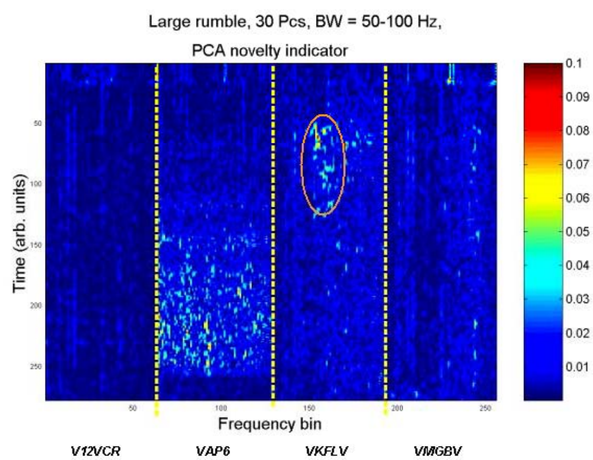

Figure 13. PCA novelty for the test run with large rumble instability. Data for all available (four) vibration sensors are shown side-by-side. Novelty associated with rumble is seen in the third sensor (VKFLV) signal in the range 68.5 – 79 Hz (circled area).

The origin of VAP6 anomaly signature is not clear. It appears to occur over a broad frequency range and thus, it may be associated with broadband noise. It is worth mentioning that there is a qualitative difference between the novelty in VAP6 and VKFLV signals. We were able to detect the VKFLV signature using purely statistical approach, in which the statistical (Gauss) distribution is evaluated for each of the frequency bins for the normal data, and the novel features are detected as those that deviate substantially from this nominal distribution. This approach, unlike the PCA, does not take into account possible correlations between diagnostic features (frequency bins). The VAP6 novelty is not detected by the statistical approach. Thus, for the present data, the VKFLV rumble novelty seems to have predominantly 'statistical' character while the VAP6 novelty is associated with correlations between the frequency bins.

Rumble Indicator

The rumble indicator signal is obtained as an integral of the novelty indicator signal over the frequency range 68.5 – 79 Hz. Figure 14 shows the amplitude of this signal for the three experimental runs: normal (no rumble), low level rumble and large rumble.

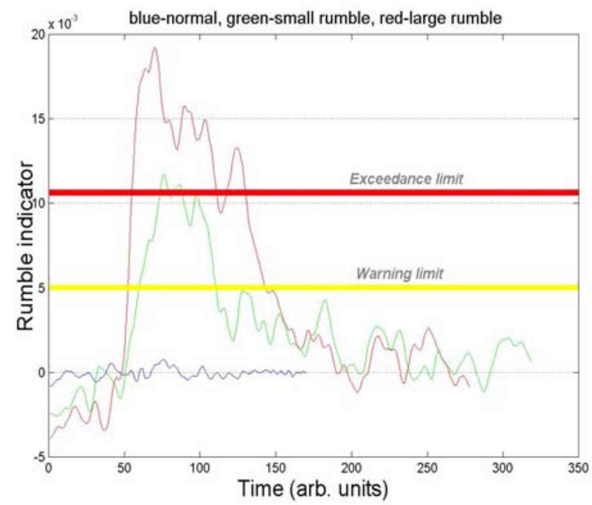

Figure 14. Amplitude of PCA novelty based rumble indicator as a function of time for the three available data sets. The exceedance and caution limits are arbitrary and indicate how this signal may be used in a real-time diagnostic system.

6. Concluding Remarks and Recommendations

The development of advanced diagnostic and prognostic techniques is challenging and an area of active research. Successful development requires that a number of issues be addressed. Among these issues is need for continuous collection of data during the lifetime of an engine. As the life of an engine progresses along there will be evolving, changing classes of engine faults and accompanying changes in diagnostic and prognostic systems are needed to detect them.

Improvements and advances in algorithm development to process data from new and improved sensors will be needed. This will allow for improved detection and classification of faults. These health management systems will also need to be modular in design and easy to maintain. This follows from need for the system to continually adapt to a changing environment.

This paper describes a *Distributed Health Management System* (DHMS) testbed for development of advanced engine prognostics systems. The DHMS is currently under development at IAC. It will be used to perform collection and monitoring of helicopter vibration and engine performance data. It is also used for the development and application of data fusion techniques for health monitoring of gas turbine engines. The system uses a combination of signal and information processing algorithms to perform data fusion for engine fault diagnostics and prognostics to support individual aircraft field maintenance, fleet maintenance, as well as the development of new diagnostics and prognostics algorithms using real data.

7. REFERENCES

[1] Grabill, P., Berry, J., Grant, L., and Porter, J, "Automated Helicopter Vibration Diagnostics for the US Army and National Guard",

[2] Brotherton, T.W., T. Johnson, and G. Chadderdon, "Classification and novelty detection using linear models and a class-dependent elliptical basis function neural network," *Proc. of The Int'l Joint Conf. on Neural Networks* (IJCNN), Anchorage, AK, May 1998.

[3] Brotherton, T. and T. Johnson, "Anomaly Detection for Advanced Military Aircraft using Neural Networks", *The 2001 IEEE Aerospace Conference,* Big Sky, MT, March 2001.

[4] Brotherton, T.W. and R. Mackey, "Anomaly Detector Fusion Processing for Advanced Military Aircraft", *The 2001 IEEE Aerospace Conference,* Big Sky, MT, March 2001.

[5] Zakrajsek, J, "An Investigation Of Gear Mesh Failure Prediction Techniques," NASA Technical Memorandum 102340, AVSCOM Technical Memorandum 89-C-005, Nov 1989.

[6] Braun, S., "Mechanical Signature Analysis Theory and Applications", Academic Press INC. ISBN 0-12-127255-9 1986.

[7] Marple S.L., T. Brotherton and D.L. Jones, "The Application of Time-Frequency and Wavelets to Medical Ultrasound" in *Time-Frequency and Wavelets in Biomedical Engineering*, IEEE Press, 1997

[8] Powell, A., "On the Noise Emanating from a Two dimensional Jet Above the Critical Pressure," Aeronautical Quarterly, Vol. 4, Feb. 1953, pp. 103-122

Tom Brotherton *received his B.S. degree from Cornell University (1974), an M.A.Sc. from the University of Toronto (1976) and the Ph.D. from the University of Hawaii (1982) all in electrical engineering. He was an assistant professor in the Information and Computer Sciences Department at the University of Hawaii in 1983 and with the Orincon Corp from 1983-1999. He is currently a VP of the Intelligent Automation Corp. (IAC) in San Diego, CA. IAC is involved with the development of aircraft and related equipment monitoring software and hardware. Dr. Brotherton is also on the editorial board for the IEEE Press series on Biomedical Engineering. His interests are in the development of adaptive signal and data fusion techniques for machine condition, fault monitoring, and prognostics as well as medical systems applications.*

Paul Grabill *received his B.S. degree in mechanical engineering in 1986 at the University of Cincinnati. He worked for 10 years at General Electric Aircraft Engines as a dynamics engineer where he was involved in gas turbine vibration testing, engine balancing, and system dynamic analysis. He received his M.S. degree in mechanical engineering in 1992 at the University of Cincinnati where he was sponsored by the NASA Health Monitoring Center and UC's Structural Dynamics Research Laboratory. He is currently Vice President of Engineering at Intelligent Automation Corporation in San Diego where he is currently developing advanced diagnostic and prognostic systems for helicopters, gas turbines, and aircraft.*

Dariusz Wroblewski *received M.S. degree in Physics from the Warsaw University, Warsaw, Poland, and Ph.D. degree in Electrical Engineering from the University of Wisconsin-Madison, Wisconsin, USA, in 1984. He held research positions at the Johns Hopkins University and the Lawrence Livermore National Laboratory, where he worked on the physics of high-temperature thermonuclear plasma. His contributions to the development of specialized spectroscopic diagnostics for measurement of plasma internal magnetic field resulted in major advances in the understanding of magnetic confinement of plasma. Presently, Dr. Wroblewski is Chief Scientist at Intelligent Automation Corporation in San Diego, California, where he specializes in development of numerical methods for modeling of complex systems. His recent work has dealt with applications of artificial intelligence methods to* *problems in machine health monitoring, medical diagnostics, bio-informatics, and monitoring of biological systems.*

Wing Commander Richard Friend *is a Royal Air Force engineering officer serving on an exchange program in the Propulsion Directorate at the United States Air Force Research Laboratory, Wright Patterson Air Force Base, Ohio. He is the research group leader for USAF Engine Health Management and Control, the Secretariat for the Steering Committee for the High Cycle Fatigue Science and Technology Initiative, and is the USAF lead for the Joint Strike Fighter engine Prognostic Health Management systems. An aeromechanical engineer, he is specialized in gas turbine engines and has accrued over 24-years experience in the maintenance and management of fighter aircraft, engines and engine test facilities. For the last 3 years he has been involved with specifying a new initiative on intelligent engines under the USAF's new research thrust, the Versatile Affordable Advanced Turbine Engine; VAATE.*

No picture available

Bill Sotomayer *is currently a Ph.D. candidate at the University Of Dayton. He became actively involved in Engine Health Management (EHM) approximately one year ago. Previously, he was involved in friction modeling for blade/disk interfaces on high performance jet engines.*

John Berry *received his B.S., M.S. and Ph.D. degrees from the Georgia Institute of Technology (1974, 78, 90), all in aerospace engineering. He served for 8 years as a signal officer in the US Army. From 1982 until 1998, Dr. Berry served the US Army as a civilian research engineer in experimental and analytical rotorcraft aerodynamics at the NASA Langley Research Center. Since 1998 he has been an aeromechanics branch chief for the Aviation and Missile Research, Development, and Engineering Center. Dr. Berry is a past associate-editor for the Journal of the American Helicopter Society and contributes actively to that professional society. His interests are in the development of innovative technologies that contribute to improved rotorcraft aeromechanical systems.*