Proceedings of ASME Turbo Expo 2011
GT2011
June 6-10, 2011, Vancouver, British Columbia, Canada

GT2011-46019

A DATA ANALYTICS APPROACH TO FAILURE PRECURSOR DETECTION OF GAS TURBINE

Ousmane Diallo
Georgia Institute of Technology
Atlanta, GA, USA

Dimitri Mavris, PhD
Georgia Institute of Technology
Atlanta, GA,USA

ABSTRACT

In this paper, a novel approach is proposed to detect precursory events that lead to catastrophic systems failures. This approach is applied to investigating failures of heavy duty gas turbines. Current industry standards rely on either vibration sensors or gas path performance measurement sensors to identify system anomalies, but this proposed process is based on a combination of information from both type of monitoring sensors. This process is built on a systematical multi-step concept developed by assembling proven mathematical and statistical signal processing techniques to achieve a robust and more precise failure precursor detection methodology. The first step includes performing a multi-resolution analysis of gas turbines gas path performance measurement parameters, condition monitoring and vibration sensors data using wavelet packet transform to extract their signal features. Then, the probabilistic principal component analysis is utilized to fuse data of different types into a set of uncorrelated principal components. Next, a one-dimensional signal representing the multi-variable data is computed. After that a statistical process control technique is applied to set the anomaly threshold. Finally, a Bayesian hypothesis testing method is applied to the monitored signal for abnormality detection. As a proof of concept, the proposed process is successfully applied to a gas turbine compressor failure precursor detection problem.

1 INTRODUCTION

As a consequence of the recent deregulation in the electrical power production industry, there has been a shift in the traditional ownership of power plants and the way they are operated. Many new private entrepreneurs with no prior experience in power plant operation have entered into the power generation business. To hedge their business risks, those private entrepreneurs enter into long-term service agreements (LTSA) with third parties service providers for their operation and maintenance (O&M) activities with whom they share both the risks and the eventual rewards of plant performance.

As a result, the original equipment manufacturers (OEMs) become the natural choices as third party O&M providers because they are experts of their designed products and will be willing to guarantee their operation. In return the OEM becomes responsible for the majority of the costs associated with unplanned outages like usual insurances contracts. Because the estimated cost-benefit of preventing such unplanned outages as a gas turbine compressor failure is very high, techniques for detecting failure precursors to avoid or limit the number of systems catastrophic failures are an absolute necessity. Each of the main gas turbine OEMs (together they represent about 94% of the global market [1]) has its own set of definitions and foreseeable benefits to the plant owners of their LTSA offerings. The major OEMs have been developing preventive maintenance strategies to minimize the occurrence of the unplanned outages resulting from failures of equipment covered under LTSA contracts. The high potential for cost benefits to gas turbine OEMs when failures can be prevented raises the importance of techniques for detecting faults in gas turbines. In this paper, a systematic process is proposed that can successfully detect failure precursory events.

The remaining of the paper is organized as follow: Section 2 sets the context and background regarding power plant O&M and the background for the problem addressed. Section 3 presents the steps of the proposed approach to detect catastrophic failure precursors. Illustrative examples of application to a gas turbine compressor failure problem are presented in Section 4, followed by a brief conclusion in Section 5.

2 POWER PLANT O&M BACKGROUND

Typically, the LTSA contracts work like insurance policies where the manufacturer guarantees a given level of power output and/or efficiency over several years. They may provide repair, replacement, and upgrade parts to the degrading power plant. Overall, it is supposed to be a "win-win" partnership for both parties, as they share the operational risks as well as the rewards of extra performance generated by the power plant. There are several other advantages for both parties to entering into LTSA. For the plant owner, an advantage is that it is well accepted in the power generation field that LTSA contracts raise the plant re-sale value. While for the OEM, the equipment under contract provides unprecedented access to "a live laboratory" that should allow the OEM to learn from eventual design shortcomings of previous gas turbine designs in order to improve upon future designs.

2.1 Power plant operation and maintenance

The O&M expenditures of a typical power plant are significant and consist of 15% to 20% of total life cycle cost, while equipment maintenance costs account for approximately 10% to 15% [2]. To be clear, there is always a cost associated with an outage whether it is planned or unplanned. Hence, to make the LTSA contracts profitable, the providers need to reduce the number of unplanned outages. Typically under a LTSA contract, the provider has to pay the plant owners a liquidated damage for each forced outage. In general, the liquidated damage cost for a forced outage includes: the loss of production cost, the repair cost, the cost of buying the equivalent power to meet the quantity that the forced outage plant was dispatched for at usually higher prices, and eventual regulatory penalties.

2.2 Problem Background

With the high cost of liquidated damage associated with not meeting the reliability and performance requirement, LTSA providers need to develop strategies so that the revenues from the contracts exceed the cost of the involved risks. In fact, according to a report of the Electric Power Research Institute (EPRI), the cost benefit from preventing a typical a gas turbine compressor failure is estimated between ten and twenty million dollars [3]. Accordingly, OEMs have spent huge resources to develop strategies to avoid unplanned plant outages. For example, OEMs like GE Energy created a Power Answer Center in Atlanta, GA, where all power plants under its LTSA contract are continuously monitored using installed sensors on gas turbine. The illustrative Figure 1 shows the GE Power Answer architecture wherein the on-site monitor compares the actual unit performance with baseline predictions and provides the first level of anomaly detection and notification.

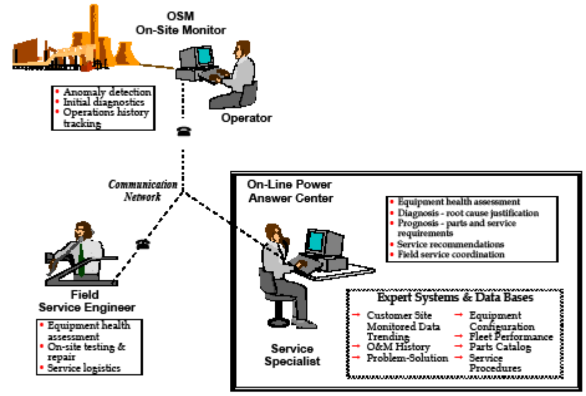


Figure 1: GE Monitoring & Diagnostics concept [1].

Major OEMs have the ability to monitor hundreds of units throughout the world in real time in order to establish knowledge to detect faults before they can develop into failure. This is both challenging and can yield some advantages toward sustaining the technological competitive advantage of an OEM in the long run. Despite all of the efforts to avoid forced outages, there are still undetected failure precursors that led to catastrophic failure as reported by EPRI in its 2007 updated report [4].

3 PROPOSED APPROACH FOR FAILURE PRECURSORS DETECTION

Though in recent years, there have been new and improved techniques such as condition-based monitoring (CBM) to help detect anomalies in their early stages of development. However, the new techniques are not able to totally resolve the issue of missed detections of all the anomalies. Their merit is well accepted, their practical implementation is still inefficient because these techniques tend to be theoretical, difficult, and/or expensive to apply to real world problems. Therefore, the method proposed herein intends to take advantage of the two types of monitoring sensors to combine the information from the performance sensors and the vibration sensors to capture catastrophic failure precursors.

In general, the gas path performance and condition of power plants are monitored using two types of sensors: the static or process-related sensors (used to measure temperature, pressure, and flow), and the sensors characterized by their high-bandwidth used for high-frequency signals like the vibration measurements. Although, there are many time-frequency techniques reported in the literature such as the Wigner-Ville distribution, the Choi-Williams distribution, the short time Fourier transforms; the wavelet transform is the best one to deal with short lasting anomalies and sharp discontinuities [5]. The following subsections provide a brief overview of the wavelet transform followed by a presentation of a step-by-step explanation of the proposed approach.

3.1 Fourier transforms overview

The time-frequency analysis techniques are appropriate when dealing with identifying anomalies in time series signals because more information can be extracted about small variations of a signal in the combination of the time and the frequency domains than can be extracted in the time domain alone. The most popular frequency domain analysis technique is the Fourier transform because of its ability to decompose an energy limited signal time domain signal f(t) into its frequency domain contents $F(\omega)$ as defined by Eqs 1 and 2:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega \qquad (1)$$

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt \qquad (2)$$

However, the Fourier transform provides only the global information on the frequencies of a signal, it cannot provide local information if the spectral composition of a signal that changes rapidly with time [6]. In other words, once a signal is Fourier transformed, all the time domain information is lost, while the wavelet transform conserves both the time and the frequency information. Thus, the wavelet transform is an improvement over Fourier transforms for time-frequency analysis in that context. Wavelet transforms decompose a given signal through two filters: a low-pass filter that provides a low frequency part which trends and smoothes the original signal (i.e., approximation), and a high-pass filter that provides the high frequency part (i.e., details) which reveals local properties such as anomalies.

3.1.1 Mathematical overview of Wavelet Transforms

There is plenty of literature describing the theory of wavelet transforms and its applications [7, 8]. Just like the Fourier transforms, the wavelet transform can be defined for any square-integrable function $L^2(\Re)$ [9]. But instead of using the harmonics, $e^{i\omega t}$, the wavelet basis, ψ, called a mother wavelet function, is used and defined as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \qquad (3)$$

Where a is the dilation or scaling parameter and b is the time location or translation parameter. Thus the wavelet transform of a signal f(t) is computed as follows [8]:

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt \qquad (4)$$

3.1.2 Wavelet Packets

The standard wavelet transforms has limitations because it can only decompose the low-frequency part of a signal. To remedy that limitation, the wavelet packet transform was introduced. It has the ability to decompose both the approximation part as well as the detail part. The wavelet packet transform decomposes a signal into more detailed components than the standard wavelet transform could, thereby yielding more information about the signal. For that reason, it is more advantageous to use the wavelet packet transform to realize the multi-resolution analysis (MRA) by decomposing both the low frequency and high frequency components of a signal into subspaces so as to obtain finer and adjustable resolution [10]. Figure **2** illustrates a wavelet packet decomposition of a signal S. For a given signal original S(t) at level (0,0), the standard Wavelet decomposition would yield equation (5), while the Wavelet packet decomposition provides equation (6). On the tree decomposition, the nodes with the second entry of “0” represent approximation, whereas the nodes that have both entries as non-zero represent the detail components of the decomposition. Clearly, the Wavelet packet provides a higher level of detail (e.g. eight wavelet components) than the standard wavelet (i.e. five wavelet components).

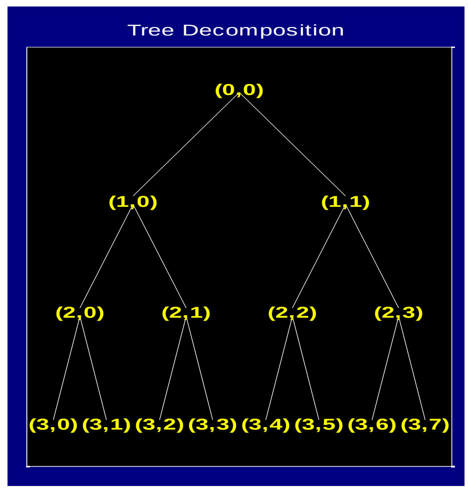


Figure 2: Wavelet packet decomposition

$$S(t) = (1{,}0) + (3{,}4) + (3{,}5) + (3{,}6) + (3{,}7) \qquad (5)$$

$$S(t) = (3{,}0) + (3{,}1) + (3{,}2) + (3{,}3) + (3{,}4) + (3{,}5) + (3{,}6) + (3{,}7) \qquad (6)$$

3.2 Steps for failure precursor detection

As mentioned above, the proposed approach intends to take advantage of monitoring sensors to capture catastrophic failure precursors. Figure 3 shows a flowchart of the proposed methodology for intelligent failure precursor detection using multi-resolution analysis. A step-by-step explanation of each block in the flowchart is presented in the subsections below. The different steps of the proposed approach have been explained in [11].

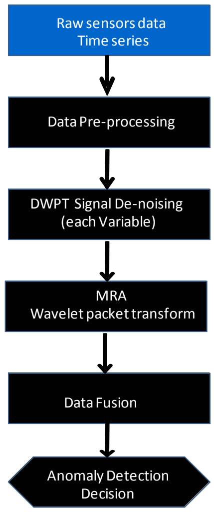


Figure 3: Intelligent failure precursor detection

3.2.1 Raw time series data collection

The systems performance and operating condition parameters are continuously monitored and collected using installed sensors and stored for potential post-processing. The installed sensors for heavy-duty gas turbines typically include the two types mentioned previously, static or process-related sensors (used for pressure, temperature, and flow rate measurements) and high-bandwidth sensors used to measure high-frequency measurements (e.g., vibration measurement).

3.2.2 Data Pre-Processing

The pre-processing of the raw data is a necessary step for a couple of important reasons. This step is necessary to allow an apples-to-apples comparison of a system overall health regardless of the ambient condition and the operating condition. First, the OEMs will not want to share their proprietary data on equipment malfunctioning because that may affect their competitive advantages due to risk of possibility of reverse engineering. Secondly, the sensors monitor different engine parameters (e.g., temperature, pressure, vibration, etc) that are recorded in different units and more importantly in different orders of magnitude. For instance, a typical normal base load operation of a gas turbine can have a compressor discharge temperature measurement in the range of 600 to 800 degrees Fahrenheit, while the vibration sensor measurements could be on the order of 1/10 of an inch per second. Therefore, an analysis with the raw measurement could be artificially skewed towards the variables with higher absolute values. Thus, the pre-processing step consists of normalizing each measured parameter value by the mean value of that variable measurement, and eliminating the visual outliers that would misrepresent the finding and affect the accuracy of the conclusion.

3.2.3 DWPT signal de-noising

The de-noising step is essential because a sensor measurement signal is always tainted by noise. In [12] the authors presented a de-noising technique that adequately removes the noise by combining the discrete wavelet packet transform (DWPT) and Bayesian thresholding. The result is the removal of just the noise without the drawbacks of many other de-noising algorithms that either remove useful information along with the noise or remove too little noise thus leaving some noise in the signal.

3.2.4 Multi-resolution analysis using discrete wavelet packet decomposition

In this step, the de-noised signal is decomposed at an appropriate level (3-level) of resolution (as in Figure 2) to get the approximation and the detail components. The content of each component resulting from the decomposition can be analyzed. Once the decomposed tree is obtained, the energy content of the scaling function (approximation) and the wavelet functions (details) representing the nodes of the tree is calculated as:

$$E_{j,m} = \int_{-\infty}^{\infty} (W_{j,m,s}(dt))^2 \, dt = \sum_s W_{j,m,s}^2 \qquad (7)$$

Where: $W_{j,m,s}$ is the wavelet packet transform coefficient, j is the level, s is the translation, and m is the modulation parameter (approximation or detail). The energy content of each node will then be used as the signal features.

3.2.5 Data Fusion using Probabilistic Principal Components Analysis

The goal of the data fusion step is to combine pieces of information from a system with potentially correlated multi-sensory data set into fewer uncorrelated variables that allow for drawing a more adequate conclusion than one could get from each individual sensor. Thus, the probabilistic principal component analysis (PPCA) is used to merge the information from the sensors of interest. To perform the PPCA, the steps of the principal components analysis (PCA) are executed. Then the maximum likelihood and the variance of the reduced data are calculated. Only the most significant weights obtained from the standard PCA are used as entries in a maximum likelihood matrix. The PPCA is an improved version of the standard PCA because it takes advantage of data uncertainty [13].

3.2.6 Anomaly detection decision

The anomaly detection decision is a multi-step process. It is important to get the anomaly detection decision right as it represents the core of the proposed methodology.
After completion of the PPCA step, the different principal components of the signal as obtained from PPCA are converted into a one-dimensional signal calculated as follows:

$$RS = \sum_{i=1}^{r} \lambda_g P_g^*(i) \qquad (8)$$

r : The number of retained principal components (PC)

λ_g : The contribution of the g^{th} Eigenvalues.

$P_g^*(i)$: The signal corresponding to the g^{th} principal component of the data matrix.

RS: reconstructed signal.

Statistical Process Control (SPC) for threshold

The authors presented the steps of the SPC in [14] as follows:

Step1: The obtained reconstructed signal RS is decomposed using the discrete Wavelet packet decomposition. Then, the energy content of each node is calculated (similar to the multi-resolution analysis step) using equation 7.

Step 2: Calculate damage indicators SAD and SSD to be monitored instead of directly monitoring the change of the energy content as suggested in [15]. Instead of using the SAD and SSD as defined in [16]. The authors proposed the use of the modified version of SAD and SSD as introduced in [14]:

Sum of Absolute Difference (SAD) and computed as:

$$SAD(k) = \left| E(k) - E_{ref} \right| \qquad (9)$$

Square Sum of Difference (SSD) and computed as:

$$SSD(k) = (E(k) - E_{ref})^2 \qquad (10)$$

With: E_{ref} : is the reference signal energy content at the approximation node of the decomposed tree, calculated as the mean value of the energy content over a healthy period before the monitoring period.

$E(k)$: Energy content of the RS at the monitoring time step k

k : Indice that starts right after the interval over which E_{ref} is computed.

The damage indicators are the deviation of energy content from the reference energy E_{ref} .

Step 3: Apply SPC (Statistical Process Control)

The X-bar control chart concept [17] is used to established the threshold of damage indication. Thus, the threshold for a one-sided upper (1-α) upper confidence limit for the SAD damage indicator monitor separately can be calculated as follows (a similar threshold is calculated for SSD) [18]:

$$UL_{SAD}^{\alpha} = \mu_{SAD} + Z_{\alpha}\left(\frac{\sigma_{SAD}}{\sqrt{q}}\right) \qquad (11)$$

Where: UL_{SAD}^{α} : Upper Confidence Limit

μ_{SAD} : is the value toward which the mean value of the parameter SAD converges

Z_{α} : is the value of standard normal distribution with zero mean and unit variance, so that the cumulative probability is 100*(1-α)

σ_{SAD} : is the value toward which the standard deviation of the parameter SAD converges

q : Number of interval of monitoring time step

α : is the acceptable error or type I error

Then, X-bar control chart upper limit is used to monitor of the damage indicators over a given period of time. The different statistical parameters are obtained after the system stabilized (see section 4 for illustration).

Bayesian Hypothesis for monitoring time

In this paper, it is proposed to apply the Bayesian evaluation method to the modified threshold value $\varepsilon = UL_{SAD\_MX}^{\alpha}$, which corresponds to the use of the maximum value reached by μ_{SAD} in the equation (11) instead of its converged value. It is important to use the modified threshold because it has been observed that there is an overshoot before convergence occurred. Thus, the Bayesian evaluation method for hypothesis testing is conducted with a binary outcome over a given period of monitoring time to facilitate the monitoring process overtime. The anomaly function is defined as $h(t)$ which is the vector of the Bayesian hypothesis testing result and is based on answering the question, is the system healthy? The null and the alternative hypotheses are defined as follows:

- Null hypothesis H0:

$$SAD(t) \le \varepsilon, h(t) = 1 \qquad (12)$$

- Alternative hypothesis H1:

$$SAD(t) > \varepsilon, h(t) = 0 \qquad (13)$$

The function $h(t)$ has values of 1 or 0 and can be plotted over time for visualization. Where a $h(t)$ value of 1 corresponds to a healthy state while a $h(t)$ value of 0 is an abnormal one. Therefore, the appearance of the value of $h(t) = 0$ can be considered as a sign of failure precursor. It is important for practical purposes to reset (recalculate) the threshold value (i.e. all the parameters used to calculate the threshold) after any exterior performance change as offline compressor water-wash, or installation of new parts or components.

Type I and type II errors calculation

Recall that the probability of a type I error or false-positive is defined as: α = P{reject H0|H0 is true}; this is the probability of detecting a failure precursor while there is no defect. Whereas, the type II error or false-Negative is defined as β = P{fail to reject H0|H0 is false}, that is the probability of missing a defect while one is present. In the proposed process the statistical confidence level is decided by the system operator; it has a probability of 100*(1-α)%.

To compute the type II error we assume H0 is false and H1 is true, and that the difference between the mean values of the H0 distribution and the H1 distribution is δ. The type II error is the probability that the test statistic will fall between $-Z_{\alpha/2}$ and $Z_{\alpha/2}$ under H1 being true, as illustrated in figure 4. A more detail explanation of the concept of determination can be seen in [17].

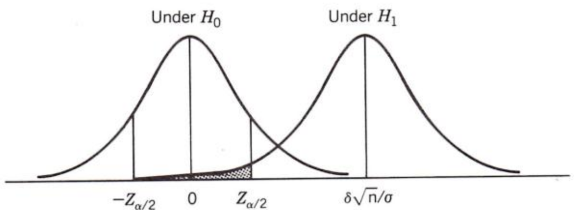


Figure 4: Graphical representation of type II Error

The type II error can be calculated for each damage indicator as follows:

$$\beta = \Phi\left(Z_{\alpha/2} - \frac{\delta\sqrt{q}}{\sigma} \right) - \Phi\left(-Z_{\alpha/2} - \frac{\delta\sqrt{q}}{\sigma} \right) \qquad (14)$$

Where: Φ : is the cumulative standard normal distribution

δ : is the difference between the mean value used to calculate the threshold value and the mean value of the monitored interval of time of the damage indicator SAD and SSD.
σ : is the standard deviation
Another interesting statistical parameter is the process power defined as 1–β = P{reject H0|H0 is false}; it is the probability of correctly rejecting H0.
The steps of the anomaly decision are summarized in Figure **5**.

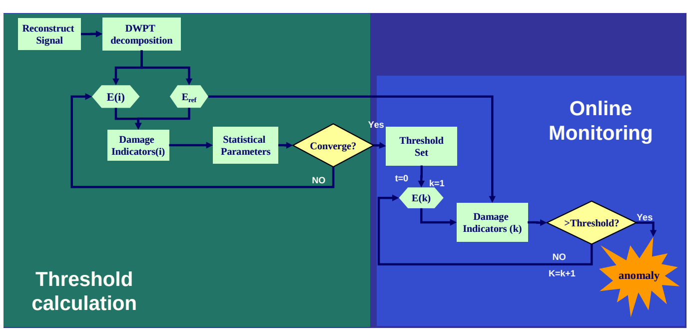


Figure 5: Anomaly Detection Decision Flowchart

4 EXAMPLE APPLICATION TO A GAS TURBINE COMPRESSOR FAILURE

To demonstrate its practical value, the proposed methodology is applied to a set of base load operating condition of heavy-duty gas turbine compressor failure precursor detection.

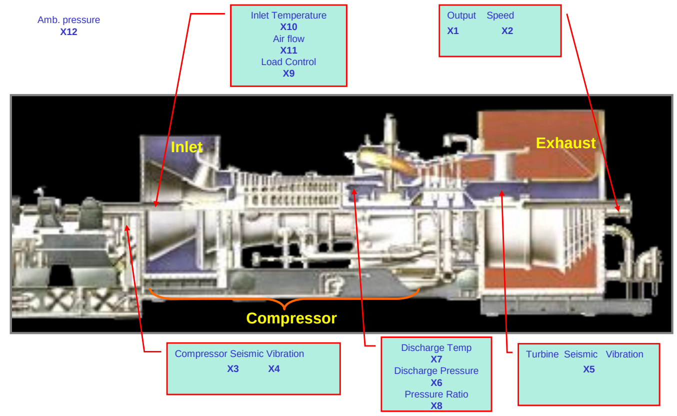


Figure 6: Layout of monitored sensors for gas turbine compressor anomaly detection

The layout of sensors that are monitored is shown on Figure **6**. Table **1** provides the description of the twelve monitored sensors (variables).

Table 1: Monitored GT sensors for compressor failure

Variable	Variables Description
X1	Overall system health parameter 1
X2	Overall system health parameter 1
X3	Compressor seismic vibration 2
X4	Compressor seismic vibration 2
X5	Turbine seismic vibration 3
X6	Compressor health parameter 1
X7	Compressor health parameter 2
X8	Compressor effectiveness health parameter 1
X9	operating condition (load)
X10	operation condition 1 (environment)
X11	Compressor effectiveness health parameter 2
X12	operation condition 2 (environment)

The remainder of this section is divided into two parts:

1. Illustration of the proposed method applied to a case of compressor failure with only eight available sensors measurement
2. Illustration of the benefit of combining both the vibration and the gas path performance measurement sensors.

4.1 Illustration of Failure Precursor Detection Method

4.1.1 Background of test unit

The proposed process is applied to a gas turbine compressor failure problem where the working sensors are X1, X3, X4, X5, X7, X8, X9 and X10 (as shown on Figure 6). The test unit failed on June 24, 2006 at 18:18. The gas turbine manufacturer found through a post compressor failure analysis that there was a failure precursor event (artificially big increased in sensor data) on June 20, 2006 at 23:30. Also, the manufacturer indicated that the operated hours of the unit were about one half the number hours required for inspection and that there were no major events prior to the compressor failure.

4.1.2 Proposed methodology steps

Step 1*:* The sensor measurements for the eight sensors of interest at 5-second intervals from June 19, 2006 at 00:00 to June 25, 2006 at 00:00 are obtained and presented in this study because there were no prior noticeable events.

Step 2*:* The raw data are normalized using the mean value of each variable. As a result, the normalized sensor readings are within the same order of magnitude with a mean value of 1 for each variable.

Step 3*:* All normalized raw sensors data are de-noised using DWPT.

Step 4*:* Each sensor signal is decomposed into a 3-level tree as shown on Figure **2** using the DWPT and the "Daubechies 4" wavelet mother function. The energy content of each of the 8 nodes ((3,0), (3,1), etc) representing the wavelet component at the level 3 is calculated and used as the signal feature characteristic.

It observed that each of the eight sensors has over 99.9% of its energy content at the approximation node (3, 0) which signal $E_{3,0}$ is shown on the right hand plot of Figure **7** for the sensor X8. Therefore, the approximation will be used as a representative of the actual signal in the subsequent steps.

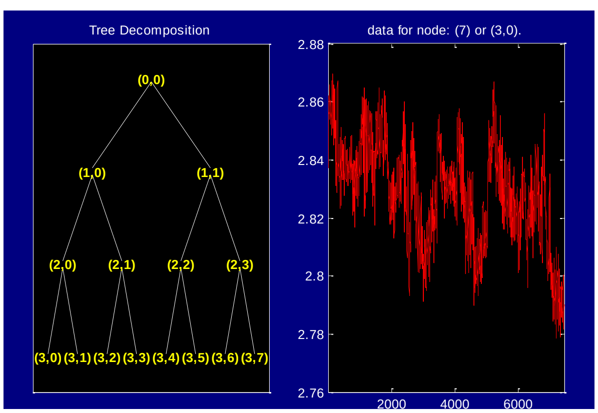
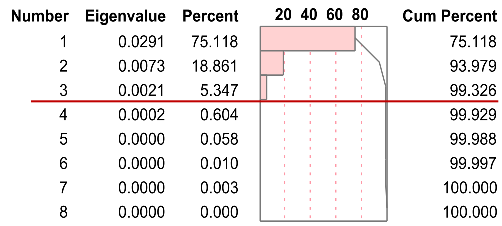


Figure 7: Example tree decomposition of sensor X8

Step 5*:* The standard PCA steps are executed to determine the principal components which are the eigenvectors corresponding to the most significant eigenvalues of the covariance matrix formed with the sensor data. As shown on Figure **8**, to maintain at least 95% of the original information in the model, the first 3 PCs representing 99.326% of the original information should be retained.



Number	Eigenvalue	Percent	20 40 60 80	Cum Percent
1	0.0291	75.118		75.118
2	0.0073	18.861		93.979
3	0.0021	5.347		99.326
4	0.0002	0.604		99.929
5	0.0000	0.058		99.988
6	0.0000	0.010		99.997
7	0.0000	0.003		100.000
8	0.0000	0.000		100.000

Figure 8: Pareto chart of eigenvalues contribution

Then the PPCA parameters are calculated with the maximum likelihood weight matrix first by setting to 0 any PC weight that is less than 0.1 as shown in Table **2**.

Table 2: Maximum likelihood weight matrix

	3 Principal Components for 99.3%		
Variables	*PC1 (75.1%)*	*PC2 (18.9%)*	*PC3 (5.3%)*
X1	*0*	*0*	*0*
X3	*0.59281*	*0.7203*	*-0.35998*
X4	*0.46975*	*0*	*0.88042*
X5	*0.65235*	*-0.69038*	*-0.30792*
X7	*0*	*0*	*0*
X8	*0*	*0*	*0*
X9	*0*	*0*	*0*
X10	*0*	*0*	*0*

Step 6*:* This step deals with the anomaly detection:

- Computation of reconstructed signal

Since only the 3 most important principal components are kept, the reconstructed 1-dimensional signal is obtained as*:*

$$RS(t) = \lambda_1 \Phi_1^*(t) + \lambda_2 \Phi_2^*(t) + \lambda_3 \Phi_3^*(t) \qquad (15)$$

With λ_1 =75.1%, λ_2 =18.9% and λ_3 =5.3%, which are the percentage of total information content in the 3 major eigenvalues.

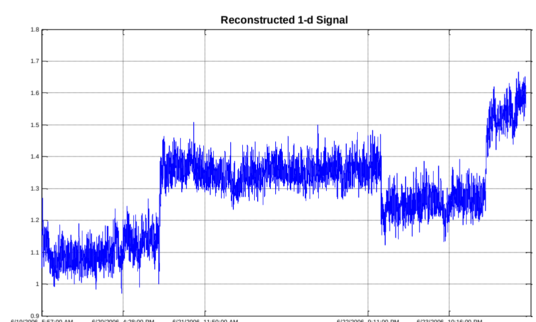


Figure 9: Reconstructed 1-d signal

The remaining analysis is based on the reconstructed 1-dimensional signal (RS) on Figure **9**, which is a representation of the original eight sensors. The RS in turn is decomposed using the DWPT up to the level-3 decomposition, since higher level of decomposition did not yield any additional information. The remaining sections of the anomaly detection decision are:

- Threshold calculation

To compute the damage indicators SAD and SSD, E_{ref} needs to be established first. E_{ref} is calculated as the mean value of the energy of the approximation node (3, 0) (representing more than 99.9% of the RS energy content) from the decomposed tree on Figure **7** of the RS signal over an hour period using a 5-second sample interval. Through this process a value E_{ref} =1.1328 is obtained. Next, the mean and standard deviation values of SAD and SSD are needed to calculate the anomaly threshold.

Therefore, $SAD(k)$ is calculated at each time step and its value is added to a set to compute the mean and standard deviation value of that set. Similar to the E_{ref} calculation, $E(k)$ represents the energy content of the node (3,0) at each subsequent time step k. The calculation is repeated over time until the mean value and standard deviation of the set SAD values converge towards μ_{SAD} and σ_{SAD} respectively as shown on Figure **10**.

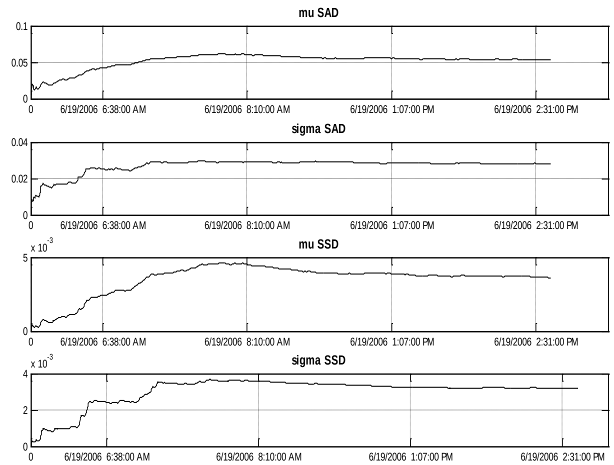


Figure 10: Convergence of mean and standards deviations of SAD & SSD

The convergences of the statistical parameters are based on the hours of operation which may be much shorter the calendar hours shown on the x-axis on Figure **10**. After the statistical parameters convergence, the modified threshold of SAD is calculated as $\varepsilon = 0.1337$ using a value of α of 2%. The idea is that, once the threshold is set, a system operator monitors the SAD (or the SSD) signal (green curve) instead of the original eight sensors in this case until system failure. Any time the value of SAD goes above the calculated threshold, it is considered an anomaly. Also, since the monitoring SAD or SSD yield the same conclusion, only the results obtained from SAD are presented in this paper. The Figure **11** shows the threshold (red dash line), the magnitude, length of anomalies and the point of the catastrophic failure. The blue portion of the curve represents the time history used to established the threshold parameters (E_{ref} , μ_{SAD} , σ_{SAD} and ε) .

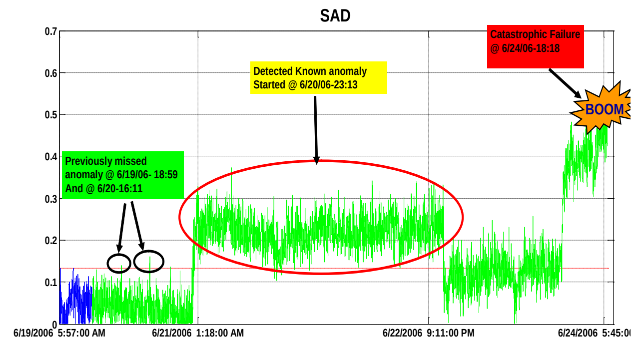


Figure 11: SAD parameters with detected anomalies

- Bayesian Hypothesis Testing

The final step is the Bayesian hypothesis testing. The result is the binary function h(t) with entry values of "1" and "0". As shown on the Figure **12**, there are four abnormal events during the almost 17 hours of monitoring. In a post-failure analysis, the gas turbine manufacturer established that the initial indication of a precursory event that led to the compressor failure was on June 20, 2006 at 23:30.

Indeed, the proposed technique successfully detected that event and found that it has started precisely at 23:13. The difference in the exact time of initiation may be explained by the fact that the manufacturer's analysis had a 30-minute sampling time. Furthermore, the proposed process has detected three other less severe and short lasting events as marked on Figure **12**. Although these events may be part of the accepted 2% type I error, they are marked as warning signs because their damage indicator values are higher than the threshold value (one of them has a damage indicator about 10% over the threshold).

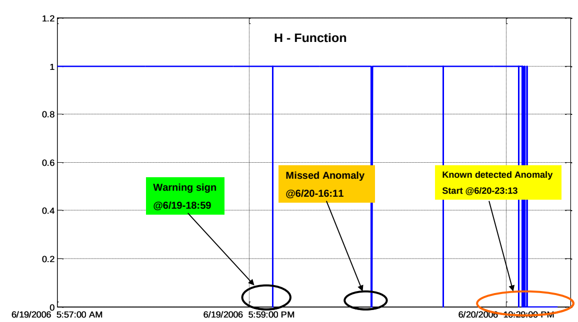


Figure 12: Result Bayesian hypothesis testing

Finally, the errors associated with the precursor detection are calculated. Recall that the probability of false-positive is an input that is decided by the analysis and it represents 2% in this example. The corresponding probability of type II error β or false-negative is calculated to be less than 10e-4. Thus, the proposed methodology meets the goal of any practical process that to be implemented has to have a much smaller type II error for given type I error. Because in the case of high cost catastrophic failure systems as a gas turbine, erroneously stopping a system operation while thinking there exists a anomaly is less expensive than missing a precursor that leads to failure.

4.2 Illustration of Advantage of combining performance and vibration sensors

To illustrate the robustness of combining both types of sensors, data from a unit that failed on September 10, 2009 at 12:09 with eleven working sensors is tested. First, three vibration sensors (X3, X4, X5) and eight gas path performance measurements sensors (X1, X2, X6,X7,X8, X9,X10, X11) are used separately. Then the eleven sensors are put together to carry out all the steps of the proposed methodology. To allow a fair comparison the same analysis inputs assumption are for each of the three cases (i.e. value of α =2%, maximum likelihood for weight matrix and with entry (W>=0.1)).

4.2.1 Vibration sensors only

After the preprocessing and de-noising of each of the three vibration sensors, the PPCA step is done. The eigenvalues contribution is shown on Figure **13** .

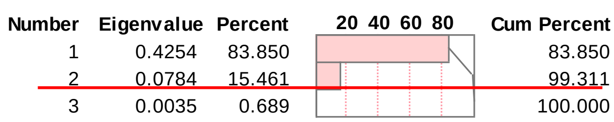


Figure 13: Vibration sensors Pareto plot of eigenvalues contribution

The PPCA is performed, which yields two retained principal components (99.31% of information content). The retained principal components are shown in**Table 3**.

Table 3: Retained PC for vibration sensors only

Sensors	PC1	PC2
x3	0.67079	-0.10819
x4	0.7126	-0.1801
x5	0.2055	0.97768

The anomaly detection step starts with the reconstructed signal is shown in Figure 14. Then follow the calculation of the damages indicators SAD and SSD.

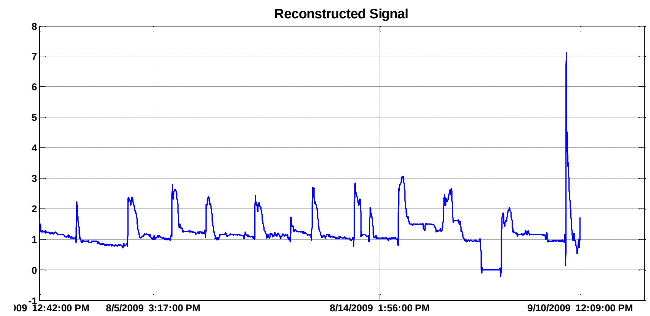


Figure 14: Reconstructed signal for vibration sensors only

Finally, the fault threshold is calculated and the hypothesis testing is performed and shown in Figure 15, where three failure precursor anomalies are detected.

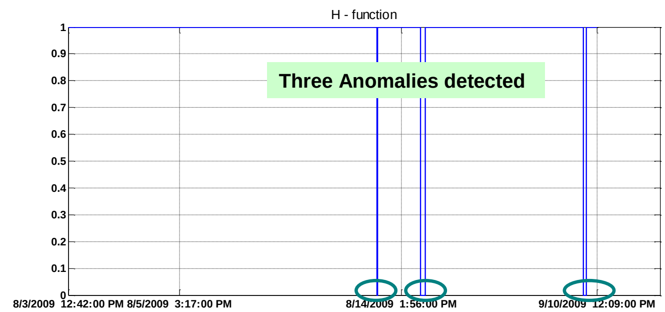


Figure 15: H-function for vibration sensors only

4.2.2 Gas path performance sensors only

In a similar fashion, all the steps are repeated with the eight gas path performance sensors. The eigenvalues contribution is shown on Figure **16**.

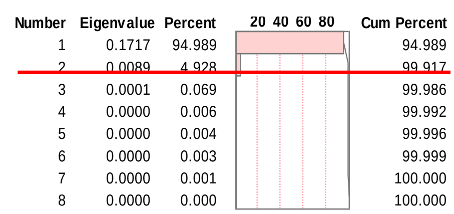


Figure 16: Gas path performance Pareto plot of eigenvalues contribution

After the conventional PCA, the PPCA is carried-out, and two principal components are retained (99.917% of information content). This is shown in Table 4.

Table 4: Retained PC for gas path performance sensors

Sensors	PC1	PC2
X1	0.4769	-0.13152
X2	0.40805	0
X6	0.47494	0
X7	0.4299	0.18129
X8	0.44293	0
X9	0	0
X10	0	0.96825
X11	0	0

The reconstructed 1-d signal for gas path performance sensors is shown on Figure **17**.

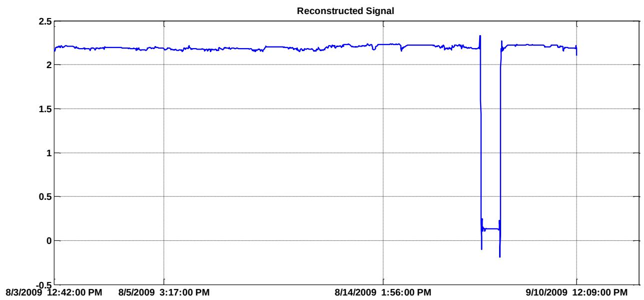


Figure 17: Reconstructed signal for performance sensors

The fault threshold is calculated and the hypothesis testing is applied to the damage indicator for gas path performance sensors and one failure precursor event is detected as shown on Figure 18.

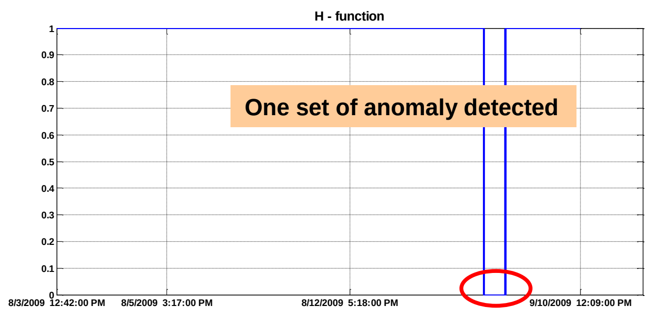


Figure 18: H-Function for performance sensors only

4.2.3 Combination of Vibration and Performance sensors

In this section both the gas path performance measurement and vibration sensors are combined and all the steps of the anomaly detection process are performed. The eigenvalues contribution of the eleven variables are shown on Figure **19**.

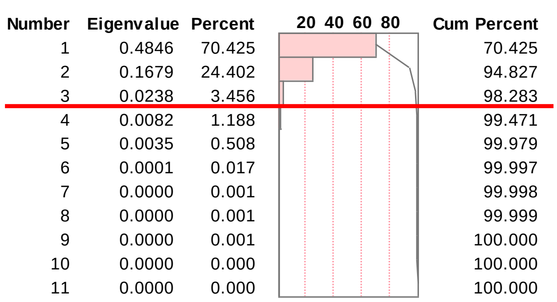


Figure 19: Pareto plot of Eigenvalues contribution for combined sensors

The result of the PPCA is three retained principal components representing 98.28% of information content as shown in Table **5**.

Table 5: Retained PC for ALL sensors combined

Sensors	PC1	PC2	PC3
X1	0.1897	0.34279	0.28134
X2	0.1664	0.28797	0.22198
X3	0.60303	-0.30941	0
X4	0.63522	0.36886	0
X5	0.26122	0.5137	0.80535
X6	0.19144	0.33751	0.28019
X7	0.17847	0.29981	0.21548
X8	0.1778	0.31576	0.26381
X9	0	0	0
X10	0	0	0
X11	0	0	0

The reconstructed 1-d signal is obtained using the three PC and shown for all the sensors combined in Figure 20.

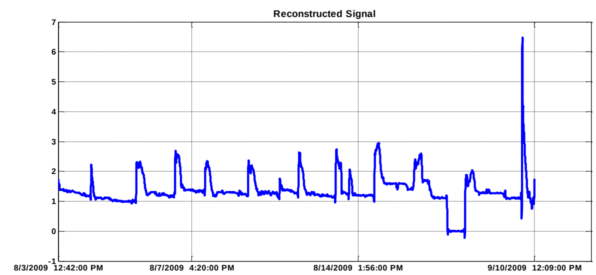


Figure 20: Reconstructed signal for ALL sensors combined

The hypothesis testing is applied to the damage indicators for the combined sensors case and shown on Figure 21. In this case, there are four failure precursory events detected, three of which have signatures within the vibration sensors (green circle) and one within the gas path performance sensor (red circle).

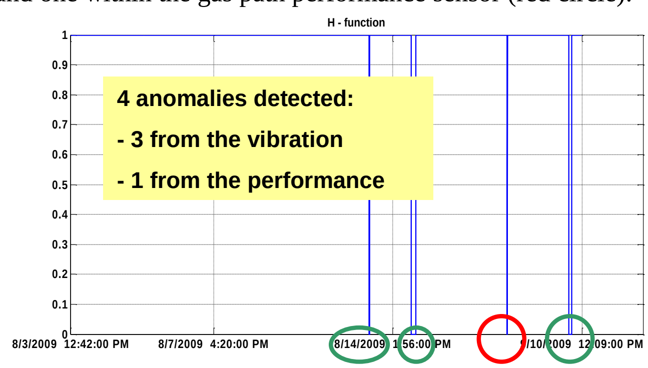


Figure 21: H-Function for ALL sensors combined

In this case the combination of both types of sensors allowed the detection of the failure precursor confirmed by the manufacturer’s past-failure analysis as initiated on September 8, 2009 at 18:30. However, any failure detection procedure which relies on anomaly within the vibration domain anomaly would have missed the precursor which had a more visible signature within the gas path performance domain.

5 CONCLUSION

In this paper, a systematic approach that combines both types of monitoring sensors in a heavy duty gas turbine is offered to detect precursory anomalies that could lead to catastrophic gas turbine compressor failure in an effort to reduce or even eliminate unplanned power plant outages. The proposed approach is promising as it has successfully detected previously known failure precursory anomalies as well as potential missed warning signs.

The proposed methodology of combining both the vibration and gas path performance sensor data appear to be robust as it can detect failure precursors which have signatures contained within the vibration sensors domain or within the gas path performance domain. Thus, it can be stated that the combination of the two types of sensors through the proposed approach can decrease the number of missed precursory anomaly.

Also, the proposed approach can easily be implemented to the failure detection of other sub-systems of the gas turbine or systems that are monitored with installed sensors. The use of this statistical approach allowed the handling of practical issues specific to heavy-duty gas turbines such as machine-to-machine variation and the wide variation in operation condition. Importantly, the proposed methodology has the ability to not only detect an anomaly, but also its severity and its length which can help trained technicians make the right decisions.

Overall the proposed approach is a novel as it is based on the fusion of information from both the gas path performance measurement sensors as well as the vibration sensors. Current practical industry standard relies on either vibration sensors or gas path performance sensors. Consequently whenever a sensor monitored system is based on either one of the sensor types, any failure signature that lies entirely within the other domain would be missed.

However, it is important to note that the successful implementation of this approach has been shown for the base load operation only. The authors are aware of the fact that this technique may most likely need some fine-tunings to be applicable to other load operating conditions (e.g. part-load or peak load). Also, the authors plan to extend the presented methodology to the case of transient engine operation as future work.

ACKNOWLEDGMENTS

The authors would like to thank GE Energy for providing gas turbine data.

REFERENCES

1. Thaler, H. *Global Gas Turbine Markets and Trends - What does the future hold?* in *The Future of Gas Turbine Technology - 3rd International Conference*. 2006. Brussels.
2. Stoll, H., *Creating owner's competitive advantage through contractual services.* GER-4208, GE Power Systems, 2001.
3. EPRI, *GE FA Compressor Dependability (phase 2)*. 2008, Electric Power Research Institute (EPRI).
4. EPRI, *Gas turbine compressor dependability: investigation of the MS7001F /9001F R0 compressor stage blade designs and failures*. 2007, Electric Power Research Institute (EPRI).
5. Graps, A., *An introduction to wavelets.* IEEE Computational Science & Engineering, 1995. **2**(2): p. 50-61.
6. Misrikhanov, A., *Wavelet transform methods: Application in electroenergetics.* Automation and Remote Control, 2006. **67**(5): p. 682-697.
7. Daubechies, I., *Ten lectures on wavelets*. 1992: Society for Industrial Mathematics.
8. Chui, C., *An introduction to wavelets*. 1992: Academic Pr.
9. Wu, Y. and R. Du, *Feature extraction and assessment using wavelet packets for monitoring of machining processes.* Mechanical systems and signal processing, 1996. **10**(1): p. 29-53.
10. Jiang, X. and H. Adeli, *Wavelet packet-autocorrelation function method for traffic flow pattern analysis.* Computer-aided civil and infrastructure engineering, 2004. **19**(5): p. 324-337.
11. Diallo, O. and D. Mavris. *A Concept for Intelligent Fault Detection Using a Multi-resolution Analysis*. in *Aerospace Systems Conference*. 2010. Los Angeles, CA.
12. Jiang, X., S. Mahadevan, and H. Adeli, *Bayesian wavelet packet denoising for structural system identification.* Progress in Structural Engineering and Materials, 2006. **14**(2): p. 333-356.
13. Tipping, M. and C. Bishop, *Probabilistic principal component analysis.* Journal of the Royal Statistical Society. Series B, Statistical Methodology, 1999: p. 611-622.
14. Diallo, O. and D. Mavris, *A Statistical Wavelet-Based Process for Systems Catastrophic Failure Precursor Detection.*
15. Sun, Z., *Structural damage assessment based on wavelet packet transform.* Journal of structural engineering, 2002. **128**: p. 1354.
16. Sun, Z. and C. Chang, *Statistical wavelet-based method for structural health monitoring.* Journal of structural engineering, 2004. **130**: p. 1055.
17. Montgomery, D., *Introduction to statistical quality control*. Vol. 3rd. 1996: Wiley- New York.
18. Ang, A. and W. Tang, *Probability concepts in engineering planning and design.* 1975.

