THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS
345 E. 47th St., New York, N.Y. 10017

98-GT-101

THE APPLICATION OF EXPERT SYSTEMS AND NEURAL NETWORKS TO GAS TURBINE PROGNOSTICS AND DIAGNOSTICS

Hans R. DePold
United Technologies
Pratt & Whitney
East Hartford, CT 06108
(860) 565-1749, FAX (860) 565-1215, depoldhr@pweh.com

F. Douglas Gass
United Technologies
Pratt & Whitney
West Palm Beach, FL 33410
(305) 796-366, FAX (561) 796-3667, gassfd@pwfl.com

ABSTRACT

Condition monitoring of engine gas generators plays an essential role in airline fleet management. Adaptive diagnostic systems are becoming available that interpret measured data, furnish diagnosis of problems, provide a prognosis of engine health for planning purposes, and rank engines for scheduled maintenance. More than four hundred operations worldwide currently use versions of the first or second generation diagnostic tools.

Development of a third generation system is underway which will provide additional system enhancements and combine the functions of the existing tools. Proposed enhancements include the use of artificial intelligence to automate, improve the quality of the analysis, provide timely alerts, and the use of an Internet link for collaboration. One objective of these enhancements is to have the intelligent system do more of the analysis and decision making, while continuing to support the depth of analysis currently available at experienced operations.

This paper presents recent developments in technology and strategies in engine condition monitoring including:

1) application of statistical analysis and artificial neural network filters to improve data quality;
2) neural networks for trend change detection, and classification to diagnose performance change; and
3) expert systems to diagnose, provide alerts and to rank maintenance action recommendations.

LIST OF SYMBOLS

EPR	engine pressure ratio	μ	exponential average
EGT	cor. turbine temperature	σ	standard deviation about μ
WF	cor. engine fuel flow	ANN	artificial neural network
N2	cor. high rotor speed	B	bias applied to a neuron
N1	cor. Low rotor speed	W	weighting of neuron input
T3	cor. HPC exit temp	TAT	Total Air Temperature
HPT	High Pressure Turb.	HPC	High Pressure Compres.
LPC	Low Pressure Compres.	LPT	Low Pressure Turb.
		TCC	Turbine Case Cooling

INTRODUCTION

Years of accumulation of knowledge of jet engine diagnostics has led to an understanding of the processes, rules of thumb, diagnostic fingerprints, and hierarchies for ranking, and fault isolation techniques. Effective and timely diagnostics and prognostics now necessitates that the users of diagnostic tools exercise considerable judgment and experience in applying this accumulated knowledge. A minimum of one week of intensive training is required to apply this knowledge, but effective utilization of current tools requires years of experience.

A critical mass of diagnostics knowledge has been achieved to permit the effective use of artificial intelligent systems such as neural networks and knowledge based systems, to emulate much of the required judgment and experience.

OVERVIEW OF DIAGNOSTIC PROCESSES

The diagnostic process (Figure 1) is modular and begins with propulsion system data, and applies data validity analysis to convert the data to more usable information.
Then it extracts from the information the knowledge of performance and mechanical trends. It then compares the extracted knowledge to several knowledge bases, and completes a diagnosis or prognosis of the propulsion system's health. Finally it alerts the operator to any important findings, and constructs a hierarchy of potential actions to correct any problems it uncovered.

Presented at the International Gas Turbine & Aeroengine Congress & Exhibition
Stockholm, Sweden — June 2–June 5, 1998
This paper has been accepted for publication in the Transactions of the ASME
Discussion of it will be accepted at ASME Headquarters until September 30, 1998

The Diagnostic Process

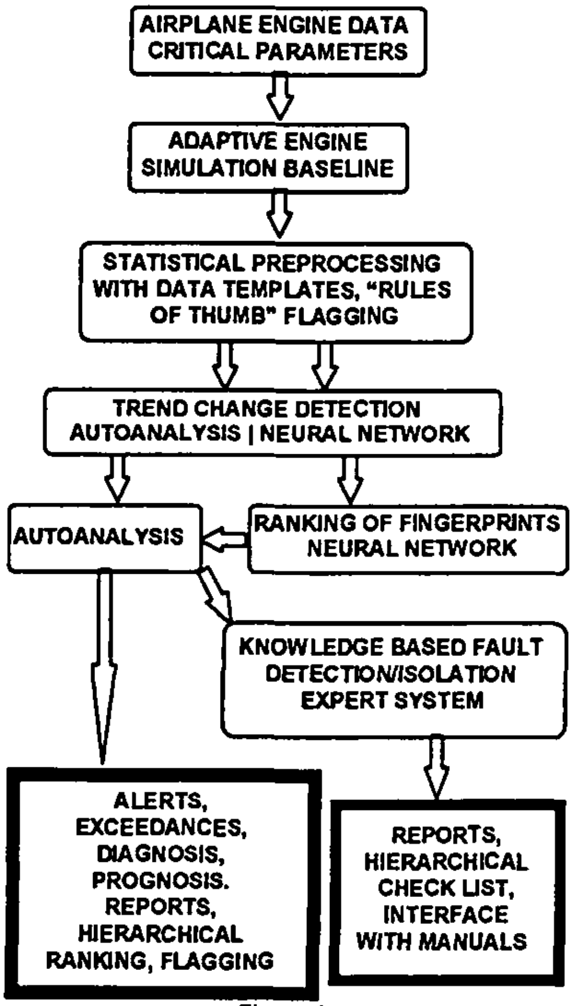

Figure 1

A robust diagnostic process uses competing strategies with alternative analytical processes that operate in parallel. It considers from the outset, more than one possible explanation for any trends it discovers. By using knowledge of the measurement uncertainties and the fingerprints for the possible explanations, it identifies and ranks the suspected root causes. It can carry the recommendations further by consulting additional knowledge bases and providing check procedures and maintenance procedures.

ADAPTIVE DIAGNOSTICS

The computational engine of a diagnostic system is an adaptive performance model of the gas turbine being analyzed. The model's primary function is to normalize data so that every data point is evaluated at effectively the same flight condition, power setting, system power off-takes and bleed off-takes. This is accomplished by running the model to each input data condition and computing the difference between the model and the data.

Adaptive diagnostics requires that the model be able to adapt itself to reasonably match revenue service data. Models now have the capability of thermodynamically scaling their component maps to close with early engine revenue service data. The closer the thermodynamic match of the baseline model, the better the normalization of the data.

DATA VALIDITY

Statistical analysis is essential for evaluating the quality of the data. Rolling averages typically waste the initial data points and are slow in responding to trend changes. Several data validation improvements have been developed.

First the rolling average method is replaced by an exponential memory retention method. An exponential average equivalent of a ten point rolling average requires the storage of only the exponential average, not the ten preceding points. With each new data point 15% of the remembered average is replaced by new data. Therefore the old data is forgotten 15% at a time, resulting in an exponential decline in its usage. In a ten point rolling average the old data is 100% recalled until it is 11 points old. Then it is 100% forgotten. The exponential equivalent of the 10 point moving average is given by:

$$\text{EXP\_Average}(t)_{10} = \text{Exp\_Average}(t-1)_{10} * 0.85 + \text{New\_Data} * 0.15$$

With each new data point a fraction of the memory of the older data is replaced. This means that since only the last average is retained, the exponential averages can respond instantaneously to step changes in trends.

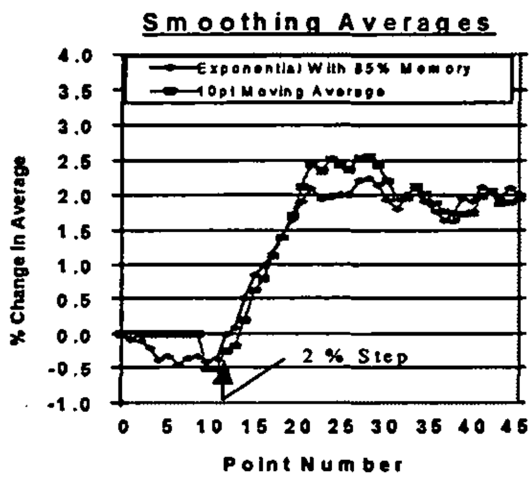

Figure 2

Figure 2 shows the response of the two methods to a 2% step increase with +-1.5% random scatter. The exponential has an equivalent or better response and definition. As soon as a trend is detected the exponential average can be incremented to show the step change.

Similarly, the exponential averaging method allows statistical bands to be carried with little overhead. That allows the statistical analysis to adapt quickly to changes in data quality. Therefore the diagnostics automatically adapt to the data quality of each airline reducing the likelihood of false alarms. Unlike a moving average, an exponential average can be changed at the moment of the trend change detection, so the statistical bands also show the discontinuity.

If care were not taken, pure statistical analysis disregards trend changes presuming them to be bad data. Therefore earlier diagnostic systems kept the bad data and gave the analyst a waming flag and a few rules to evaluate the data validity. A neural network (Figure 3) is a system of computations plus logical tests that can be used to recognize pattems. The input to the network is the instantaneous average value and standard deviation of each critical parameter. The neurons apply the knowledge and experience that the analyst would normally consider. The network can act as templates to eliminate bad data much the same as analysts would with their rules of thumb.

Filtering Data With Six Templates

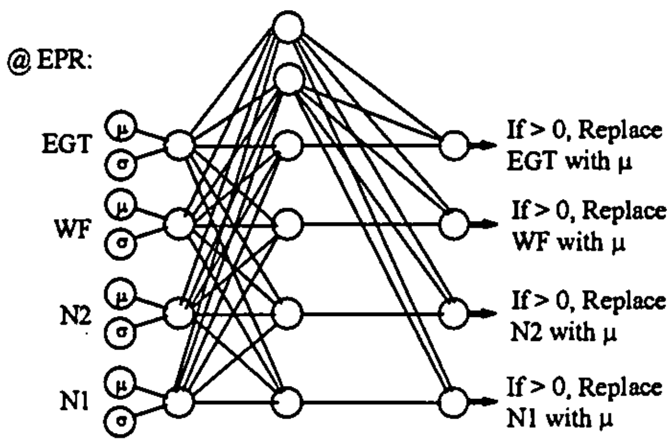

Figure 3

These templates (Figure 3) identify spurious TAT, EPR, input errors, and larga instrumentation errors. The first two ANN templates identify errors that result in all four parameters being either high out of statistical limits or low out of statistical limits. The remaining four ANN templates detect single parameter limit exceedences that are not physically possible.

When five key jet engine parameters (EPR, EGT, WF, N2, N1) are considered, six basic templates can be defined which correct approximately 9% of typical data (example Figura 4). In this typical example, 9% of the data caused 26% to 63% of the raw data measurement uncertainty. A separate report is generated so that the raw and corrected data are both available.

The test case selected was a recent event where a JT3D combustor crack occurred. The weightings for each of the four engine parameters entering the first two template ANN nodes were set equal to one another and then optimized. The optimum node threshold limit was found to be 1.4 sigma for four parameters. That is the equivalent of more than 2.5 sigma for a single parameter statistical test.

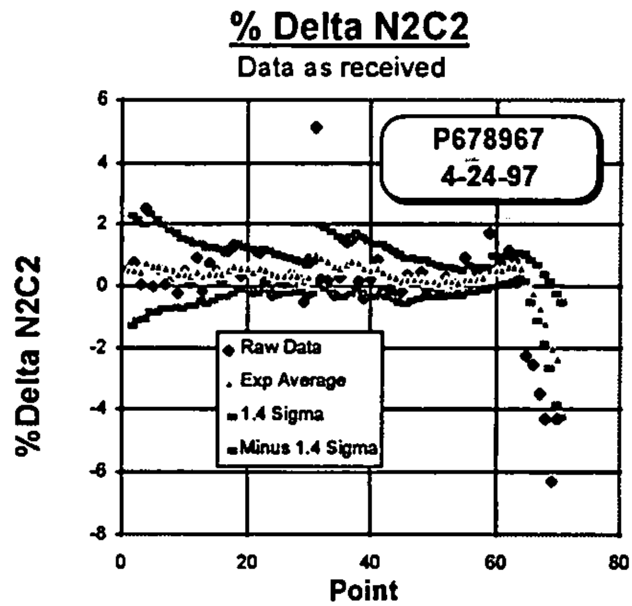

Figure 4

The weightings for the other four template nodes were individually optimized because JT3D experience showed N2 to be the most reliable parameter and WF to be the least reliable. The optimum relative weightings for this case were found to be, N2 (1.30), EGT (0.96), N1 (0.93), and WF (0.81).

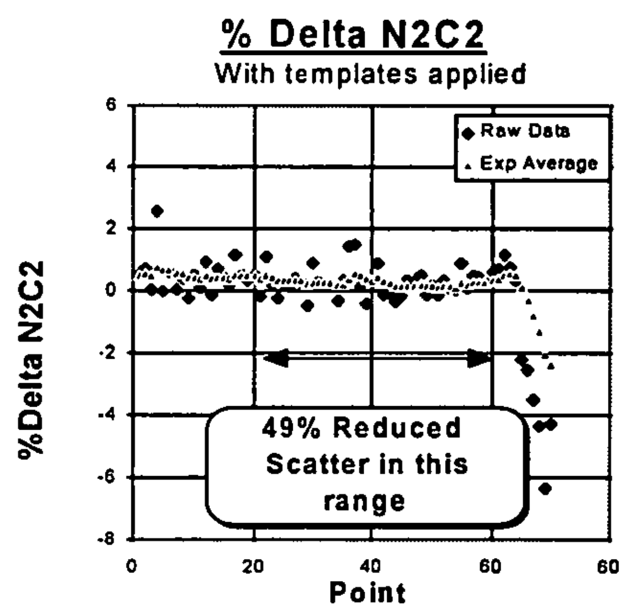

Figure 5

Several training cases are typically used to determine the ANN weightings. However, the direct knowledge of the individual engine's current data scatter is an available and useful alternative for a diagnostic system that adapts to the data being analyzed. As processes are improved and data quality improves, the improvement is measured and the sensitivity of the diagnostic system adapts.

Over the range of the 44 points before the event, the templates reduced N2 scatter 49% (Figure 5), EGT scatter by 34%, N1 scatter by 26%, and WF scatter by 63%. The pre processing of the data with statistical analysis applied through ANN templates improved the data validity significantly thereby improving the accuracy of all the downstream processes.

TREND CHANGE DETECTION- NEURAL NETWORKS

Figure 5 shows that the current system required six data points collected over a period of six days before the sample trend change was detected and the engine was removed. N2 was the most reliable parameter for detecting this trend change. That is characteristic of the data acquired from JT3D and JT8D engines. EGT is the most reliable parameter for high bypass engines such as PW2000, PW4000, and JT9D engines. Having four or more parameters for detection improves the reliability of the prediction. Therefore the detection system can be used on all the key parameters to improve the confidence level. It is quite likely that as experience is gained with intelligent diagnostic systems, that calculated parameters such as efficiencies and flow capacities will be used for detecting trend changes. Synthesized non-dimensional parameters may also come into use.

In this example (Figure 6) we use an ANN that evaluates three different exponential averages for corrected N2, the equivalents of 3 pt, 5 pt, and 10 pt moving averages.

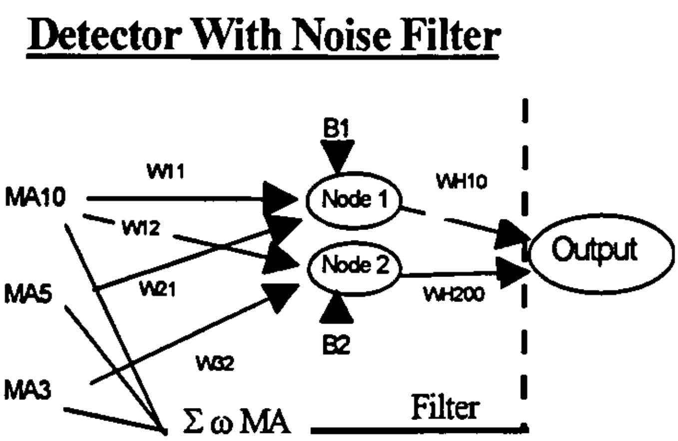

Figure 6

The ANN is trained with trend changes in the positive direction and then used to process the trend twice to account for both positive and negative trend changes. A separate calculation of the magnitude is computed based on the relationships between the moving averages and the response of the detection system. The calculation is a weighted sum of the three averages trained separate from the neural network. The magnitude of the change is a criteria to filter out spurious detection. The filter is applied to inhibit trend change detection when the change detected is less than one third the data's noise level. It is very important to inhibit potential false alerts. It is preferable to miss a detection than to give a premature alert. For that reason three detection systems are used, the first to detect significant changes within two to three data points, the second to detect subtler significant changes within five or six data points, and the third to detect 1% or greater moves in ten data points.

Figure 7 shows the trend detection system applied to the filtered N2 data. A 66% improvement in detection time response was achieved using the ANN/filter detection system in this example. The initial estimate of the step change was -3% and settled at about a 4% decrease in corrected N2.

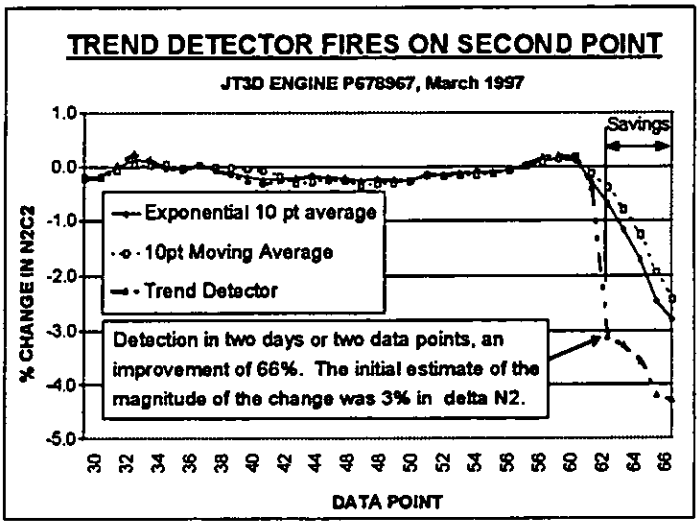

Figure 7

TREND CHANGE DETECTION- AutoAnalysis

AutoAnalysis is an analytical tool developed by Pratt & Whitney to assess changes in the performance of engine components. The AutoAnalysis process is an independent, parallel analytical process for detecting trend changes. For gradual changes it does a modular analysis of performance retention. It uses a Kalman filter which makes it adaptive to the quality of the data being analyzed [2].

The Kalman filter attempts to model and explain the data. When an abrupt change in performance occurs, the change is first perceived as a large sensor error. A persistent sensor error term is therefore used for trend change detection.

TREND CHANGE ROOT CAUSE DETERMINATION- NEURAL NETWORKS

The root cause of performance trend shifts needs to be determined to make recommendations for corrective action. Measurement uncertainty and the similarities of some problem symptoms does not always yield high confidence in the most probable root cause. Typically three alternative assessments are needed to encompass the true root cause. Neural networks provide easy access to other possible root causes. The computations of the output neurons of a neural network can be intercepted. That reveals the most probable alternatives and allows them to be ranked in the order of the strength of their signals.

Each neural network node represents a line that partitions the data [1]. The slope of the partitioning line is given by the ratio of the two input parameter weightings (Figure 8). That ratio can be determined from gas generator influence coefficients. Therefore all the node weights for hidden layers in this particular network can be calculated and require no training or optimization. Each data phase space was partitioned into high ratio (>), low ratio (<), and intermediate ratio regions. Nodes based on ambivalent gas generator parameter ratios that could easily change signs were removed.

A NEURAL NETWORK NODE, CREATES A PARTITION OF THE PHASE SPACE DEFINED BY THE INPUT

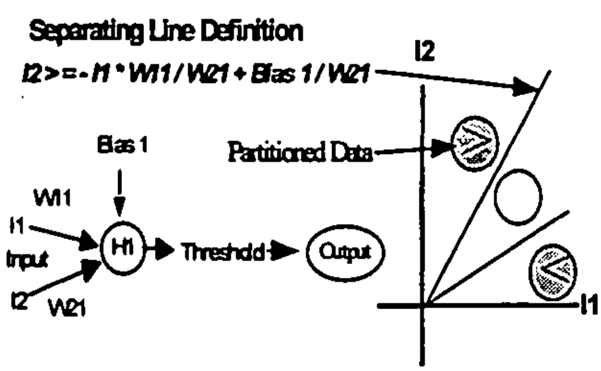

Figure 8

There are two design strategies for neural networks. The most common strategy is to use a symmetric matrix of neurons and train them with specific cases using an error back propagation technique. This method can provide the best fit of the data with the least number of neurons. However this method provides solutions that are for the most part incomprehensible, and non adaptive. This method can also provide the best fit of the data with the greatest number of neurons. However, again the network tends to be incomprehensible and the data over-fitted.

Another strategy is to partition the problem into key elements and independently optimize those neural network sub functions. Adaptability then becomes a part of the solution. For instance, in the detection system previously discussed (Figure 6), the neuron bias terms were made a linear function of the standard deviation of the data, and of the magnitude of the deviation of the data from the model baseline. The network was then first optimized for one signal to noise ratio, with no deviation of the data from the baseline, and with constant neuron bias terms. The neuron thresholds were then independently optimized as a variable accounting for different signal to noise ratios, different moving averages, and different deviations between the data and the baseline. It was found that as uncertainty increased, the threshold required for firing the neurons also increased. This method that applies first principles and physical insight, is easier to understand and to optimize [3].

A useful feature applied in these networks was the implementation of gains as transfer functions, so that the use of classical sigmoid functions were not necessary. The sigmoid functions are continuous functions that increase the uncertainty of the input data. The gains on the other hand are analogous to the biological neuron function and allow the output of the nodes to be variables. Therefore if the node's input barely meets the threshold, the node's output is calculated to be less than 1.0. If the input greatly exceeds the threshold then the node output carries a percentaga (typically 20% used) of the amount by which the input exceeded the threshold. The gain is kept small so that the network is stable.

The input parameters were normalized for the data uncertainty by dividing each input parameter by its respective standard deviation. The standard deviations were exponential averages that continuously adapted to changing data quality.

The network weights between the hidden layer and the output root causes, are independent and are trained one root cause at a time.

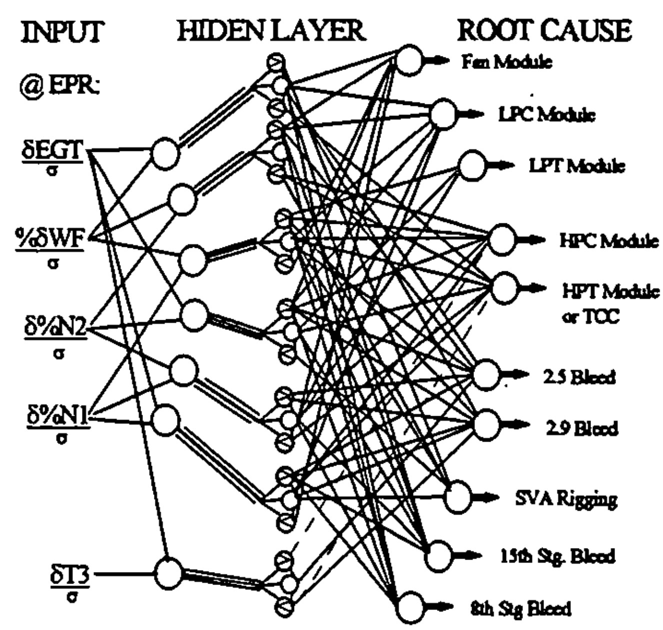

Figure 9

The symptoms or fingerprints of the root causes of the performance trend shifts have the efficiencies, flow capacities, and effective areas of the modules coupled the way the problems usually occur. That improves the effectiveness of separation of root causes such as the HPC versus HPT deterioration. During development, false positive indications are minimized by treating no-fault found (NFF) results as faults.

The focus of this paper is on diagnostics for a minimum case of four measured key engine parameters (EGT, WF, N1 and N2) at constant EPR. However, typically as many as nine parameters are monitored. T3 is shown to indicate how additional measured data is used in areas where they can improve the separation of root causes of performance changes.

In Figure 9, T3 is shown to be used to separate a 2.5 bleed from a 2.9 bleed problem, or an HPC from an HPT problem.

In addition to engine modular changes, it is necessary to identify systematic sensor errors or changes to instrumentation. They too can trigger trend detection logic. These changes are not quickly detected by the rules of thumb data filter templates because they are repeated and are gradually accepted as significant. Systematic sensor errors can be of three types, aircraft instrumentation errors (total air temperature, Mach number, and altitude), engine instrumentation errors (all measured parameters), and changes that can result from installation of new sensors or equipment.

Systematic airplane related instrumentation changes and errors affect all engines equally. When the root cause of a trend shift is found to be the same for all the engines on the aircraft, then the aircraft instrumentation is identified as the actual root cause.

Systematic engine related instrumentation changes and errors usually affect one engine, and often only one parameter. When an engine instrumentation change or error is responsible for a trend change, the affected instrumentation shows the change and is therefore identified.

The neural network can determine the alternative root causes, but the magnitudes of the root causes are best determined using AutoAnalysis. The aerothermal diagnosis is ranked in order of probability.

Trend change detection 5-18-87, JT9D-7Q, P702143

	Probability
Boroscope HPT	*74%*
Check TCC system	*62%*
Boroscope Burner	*30%*

TREND CHANGE ROOT CAUSE DETERMINATION-AutoAnalysis

AutoAnalysis includes a Kalman sensor error analysis followed by a statistical determination of the root causes. It uses the instrumentation uncertainties and an adaptive engine model to generate a matrix of influence coefficients. AutoAnalysis independently assesses the most probable root cause as well as its magnitude [4].

In addition, AutoAnalysis accepts the alternative root causes determined from the neural network, and computes the magnitudes of the alternative root causes. In the following example, 2.3% high turbine distress is the most probable root cause to explain the performance loss. A 100% TCC failure or a 2% burner pressure loss are possible alternative root causes.

Trend change detection 5-18-87, JT9D-7Q, P702143

	Probability	*Magnitude*
Boroscope HPT (%EFF/FC)	*74%*	*2.0/1.5%*
Check TCC system (%TCA)	*62%*	*100%*
Boroscope Burner (%DP)	*30%*	*2%*

The AutoAnalysis assessments provide detailed modular breakouts suitable for engineering analysis. In addition, a knowledge based expert system is used to provide the flight line recommended checks and maintenance procedures.

KNOWLEDGE BASED EXPERT SYSTEM FOR FAULT DETECTION AND ISOLATION

The knowledge based expert system diagnostics provides the maintenance mechanic with "What's wrong and how to fix it," value analysis based task optimization.

The expert system integrates the aerothermal input from AutoAnalysis with onboard control status and maintenance words (codes) for additional fault detection and isolation with maintenance recommendations. It is an object oriented system [6] designed to emulate the human thought process. The control identifies faults in scheduled system switches, valves, and indicators. The root cause assessment of AUTOANALYSIS is compared and linked with the control fault analysis providing a specific check list. The reasoning is temporal in that it considers not only the facts, but the order in which the facts occur [7]. A Bayesian type statistical evidence approach is used to reflect the uncertainties in the rule based analysis.

Value analysis considers the probability that the diagnosis is correct, and the cost required to check the hardware and verify the diagnosis. An alert is issued with the recommended corrective action whenever a significant trend change occurs.

Trend change detection 5-18-87, JT9D-7Q, P702143
(For illustration only)

	Probability	*Magnitude*	*Cost*	*Rank*
Boroscope HPT (%EFF/FC)	*74%*	*2.0/1.5%*	*$260*	*2*
Check TCC system (%TCA)	*62%*	*100%*	*$60*	*1*
Boroscope Burner (%DP)	*30%*	*2%*	*$220*	*3*

The maintenance recommendations provide 'Answers-not-Data' type of results. The information will be available in a hand held computer for powerplant analysis or line maintenance groups. Substantial maintenance cost saving and reduced down time are predicted for the quicker, and more accurate diagnostics [5].

Symptoms of open clearances or Hot Section Distress

Maintenance word indicates TCC valves powered.
Action: Check function of valves

If valves are functional, then
Action: Check for system cracks and loose fittings

If TCC system OK, then
Action: Boroscope HPT looking for distress

If HPT system OK, then
Action: Boroscope burner looking for distress

The expert system prognostics module integrates the aerothermal input from AUTOANALYSIS with module configuration information and the fault detection maintenance information from the control.

KNOWLEDGE BASED EXPERT SYSTEM FOR PROGNOSIS

The knowledge based expert system prognostics provides the quality assurance, reliability engineering, maintenance support, flight operations, and fleet management teams with projections and advice based on the expected long term behavior of the engine with the diagnosed condition.

Having isolated the cause of the fault(s), the expert system presents a prognosis. A critical fault could result in the prognosis that requires the engine immediately be removed from the wing or that the engine be grounded until a critical part is replaced.

If the isolated fault is non-critical such as a missing bleed seal or a TCC valve failure, the prognosis may (depending on other conditions) be, that the fault need only be corrected at the next scheduled maintenance, or that the aircraft can safely fly to the overhaul facility.

A fault occurring as the normal maintenance cycle approaches may result in the prognosis that the aircraft can continue to operate between city pairs in cooler climates where the engine operates at cooler internal temperatures.

Another prognosis could advise if overhaul can be postponed or if overhaul should be pulled forward.

With a prognosis for all engines in the fleet, scheduling of fleet maintenance as well as the ordering of spare parts can be planned. This information can be used for asset management to assure that spare parts are available on site on time.

EXPERT SYSTEM FOR INTEGRATION OF FLIGHT AND TEST CELL DATA ANALYSIS

Flight data can be used with or in lieu of tested-as-received-data for workscope value analysis. The probabilities are set to 100% when inspections and checks verify the causes of the trend changes. In the following case the certainty of the TCC and 2.5 bleed seal involvement make them highest ranked for maintenance and repair.

Workscope 5-19-87, JT9D-7Q, P702143
(For illustration only)

	Probability	*Value*	*Cost*	*Rank*
Replace TCC valve	*100%*	*1.5%/15C*	*$1523*	*1*
Replace 2.5 Bseal	*100%*	*0.3%/4C*	*$2022*	*2*
Waterwash engine	*20%*	*0.5%/5C*	*$535*	*3*

CONCLUSIONS:

A modular intelligent and adaptive system is presented for gas turbine diagnostic and prognostic analysis.

The system uses adaptive intelligent diagnostics to improve data quality, detect trend changes, interpret data, furnish diagnosis of problems, provide a prognosis of future engine behavior for planning purposes, and provide corrective actions on the basis of value. Artificial intelligence is used to automate, improve the quality of the analysis, provide timely alerts, and to integrate flight data and test cell data analysis for workscope and value analysis.

A robust dual approach is taken combining the attributes of neural networks, Kalman filters, statistical analysis, Bayesian/Evidence based decision making, and rule based analysis. An expert system is used to integrate the analysis and to perform value analysis in making recommendations.

Significant improvements in accuracy, quality of analysis, timeliness and usefulness of reporting are shown.

Further development is underway to provide additional system enhancements. Diagnostics will be expanded to include vibration and oil systems to detect blade damage, schedule trim balance/lubrication, detect shifts in SVS schedule or rigging, detect bearing thrust balance changes, improve fatigue analysis, and possibly enable safe life analysis.

Other enhancements include the use of artificial intelligence and value analysis to do more of the decision making for asset management, and to predict part lives, shop visits, work scopes, and inventory requirements. Improvements are also planned for sensors, real time telemetry of data, data links, and secure Internet collaboration.

Maintenance costs can be reduced by:

- Early detection for prevention of collateral damage.
- Reduced in flight shut downs
- Reduced down time
- Improved planning and asset management

REFERENCES:

[1] G. Vachtsevanos, "Neuro-Fuzzy Tools For Modeling, Optimization, and Control of Complex Systems," Georgia Institute of Technology, Intelligent Control Laboratory, 1995.

[2] A. Volponi, "Sensor Error Compensation In Engine Performance Diagnostics," ASME 94-GT-58, 1994

[3] L. Jaw, "Neural Network Modeling of Engine Tip Clearance," 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Session 24-ASME-7 Smart Engine Technology, 1997

[4] L. Urban & A. Volponi, "Mathematical Methods Of Relative Engine Performance Diagnostics," SAE Technical Paper 922048, 1992

[5] B. Miller, "AMOSS May Stem Rising Maintenance Costs," Overhaul & Maintenance, p 51-56, September-October, Volume 3 number 5, McGraw-Hill Companies, 1997.

[6] J. Rumbaugh, "Object-Oriented Modeling and Design," Prentice-Hall Inc, 1991.

[7] R. Fikes, T. Keller, "The Role of Frame-Based Representation in Reasoning," Communications of the ACM, Volume 28 number 9, September, 1985.