JOURNAL OF PROPULSION AND POWER
Vol. 18, No. 5, September–October 2002

Gas Turbine Engine and Sensor Fault Diagnosis Using Optimization Techniques

Marco Zedda*
Rolls–Royce, plc., Derby, England DE24 8BJ, United Kingdom
and
Riti Singh†
Cranfield University, Cranfield, England MK43 OAL, United Kingdom

A diagnostic system is described for performance analysis of gas turbine engine components and sensors. The system estimates the performance parameters expressing the fault condition of the engine components in the presence of measurement noise and biases. The measurement uncertainty is supposed to affect even the parameters setting the operating condition of the engine. Estimation is performed through optimization of an objective function by means of an ad hoc genetic algorithm. The genetic algorithm uses an accurate nonlinear steady-state performance model of the engine. The only statistical assumption required by the technique concerns the measurement noise and the maximum allowed number of faulty sensors and engine components, which is enclosed as a constraint. The technique has been thoroughly tested with the model of a low bypass ratio turbofan, and the results show the high level of accuracy achieved.

Nomenclature

$\boldsymbol{b}$ = bias vector of monitoring measurements
$\boldsymbol{b}_w$ = bias vector of the environment and power setting parameters
C_D = nozzle discharge coefficient
e_i = estimation error of the ith setting parameter
F = cumulative distribution function
F_G = cumulative Gaussian distribution function
FN = thrust
$\boldsymbol{h}$ = simulated measurement vector
J = objective function
M = number of measurements
M_{bias} = assumed number of biases
N = overall number of performance parameters
N_{perf} = number of fault-affected performance parameters
N_H = high-pressure spool speed
N_L = low-pressure spool speed
P = number of environment and power setting parameters
P_1 = inlet total pressure
P_3 = high-pressure compressor exit total pressure
P_5 = low-pressure turbine exit total pressure
P_{13} = fan outer exit total pressure
P_{21} = fan inner exit total pressure
Q = number of evaluations of J for each $\boldsymbol{x}$ and $\boldsymbol{w}$
T_1 = engine inlet total temperature
T_3 = high-pressure compressor exit temperature
T_5 = low-pressure turbine exit temperature
T_{13} = fan outer exit temperature
T_{21} = fan inner exit temperature
$\boldsymbol{u}$ = vector of measured environment and power setting parameters
W_{FE} = main burner fuel flow
W_{1a} = engine inlet airflow
W_{21} = inner fan mass flow
$\boldsymbol{w}$ = vector of actual environment and power setting parameters
$\boldsymbol{x}$ = performance parameter vector
$x_{\text{est}\,j}$ = estimate of the jth performance parameter
$\boldsymbol{z}$ = monitoring measurement vector
$\boldsymbol{z}_{\text{od}}$ = vector of simulated undeteriorated measurements
Γ_{fan} = overall fan flow capacity
Γ_{HPC} = high-pressure compressor flow capacity
Γ_{HPT} = high-pressure turbine flow capacity
Γ_{LPT} = low-pressure turbine flow capacity
ε = probability of deviation from the Gaussian distribution
η_{fanin} = fan inner efficiency
η_{fanout} = fan outer efficiency
η_{HPC} = high-pressure compressor efficiency
η_{HPT} = high-pressure turbine efficiency
η_{LPT} = low-pressure turbine efficiency
$\boldsymbol{\nu}$ = noise vector of monitoring measurements
$\boldsymbol{\nu}_w$ = noise vector of environment and power setting parameters
σ_j = standard deviation of the jth measurement noise

Introduction

ANALYSIS of the performance of gas turbines is important for both testing of development engines and condition monitoring of operating engines. Moreover, a thorough analysis is required when an engine fails a pass-off test. The economic and operating gains produced by application of accurate and reliable diagnostics are relevant. In particular, identification of the component(s) responsible for the loss of performance makes it possible to choose the recovery action to be undertaken. Gas path analysis (GPA) is a diagnostic technique able to focus on the health of gas turbine components by calculation of performance parameters, for example, efficiency and flow function, using measurements as input, for example, pressures, temperatures, and spool speeds. Basic aerothermodynamic equations and component characteristics are used to model engine performance. Once the variations of performance parameters with respect to their reference values are calculated, a further step is necessary to relate those changes to actual component shortfall. Even though GPA shows interesting potential, when used in practice it is affected by several drawbacks. Some of them are listed here.

Received 23 July 2001; revision received 15 May 2002; accepted for publication 16 May 2002.

*Principal Technologist, Performance Department, P.O. Box 31.
†Professor of Gas Turbine Engineering, University Technology Centre, Bedfordshire.

1) Measurement noise level is large, and the effect on measurements can be of the same order of magnitude as the changes induced by a real engine fault.

2) Measurement biases are frequent, and their effect on diagnostics can be dramatic. An undetected sensor bias can either point at a nonexistent fault or neglect a real fault in an engine component.

3) The relationship between measurements and performance parameters is nonlinear, and this complicates solution of the first two problems.

4) A large number of performance parameters usually have to be estimated to assess the health of a modern gas turbine engine, and this calls for a large number of sensors, which may simply not be available, especially for analysis of on-wing engines.

To cope with the problem of measurement noise and biases, a great deal of research has been done to apply estimation techniques to deterministic GPA. Weighted least-squares- and Kalman-filter- (KF-) based techniques[1−3] have seemed to be perfect candidates to make GPA effective. However, their performance is not satisfactory for several reasons. Some of them are the following.

First, the standard KF should be used with linear processes, whereas gas turbines perform in a nonlinear way. Process linearization causes inaccuracy due to the low robustness of the filter to model errors.[4] Unrealistic statistical assumptions about the performance parameters have to be accepted to apply KF. 1) The prior probability density function of the performance parameters is assumed to be Gaussian. 2) The KF-based estimation techniques must be given the standard deviations of the performance parameters that are usually known only approximately. When applied to gas turbines, KF often fails because it tries to produce an answer by relying on poor information about the probability of fault occurrence. 3) Even though sensitivity techniques can be applied, tuning of the filter is not straightforward and can bias the diagnostic answer.

Second, measurements are processed to estimate both performance parameters and measurement biases. Because the problem is largely underdetermined and a high level of noise is present, the proposed solution is likely to be inaccurate. The maximum likelihood basis of the filter entails that inaccuracy takes the form of "smearing": Every time a single component or sensor is faulty, the filter's typical answer underestimates the variations in the parameters actually affected by faults and, moreover, detects variations in the other parameters. This occurrence is rather common and makes it difficult to understand the real fault condition of the engine.

Awareness of the aforementioned flaws has led to the development of modified estimation techniques. The main engine manufacturers have devised different methods to focus on a limited number of parameters to reduce the smearing effect and to cope with biases.[5−8] However, the effectiveness of those techniques seems to be limited.[9]

As an alternative, application of computational intelligence techniques and especially neural networks (NNs) have been proposed to cope with measurement noise and biases.[10−12] When applied to gas turbine fault diagnosis, the main drawback of NNs is that a rather complex structure of nets is necessary to solve real-world, complex problems. However, if this lack of flexibility is accepted, NNs can provide accurate solutions.[13,14]

In this paper the gas turbine fault diagnosis problem is solved through optimization of an objective function by means of a genetic algorithm. The proposed diagnostic system is suitable to deal with most of the issues presented earlier. Tests with simulated engine data confirm the effectiveness of the technique.

Gas Turbine Engine Fault Diagnosis as an Optimization Problem

If neither noise nor biases were present, the following relationship would hold:

$$z = \boldsymbol{h}(\boldsymbol{x}) \tag{1}$$

where $z \in R^M$ is the measurement vector and M is the number of measurements, $\boldsymbol{x} \in R^N$ is the performance parameter vector and N is the number of parameters, and $\boldsymbol{h}(\cdot)$ is a vector-valued function. Here $\boldsymbol{h}(\cdot)$ is provided by the simulation program and is nonlinear. A further assumption for Eq. (1) to apply is the absence of modeling errors.

Because measurement noise is present, Eq. (1) must be modified as follows:

$$z = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{\nu} \tag{2}$$

In presence of biases, Eq. (2) becomes

$$z = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{b} + \boldsymbol{\nu} \tag{3}$$

Equation (3) defines the relationship for a certain operating point. If the dependance on the operating point is written explicitly,

$$z = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{w}) + \boldsymbol{b} + \boldsymbol{\nu} \tag{4}$$

where $\boldsymbol{w} \in R^P$ is the vector of the environment and power setting parameters, for example, inlet condition parameters and fuel flow.

Usually $\boldsymbol{\nu}$ is assumed to have a Gaussian probability density function (PDF) and, moreover, to have independent components. Therefore, the joint PDF is

$$p(\boldsymbol{\nu}) = \frac{1}{\left(\sqrt{2\pi}\right)^M} \prod_{j=1}^{M} \left(\frac{1}{\sigma_j}\right) \exp\left[-\frac{1}{2}\sum_{j=1}^{M}\left(\frac{\nu_j}{\sigma_j}\right)^2\right] \tag{5}$$

where σ_j is the standard deviation of the jth measurement's noise. Note that $\boldsymbol{w}$ is affected by noise as well as biases like the other measurements,

$$\boldsymbol{u} = \boldsymbol{w} + \boldsymbol{b}_w + \boldsymbol{\nu}_w \tag{6}$$

In the proposed method, performance parameters are estimated by minimizing an objective function by means of genetic algorithms (GAs). Figure 1 shows the structure of the diagnostic system. The basic requirements for the objective function are as follows:

1) It should be a measure of the consistency between actual and predicted measurements.

2) Measurement noise should be taken into account.

3) Measurement biases should be taken into account.

4) Its minimization should reduce the smearing effect.

5) Evaluation of the function should not be too burdensome from a computational point of view. Note that no statistical assumption is made about the performance parameters' PDF. The only assumption regards the measurement noise, which is usually statistically well known. A classic choice for the objective function would be, given a certain operating point,

$$J(\boldsymbol{x}) = \sum_{j=1}^{M} \frac{[z_j - h_j(\boldsymbol{x})]^2}{(z_{\mathrm{od}\,j}\sigma_j)^2} \tag{7}$$

where $z_{\mathrm{od}\,j}$ is the value of the jth measurement in the off-design undeteriorated condition. The terms are conveniently expressed with respect to the reference condition. Minimization of function (7) provides the maximum likelihood solution for the nonlinear problem at hand.[15]

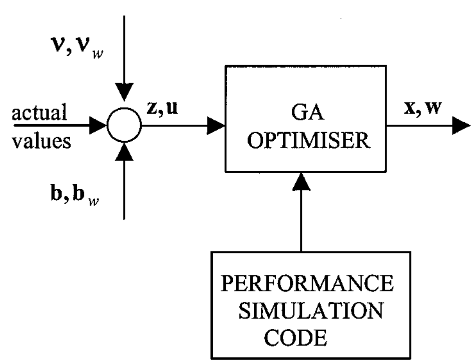

Fig. 1 Layout of the diagnostic system.

Another suitable function is the absolute deviation,

$$J(\boldsymbol{x}) = \sum_{j=1}^{M} \frac{|z_j - h_j(\boldsymbol{x})|}{z_{\mathrm{od}\,j}\sigma_j} \tag{8}$$

Equations (7) and (8) are simple functions expressing the consistency between actual and predicted measurements. Measurement noise is taken into account by the standard deviations σ_j.

Because uncertainty affects the environment and power setting parameters as well, function (7) will be modified as follows:

$$J(\boldsymbol{x}, \boldsymbol{w}) = \sum_{j=1}^{M} \frac{[z_j - h_j(\boldsymbol{x}, \boldsymbol{w})]^2}{[z_{\mathrm{od}\,j}(\boldsymbol{w}) \cdot \sigma_j]^2} \tag{9}$$

whereas function (8) becomes

$$J(\boldsymbol{x}, \boldsymbol{w}) = \sum_{j=1}^{M} \frac{|z_j - h_j(\boldsymbol{x}, \boldsymbol{w})|}{z_{\mathrm{od}\,j}(\boldsymbol{w}) \cdot \sigma_j} \tag{10}$$

Measurement noise is usually assumed to be Gaussian. Whenever the Gaussian PDF can be considered an accurate model of the noise, function (9) should be used. As a matter of fact, the actual measurement noise PDF may not be perfectly Gaussian shaped, for a number of reasons. First, in practice, the occurrence of readings out of the three standard deviation range is much more common than what is supposed by the Gaussian model. Second, modeling errors are inevitably present, especially in the simulation of gas turbine performance. For these reasons, other objective functions, such as Eq. (10), are more suitable because they provide a robust estimation.[15] Straight minimization of any of the preceding functions does not account for measurement biases and the smearing problem. The classic approach to cope with measurement biases is to assume that all measurements may be affected by systematic errors, which have to be estimated (see vector $\boldsymbol{b}$). It also assumes that all performance parameters may be fault affected. As mentioned earlier, this may lead to inaccuracy of the estimation. The goal would be to concentrate on a subset of performance parameters actually responsible for the engine faults after the subset of biased measurements has been rejected.

In the diagnostic system presented here the following assumptions are made:

1) A maximum number of biased measurements is allowed, M_{bias}.

2) A maximum number of faulty engine components is allowed (and then of fault-affected performance parameters, N_{perf}).

These constraints seem to be sensible from a diagnostic point of view because the most common case in fault diagnosis of gas turbine engines is that just a single engine component is faulty in presence of one or two biased measurements. From an analytical point of view, they are definitely helpful in reducing the estimation inaccuracy.

The way the diagnostic system applies these constraints is described in the next two sections. Note that the method presented here is not suitable for analysis of deterioration of on-wing engines, where the performance shortfall is usually spread over a large number of components and the available instrumentation is less comprehensive. In that case, smearing should be allowed, and, thus, more conventional methods can be used.

Sensor Validation

The functions described can also be modified to deal with measurement biases according to the following criterion: The presence of a bias will introduce inconsistency between actual and predicted measurements. When dealing with measurement biases, the possible diagnostic tasks are the following: 1) detection, which is the capability to realize if there is something wrong with the instrumentation set, no matter where the problem is located; 2) isolation, which is the capability to identify the fault-affected sensor; and 3) accommodation, which is the capability to produce an estimate of the biased measurement. A complete treatment of measurement biases is called sensor failure detection, isolation, and accommodation (SFDIA).

The way the GA handles measurement biases is based on the concept of relative redundancy. If no bias affected the measurements, then minimization of function (10) would estimate $\boldsymbol{x}$ and $\boldsymbol{w}$ so that the equations used in the terms of the summation would be mutually consistent. The inconsistency due to a biased measurement would manifest itself with larger values of the objective function because no $\boldsymbol{x}$ or $\boldsymbol{w}$ can be found to correspond to predicted measurements fitting sufficiently well the real ones. The problem can be overcome by elimination in the summation of function (10) of the M_{bias} terms corresponding to the biased measurements. Then the remaining terms are mutually consistent, and the optimized function will reach a low value. For the technique to apply, it is necessary that

$$M - M_{\mathrm{bias}} > N_{\mathrm{perf}} + P \tag{11}$$

It is also assumed that the $N_{\mathrm{perf}} + P$ parameters are function of the $M - M_{\mathrm{bias}}$ remaining measurements. The assumption is usually acceptable due to the coupling of the equations defined by Eq. (1). This redundancy relative to the number of fault-affected performance parameters is required because, if $M - M_{\mathrm{bias}} = N_{\mathrm{perf}} + P$, then any choice as to the identity of the biased measurements would produce a consistent solution. However, if the equations defining the terms used in the objective function are more numerous than the parameters to be estimated, the problem becomes overdetermined, and the lowest value of the function should provide the more consistent solution. In the case considered, the relative redundancy is guaranteed by the assumption about the maximum number of faulty engine components, which is acceptable for fault diagnosis. In this respect, note that the sensor validation task can be feasible even if the number of measurements is smaller than the number of performance parameters ($M < N$), provided inequality (11) holds.

Because of the large level of measurement noise, the larger the left-hand side is with respect to the right-hand side the better, as the redundancy is larger. The way just N_{perf} out of N performance parameters are allowed to vary is explained in the next section.

Typical SFDIA techniques proceed in three sequential steps: Whenever a fault in the instrumentation set is detected, the faulty sensor is isolated, and then the measurement is possibly accommodated. The approach described here is different: M_{bias} is chosen at the onset of the analysis, so that an M_{bias} number of measurements are not used in the objective function. The optimization produces an estimation of the performance parameters $\boldsymbol{x}$ and the environment and power setting parameters $\boldsymbol{w}$ (Fig. 1). Given the estimated $\boldsymbol{x}$ and $\boldsymbol{w}$, the predicted values of the M_{bias} measurements not used in the optimization are automatically obtained. If the difference between the actual and predicted measurements is larger than a threshold based on the noise standard deviations, then SFDIA is achieved; otherwise, the outcome of the analysis is that no bias is present and estimation of $\boldsymbol{x}$ and $\boldsymbol{w}$ has been done by using the subset of measurements providing the best consistency.

Because the identity of the faulty sensors is unknown, a combinatorial search has to be done for every $\boldsymbol{x}$ and $\boldsymbol{w}$ to find the selection providing the lowest value. If, for example, two biases are assumed to be present, every time the objective function $J(\boldsymbol{x}, \boldsymbol{w})$ has to be evaluated during the optimization procedure, the following set of functions are calculated by sequential elimination of two measurements:

$$J_{kl}(\boldsymbol{x}, \boldsymbol{w}) = \sum_{\substack{j=1 \\ j \neq k,l}}^{M} \frac{|z_j - h_j(\boldsymbol{x}, \boldsymbol{w})|}{z_{\mathrm{od}\,j}(\boldsymbol{w}) \cdot \sigma_j} \tag{12}$$

The value assigned to the objective function will be

$$J(\boldsymbol{x}, \boldsymbol{w}) = \min_{k,l} J_{kl}(\boldsymbol{x}, \boldsymbol{w}) \tag{13}$$

In general, the number Q of evaluations of $J(\boldsymbol{x}, \boldsymbol{w})$ to choose from is equal to the number of possible combinations:

$$Q = \binom{M}{M_{\text{bias}}} = \frac{M!}{M_{\text{bias}}!(M - M_{\text{bias}})!} \tag{14}$$

Environment and power setting parameter biases are dealt with in a different way because they basically affect most of the terms in the summation. A bias in any of these parameters is likely to increase the value of most terms and, therefore, of the overall sum. Therefore, SFDIA is automatically performed for the environment and power setting parameters because they have to be guessed by the optimizer.

Estimation of Performance Parameters by GAs

A large level of inaccuracy is likely to affect the solution if all performance parameters are allowed to be responsible for the shift in the measurements that are actually due to engine faults located in one or two components. Application of the constraint on the maximum number of faulty engine components enables reduction of inaccuracy and smearing. Moreover, the sensor validation technique based on relative redundancy can be used. Concentration on the faulty engine components is effected by use of a tailored GA.

With typical calculus-based optimization methods (with or without computation of gradients) it is difficult to minimize functions such as Eqs. (12) and (13), which are remarkably "non-well behaved." In the present work GA[16,17] have been used to minimize the chosen objective function. Several distinctive features make them different from other optimization techniques.

1) No derivatives need be calculated, and in principle, any nonsmooth function can be optimized. Calculation of gradients for complex real-world gas turbine engine performance models is basically unfeasible analytically and burdensome and error prone numerically.

2) Constraints can be dealt with in a very flexible way. There are two different ways of embedding constraints into a GA-based optimization: by means of penalty functions or design of specific operators.[18]

3) GAs use a population of possible solutions (strings) rather than a single solution to be updated iteration by iteration. They are global search techniques and, therefore, less likely than calculus-based methods to get stuck in a local rather than the global minimum.

4) Probabilistic rather than deterministic transition rules are used to create the next generation of strings from the current one.

Points 1 and 2 highlight the suitability of GAs for gas turbine performance analysis. The underlying idea to reduce inaccuracy through optimization is as follows: Because the performance parameter vector solution is supposed to have just few nonzero elements due to faults affecting one or two engine components at most, the optimization technique is only allowed to propose solutions satisfying this constraint. In fact, a reduction of the number of degrees of freedom produces more accurate results. Achievement of this objective is much easier if a GA-based optimizer is used.

The constraint as to the maximum number of fault-affected performance parameters is applied by guided initialization of the population of strings and design of tailored genetic operators rather than use of a penalty term.

A typical GA is based on three operators: selection, crossover, and mutation. The selection operator chooses the strings to be used in the next generation according to a survival of the fittest criterion: The objective function to be minimized is mapped onto a fitness function to be maximized, and the larger the fitness the higher the probability of survival. The crossover operator allows information exchange between strings, in the form of swapping of parts of the parameter vector in an attempt to generate fitter strings. Mutation is used to introduce new or prematurally lost information in the form of random changes applied to randomly chosen vector components.

The GA here used for diagnostics is somewhat different from the typical one just outlined. Its main features are the following.

1) The objective function is a function of the performance parameter vector $\boldsymbol{x}$ and the environment and power setting parameter vector $\boldsymbol{w}$.

2) Whereas most GAs use a binary coding of the parameter space, here a real coding is chosen as faster and more accurate convergence has been achieved.[19,20] The creation of problem-specific operators is remarkably simplified as well.

3) Initialization of the population is guided. As a calculus-based iterative technique for solving an optimization problem is more likely to converge successfully if the initial guess it starts with is sufficiently close to the solution being sought, the population of strings is initialized according to the constraint on the maximum number of fault-affected performance parameters. In this way, various fault classes are generated: Some of them will represent faults in a single engine component, others faults affecting two components. These classes are then processed by the three operators described next.

4) The selection operator uses the stochastic universal sampling algorithm.[21] Selection is extended to all classes, so that iteration after iteration those classes will survive that on average show lower values of the objective function. After a few iterations, the competition is over because the whole population is made of a single fault class. Concentration on one or two engine components is then achieved.

5) The crossover operator is applied to pairs of strings that are members of the same fault class. Crossmating between different fault classes is not allowed.

6) The mutation operator is dynamic,[22] in that size of the random mutations applied to the parameters is exponentially reduced to foster fine convergence to the solution. The operator just modifies the fault-affected parameters for each fault class, so that the constraint is satisfied. In simple terms, selection extended to the whole population forces concentration on the faulty engine components, whereas crossover and mutation help selection to refine the solution.

When GAs are dealt with, a problem can be the computing power necessary to achieve convergence. Every iteration involves evaluation of the objective function for most of the strings, and from the point of view of computing power, the main burden is running the performance simulation code [calculating $\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{w})$ and $z_{\text{od}}(\boldsymbol{w})$]. The code requires application of an efficient Newton–Raphson based method to reach convergence. Nowadays, all major gas turbine manufacturers have developed quick and robust methods for iteratively solving the set of equations defining the performance simulation codes.

The tests described in the next section show that running the GA-based optimizer to convergence is definitely feasible from a computational point of view. The optimization task demanded, though, is certainly difficult inasmuch as the objective function is very nonsmooth and a combinatorial problem is actually underlying the search of the solution because a choice of faulty sensors and engine components is required. As a matter of fact, the method is much more burdensome than classic KF-based techniques. However, it is believed that the achievable gains justify the use of the computing power required, especially for performance analysis of development engines. In this case, performance engineers may spend a lot of time trying to solve the difficult diagnostic task at hand.

Note that no linearization is ever done during any stage of the fault diagnosis. The fully nonlinear approach provides accurate results, even though estimation is more complicated.

Testing

The diagnostic method has been tested with simulated data of a low bypass ratio turbofan, the EJ200,[23] powerplant of the European Fighter Aircraft (Typhoon). A Rolls–Royce steady-state performance simulation model (named RRAP) has been used. The engine is mainly modeled by six components: inner and outer fan, high-pressure (HP) compressor, HP turbine, low-pressure (LP) turbine, and propelling nozzle. To monitor the health of the engine components, 10 performance parameters are used: outer and inner fan efficiency η_{fanout} and η_{fanin}, overall fan flow function Γ_{fan}, nozzle discharge coefficient C_D, HP compressor (HPC), HP turbine (HPT), LP turbine (LPT) efficiencies, and flow functions η_{HPC}, Γ_{HPC}, η_{HPT}, Γ_{HPT}, η_{LPT}, and Γ_{LPT}. The engine is supposed to be analyzed in a typical development test bed. The instrumentation suite supposed to be available is made of 16 measurements: inlet total pressure and temperature P_1 and T_1, spool speeds N_H and N_L, outer fan exit total pressure and temperature P_{13} and T_{13}, inner fan exit total pressure and temperature P_{21} and T_{21}, inner fan mass flow W_{21}, HP compressor exit total pressure and temperature P_3 and T_3, LPT exit total pressure and temperature P_5 and T_5, main burner fuel flow W_{FE}, engine inlet airflow W_{1a}, and thrust FN. P_1 and T_1 are used to set the ambient condition, W_{FE} is given, and thrust is measured. Thus 3 measurements are used to set the operating point and 13 to monitor the engine components. Real measurement noise levels have been used, and they are similar in magnitude to what is available in the public literature.[24]

One or two engine components are supposed to be faulty in the presence of two or four biases affecting the measurements. The three environment and power setting parameters are both noisy and biased. HP and LP spool speeds are assumed to be unbiased. This situation has actually been chosen to stay on the safe side because in practice such a large number of faulty sensors is unlikely in a real test bed. The level of deterioration is defined by the ranges of variation of the performance parameters: A maximum of 3% is allowed. Flow functions are allowed to both increase and decrease.

The biases superimposed to measurements have purposely been chosen small because small values are particularly undesirable in that their detection is difficult. Undected biases can affect the estimation of the performance parameters. The magnitude of the biases has been set to 1% for all measurements; apart from those whose assumed measurement nonrepeatability range, equal to $3^*\sigma$, is comparable to 1% variation. For these, a level of 2% has been chosen. The biases affecting the environment and power setting measurements have been set to the large value of $50^*\sigma$ as accommodation is automatically carried out. The GA population is set to 4000 strings, and 200 iterations are sufficient to reach convergence.

Table 1 shows the comparison between typical results provided by the proposed diagnostic system and a straightforward maximum likelihood-based optimization minimizing function (9) with allowance for noise and biases in $\boldsymbol{w}$, without constraint on the number of fault-affected parameters. The comparison shows that the GA-based approach produces a more accurate estimation, and in particular the smearing effect is ruled out. Interpretation of the results is much easier because the fault-affected components are easily isolated. Note that with both approaches the estimation of the environment and power setting parameters is accurate.

Table 2 shows the values of all of the terms that could be used to sum up in the objective function in the case of faulty outer fan and HPC introduced earlier. In the considered case, the W_{1a}, P_{13}, T_{13}, and P_{21} are biased. Correctly, the GA identified the first four measurements as faulty, and then the corresponding terms have not been included in the objective function.

Figure 2 shows the typical convergence of the population of strings to a single fault class. Concentration on the faulty components is easily achieved after few iterations. Table 3 summarizes the 150 test cases that have been run as well as the average results.

Table 1 Comparison in a two faulty components case

Parameters	Actual	Predicted	Maximum likelihood
$\Delta\eta_{\text{fanout}}$	−3	−2.99	−2.9
$\Delta\Gamma_{\text{fan}}$	0	0	−0.27
$\Delta\eta_{\text{fanin}}$	0	0	−0.33
$\Delta\Gamma_{\text{HPC}}$	3	2.99	2.91
$\Delta\eta_{\text{HPC}}$	−1	−1.03	−0.41
$\Delta\Gamma_{\text{HPT}}$	0	0	−0.56
$\Delta\eta_{\text{HPT}}$	0	0	−0.11
$\Delta\Gamma_{\text{LPT}}$	0	0	−0.72
$\Delta\eta_{\text{LPT}}$	0	0	−0.2
ΔC_a	0	0	−0.27
Setting parameters			
Wf	811.728	811.603	811.231
P_1	83.688	83.773	83.574
T_1	312.02	312.094	311.89
RMS	——	0.01	0.39

Table 2 Bias isolation

Parameter	Actual	Predicted
W_{1a}	6.59	6.65
P_{13}	13.07	13.28
T_{13}	11.19	11.38
P_{21}	11.84	12.09
T_{21}	0.16	0.02
W_{21}	0.06	0.13
P_3	0.73	0.48
T_3	0.1	0.19
T_5	0.12	1.02
P_5	0.81	0.16
F	0.63	0.55
N_{hp}	0.92	0.77
N_{lp}	0.27	0.27

Table 3 Test case results

Head	Two biases	Four biases
	One faulty component	
rms	0.037	0.106
e_1, g/s	0.44	0.87
e_2, kN/m^2	0.11	0.13
e_3, K	0.18	0.23
SCC, %	93.2	95.2
SSC, %	98.3	97.6
	Two faulty components	
rms	0.181	0.272
e_1, g/s	0.86	1.46
e_2, kN/m^2	0.22	0.27
e_3, K	0.32	0.38
SCC, %	98.1	96.3
SSC, %	91.9	91.7

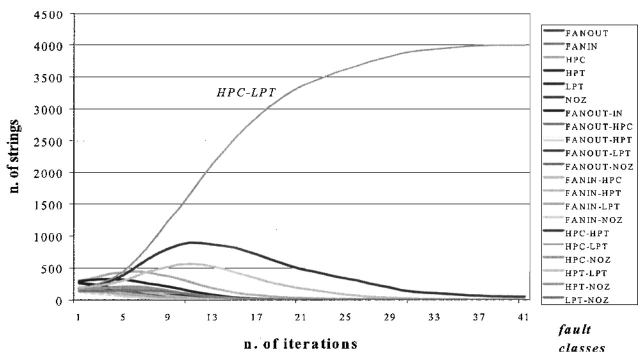

Fig. 2 Convergence of the fault classes.

The root mean square error,

$$\text{rms} = \sqrt{\sum_{j=1}^{N}(x_j - x_{\text{est}\,j})^2 \Big/ N} \tag{15}$$

where x_j and $x_{\text{est}\,j}$ are the actual and estimated values of the jth performance parameter, respectively and e_1, e_2, and e_3 are the average errors of estimation of the ambient and power setting parameters. Successful component classification (SCC) rate and successful sensor classification (SSC) rate are the percentage of correct identification of faulty engine components and sensors, respectively. The results show the high level of accuracy achieved, especially when data are analyzed in light of the overall number of biases (5 and 7, respectively, out of 16 measurements). As expected, an increase in the number of biases produces larger estimation errors in terms of rms and e_1, e_2, and e_3.

An increase in the number of faulty components also produces larger estimation errors. Note that a larger value of $(N_{\text{perf}} + P)/(M - M_{\text{bias}})$ implies less relative redundancy. Conversely, the capability to isolate faulty engine components and sensors does not generally seem to be dependent on the number of biases. The percentages of successful component classification show how much the smearing effect has been reduced. The better SCC performance provided in the two faulty component cases is due to a number of runs where a single actually faulty component has been identified, but a small deterioration has wrongly been found in the nozzle as well. Usually these errors affect the estimation rms only marginally. Note that, even though any fault relating to a single faulty component can be represented by a string of a two-component fault class provided that the fault-free component's performance parameters are zero, this occurrence is very rare except for the nozzle. Whereas the GA automatically identifies during the convergence whether one or two engine components are faulty, the number of supposedly biased measurements has to be set from the beginning.

If the actual biased measurements are greater than the assumed ones, M_{bias}, then the optimizer is likely to converge to an incorrect solution due to the effect of the undetected biases. However, if M_{bias} is larger than the number of actual biases, the optimizer is likely to estimate the engine faults correctly because the biased measurements will be isolated together with other fault-free sensors to get the most consistent solution. M_{bias} should be chosen on the basis of specific knowledge about the typical occurrence of sensor faults for the given engine-sensor suite. In the considered case of the test bed for a two-spool development engine, occurrence of faults in more than two sensors has to be regarded as unlikely. However, the choice of M_{bias} is usually not crucial because the number can safely be chosen to overestimate the real number of biases.

If no clue about a good choice of M_{bias} is available, a sensible approach is to run a number of optimizations with different values of M_{bias}. For every value of M_{bias}, an average value can be fixed, which represents the expected magnitude of the optimized objective function when no bias has gone undetected. Comparison of the results allows the guess of the actual number of biases. The results shown in Table 3 confirm that the optimizer performs well even with a relatively large number of biases. Good performance should be achieved as long as relative redundancy is ensured. This condition is usually satisfied in comprehensively instrumented development engine test beds. Pass-off tests are usually dependent on a much more restricted number of sensors, and then multiple operating point analysis[25] should be used to maximize the amount of information extractable from the sensor suite.

The results presented so far have been obtained with function (9) applied to Gaussian noise measurements. A set of test cases have been run with both Eqs. (10) and (9) to compare the diagnostic accuracy when the measurement noise is not perfectly Gaussian. It is common to simulate the discrepancy from the Gaussian noise model by superimposition (with probability ε) to the correct Gaussian PDF of a similar PDF with a standard deviation that is three times larger. The noise cumulative distribution function is then assumed to be made of two terms:

$$F(z;\sigma) = (1-\varepsilon)F_G(z;\sigma) + \varepsilon F_G(z;3\sigma) \tag{16}$$

where z is a measurement, σ is its standard deviation, ε is a small number ($\sim$0.1), and F_G is the cumulative Gaussian distribution function.

Using the objective function (10) is like assuming the measurement noise is to be defined by a double exponential joint PDF:

$$p(\nu) = \frac{1}{\left(\sqrt{2}\right)^M} \prod_{j=1}^{M} \left(\frac{1}{\sigma_j}\right) \exp\left(-\sqrt{2}\sum_{j=1}^{M} \frac{|\nu_j|}{\sigma_j}\right) \tag{17}$$

which has much wider tails than the Gaussian PDF for the same value of the standard deviation. Figure 3 shows a double exponential and two Gaussian PDFs.

From a statistictal point of view, utilization of Eq. (10) is sound if σ_j is the mean absolute deviation and not the standard deviation. However, measurement nonrepeatability of gas turbine sensors is usually expressed by standard deviations. Moreover, if the noise is Gaussian, the absolute mean deviation converges to $\sqrt{(2/\pi\sigma)}$ if the number of observations to estimate the noise scatter is large.

Table 4 Comparison of results with different objective functions

Head	Function (9)	Function (10)
rms	0.487	0.198
e_1, g/s	3.30	1.80
e_2, kN/m^2	0.51	0.27
e_3, K	0.83	0.58
SCC, %	63.6	72.7
SSC, %	86.4	86.4

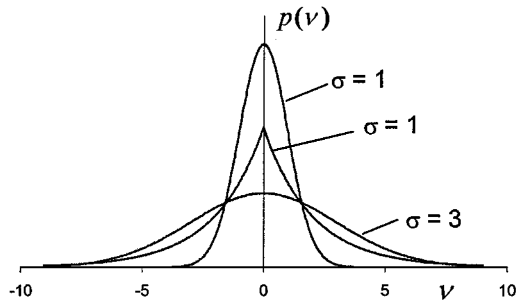

Fig. 3 PDFs.

With $\varepsilon = 0.3$,50 mixed test cases have been run to compare squares and absolute values [functions (9) and (10)] . Table 4 summarizes the results. As expected, the objective function (10) provides remarkably more accurate results due to its robustness to small deviations from the assumed model. The estimation errors are smaller, and the capability to isolate faulty components better. However, the two functions show the same capability to isolate faulty sensors. As far as the optimization process is concerned, use of the smoother function (9) allows a faster convergence, whereas function (10) is harder to minimize.

Conclusions

Since its introduction, the successful application of GPA has been hindered by many factors. KF-based estimation techniques have long been used to deal with real-world issues (such as measurement noise and biases) and make GPA effective. However, a number of unrealistic assumptions have to be accepted to apply the technique, and the outcome is that the diagnostic capability is often unsatisfactory when smearing is not required.

A novel diagnostic approach has been developed and tested. The gas turbine engine and sensor fault diagnosis problem is solved by means of an optimization technique. Performance parameters are estimated by minimizing an objective function. The smearing effect, typical of classic maximum-likelihood-based techniques, makes it difficult to identify clearly the component(s) responsible for the engine's loss of performance. This problem is successfully overcome by application of a constraint on the maximum number of fault-affected performance parameters. Another consequence of the application of this constraint is that relative redundancy is available and, thus, SFDIA is carried out for the monitoring measurements, as well as for the ambient and power setting parameters. The remarkably nonsmooth objective function is minimized by a specifically designed, real coded GA, which isolates the fault-affected engine components and sensors.

The proposed diagnostic system has been thoroughly tested with a large number of cases produced with a Rolls–Royce accurate steady-state performance simulation model of a two-spool, low bypass ratio turbofan, the EJ200. The results show the high level of accuracy achieved even in presence of such a large number of sensor faults.

Furthermore, the diagnostic system is flexible: In case sensible guesses on the maximum number of faulty sensors are available, the optimizer can be tailored accordingly.

Use of an objective function different from the classic quadratic one allows to cope with discrepancies from the assumed Gaussian-shaped measurement noise, as well as with small modeling errors.

Acknowledgments

The authors would like to thank John Martin (Chief of Corporate Performance Technology, Rolls–Royce, plc., Derby, England, U.K.) and Peter Robinson (Performance Consultant, Rolls–Royce Marine Power, Bristol, England, U.K.) for their help, support, and guidance during the development of the present work.

References

[1]Bryson, A. E., Jr., and Ho, Y., *Applied Otimal Control: Optimization, Estimation, and Control*, rev., Hemisphere, Washington, DC, 1975.

[2]Jazwinski, A. H., *Stochastic Process and Filtering Theory*, Academic Press, New York, 1970.

[3]Gelb, A., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

[4]Haupt, G. T., Kasdin, N. J., Keiser, G. M., and Parkinson, B. W., "An Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation," AIAA Paper 95-3218, 1995.

[5]Doel, D. L., "TEMPER4: Gas-Path Analysis Tool for Commercial Jet Engines," *Journal of Engineering for Gas Turbines and Power*, Vol. 116, No. 1, 1994; also American Society of Mechanical Engineers, ASME Paper 92-GT-315.

[6]Provost, M. J., "The Use of Optimal Estimation Techniques in the Analysis of Gas Turbines," Ph.D. Dissertation, School of Mechanical Engineering, Cranfield Univ., Cranfield, England, U.K., 1995.

[7]Urban, L. A., and Volponi, A. J., "Mathematical Methods of Relative Engine Performance Diagnostics," Society of Automotive Engineers, SAE Paper 922048, 1992.

[8]Volponi, A. J., "Sensor Error Compensation in Engine Performance Diagnostics," American Society of Mechanical Engineers, ASME Paper 94-GT-58, 1994.

[9]Doel, D. L., "An Assessment of Weighted-Least-Squares-Based Gas Path Analysis," *Journal of Engineering for Gas Turbines and Power*, Vol. 116, No. 2, 1994; also American Society of Mechanical Engineers, ASME Paper 93-GT-119.

[10]Junk, R., and Lunderstaedt, R., "Model Optimization of a Gas Turbine with Fault Affected Reference Measurements by Neuronal Networks for State and Sensor Diagnosis," *Proceedings of the AAIDA/AAAF/DGLR/RAeS 5th European Propulsion Forum*, 1995.

[11]Eustace, R., "Neural Network Fault Diagnosis of a Turbofan Engine," *Proceedings of the 11th International Symposium on Air Breathing Engines*, 1993, pp. 937–946.

[12]Eustace, R., and Merrington, G., "Fault Diagnosis of Fleet Engines Using Neural Networks," *Proceedings of the International Symposium on Air Breathing Engines*, 1995, pp. 926–936.

[13]Zedda, M., and Singh, R., "Fault Diagnosis of a Turbofan Engine Using Neural Networks: A Quantitative Approach," AIAA Paper 98-3062, 1998.

[14]Zedda, M., and Singh, R., "Neural Network Based Sensor Validation for Gas Turbine Test Bed Analysis," *Proceedings of the Institution of Mechanical Engineering*, Vol. 215, Pt. 1, 2001, pp. 47–52.

[15]Huber, P.J., *Robust Statistics*, Wiley, New York, 1981.

[16]Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley Longman, Reading, MA, 1989.

[17]Krishnakumar, K., "Genetic Algorithms: A Robust Optimization Tool," AIAA Paper 93-0315, Jan. 1993.

[18]Michalewicz, Z., and Janikow, C. Z., "Handling Constraints in Genetic Algorithms," *Proceedings of the Fourth International Conference on Genetic Algorithms*, 1991.

[19]Antonisse, J., "A New Interpretation of Schema Notation that Overturns the Binary Encoding Constraint," *Proceedings of the Third International Conference on Genetic Algorithms*, 1989.

[20]Goldberg, D. E., "Real Coded Genetic Algorithms, Virtual Alphabets and Blocking," Dept. of General Engineering, IlliGAL Rept. 90001, Univ. of Illinois, Urbana–Champaign, IL, 1990.

[21]Baker, J.E., "Reducing Bias and Inefficiency in the Selection Algorithm." *Proceedings of the Second International Conference on Genetic Algorithms*, 1987.

[22]Janikow, C. Z., and Michalewicz, Z., "An Experimental Comparison of Binary and Floating Point Representations in Genetic Algorithms," *Proceedings of the Fourth International Conference on Genetic Algorithms*, 1991.

[23]Lane, R. J., and Behenna, J., "EJ200: The Engine for the New European Fighter Aircraft," *Journal of Engineering for Gas Turbines and Power*, Vol. 113, Jan. 1991; also American Society of Mechanical Engineers, ASME Paper 90-GT-119.

[24]Walsh, P. P., and Fletcher, P., *Gas Turbine Performance*, Blackwell Science, Boston, 1998.

[25]Stamatis, A., Mathioudakis, K., Berios, G., and Papaliou, K., "Jet Engine Fault Detection with Differential Gas Path Analysis at Discrete Operating Points," International Symposium on Air Breathing Engines, ISABE Rept. 89-7133, 1989.