2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Anomaly Detection for Health Management of Aircraft Gas Turbine Engines†

Devendra Tolani‡, Murat Yasar, Shin Chin, Asok Ray
Mechanical Engineering Department
The Pennsylvania State University
University Park, PA 16802

‡**Corresponding Author**: Email tolani@psu.edu, Tel: (814) 865-8427

Keywords: Anomaly Detection, Gas Turbine Engine, Health Monitoring, Pattern Recognition, Symbolic Dynamics

Abstract†

This paper presents a comparison of different pattern recognition algorithms to identify slow time scale anomalies for health management of aircraft gas turbine engines. A new tool of anomaly detection, based on Symbolic Dynamics and Information Theory, is compared with traditional pattern recognition tools of Principal Component Analysis (PCA) and Artificial Neural Network (ANN). Time series data of the observed variables on the fast time scale are analyzed at slow time scale epochs for early detection of anomalies. The time series data are obtained from a generic engine simulation model. Health monitoring of gas turbine engines based on these techniques is discussed.

1 Introduction

Anomalous operation of gas turbine engines is undesirable from the perspectives of both engine operation and aircraft mission management. Early detection of anomalies and their characterization are essential for health management, which includes prognosis of impending failures in critical components and mitigation of their detrimental effects on the engine operation. For detection of these slow time scale deviations, it might be necessary to rely on the time series data generated from the available sensors as well as other relevant information regarding pilot's experience. Since sufficiently accurate and computationally tractable modeling of thermo-fluid and structural system dynamics is often infeasible solely based on the fundamental principles of physics, small changes in the system behavior may be inferred from both time series analysis of the sensor data and model-based information [A96] [KRCRK04]. By taking advantage of the available model-based information, the time series data can be converted, by phase-space partitioning, into a symbol sequence [BP97] [BS93] that, in turn, will generate a finite-state machine model of the dynamical system behavior.

Health monitoring of gas turbine engines can be viewed as a class of slow time scale problem [VDGD00]. For gas turbine engines, health condition of an engine changes in hundreds of hours, whereas the engine runs at a much faster time scale, usually in the order of seconds. Identification of the current state of the engine health is very important for maintenance engineers because necessary repairs must be carried out before the engine becomes permanently non-operable. Thus, it is essential to monitor slow-time-scale anomalies for gas turbine engines from the time series data of the engine response. A generic gas turbine engine simulation [FYR04] test bed has been used to validate the anomaly detection techniques.

The anomaly detection algorithm is built upon two-time-scale analysis of stationary behavior of dynamical systems using the principles of Symbolic Dynamics [K98] [DM95], Information Theory [AS91], Automata Theory [SSC02], and Pattern Recognition [DHS01]. Symbolic Dynamics captures the essential dynamical features of the physical process through phase-space partitioning. Information Theory allows modeling of incipient catastrophic failures and chaotic behavior that are analogous to thermodynamic phase transitions. Automata Theory generates finite-state machine models of the dynamical system behavior under nominal and anomalous conditions. Pattern Discovery methods infer anomalies through quantitative evaluation of the deviations in statistical patterns of the respective state machines from those under the nominal condition. These features are described briefly in the following section.

The paper is organized in six sections including the present one. Section 2 reviews the concepts of the symbolic dynamics and finite state machine based anomaly detection. Section 3 summarizes the radial basis function approach to solve the anomaly detection problem. Section 4 discusses principle component analysis technique. In section 5, the simulation results are

† This work has been supported in part by NASA Glenn Research Center under Grant No. NNC04GA49G and the United States Army Research Office under Grant No. DAAD19-01-1-0646.

discussed. The paper is summarized and concluded in Section 6.

2 Symbolic Dynamics

Due to the simplicity and ability to capture the state dynamics in a lower dimensional space, Symbolic dynamics is a widely used tool in the pattern recognition applications. In general, state vectors of a process may form a very high dimensional space which is difficult to handle both mathematically and physically. However, symbolic dynamics uses the phase trajectory partitioning of this high order space to represent the continuous dynamics with a predefined alphabet of symbols [C94].

These underlying concepts of symbolic dynamics and partitioning of phase space for anomaly detection of gas turbine engine are briefly introduced in a previous publication [KRCRK04]. Further details on symbolic dynamics are available in [SSC02] and on phase-space partitioning in [KB03].

Time series data of combustor outlet temperature, which is used for early detection of incipient faults is converted to a symbol sequence by partitioning the finite region in the phase-space of engine operation dynamics (over which the time series data evolves) into finitely many discrete blocks [A96] [BP97]. Each block is labeled as a symbol $\sigma \in \Sigma$, where the symbol set Σ is called the *alphabet* consisting of m different symbols. In this way, a data sequence, obtained from a trajectory of the dynamical system, is converted to a symbol sequence $\{\sigma_i, \sigma_j, \sigma_k, \cdots\}$ that characterizes the system dynamics represented by the data sequence. Critical steps in the symbol generation process are: (i) partitioning of a finite region in the phase space; and (ii) construction of a mapping from the partitioning into the symbol alphabet, which becomes a representation of the system dynamics defined by the trajectories.

Partitioning of the phase space of relevant system dynamics can be difficult especially if the time series data are noise-corrupted [A96] [KB03]. There are potentially a number of ways to do this. Two such methods to perform the partitioning are described below.

Kennel and Buhl [KB03] have formulated a phase-space partitioning method that is built upon the concept of symbolic false nearest neighbors (SFNN), where a statistic and algorithm is introduced to define empirical partitions for symbolic state reconstruction. This method avoids topological degeneracy where avoiding the degeneracy is an essential feature of a generating partition [BP97]. The major advantage of this method is that the partitioning using the algorithm is entirely based on the time series data. The partitions are defined with respect to a set of radial-basis influence functions, $f_k(x) = \frac{\alpha_k}{||x - z_k||^2}$, each associated with a symbol s_k with the center z_k and weight α_k . For each element x of the time series data set, one $f_m(x)$ is generally expected to be greater than other $f_k(x)$ with $k \neq m$. Then, the data point x in the phase space is transformed to a symbol s in the symbol space. The parameters z_k and α_k are the free optimization variables, with the constraint $\alpha_k \geq 0 \ \forall k$. There may be one or more influence functions assigned to each of the symbols in the alphabet. The partitions remain invariant at all epochs of the slow time scale.

An alternative scheme for obtaining partitions is based on wavelet transform of time series data, which yields a graph of coefficients versus scale at each time shift. After the wavelet transform [K94] is applied to the data, we partition the space of wavelet coefficients that is a function of scale and time. These graphs are stacked from end to end starting with the smallest value of scale and ending with the largest value. For example, the wavelet coefficients versus scale at time shift t_k are stacked after the ones at time shift t_{k-1} to obtain the so-called *scale series data* in the wavelet space, which is analogous to the time series data in the phase space. In this paper, the wavelet space is partitioned into horizontal slabs; future work will focus on optimized partitioning of the wavelet space. The number of blocks in a partition is equal to the size of the alphabet and each block of the partition is associated with a symbol in the alphabet. For a given stimulus, the partitioning of wavelet space must remain invariant at all epochs of the slow time scale.

After the partitioning is finalized, the sequence of symbols is generated from the time series (or scale series) data at different epochs of the slow time at which anomalies may take place. Then, a probabilistic finite state automaton is constructed from the symbol sequence at each time epoch. The anomaly measure at a given epoch is obtained as a distance between the state probability vector of the finite state machine at that epoch and the state probability vector of the finite state machine at the nominal condition. Thus, the above vector measure quantifies growth of anomaly from the nominal condition in the slow time scale.

Finite state machines, generated from the symbol sequences of a dynamical system identify its behavioral pattern. As the system trajectory evolves, different states are visited with different frequencies. The number of times a state is visited as well as the number of times a particular symbol is received, while sliding the window from a state leading to another state, is counted. In this way, the state probability vector is calculated from the time series data associated with different anomalous conditions. Having obtained the state probability vectors, the next step is to calculate the anomaly measure that signifies the change in stationary behavior of the dynamical system as the fault progresses. The state probability vector under the nominal case serves as a benchmark. The anomaly measure is obtained as the norm of the difference between the state probability vector associated with faulty behavior and the state probability vector for the benchmark condition. Obviously, the deviation measure at the benchmark condition is zero. The step by step process is depicted in Figure 1.

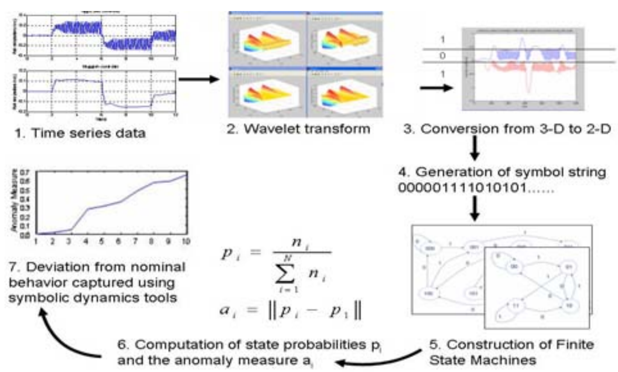

Fig. 1 Wavelet transforms of time series data

3 Radial Basis Function Neural Network

Neural networks learn complex input-output relationships, use sequential training procedures, and adapt themselves to the data. The most commonly used family of neural networks for pattern classification tasks is the feed-forward network. The learning process involves updating network architecture and connection weights so that a network can efficiently perform a specific classification/clustering task. Neural networks provide a new suite of nonlinear algorithms for feature extraction and classification. However, most of the well-known neural network models are implicitly equivalent or similar to classical statistical pattern recognition methods. Despite these similarities, neural networks do offer several advantages such as, unified approaches for feature extraction and classification, and flexible procedures for finding good, moderately nonlinear solution.

A major class of neural network model is the radial basis function (RBF) neural network (NN) [B95] in which the activation of a hidden unit is determined by the distance between the input vector and the prototype vector. The RBF NN is essentially a nearest neighbor type of classifier. The technique used in the anomaly detection is a variation of RBF NN to form the statistical model of nominal data. As new data enters into the system, it is compared with the RBF NN model. If it falls within the boundaries defined by the model than it is considered as a nominal data, if does not than the data is considered as anomalous. The approach is generic and has been applied to a variety of problems including advanced military aircraft subsystems [BJ01]. The key aspect of Neural Net approach is the appropriate selection of the radial basis function and the order of the statistics of the model. From this perspective, a radial basis function is introduced as:

$$f(x) = \exp\left(-\frac{\sum_k |x_k - \mu|^\alpha}{\theta_\alpha} \right)$$

where the parameter $\alpha \in (0, \infty)$; and μ and θ_α are the center and α^{th} central moment of the data set, respectively. For $\alpha = 2$, $f(\bullet)$ becomes Gaussian, which is the typical radial basis function used in the neural network literature.

From a sampled time series data under the nominal condition, the mean μ and the central moment θ_α are calculated as:

$$\mu = \frac{1}{N} \sum_{k=1}^{N} x_k \ ; \text{ and } \ \theta_\alpha = \sum_{k=1}^{N} |x_k - \mu|^\alpha .$$

The distance between any vector $\boldsymbol{x}$ and the center μ is obtained as: $\|x - \mu\|_{\ell_\alpha} = \left(\sum_k |x_k - \mu|^\alpha \right)^{1/\alpha}$. Hence, at the nominal condition, the radial basis function $f_{nom} = f(x)$. For different anomalous conditions, the parameters, μ and θ, are kept fixed; and f_k is evaluated from the data set under the (possibly anomalous) condition at the k^{th} epoch at the slow time scale. Then, the anomaly measure is defined as the distance function

$$\mathcal{M} = d(f_{nom}, f_k)$$

4 Principal Component Analysis

In the statistical approach, each pattern is represented in terms of d features or measurements and is viewed as a point in a d-dimensional space. The goal is to choose those features that allow pattern vectors belonging to different categories to occupy compact and disjoint regions in d-dimensional feature space. The effectiveness of the representation space (feature set) is determined by how well patterns from different classes can be separated. Given a set of training patterns from each class, the objective is to establish decision boundaries in the feature space which separate patterns belonging to different classes. In the statistical decision theoretic approach, the decision boundaries are determined by the specified or learned probability distributions of the patterns belong to each class.

Feature extraction methods determine an appropriate subspace of dimensionality m in the original feature space of dimensionality d ($m \leq d$). The best known linear feature extractor is the Principal Component Analysis (PCA) [DHS01] that makes use of Karhunen-Loéve expansion to compute the m largest eigenvectors of the $d \times d$ covariance matrix of the N d-dimensional patterns. Since PCA uses the most expressive features (eigenvectors with the largest eigenvalues), it effectively approximates the data by a linear subspace using the mean squared error criterion.

To detect growth in anomaly from time series data, the PCA can be performed for dimensionality reduction. For a time series data with length L, a data matrix of size M x N is to be created by dividing the length L time series into M=L/N length N subsections. Each row of the data matrix is a length N subsection obtained above. Then the N x N covariance matrix can be obtained from the data matrix. After determination the orthonormal eigenvectors $v_1 \ldots v_N$ and eigenvalues $\lambda_1 \ldots \lambda_N$ of the covariance matrix, where the eigenvalues are in increasing orders of magnitude, it is possible to choose the m largest eigenvalues and associated eigenvectors such that $\sum_{i=1}^{m} \lambda_i > \eta \sum_{i=1}^{N} \lambda_i$ where $\eta < 1$ is a real positive number close to 1 (e.g., $\eta = 0.95$).

The feature matrix F is defined as:

$$F = \left[\sqrt{\frac{\lambda_1}{\sum_{k=1}^{N} \lambda_k}}\, \upsilon_1 \quad \cdots \quad \sqrt{\frac{\lambda_m}{\sum_{k=1}^{N} \lambda_k}}\, \upsilon_m \right]$$

The feature matrix F_{nom} for nominal condition and feature matrices F_1, F_2 … corresponds to different health conditions are formed. The anomaly measure for health condition k defined as:

$$\mathcal{M} = d(F_{nom}, F_k)$$

where d is a metric signifying the difference between the nominal and the anomalous condition. For the results shown in the next section, metric was chosen to be the Euclidian norm.

5 Experimental Results

In the health monitoring of the engine, usual quantitative approach to identify anomalous condition is to measure the deviation of the efficiency values from nominal state (brand new engine). In the simulation setup, the high and low pressure turbine efficiencies, fan and compressor tip velocity ratios are reduced to observe the effect on the time series data. Here, the drop in the efficiencies of the different components of the engine is assumed to be same. This assumption is reasonable for testing various methods, although it may not hold true for a real engine because each component may deteriorate at different rates.

Health condition (i.e. the efficiencies) of an engine changes in a very slow time scale therefore, for a short time period (in the order of minutes) the efficiency values are constant for all practical purposes. To replicate these conditions, hundreds of hours of engine simulation are required. This is not practically feasible, so the efficiencies are reduced for each run of the simulation and a certain period of operation is observed for each health condition. This is very similar to experimental methods where engine is tested in extreme conditions to simulate years of operation in a few days. Also, transient data were not used for the analysis.

It was noticed that the output time series data was almost impossible to distinguish with a simple threshold check or visual inspection, or a peak pass test in frequency domain. However, when the saturation occurs , high fluctuations in temperature are observed and visual inspection is sufficient for detection of anomaly, but our goal here is to capture the anomalous behavior much before saturation sets in, and a permanent damage occurs to the engine. Time series data was gathered by running the simulation in different anomaly conditions, by altering the efficiency values and flow constants to simulate faults. The nominal values are 1 and the anomalous condition is simulated by decreasing the efficiencies up to 0.97.

To observe the effects of the decreasing efficiency on the output data persistent excitation is needed. This excitation could either be given to the system as an external reference input or from an inner summation point of the continuous gain scheduling controller. In these experiments, the inner control loops were modified to simulate anomaly effects under the excitation. For this particular time series the perturbed parameter is the booster vane angle. The perturbation is square wave with amplitude 10% of the nominal value.

The time series data used for analysis is the combustor outlet temperature and the main burner fuel flow. From thermodynamic perspective, due to decreasing efficiency, it is expected the anomaly will mainly affect these two variables. The results of the analyses corroborate this fact, however only the temperature output gave distinctive results. Thus, in this paper the results obtained from the temperature of the combustor outlet are presented.

Figure 2 was obtained by the Symbolic Dynamics based method using wavelet space partitioning. The alphabet size used in partitioning the wavelet coefficient space is chosen to be 8.

When times series data is compared it is seen that wavelet transform method shows reasonable and meaningful results since the saturation occurs at the point where this method predicts. Figure 2 is a normalized anomaly measure curve where the increasing anomaly can be detected much before saturation occurs.

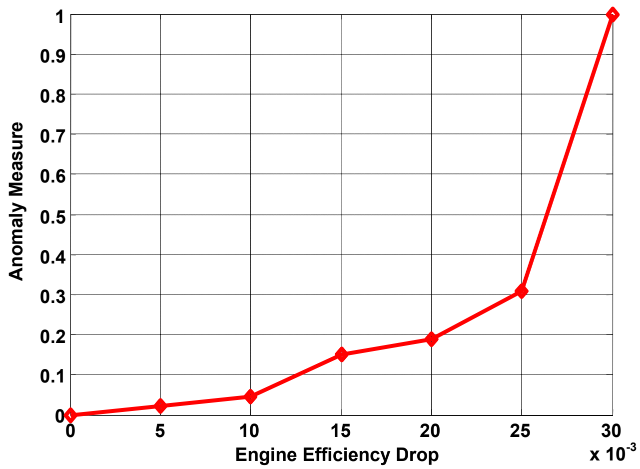

Fig. 2 Normalized anomaly measure up to saturation

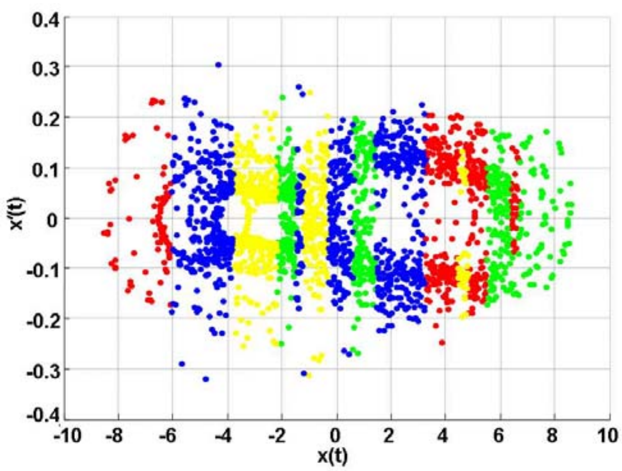

Fig. 3 Phase space of the nominal time series data

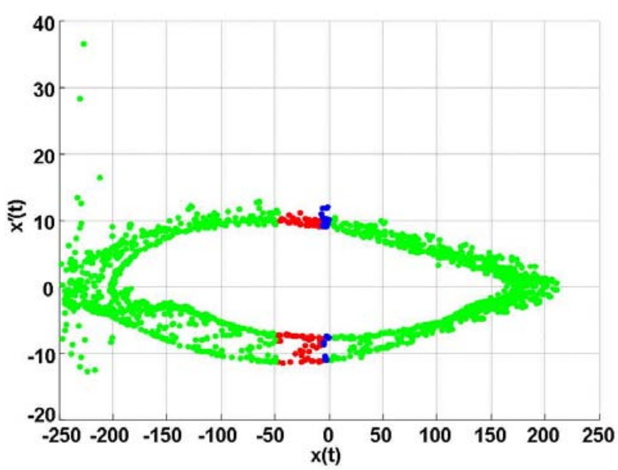

Fig.4 Phase space (for efficiency drop of 3%)

The phase space partitioning using symbolic false nearest neighbors (SFNN) is shown in Figure 3 and Figure 4. The x-axis of the phase plane corresponds to the unbiased time series data and y-axis is the difference between two successive x-axis data. As shown in Figure 5, this method provides a better anomaly measure. In this partitioning method the alphabet size to generate the finite state machine is 4.

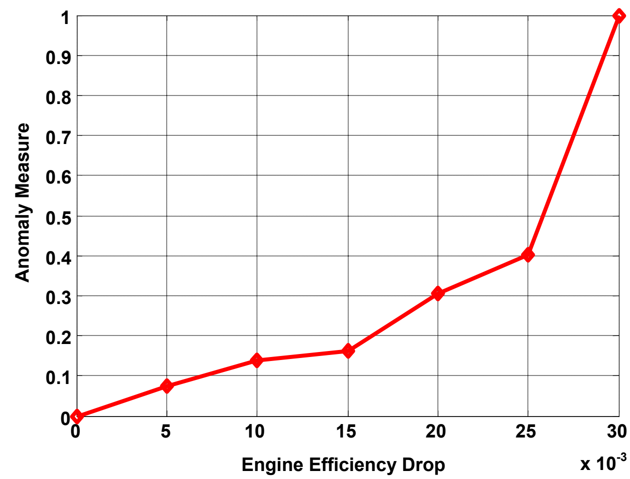

Fig.5 Anomaly measure using phase space partitioning

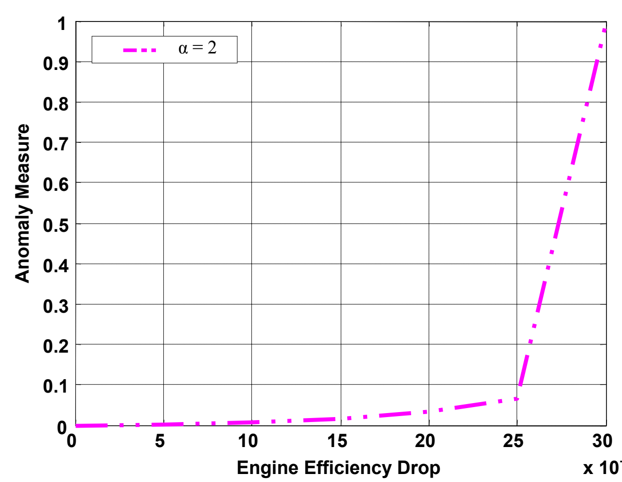

Fig.6 Anomaly measure using RBF NN

Comparison of the two partitioning techniques shows that the SFNN method performs marginally better than the wavelet based method in predicting the onset of anomaly.

Figure 6 shows the result using the radial basis function Neural Net method. Parameter "α", described in section 3, determines the order of statistics used. Results are normalized with respect to maximum anomaly condition. $\alpha = 2$ creates a widely used Gaussian RBF in pattern discovery.

Figure 7 shows the results of the anomaly measure using PCA technique. Time series data is divided into 10 subsections to form the data matrix used in the PCA. PCA technique also agrees with the other ones predicting the health condition of the gas turbine engine before saturation; however these results are comparatively inferior to the methods discussed previously.

Figure 8 compares the four different methods of anomaly detection. It clearly shows that the results for RBFNN and PCA based methods are comparatively inferior to the symbolic dynamics based methods in terms of early detection of anomalies.

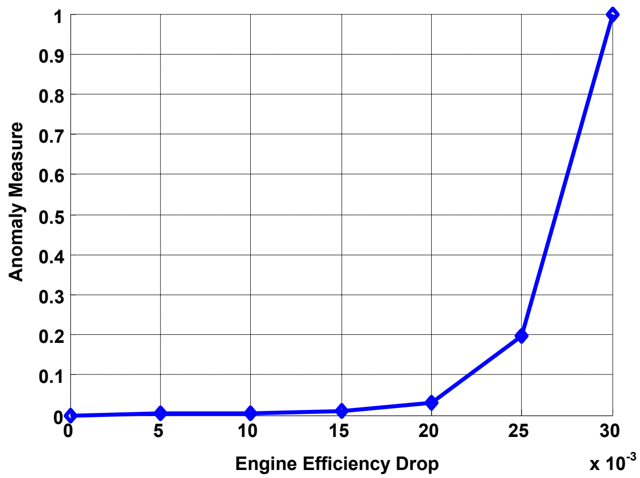

Fig.7 Anomaly measure using PCA

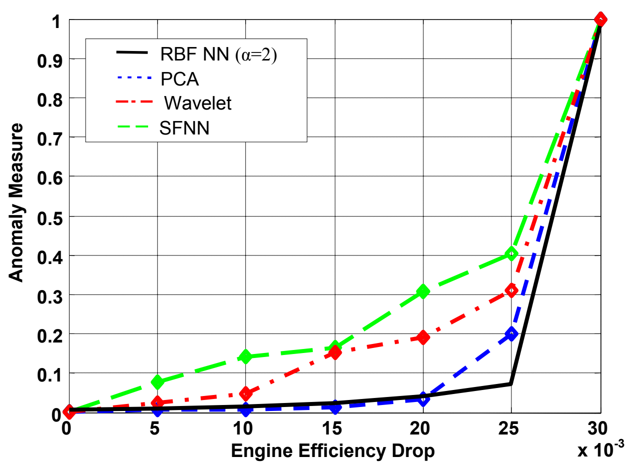

Fig.8 Comparison of the four methods

6 Summary and Conclusions

This paper compares the latest anomaly detection techniques for safety and performance enhancement of aircraft gas turbine engines. A test bed based on a generic gas turbine engine simulation model has been established to validate the various anomaly detection techniques. In this paper, four methods for anomaly detection are compared quantitatively. Although all techniques were able to predict the saturation point correctly, the best results are obtained using phase space partitioning using SFNN and wavelet methods. Computational complexity is an important issue for online implementation of these methods. Symbolic dynamics method based on both, wavelet space partitioning and SFNN, are compared with more traditional techniques using RBF NN and PCA. The wavelet space partitioning technique has been investigated in detail. This method is promising but more work needs to be done for optimal partitioning. Future research will also involve Health monitoring of aircraft propulsion systems based on this work

REFERENCES

[A96] H.D.I. Abarbanel, *The Analysis of Observed Chaotic Data,* 1996, Springer-Verlag, New York, 1996

[AS91] H.A. Atmanspracher,and H. Scheingraber (1991). Reconstructing language hierarchies, *Information Dynamics*, eds., New York, p.45, Plenum, 1991

[B95] C.M. Bishop, *Neural networks for pattern recognition*, Oxford University Press Inc., New York, 1995

[BJ01] T. Brotherton and T. Johnson, *Anomaly Detection for Advanced Military Aircraft Using Neural Networks,* IEEE vol. 6, pp. 3113-3123, 2001

[BP97] R. Badii, and A. Politi, *Complexity, Hierarchical Structures and Scaling in Physics,* Cambridge University Press, Cambridge, U.K. 1997.

[C94] J.P. Crutchfield, The Calculi of Emergence: Computation, Dynamics and Induction," *Physica D,* 75, pp. 11-54, 1994

[CY89] J.P. Crutchfeld and K. Young, *Inferring statistical complexity*, Physical Review Letters 63, pp. 105-108, 1989

[DHS01] R. Duda, P. Hart, and D. Stork, *Pattern classification*, John Wiley & Sons Inc., 2001

[DM95] D. Lind and M. Marcus, *A introduction to symbolic dynamics and coding*, Cambridge University Press, United Kingdom, 1995

[FYR04] J. Fu, M. Yasar, A. Ray, *Optimal Discrete Event Supervisory Control of Gas Turbine Engines,* American Control Conference, Boston, MA, 2004

[K94] G. Kaiser, *A friendly guide to Wavelets*. Birkhäuser 1994

[K98] B. P. Kitchens, *Symbolic dynamics: One sided, two sided and countable state Markov shifts*, Springer-Verlag, 1998

[KB03] M.B. Kennel and M. Buhl, *Estimating good discrete partitions from observed data: symbolic false nearest neighbors Physical Review Letters*, Vol. 91, 084102 , 2003

[KRCRK04] A. Khatkhate, A. Ray, S. Chin, V. Rajagopalan and E. Keller, *Early Detection of Fatigue Crack Anomaly: A Symbolic Dynamic Approach*, Preprints American Control Conference, Boston, MA, June-July 2004.

[SSC02] C. R. Shalizi, K.L. Shalizi, and J.P. Crutchfeld, *An algorithm for pattern discovery in time series*, SFI Working Paper 02-10-060, 2002

[VDGD00] A. Volponi, H. DePold, R. Ganguli and C. Daguang, *The Use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics: A Comparative Study*, ASME Turbo Expo 2000, 2000-GT-547, Munich, 2000