Benchmarking Gas Path Diagnostic Methods:
A Public Approach

Donald L. Simon
U.S. Army Research Laboratory, Glenn Research Center, Cleveland, Ohio

Jeff Bird and Craig Davison
Institute for Aerospace Research, National Research Council of Canada, Ottawa, Ontario, Canada

Al Volponi
Pratt & Whitney, East Hartford, Connecticut

R. Eugene Iverson
The Boeing Company, Seattle, Washington

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

* TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

* TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

* CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

* CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

* SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

* TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, organizing and publishing research results.

For more information about the NASA STI program, see the following:

* Access the NASA STI program home page at *http://www.sti.nasa.gov*

* E-mail your question via the Internet to *help@sti.nasa.gov*

* Fax your question to the NASA STI Help Desk at 301–621–0134

* Telephone the NASA STI Help Desk at 301–621–0390

* Write to:
  NASA Center for AeroSpace Information (CASI)
  7115 Standard Drive
  Hanover, MD 21076–1320

NASA

Benchmarking Gas Path Diagnostic Methods: A Public Approach

Donald L. Simon
U.S. Army Research Laboratory, Glenn Research Center, Cleveland, Ohio

Jeff Bird and Craig Davison
Institute for Aerospace Research, National Research Council of Canada, Ottawa, Ontario, Canada

Al Volponi
Pratt & Whitney, East Hartford, Connecticut

R. Eugene Iverson
The Boeing Company, Seattle, Washington

Acknowledgments

The authors graciously acknowledge the support of the following individuals in conducting this effort: Link Jaw, Richard Eustace, Tak Kobayashi, Jonathan Litt, Jon DeCastro, Eric Gillman, and Christopher Heath.

Level of Review: This material has been technically reviewed by technical management.

Benchmarking Gas Path Diagnostic Methods: A Public Approach

Donald L. Simon[*]
U.S. Army Research Laboratory
Glenn Research Center
Cleveland, Ohio 44135

Jeff Bird and Craig Davison
Institute for Aerospace Research
National Research Council of Canada
Ottawa, Ontario, Canada K1A 0R6

Al Volponi
Pratt & Whitney
East Hartford, Connecticut 06108

R. Eugene Iverson
The Boeing Company
Seattle, Washtington 98124

Abstract

Recent technology reviews have identified the need for objective assessments of engine health management (EHM) technology. The need is two-fold: technology developers require relevant data and problems to design and validate new algorithms and techniques while engine system integrators and operators need practical tools to direct development and then evaluate the effectiveness of proposed solutions. This paper presents a publicly available gas path diagnostic benchmark problem that has been developed by the Propulsion and Power Systems Panel of The Technical Cooperation Program (TTCP) to help address these needs. The problem is coded in MATLAB (The MathWorks, Inc.) and coupled with a non-linear turbofan engine simulation to produce "snap-shot" measurements, with relevant noise levels, as if collected from a fleet of engines over their lifetime of use. Each engine within the fleet will experience unique operating and deterioration profiles, and may encounter randomly occurring relevant gas path faults including sensor, actuator and component faults. The challenge to the EHM community is to develop gas path diagnostic algorithms to reliably perform fault detection and isolation. An example solution to the benchmark problem is provided along with associated evaluation metrics. A plan is presented to disseminate this benchmark problem to the engine health management technical community and invite technology solutions.

Introduction

A recent technology review has revealed that while Engine Health Management (EHM) related research and development has increased significantly in recent years, there exists a fundamental inconsistency in defining and representing EHM problems (ref. 1). Currently many of the EHM solutions published in the open literature are applied to different platforms, with different levels of complexity, addressing different problems, and using different metrics for evaluating performance. As such it is difficult to perform a one-to-one comparison of candidate approaches. Furthermore, these inconsistencies create barriers to effective development of new algorithms and the exchange of EHM-related ideas and results.

Past efforts have made progress towards addressing these issues. Several authors have presented results from their self-conducted comparative assessments of gas path diagnostics methods (refs. 2 to 4). On a broader scale, the On Board Identification, Diagnosis and Control of Gas Turbine Engines (OBIDICOTE) Project conducted by the European research community defined a common set of gas turbine engine fault cases which were used by several researchers for diagnostic method development and evaluation (refs. 5 to 7). To help further address these issues, and to facilitate international cooperation, an Engine Health Management Industry Review (EHMIR) effort has been initiated under the auspices of The Technical Cooperation Program (TTCP), Aerospace Systems Group, Propulsion and Power Systems Panel. TTCP is a forum for defense science and technology collaboration between Australia, Canada, New Zealand, the United Kingdom and the United States (ref. 8). The EHMIR will provide reference, or theme problems, to aid in technology development and evaluation. Gas path and vibration sub-teams are currently developing theme problems that will provide relevant challenging problems. The objective of this effort is to construct and disseminate EHM theme problems, and invite solutions from the EHM community. Following a period of individual development, it is proposed that a conference be held to present the

---

[*]Currently employed by the NASA Glenn Research Center.

results with blind test case evaluations. The overall goal is to provide a set of environments that will truly facilitate the development and evaluation of significant EHM capabilities.

This paper specifically covers the progress on defining the gas path diagnostic theme problem and the associated metrics for benchmarking the performance of diagnostic solutions. It is organized as follows. First, the proposed public approach towards benchmarking gas path diagnostic methods is introduced. This consists of: 1) an Engine Fleet Simulator (EFS) to generate simulated test cases with implanted faults and degradation; 2) a description of the interface between EFS outputs and user developed diagnostic solutions; and 3) a description of proposed metrics for evaluating the performance of candidate diagnostic solutions. Next, an example diagnostic solution is given, and its performance against the defined metrics is presented. Finally, the follow-on steps for disseminating the problem to, and inviting technology solutions from, the EHM community are discussed. Technology developers, original equipment manufacturers (OEM) and operators are invited to provide their comments and feedback on any aspect of the proposed approach.

Nomenclature

C-MAPSS	Commercial Modular Aero-Propulsion System Simulation
EFS	Engine Fleet Simulator
EGT	Exhaust Gas Temperature
EHM	Engine Health Management
EHMIR	Engine Health Management Industry Review
FPR	False Positive Rate
GUI	Graphical User Interface
HPC	High Pressure Compressor
HPT	High Pressure Turbine
LPC	Low Pressure Compressor
LPT	Low Pressure Turbine
OBIDICOTE	On Board Identification, Diagnosis and Control of Gas Turbine Engines
OEM	original equipment manufacturer
PCC	Percent Correctly Classified
PLA	power lever angle
TPR	true positive rate
TTCP	The Technical Co-operation Program
C	confusion matrix
H	fault influence coefficient matrix
Nc	core speed
Nf	fan speed
NfR	Corrected fan speed
P_0	State covariance matrix
$P15$	total pressure in bypass-duct
$P2$	total pressure at fan inlet

$P24$	total pressure at LPC outlet
$Pamb$	ambient pressure
$PCNfR$	percent corrected fan speed
$Ps30$	static pressure at HPC outlet
R	measurement noise covariance
$T2$	total temperature at fan inlet
$T24$	total temperature at LPC outlet
$T30$	total temperature at HPC outlet
$T48$	total interstage turbine temperature
$T48R$	corrected interstage turbine temperature
VBV	Variable Bleed Valve
VSV	Variable Stator Vane
Wf	fuel flow
ema	exponential moving average
e_j	normalized measurement estimation error of hypothesized fault case j
m	number of sensor measurements
n	number of single fault types
psia	pounds per square inch absolute
rpm	revolutions per minute
$y_i(k)$	Measurement i collected on flight k
$y_{i\_baseline}(k)$	Fleet average engine value of measurement i at flight k operating conditions
α	moving average smoothing constant
β	delay in measurement delta-delta calculation
γ	component flow capacity adder
$\Delta Tamb$	difference between ambient temperature and standard atmospheric temperature
$\Delta x(k)$	fault magnitude vector
$\Delta \hat{x}(k)$	estimated fault magnitude vector
$\Delta y_i(k)$	measurement delta - difference between actual and fleet average engine in measurement i at flight k
$\Delta y_{i\_ema}(k)$	exponential moving average of measurement delta i on flight k
$\Delta\Delta y_{i\_ema}(k)$	measurement delta-delta-difference in ema of measurement i at flight k relative to some previous flight
$\Delta\Delta \hat{y}_{i\_ema}(k)$	estimated measurement delta-delta
$\Delta\Delta Y(k)$	vector of measurement delta-deltas
η	component efficiency adder
κ	kappa coefficient
σ_i	standard deviation of measurement i

Subscripts

i	measurement index
j	fault index
k	flight index

| p | confusion matrix row index |
| q | confusion matrix column index |

Superscript

| T | transpose operator |

Gas Path Diagnostic Benchmark Process

The goal of aircraft engine gas path diagnostics is to reliably assess and manage the health of gas turbine engine flow-path components. It is performed by relating observed changes in measured engine gas path variables (typically the suite of control feedback sensors) to changes in engine module performance, engine system malfunctions, or instrumentation problems. An inherent requirement is knowledge of gas-turbine parameter interrelationships. Gas path diagnostic functionality can reside both on-board and off-board the aircraft. Typical measurements include pressures, temperatures, rotor speeds, and fuel flow. The number of sensors available varies depending on engine type and model, but applications of four to eight sensors are common.

Ground-based gas-path health management systems rely on data acquired on-board the aircraft within the engine control computer. Although an engine controller may sample sensor data at 20 to 50 Hz, typically only a fraction of the data is transferred to ground-based health management systems for analysis. While new aircraft data acquisition systems are improving the quantity and quality of archived data relative to legacy aircraft, the infrastructure required to effectively archive and manage all sensed engine data generated from the flight of every engine is often not practical. A typical application will instead record "snap-shot" engine gas-path measurements collected at two to three operating points during each flight. For a commercial application, which follows a relatively consistent flight profile, this may include recordings at ground-idle, max EGT, and cruise conditions. Military applications, which exhibit much greater variability in mission profiles, may be limited to recordings at ground idle and takeoff. When possible, the data acquisition system may wait until the engine achieves quasi-steady-state operation, and then archive the average sensor measurements recorded over a window of time.

The proposed process for benchmarking aircraft engine gas path diagnostic methodologies is presented in figure 1. It specifically focuses on diagnostic methods applied to "snap shot," or discrete, engine measurements. The intent is to provide users a publicly available toolset to enable the development, evaluation, and side-by-side comparison of candidate diagnostic solutions. This includes providing the functionality to evaluate analytical, empirical or hybrid (analytical + empirical) diagnostic approaches. The elements of the benchmark process include an EFS that generates sensed parameter histories as if collected from a fleet of engines over their lifetime of use; user developed diagnostic solution(s) that interpret EFS generated parameter histories to diagnose any faults; and a routine to automatically evaluate the performance of candidate diagnostic solutions against a predefined set of metrics. Each element is discussed in more detail in the sections that follow.

Engine Fleet Simulator

Gas path diagnostic algorithm development and validation requires access to engine models and data. Ideally, this would include a rich database of information collected from engines over a broad range of operating conditions, deterioration levels, and known fault and no-fault conditions. However, to facilitate a public benchmarking approach, it was decided to generate simulated engine data utilizing a publicly available turbofan engine simulation. This approach avoids the use of engine data and analytical models that contain proprietary information. While an engine simulation will never fully capture all the nuances contained in actual engine data, it does provide some advantages. For example, it will allow the simulation of a broader range of fault types and magnitudes occurring over a broader range of engine operating conditions. It will also provide unambiguous knowledge of an engine's true fault/no-fault state, or "ground truth" condition.

The benchmark problem has been constructed as an EFS to generate histories of "snap shot" engine parameters collected from each engine on each flight. The EFS architecture, shown in figure 2, is implemented in the MATLAB environment. This architecture consists of a graphical user interface (GUI) that accepts user specified inputs regarding the number of
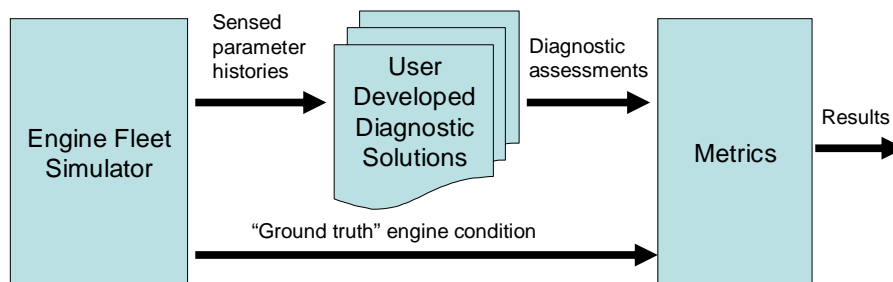

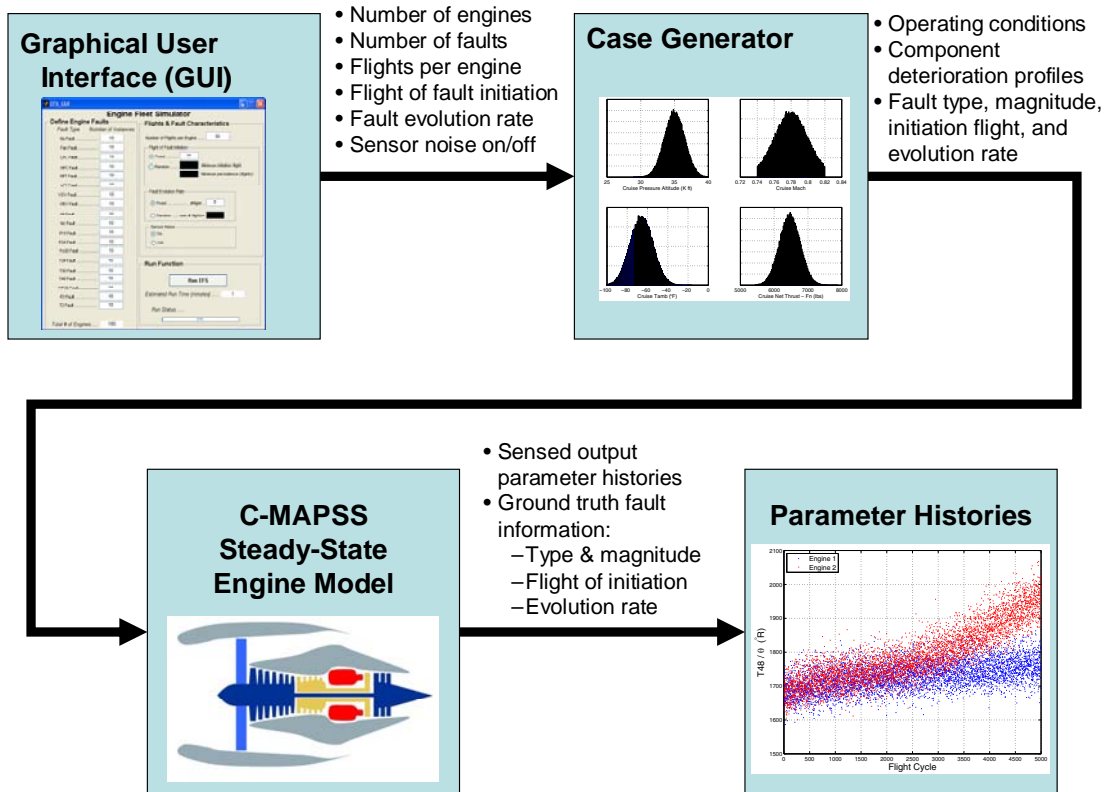Figure 1.—Gas path diagnostic benchmark process.

Figure 2.—EFS architecture.

engines and the number of occurrences of each fault type, a case generator designed to produce unique fault effects and operating profiles for each engine in the fleet, and a non-linear steady-state turbofan engine simulation which produces sensed parameter histories for each engine in the fleet. Each component of the EFS is further described below.

EFS Graphical User Interface

The EFS GUI, shown in figure 3, is designed to provide flexibility in generating data sets for diagnostic development and validation purposes. Through this interface, the type and number of faults that occur within the fleet of engines are defined. It should be emphasized that the EFS has been designed assuming that an individual engine can only experience a single fault–it will not simulate multiple faults occurring within the same engine. There are 18 possible fault scenarios plus the no-fault scenario. The sum of the number of occurrences of each scenario determines the total number of engines in the fleet. The interface also allows the user to specify the following: the number of flights over which output data will be collected for each engine; the flight of fault initiation (either at a fixed flight number or randomly within a specified window of flights); the rate at which faults evolve, either abruptly (instantaneously) or rapidly (over a number of flight cycles); and sensor noise turned on or off.

EFS Case Generator

After the EFS inputs have been specified through the GUI, the user selects the "Run EFS" button which initiates the process of generating engine parameter histories according to the number and type of scenarios specified by the user. The Case Generator randomly assigns a unique operating history and component deterioration profile to each engine in the fleet. This includes assigning the city pairs that an engine will fly between, and the calendar date on which engine data collection will commence. Altitude, Mach number, ambient temperature and power setting parameters at the takeoff and cruise operating points where data will be collected during each flight are all randomly generated from specified distributions. At takeoff, the power reference parameter is established by the Power Lever Angle (PLA) setting which will either be 100 percent, or a fixed de-rated takeoff power setting of 90 or 80 percent. At cruise, the power reference parameter is specified by net thrust. Histograms showing the distributions in operating parameters produced by the EFS Case Generator over 100,000 flights are shown in figure 4 (takeoff) and figure 5 (cruise).

Deterioration effects are simulated via adjustments to 10 health parameters within the engine simulation so that engines will continuously degrade over time. These health parameters include an efficiency scalar and a flow adder for each of the
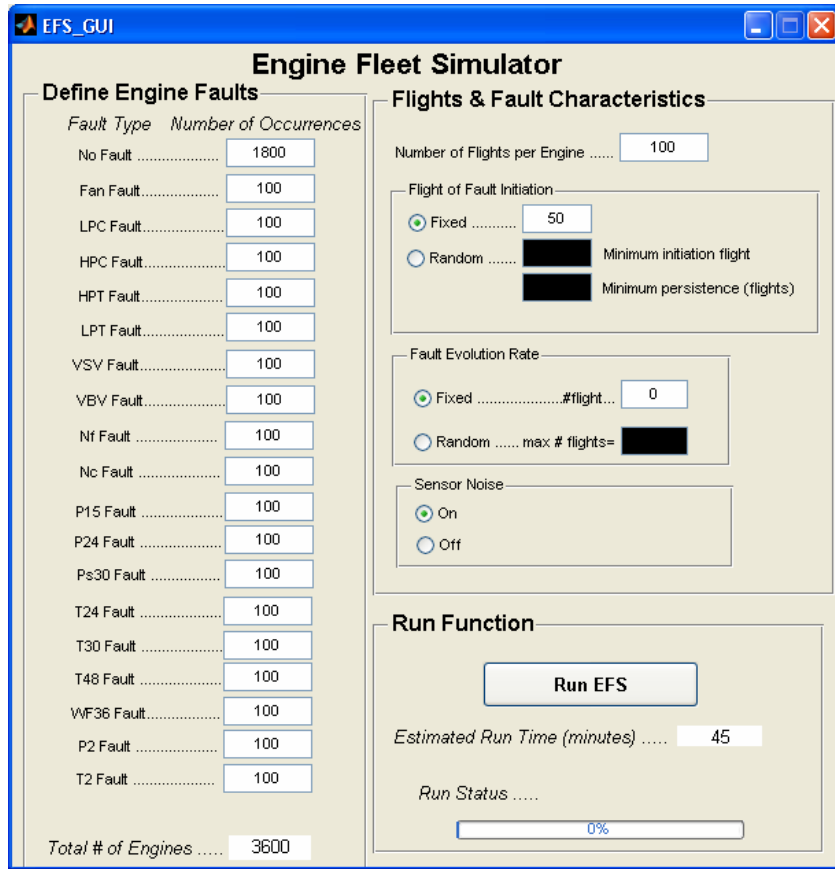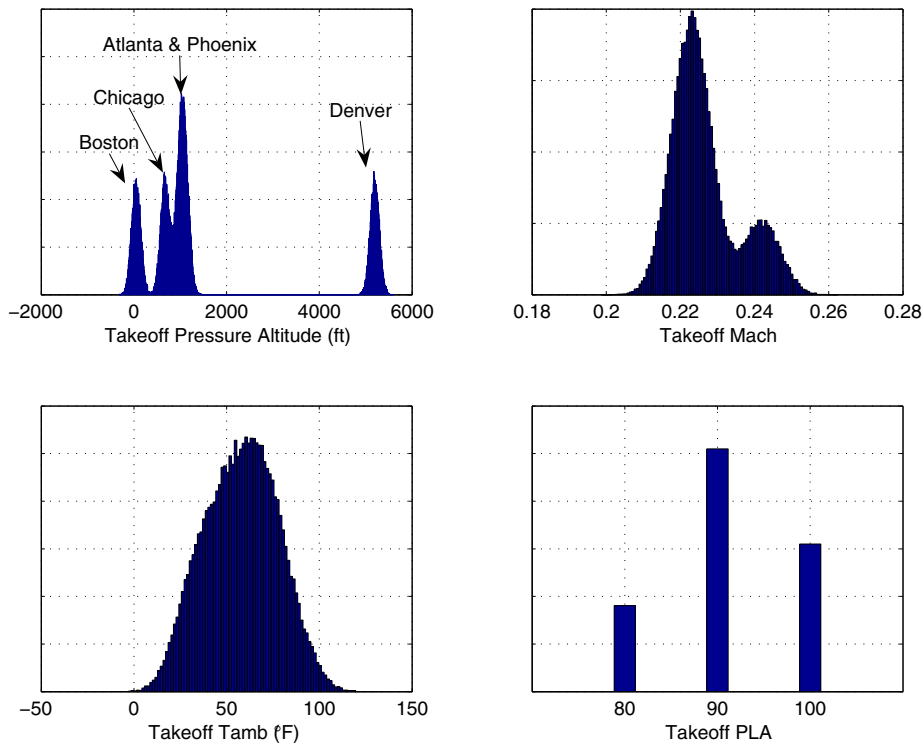
Figure 3.—EFS graphical user interface.

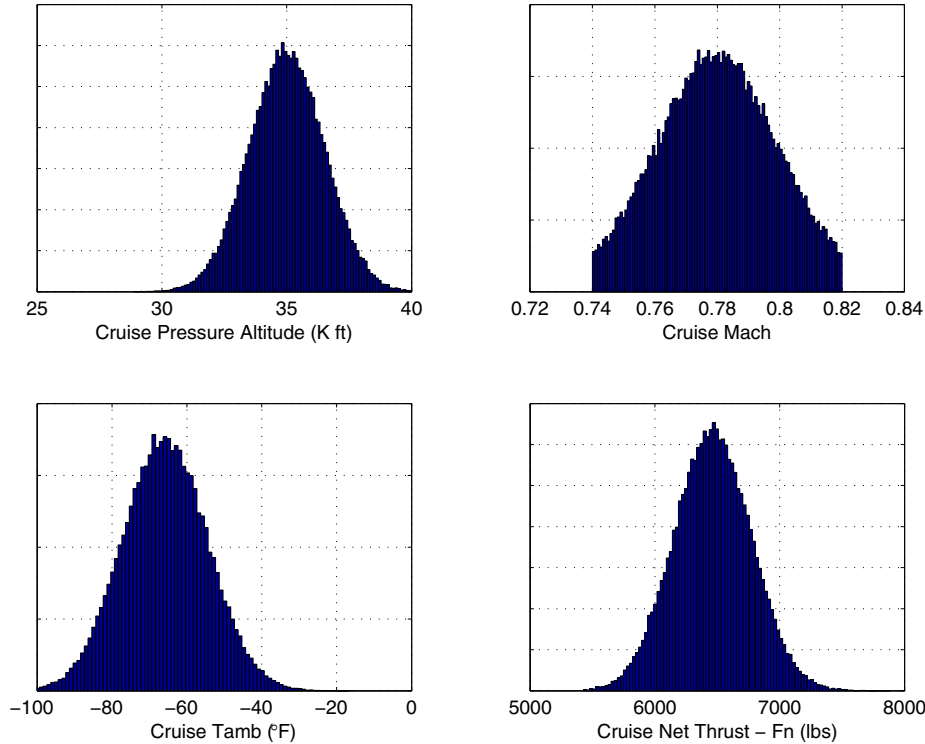Figure 4.—Takeoff operating condition distributions.

Figure 5.—Cruise operating condition distributions.

five major modules in the engine (Fan, Low Pressure Compressor (LPC), High Pressure Compressor (HPC), High Pressure Turbine (HPT), and Low Pressure Turbine (LPT)). Published information on turbofan performance deterioration profiles based upon historical data can be found in references (refs. 9 to 11). The fleet average deterioration profile implemented within the EFS is based on information provided in reference 10, although adjustments have been made to the turbine health parameters to cause them to deteriorate on a time scale consistent with the rest of the engine. Variations to the fleet average deterioration profile are included, via randomly assigned scale factors and adders, to produce a unique deterioration profile for each individual engine including: 1) More/less rapid overall engine deterioration; 2) More/less rapid individual module deterioration; 3) More/less module flow deterioration relative to efficiency deterioration; and 4) initial engine-to-engine manufacturing variation. Figure 6 shows the baseline, or average, profile for each health parameter (in red) and example variations (in cyan). The cyan points were generated by running the EFS to collect data from 100 engines over 5000 flights. 5000 flights is the maximum number of flight cycles that can be defined for any engine in the EFS. However, the user can choose fewer flights. If so, an engine's starting condition will be randomly placed somewhere along the deterioration profiles defined in figure 6.

In addition to generating the operating history and deterioration profile for each engine, the Case Generator will also define the type of fault each engine will experience, the flight of fault initiation, the fault magnitude, and the fault evolution rate. A summary of the fault types and their uniformly distributed fault magnitudes is provided in table 1. Component faults (i.e., Fan, LPC, HPC, HPT and LPT) are simulated by simultaneously adjusting the efficiency, η, and flow capacity, γ, health parameters of the component. Component fault magnitude distributions shown in the table are in terms of the root-sum-square value of the combined η and γ deviations. The uniformly distributed ratios of flow capacity to efficiency health parameter adjustment, $\gamma : \eta_{ratio}$, are also shown. For Fan, LPC, and HPC faults, both γ and η are reduced. For HPT and LPT faults, γ is reduced while η is increased. Component faults are simulated by adjusting efficiency and flow capacity as follows:

$$\eta \text{ adjustment} = \frac{\text{fault magnitude}}{\sqrt{1 + (\gamma : \eta_{ratio})^2}} \qquad (1)$$

$$\gamma \text{ adjustment} = \eta \text{ adjustment} \cdot (\gamma : \eta_{ratio})$$

Currently, Variable Bleed Valve (VBV) and Variable Stator Vane (VSV) actuator faults are modeled via adjustments to LPC and HPC flow capacity respectively. In the future, more detailed actuator models will be included to allow off-schedule VBV and VSV faults to be simulated. Sensor fault magnitudes are reflected in units of measurement standard
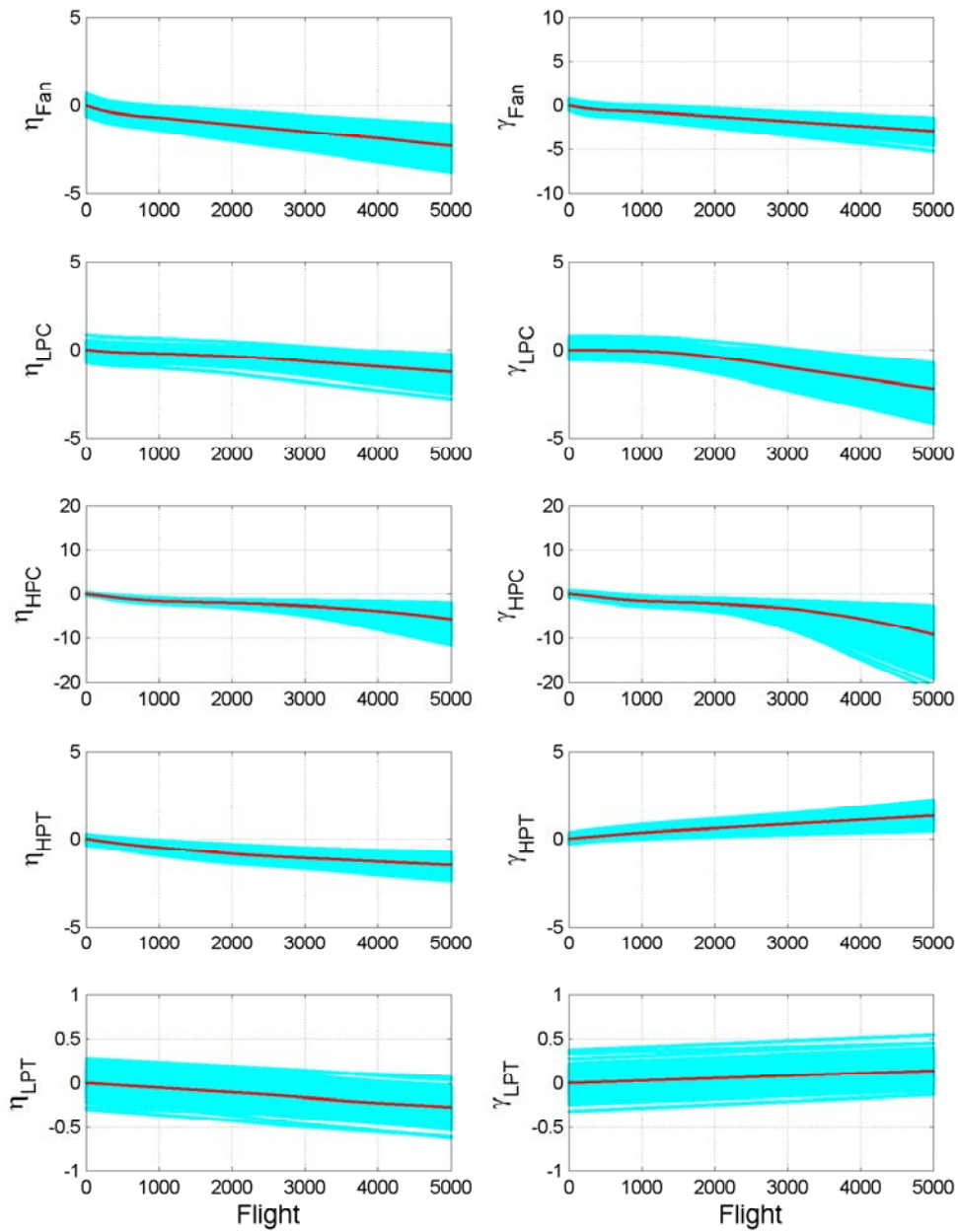
Figure 6.—Health parameter deterioration profiles,
average (red) and distribution (cyan).

TABLE 1.—EFS FAULT TYPES

Fault ID	Fault description	Fault magnitude	Fault $\gamma:\eta_{ratio}$	Takeoff average σ	Cruise average σ
0	No-fault	-----------	--------	--------	--------
1	Fan fault	1 to 7%	1 to 2	--------	--------
2	LPC fault	1 to 7%	1 to 2	--------	--------
3	HPC fault	1 to 7%	1 to 2	--------	--------
4	HPT fault	1 to 7%	−0.5 to −1	--------	--------
5	LPT fault	1 to 7%	−0.5 to −1	--------	--------
6	VSV fault	1 to 7%	--------	--------	--------
7	VBV fault	1 to 7%	--------	--------	--------
8	Nf sensor fault	± 1 to 7 σ	--------	5.59 rpm	4.64 rpm
9	Nc sensor fault	± 1 to 7 σ	--------	15.10 rpm	13.23 rpm
10	$P15$ sensor fault	± 1 to 7 σ	--------	0.042 psia	0.013 psia
11	$P24$ sensor fault	± 1 to 7 σ	--------	0.054 psia	0.016 psia
12	$Ps30$ sensor fault	± 1 to 7 σ	--------	0.889 psia	0.220 psia
13	$T24$ sensor fault	± 1 to 7 σ	--------	1.01 ºR	0.82 ºR
14	$T30$ sensor fault	± 1 to 7 σ	--------	2.47 ºR	1.92 ºR
15	$T48$ sensor fault	± 1 to 7 σ	--------	10.20 ºR	7.55 ºR
16	Wf sensor fault	± 1 to 7 σ	--------	0.058 pps	0.011 pps
17	$P2$ sensor fault	± 1 to 7 σ	--------	0.023 psia	0.008 psia
18	$T2$ sensor fault	± 1 to 7 σ	--------	0.838 ºR	0.707 ºR

deviation, σ, which varies proportional to the sensed parameter magnitude. Average σ values at takeoff and cruise are shown in the table.

C-MAPSS Engine Model

The outputs of the Case Generator are provided as inputs to the NASA Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) high-bypass turbofan engine model. C-MAPSS is a transient non-linear aerothermodynamic engine model equipped with an associated closed-loop control system. It has been developed for controls and diagnostics research and development purposes (ref. 12). A modified steady-state version of C-MAPSS is implemented within the EFS to facilitate faster convergence speeds. For each engine, at each flight operating point, the C-MAPSS engine model is run to the steady-state operating conditions specified by the Case Generator. Logic is applied to ensure that control limits are not violated in generating the steady-state solutions. End users will be provided access to the C-MAPSS steady-state model for diagnostic solution development purposes. For example, fault influence coefficient matrices can be extracted from C-MAPSS for use in the design of model-based diagnostic methods.

Parameter Histories and Fault Information

After the simulated engine parameter histories are produced by C-MAPSS, they are archived to a MATLAB file. This includes three-dimensional matrices containing sensed information collected at takeoff and cruise of dimension Number of Engines, by Number of Flights, by 12 Sensed Outputs. The 12 sensed outputs are shown in table 2. In addition to parameter histories, the EFS also stores the associated fault information for each engine in the fleet. This includes the fault type, fault magnitude, flight of fault initiation and fault evolution rate. This "ground truth" fault information will allow users to evaluate the overall performance of their developed diagnostic solutions. Once an EFS output dataset has been generated, it can be stored and used for future diagnostic algorithm development and evaluation purposes.

TABLE 2.—EFS SENSOR OUTPUT PARAMETERS

Index	Symbol	Description	Units
1	Nf	physical fan speed	rpm
2	Nc	physical core speed	rpm
3	$P15$	total pressure in bypass-duct	psia
4	$P24$	total pressure at LPC outlet	psia
5	$T24$	total temperature at LPC outlet	ºR
6	$Ps30$	Static pressure at HPC outlet	psia
7	$T30$	total temperature at HPC outlet	ºR
8	$T48$	total temperature at HPT outlet	ºR
9	Wf	fuel flow	pps
10	$P2$	total pressure at fan inlet	psia
11	$T2$	total temperature at fan inlet	ºR
12	$Pamb$	ambient pressure	psia

User Developed Diagnostic Solutions

As shown in figure 1, outputs from the EFS are provided as test case datasets for the development and validation of user developed diagnostic solutions. The challenge is to develop robust diagnostic methods to correctly diagnose the occurrence of any faults contained within the provided datasets, with minimal false alarms, missed detections, mis-

classifications, and detection latency. Diagnostic solutions shall produce an assessment for each engine, each flight, based solely on the sensed parameters shown in table 2 collected at, and prior to, the current flight. Although sensed engine outputs from upcoming flights and ground-truth fault information will also be available, it is *not permissible to utilize this information* for conducting diagnosis. As an output, each diagnostic algorithm shall produce a diagnostic assessment for each engine. These assessments are to be stored in a specified MATLAB format to facilitate automated evaluation against defined evaluation metrics. Diagnostic assessments will be stored in the form of a two-dimensional matrix of dimension number of engines by two, where each row corresponds to an individual engine, and the two columns consist of the Fault ID (as defined in table 1), and the flight of fault detection. For instances of no-fault found, zeros are to be recorded in each column.

Diagnostic Metrics

In order to enable a side-by-side comparison of candidate diagnostic approaches it is not only necessary to apply the candidate approaches to a common diagnostic problem, but also to evaluate their respective performance against a standard set of evaluation metrics. The importance of quantitative measures of effectiveness for diagnostic algorithms has been identified for gas turbine engine health management systems (refs. 13 to 16). These measures include performance metrics (e.g. thresholds, accuracy, reliability, sensitivity) as well as effectiveness metrics (e.g. complexity and cost). Metrics applications have also been developed to guide the selection of adequate data sets so that reliable statistics can be calculated (ref. 17). A complete coverage of key metrics is provided in the Society of Automotive Engineers Aerospace Recommended Practice Document 5783 (ref. 18).

For the TTCP gas path diagnostic benchmark process presented herein, a uniform set of metrics will be defined to evaluate the performance of diagnostic solutions developed and applied. Solution developers will be provided these metrics to enable them to independently evaluate the performance of their individual solutions given the ground-truth fault output information produced by the EFS. Although the complete list of metrics to be applied within the benchmark process are still being defined, they will be established to evaluate diagnostic accuracy (missed detections, false alarms, mis-classifications, and correct classifications), as well as detection latency (time required to detect a fault after fault initiation). These will include measures of overall diagnostic performance, provided by detection decision matrices and classification confusion matrices, as well as metrics which summarize overall diagnostic performance to enable a direct comparison between candidate algorithms, such as the Kappa Coefficient (ref. 18) or Normalized Product Entropy Ratio (ref. 16). The detection decision matrix, as shown in figure 7, demonstrates an algorithm's ability to discriminate between

Figure 7.—Detection decision matrix.

Figure 8.—Classification confusion matrix.

fault and no-fault cases. Here, the diagonal reflects the number of correct predictions.

From this matrix, True Positive Rate (TPR), and False Positive Rate (FPR) detection metrics can be readily calculated:

$$TPR = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \cdot 100\%$$

$$(2)$$

$$FPR = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \cdot 100\%$$

Although the detection decision matrix provides information on an algorithm's fault-detection capability, additional information is required to evaluate an algorithm's ability to discriminate, or classify, between multiple fault types. This will be reflected in a confusion matrix as shown in figure 8. The confusion matrix, denoted here as C, is a square matrix of dimension n, where n is the number of fault types. The "no-fault case" can also be included in the confusion matrix, although it has been excluded in the example shown here.

The diagonals of this square matrix reflect correct classifications. The Percent Correctly Classified (PCC) for the pth fault type, PCC_p, can be calculated by dividing the number of correct classifications of fault p, by the total number of fault p observations:

$$PCC_p = \frac{C_{pp}}{\sum\limits_{q=1}^{n} C_{pq}} \cdot 100\% \qquad (3)$$

While decision and confusion matrices provide an overall assessment of diagnostic algorithm performance, they are not in a form that enables easy comparison of different algorithms, especially when dealing with multiple fault classes. An additional metric, which summarizes the content of a confusion matrix into a single parameter, is the Kappa Coefficient (ref. 18). It is a measure of an algorithm's ability to correctly classify a fault, which takes into account the expected number of correct classifications occurring by chance. The Kappa Coefficient, denoted here as κ, can be calculated from the elements of a confusion matrix as follows:

$$\kappa = \frac{N(\text{correctly classifed}) - N(\text{expected correct by chance})}{N(\text{total}) - N(\text{expected correct by chance})}$$

where

$$N(\text{correctly classified}) = \sum_{p=1}^{n} C_{pp} \qquad (4)$$

$$N(\text{total}) = \sum_{p=1}^{n} \sum_{q=1}^{n} C_{pq}$$

$$N(\text{expected correct by chance}) = \sum_{p=1}^{n} \left\{ \sum_{q=1}^{n} \frac{C_{pq}}{N(\text{total})} \cdot \sum_{q=1}^{n} C_{qp} \right\}$$

If a diagnostic method achieves perfect fault classification performance then $\kappa = 1$. If its classification performance is worse than that expected by chance then $\kappa < 0$. The Kappa Coefficient along with detection and confusion matrices form an initial set of metrics to be included in the gas path diagnostic benchmark process. As additional metrics are defined, they will be included as well. A MATLAB routine, which auto-matically evaluates the performance of a diagnostic solution against the defined metrics, will be distributed along with the EFS.

Gas Path Diagnostic Benchmark Process—Example Diagnostic Solution

The previous sections outlined a process for developing and benchmarking candidate gas path diagnostic methodologies. In this section, an example diagnostic solution is presented and evaluated against the initial list of metrics described above. It is presented to demonstrate how a user defined diagnostic solution would integrate into the process. *This example should not be interpreted as the recommended approach for solving the problem*. It is simply provided as an example solution to illustrate the overall diagnostic process, and to serve as a template for the development of additional diagnostic solutions. The overall example solution, shown in figure 9, is partitioned into a three-step process consisting of: 1) trend monitoring; 2) anomaly detection; and 3) event isolation. Each of these steps is further discussed below.

Example Solution: Step 1—Trend Monitoring

The gas path diagnostic benchmark problem presented in this paper has been constructed assuming that engine performance changes can manifest themselves in two ways: a) gradual (long-term) deterioration or b) rapid (short-term) deterioration. The former is due to *all* of the engine components deteriorating slowly over time and is included here in an attempt to emulate physical causes such as erosion, corrosion, fouling and increased clearances within the turbomachinery. The latter is due to a *single* fault occurring. An effective gas path diagnostic solution must be able to function with both of these processes occurring and interacting simultaneously, and must be able to discriminate between the two cases, without corrupting the overall diagnostic approach.
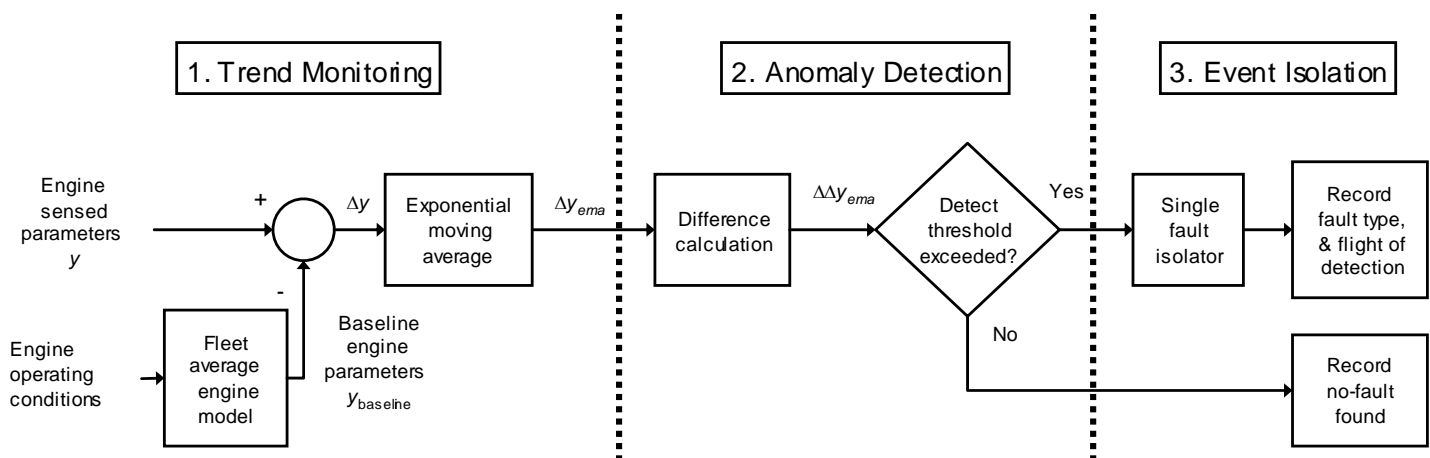
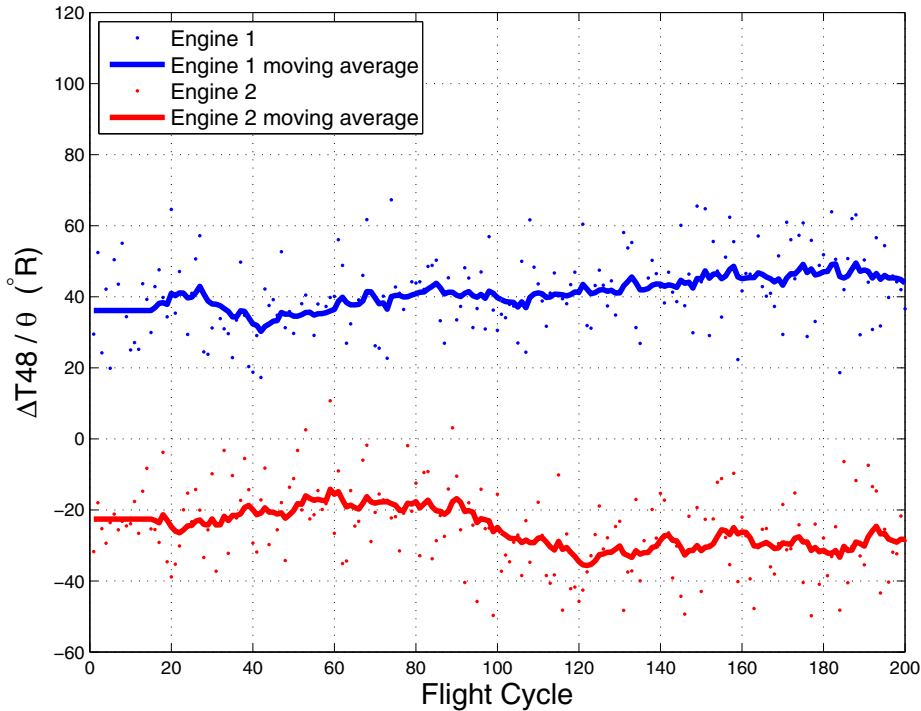

Figure 9.—Example solution process.

Figure 10.—Corrected *T48* measurement deltas and corresponding exponential moving averages.

In the example solution presented herein, a trend monitoring approach is applied to capture gradual performance changes in the form of residuals, or measurement deltas, relative to a fleet average engine. Although the defined metrics will not require the inclusion of trend monitoring functionality within a diagnostic solution, it is expected to improve overall diagnostic performance, and is thus included as part of the example solution. The C-MAPSS engine model was run over a range of altitude, Mach number, and corrected fan speed (*NfR*) settings at the 50 percent deterioration level in order to define a three-dimensional table lookup model used to represent a fleet average engine. Measured engine data collected during each flight are corrected to standard day conditions and then referenced against the fleet average engine to calculate measurement deltas, Δy_i's, as shown in equation (5)

$$\Delta y_i(k) = y_i(k) - y_{i\_\text{baseline}}(k) \qquad (5)$$

where $y_i(k)$ is the corrected value of the ith measurement collected during the kth flight, and $y_{i\_\text{baseline}}(k)$ is the fleet average engine value for the ith measurement at the corresponding pressure altitude, Mach number and *NfR* values of the kth flight. (Note: pressure altitude is the altitude corresponding to *Pamb* as defined within Standard Atmosphere tables. Pressure altitude and Mach number can be calculated from *Pamb* and *P*2). A $y_i(k)$ value is only calculated for eight of the 12 measurements shown in table 2. The parameters *Nf*, *Pamb*, *P*2, and *T*2 are used for establishing the

engine operating point and/or parameter correction purposes and are thus excluded from the $y_i(k)$ calculations. The calculated measurement delta values are trended over time applying an exponential moving average approach as described in reference 19 and shown in equation (6)

$$\Delta y_{i\_ema}(k) = \alpha \cdot \Delta y_{i\_ema}(k-1) + (1-\alpha) \cdot \Delta y_i(k) \qquad (6)$$

where $\Delta y_{i\_ema}(k)$ is the exponential moving average of the ith measurement delta on flight *k*. The moving average weighting between previous and current data is established by the constant α (where $0 < \alpha < 1$). An example of corrected *T*48 measurement deltas, ΔT48, and corresponding exponential moving average values (using an α of 0.885) from two randomly generated engines operating over 200 flight cycles is shown in figure 10. The exponential moving average can be seen to produce a smoothing effect on the overall trend shifts. Although it is not readily apparent from the plots, Engine 2 experiences a 3 percent fan fault on flight cycle 100 while Engine 1 experiences no faults. While a fan fault will cause ΔT48 to decrease, the downward trend observed in Engine 2 prior to fault initiation (between flights 90 to 100) is entirely coincidental and unrelated to the fault.

Example Solution: Step 2—Anomaly Detection

The previous sub-section presented an example approach for monitoring gradual engine performance measurement

deltas over time. Next, anomaly detection logic is applied to detect discrete events causing a rapid shift in observed measurement deltas. It is important to recognize that any anomaly detection logic must function relying only on the measurement information collected at, and prior to, the current flight cycle, k. It is also important to recognize that it would be prudent to incorporate a time latency inherent in the detection process to avoid false alarms caused by statistical outliers in the measurement data. The approach applied herein for anomaly detection uses a backwards difference calculation of the exponential moving average of each measurement delta as shown in equation (7)

$$\Delta\Delta y_{i\_ema}(k) = \Delta y_{i\_ema}(k) - \Delta y_{i\_ema}(k - \beta) \qquad (7)$$

where $\Delta\Delta y_{i\_ema}(k)$, or the measurement delta-delta, is the change in the exponential moving average of the ith measurement delta between flight k and some previous flight, k-β. The measurement delta-deltas also provide direct utility in performing the fault isolation function as will be described later. Choosing a $\beta = 10$ flight cycle distance between the compared EMA values was found to provide detection capability for both abrupt as well as rapid faults. It is anticipated that improved detection robustness for abrupt as

well as rapid faults could be obtained by combining multiple detection filters which apply different β distances within the measurement delta-delta calculations. Figures 11 and 12 show the measurement deltas (top plot), exponential moving average values (middle plot), and measurement delta-delta values (bottom plot) for eight engine parameters (*Nc*, *P*15, *P*24, *Ps*30, *T*24, *T*30, *T*48, and *Wf*) collected from the two engines previously introduced in figure 10. In these figures each signal has been normalized by the standard deviation, σ_i, of the measurement delta, Δy_i, to enable a comparison between the measurements. Anomaly detection logic is applied which monitors for a $\Delta\Delta y_{i\_ema}(k)$ exceeding a $\pm 2\sigma_i$ threshold. Engine 1, shown in figure 11, does not contain a fault and all $\Delta\Delta y_{i\_ema}$ signals remain below the anomaly detection threshold for the entire 200 flight profile. Engine 2, shown in figure 12, contains a fan fault occurring on flight 100. Here one of the $\Delta\Delta y_{i\_ema}$ signals, (*Ps*30), exceeds the anomaly detection threshold on flight 104 indicating that an anomaly is present. Note that the $\Delta\Delta y_{i\_ema}$ signals do not immediately exceed the detection threshold at the flight of fault occurrence, nor do they persistently exceed the threshold into the future. This is due to the combined effect of the exponential moving average, and the backwards difference calculations applied.
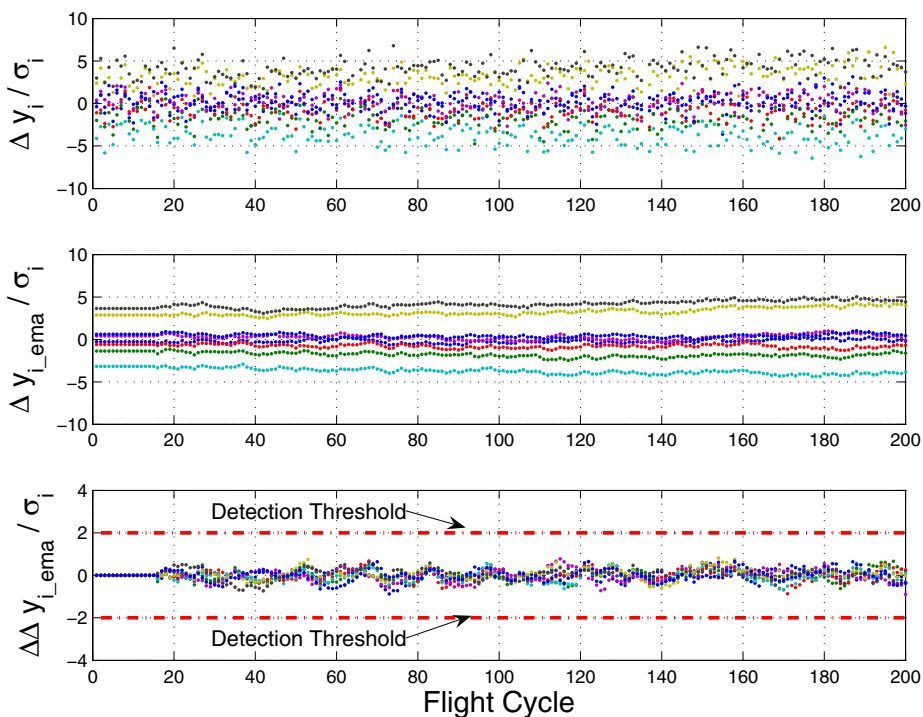

Figure 11.—$\Delta y_i/\sigma_i$, $\Delta y_{i\_ema}/\sigma_i$ and $\Delta\Delta y_{i\_ema}/\sigma_i$ values for engine 1 (no fault).
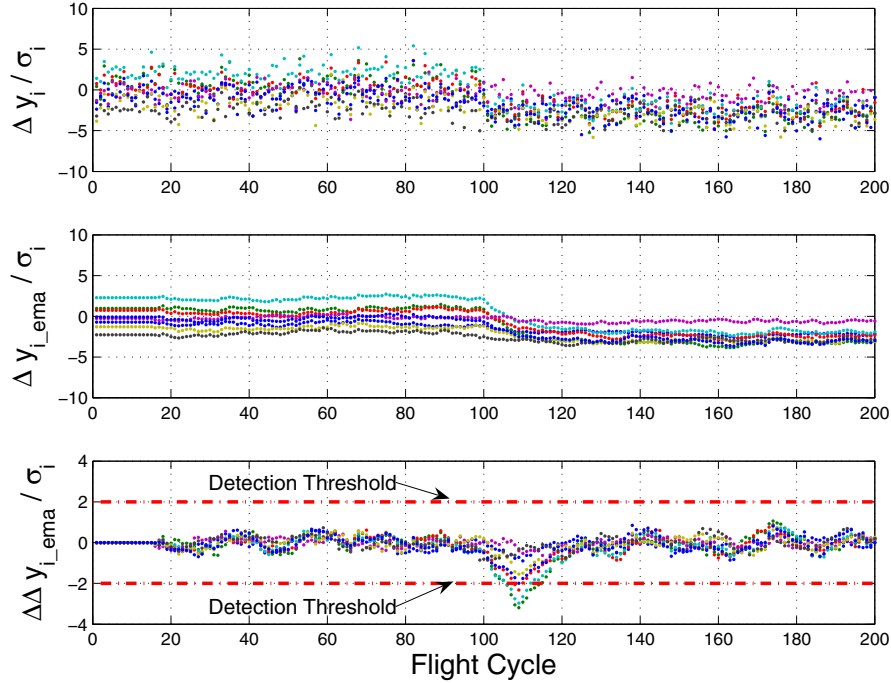
Figure 12.—$\Delta y_i/\sigma_i$, $\Delta y_{i\_ema}/\sigma_i$ and $\Delta\Delta y_{i\_ema}/\sigma_i$ values for engine 2 (fan fault occurring on flight 100).

Example Solution: Step 3—Event Isolation

The final step in the example solution is event isolation. If the diagnostic logic detects an anomaly, it classifies a root cause for the detected anomaly. Following the steps described in the previous section, an anomaly signature in the measurement delta-delta space due to the underlying fault event is obtained:

$$\Delta\Delta Y(k) = \begin{bmatrix} \Delta\Delta y_{1\_ema}(k) \\ \Delta\Delta y_{2\_ema}(k) \\ \vdots \\ \Delta\Delta y_{m\_ema}(k) \end{bmatrix} \qquad (8)$$

where m is the number of sensor measurements.

Given this information a Kalman filter can be configured, as described in references 2 and 20, to operate as a single fault isolator conducting a snapshot type of analysis. The objective of this analysis is to identify the single fault cause of the event on the basis of the $\Delta\Delta Y(k)$ measurement delta-delta vector. This requires the generation of a fault influence coefficient matrix which relates engine faults to measurement delta-delta changes in engine outputs. We will denote the ($m\times n$) fault influence matrix as H, where m = number of measurements, and n = number of single fault types. Assuming $\Delta x(k)$ is an $n\times1$ vector representing the magnitudes of the n single fault

types under consideration, the interrelationship between faults to measurement delta-delta changes can be written as:

$$\Delta\Delta Y(k) = H\Delta x(k) \qquad (9)$$

For the gas path diagnostic benchmark problem $m = 8$ measurements, and $n = 18$ single fault types (see table 1). The H matrix can be generated by running the C-MAPSS steady-state engine model to a fixed closed-loop operating condition (specified by altitude, Mach, $\Delta Tamb$ and corrected Nf) and individually introducing each of the 18 fault types. The elements of the matrix consist of the partial derivatives relating the change in corrected measured engine outputs (fault condition vs. nominal condition) to the magnitude of the implanted fault. The H^T matrix generated at the 35K ft, 0.78 Mach, 0 $\Delta Tamb$, 84 percent $PCNfR$ operating condition for a 50 percent deteriorated engine is shown in table 3. Here, rotor speeds, pressures and fuel flow delta-delta shifts are shown in percent units, and temperature delta-delta shifts are shown in degrees Rankine. These delta-delta shifts are relative to a 1 percent fault in all cases except for temperature sensor faults, which correspond to a 1 °R bias in the respective temperature sensor.

Next, a single fault estimator can be constructed as:

$$\Delta\hat{x}(k) = P_0 H^T \left(H P_0 H^T + R \right)^{-1} \Delta\Delta Y(k) \qquad (10)$$

where $\Delta\hat{x}(k)$ is an $n \times 1$ vector of estimated fault magnitudes, P_0 is the state covariance matrix, R is the measurement noise covariance, and $\Delta\Delta Y(k)$ is the vector of the observed measurement delta-deltas at the time of anomaly detection on flight k. The single fault isolation is obtained by processing (10) iteratively to provide a snapshot analysis for each of the n single fault cases under consideration. Each iteration of the above equation was made with a different P_0 matrix chosen to accentuate one of the single fault cases while zeroing out all others. Since this is a snapshot analysis, a covariance update calculation is not required. Through this iterative process an estimated fault magnitude, $\Delta\hat{x}_j(k)$, for each of the n single fault cases under consideration is produced. The index j, in $\Delta\hat{x}_j(k)$, corresponds to the fault under consideration. The corresponding estimated measurement delta-delta vector for each of the n single fault cases can be calculated as:

$$\Delta\Delta\hat{Y}_j(k) = H\Delta\hat{x}_j(k) \qquad (11)$$

It is important to emphasize that in the above equation only the jth element of the $\Delta\hat{x}_j(k)$ vector will be non-zero due to the selection of P_0 during the iterative evaluation process. The associated normalized measurement estimation error for each of the n fault cases is then calculated as:

$$e_j = \sqrt{\sum_{i=1}^{m}\left(\frac{\Delta\Delta y_{i\_ema}(k) - \Delta\Delta\hat{y}_{i\_ema}(k)}{\sigma_i}\right)^2} \qquad (12)$$

where $\Delta\Delta\hat{y}_{i\_ema}(k)$ is the ith element of the vector $\Delta\Delta\hat{Y}_j(k)$ and σ_i is the standard deviation of the ith measurement delta, $\Delta y_{i,}$. The single fault estimate, $\Delta\hat{x}_j(k)$, which produces the minimum normalized measurement estimation error is inferred to be the fault cause. Applying this isolation technique to the measurement delta-delta vector of the detected anomaly in figure 12 yielded the normalized measurement estimation error results shown in figure 13 for the 18 candidate single fault conditions. Here the estimated fan fault (fault ID 1) produced the minimum normalized measurement estimation error, and was correctly isolated as the cause of the anomaly.
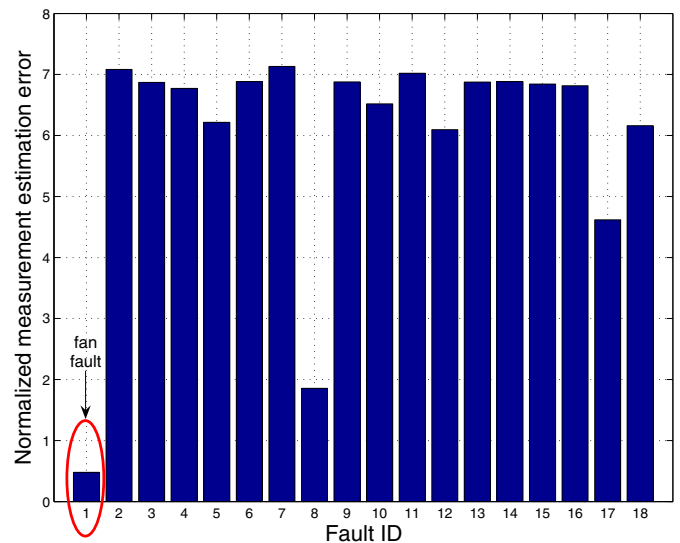

Figure 13.—Normalized measurement estimation error, e_j, for 18 single-fault candidates in the case of a fan fault.

TABLE 3.—TRANSPOSE OF FAULT INFLUENCE COEFFICIENT MATRIX, H^T

	$\Delta\Delta y_i$							
	Nc	P15	P24	Ps30	T24	T30	T48	Wf
Fan	−0.08	−0.42	−0.36	−0.81	−0.17	−1.99	−3.88	−1.02
LPC	0.00	0.00	−0.12	−0.03	−0.16	−0.02	0.32	0.01
HPC	0.09	0.02	0.08	−0.16	0.27	5.23	12.32	0.60
HPT	−0.16	0.04	0.13	−1.09	0.48	−3.51	23.47	1.09
LPT	0.15	−0.03	−0.15	0.89	−0.45	3.34	0.12	0.87
VSV	0.20	0.01	0.03	−0.03	0.08	1.82	3.51	0.17
VBV	−0.03	0.00	−0.16	−0.06	−0.53	−0.79	−0.64	−0.06
Nf	−0.34	−0.72	−1.06	−2.31	−1.93	−8.68	−16.52	−3.16
Nc	1	0	0	0	0	0	0	0
P15	0	1	0	0	0	0	0	0
P24	0	0	1	0	0	0	0	0
Ps30	0	0	0	1	0	0	0	0
T24	0	0	0	0	1	0	0	0
T30	0	0	0	0	0	1	0	0
T48	0	0	0	0	0	0	1	0
Wf	0	0	0	0	0	0	0	1
P2	0	−0.99	−0.99	−0.99	0	0	0	−0.99
T2	−0.07	0.08	0.12	0.28	−1.13	−1.98	−2.15	0.22

Example Solution—Metrics

The diagnostic performance of the given example solution was evaluated against EFS generated test cases consisting of 100 cases of each of the 18 fault types (1800 fault cases total), plus an additional 1800 no-fault cases. The corresponding metrics in the form of a detection decision matrix, classification confusion matrix, and Kappa Coefficient were calculated based on the assessed diagnostic condition produced by the example solution, and the "ground truth" conditions produced by the EFS. The detection decision matrix, along with the True Positive Rate and False Positive Rate is shown in figure 14.

The corresponding classification confusion matrix, percent correct classification rate, and average detection latency (in terms of the average number of flights) for the same 3600 test cases is shown in figure 15. The Kappa Coefficient, calculated by equation (4), was found to be 0.77. All metric calculations are implemented within a MATLAB routine which automatically calculates and archives the results given the diagnostic assessments in pre-defined format.

Future Steps

This paper and the associated ASME Turbo Expo Controls, Diagnostics and Instrumentation Committee tutorial are to invite feedback from the EHM community regarding the proposed format and content of the gas path diagnostic benchmark process. The intent is to define a publicly available benchmark problem with metrics that are of value to the entire EHM community: developers, users, and evaluators. Input received will be used to make adjustments to the benchmark process. The next step will be to disseminate the problem and invite solutions. Interested participants will be provided the benchmark problem coded in MATLAB, an example solution, the defined evaluation metrics, and identical blind test cases free of charge. In exchange, as their contribution to the EHM community, they will be asked to provide diagnostic assessments for the provided blind test cases. These blind test case diagnostic assessments will be evaluated against the defined metrics by the TTCP. Participants will receive their results plus the anonymous results of other participants. A follow-on limited access workshop will be convened to share solution results and lessons learned. The intent is not to formulate this as a competition, but rather as a means for the engine health management community to share diagnostic approaches.

Specific additional work is required prior to public release of the benchmark problem. Representative distributions for sensor noise, operating point variations and other data will be incorporated based upon available operational engine data. Assessments of metrics will be extended over a wider range of diagnostic sets representative of in service experience with single and multiple faults. For additional information, and updates on the current status of this effort, individuals are directed to visit the following website:

www.grc.nasa.gov/WWW/cdtb/software/ehmbenchmark.html

		Predicted State		
		Fault	No Fault	
True State	Fault	1343	457	TPR = 74.6%
	No Fault	0	1800	FPR = 0%

Figure 14.—Example solution metrics—detection decision matrix, true positive rate, and false positive rate.

	Fan	LPC	HPC	HPT	LPT	VSV	VBV	Nf	Nc	P15	P24	Ps30	T24	T30	T48	Wf	P2	T2	No Fault	PCC	Latency
Fan	83	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	10	83%	3.8
LPC	0	8	0	0	0	0	4	0	0	0	0	0	0	0	0	0	2	0	86	8%	9.9
HPC	0	0	86	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	12	86%	4.2
HPT	0	0	0	92	0	0	0	2	0	0	0	2	0	1	0	0	0	0	3	92%	2.1
LPT	0	0	0	0	83	0	0	0	0	0	0	0	0	0	0	0	0	0	17	83%	3.7
VSV	0	0	1	0	0	61	0	3	13	0	0	0	0	0	0	0	0	2	20	61%	5.2
VBV	0	1	0	0	0	0	32	1	0	0	0	0	0	0	0	0	0	0	66	32%	9.3
Nf	1	0	0	0	0	0	0	81	0	0	0	0	0	0	0	0	0	0	18	81%	5.5
Nc	1	0	0	0	0	2	0	5	70	0	0	0	0	0	0	0	1	0	21	70%	4.8
P15	0	0	0	0	0	0	0	0	0	92	0	0	0	0	0	0	3	0	5	92%	2.7
P24	0	16	0	0	0	0	1	0	0	0	77	0	0	0	0	0	2	0	4	77%	3.5
Ps30	1	0	0	1	0	0	0	1	0	0	0	75	0	0	0	0	0	0	22	75%	5.6
T24	0	0	0	0	1	0	0	2	0	0	0	0	69	0	0	0	0	3	25	69%	5.1
T30	0	0	0	0	1	0	0	3	0	0	0	0	0	64	0	0	0	2	30	64%	6.0
T48	1	0	0	1	0	1	0	1	0	0	0	0	0	0	73	0	0	0	23	73%	4.4
Wf	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	97	0	0	0	97%	2.0
P2	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	35	0	62	35%	8.6
T2	0	0	0	0	0	0	1	1	0	0	0	0	2	0	0	0	0	63	33	63%	6.6
No Fault	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1800	100%	N/A

Figure 15.—Example solution metrics—classification confusion matrix, percent correctly classified, average detection latency.

Conclusion

A publicly available gas path diagnostic benchmark problem has been developed and is being offered to the EHM technical community to serve as a common platform for initial development and evaluation of candidate gas path diagnostic solutions. This public approach towards benchmarking diagnostic systems is focused on providing a common basis of comparison of candidate solutions—an area of need as identified in recent engine health management technology reviews. Members of the engine health management community are invited to provide their feedback on the proposed approach to ensure that it adequately addresses their areas of interest. An international program has been developed to lead this process including a follow-on workshop to share results and lessons learned.

References

1. Jaw, L.C., "Recent Advances in Aircraft Engine Health Management (EHM) Technologies and Recommendations for the Next Step," ASME Paper GT2005–68625, 2005.
2. Volponi, A.J., DePold, H., Ganguli, R., Daguang, C., "The Use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics: A Comparative Study," Journal of Engineering for Gas Turbines and Power, vol. 125, pp. 917–924, 2003.
3. Kamboukos, Ph., Mathioudakis, K., "Comparison of Linear and Nonlinear Gas Turbine Performance Diagnostics," Journal of Engineering for Gas Turbines and Power, vol. 127, pp. 49–56, 2003.
4. Sambasivan, L. K., et al., "Comparison of a Few Fault Diagnosis Methods on Sparse Variable Length Time Series Sequences," ASME GT2007-27843, IGTI Turbo Expo 2007, Montreal, QC, May 2007.
5. Mathioudakis, K., Kamboukos, Ph., Stamatis, A., "Turbofan Performance Deterioration Tracking Using Nonlinear Models and Optimization Techniques," Journal of Turbomachinery, vol. 124, pp. 580–587, 2002.
6. Dewallef, P., Romessis, C., Mathioudakis, K., and Léonard, O., "Combining Classification Techniques with Kalman Filters for Aircraft Engine Diagnostics," ASME Paper GT–2004–52541, 2004.
7. Romessis, C., Mathioudakis, K., "Bayesian Network Approach for Gas Path Fault Diagnosis," Journal of Engineering for Gas Turbines and Power, vol. 128, pp. 64–72, 2006.
8. The Technical Cooperation Program (TTCP) Website: http://www.dtic.mil/ttcp/ , accessed Jan. 22, 2008.
9. Sasahara, O., "JT9D Engine/Module Performance Deterioration Results from Back to Back Testing," ISABE 85-7061, Proceedings of the Seventh International Symposium on Air Breathing Engines, Beijing, PRC, 1985.
10. Sallee, G.P., "Performance Deterioration Based on Existing (Historical) Data—JT9D Jet Engine Diagnostics Program," NASA CR–135448, United Technologies Corporation, Pratt & Whitney Aircraft Group Report PWA–5512–21, 1978.
11. Wulf, R.H., "Engine Diagnostics Program CF6-50 Engine Performance Deterioration," NASA CR–159867, Nov. 1980.
12. Frederick, D.K., DeCastro, J.A., Litt, J.S., "User's Guide for the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS)," NASA/TM—2007-215026.
13. Orsagh, R.F., Roemer, M.J., Savage, C.J., Lebold, M., "Development of Performance and Effectiveness Metrics for Gas Turbine Diagnostic Technologies," 2002 IEEE Aerospace Conference Proceedings, vol. 6, pp. 6-2825–6-2834, 2002.
14. DePold, H., Siegel, J., Hull, J., "Metrics for Evaluating the Accuracy of Diagnostic Fault Detection Systems," ASME Paper GT2004–54144, 2004.
15. Depold, H. R., Rajamani, R., Morrison, W. H. and Pattipati, K. R., "A Unified Metric for Fault Detection and Isolation in Engines", ASME Paper GT2006–91095, 2006.
16. Sambasivan, L. K., Mukherjee, J., and Mylaraswamy, D., "Benchmarking Diagnostic Algorithms," ASME Paper GT2007–28194, 2007.
17. Martin, B. and Meckl, P., 2007, "Input Selection for Modeling and Diagnostics With Application to Diesel Engines", Journal of Dynamic Systems Measurement and Control, Jan. 2007, vol. 129, no. 1, pp. 114–120.
18. Society of Automotive Engineers E-32, Health and Usage Monitoring Metrics, Monitoring the Monitor, Feb. 14, 2007, SAE Aerospace Recommended Practice 5783—DRAFT.
19. DePold, H.R., Gass, F.D., "The Application of Expert Systems and Neural Networks to Gas Turbine Prognostics and Diagnostics," Journal of Engineering for Gas Turbines and Power, vol. 121, pp. 607–612, 1999.
20. Volponi, A.J., et al., "Gas Turbine Condition Monitoring and Fault Diagnostics", Von Karman Institute for Fluid Dynamics, Lecture Series 2003–01, 2003.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 01-07-2008	2. REPORT TYPE Technical Memorandum	3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Benchmarking Gas Path Diagnostic Methods: A Public Approach

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Simon, Donald, L.; Bird, Jeff; Davison, Craig; Volponi, Al; Iverson, R., Eugene

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
WBS 645846.02.07.03.03.01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
John H. Glenn Research Center at Lewis Field
Cleveland, Ohio 44135-3191

8. PERFORMING ORGANIZATION REPORT NUMBER
E-16542

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORS ACRONYM(S)
NASA

11. SPONSORING/MONITORING REPORT NUMBER
NASA/TM-2008-215271; GT2008-51360

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified-Unlimited
Subject Category: 07
Available electronically at http://gltrs.grc.nasa.gov
This publication is available from the NASA Center for AeroSpace Information, 301-621-0390

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Recent technology reviews have identified the need for objective assessments of engine health management (EHM) technology. The need is two-fold: technology developers require relevant data and problems to design and validate new algorithms and techniques while engine system integrators and operators need practical tools to direct development and then evaluate the effectiveness of proposed solutions. This paper presents a publicly available gas path diagnostic benchmark problem that has been developed by the Propulsion and Power Systems Panel of The Technical Cooperation Program (TTCP) to help address these needs. The problem is coded in MATLAB (The MathWorks, Inc.) and coupled with a non-linear turbofan engine simulation to produce "snap-shot" measurements, with relevant noise levels, as if collected from a fleet of engines over their lifetime of use. Each engine within the fleet will experience unique operating and deterioration profiles, and may encounter randomly occurring relevant gas path faults including sensor, actuator and component faults. The challenge to the EHM community is to develop gas path diagnostic algorithms to reliably perform fault detection and isolation. An example solution to the benchmark problem is provided along with associated evaluation metrics. A plan is presented to disseminate this benchmark problem to the engine health management technical community and invite technology solutions.

15. SUBJECT TERMS
Aircraft engines; Systems health monitoring; Gas turbine engines; Flight safety

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON STI Help Desk (email:help@sti.nasa.gov)
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	22	19b. TELEPHONE NUMBER *(include area code)* 301-621-0390