

# A Generalized Likelihood Ratio Test for Adaptive Gas Turbine Performance Monitoring

S. Borguet

e-mail: s.borguet@ulg.ac.be

O. Léonard

e-mail: o.leonard@ulg.ac.be

Turbomachinery Group,  
University of Liège,  
Chemin des Chevreuils,  
1 4000 Liège, Belgium

*Kalman filters are widely used in the turbine engine community for health monitoring purposes. This algorithm has proven its capability to track gradual deterioration with good accuracy. On the other hand, its response to rapid deterioration is a long delay in recognizing the fault and/or a spread of the estimated fault on several components. The main reason for this deficiency lies in the transition model of the parameters that is blended in the Kalman filter and assumes a smooth evolution of the engine condition. This contribution reports the development of an adaptive diagnosis tool that combines a Kalman filter and a secondary system that monitors the residuals. This auxiliary component implements a generalized likelihood ratio test in order to detect and estimate an abrupt fault. The enhancement in terms of accuracy and reactivity brought by this adaptive Kalman filter is highlighted for a variety of simulated fault cases that may be encountered on a commercial aircraft engine. [DOI: 10.1115/1.2967493]*

*Keywords:* gas path analysis, adaptive estimation

## 1 Introduction

Predictive maintenance aims at scheduling the overhaul actions on the basis of the actual level of deterioration of the engine. The benefits are improved operability and safety as well as reduced life cycle costs. Generating reliable information about the health condition of the gas turbine is therefore a requisite and has been the subject of an intensive research in the community.

In this paper, *module performance analysis* is considered. Its purpose is to detect, isolate, and quantify the changes in engine module performance, described by so-called health parameters, on the basis of measurements collected along the gas path of the engine [1]. Typically, the health parameters are correcting factors on the efficiency and the flow capacity of the modules (fan, lpc, hpc, hpt, lpt, and nozzle), while the measurements are intercomponent temperatures, pressures, and shaft speeds.

Many techniques have been investigated to solve the diagnosis problem, ranging from optimal estimation to artificial intelligence as well as expert systems (see Ref. [2] for a detailed review). Among the various approaches, the popular Kalman filter [3] has received great attention from the researchers. This recursive, minimum-mean-squared-error algorithm has proven its capability to track engine wear with good accuracy. This *gradual deterioration* process (due to erosion, corrosion, or fouling, for instance) occurs during a normal operation of the engine and affects all major components at the same time. One reason for the ability of the Kalman filter to deal with that kind of degradation is that it assumes a “relatively” slow evolution of the health parameters to filter the measurement noise.

On the other hand, *rapid deterioration* may happen too. These faults, caused, for instance, by hot restarts or foreign/domestic object damage,<sup>1</sup> impact one (at most two) component(s) at a time. The time of occurrence of the event, as well as the impacted component and the magnitude of the fault, is typically unknown. The response of the Kalman filter to these short-time-scale varia-

tions in the engine condition is a long delay in recognizing the fault and/or a spread of the estimated fault over several components, which is termed “smearing” effect (see Ref. [4]).

To overcome these deficiencies, a strategy that combines a so-called single-fault isolator with a Kalman filter is proposed in Ref. [5]. The single-fault isolator consists of a bank of possible events that is examined whenever a residual is suspected to be representative of an abrupt fault. While very efficient, this solution implies the interaction between two different diagnosis tools. Moreover, the setup of the fault library requires significant knowledge and experience.

A broader way to tackle the problem at hand is to view it in the realm of *adaptive estimation* [6]. The basic idea consists in increasing the mobility of the health parameters momentarily in order to recognize a rapid degradation. Mathematically, it comes down to increasing the covariance of the health parameters on the basis of a statistical processing of the residuals. Several techniques have been explored in the literature, such as autocorrelation [7,8] and covariance matching [9]. In Ref. [10], the authors developed an adaptive Kalman filter for performance monitoring that integrates a covariance matching scheme.

Another attractive solution has been proposed in Ref. [11]. The transition model for the health parameters is modified so as to account for possible jumps in the parameters. This technique has been successfully applied in visual tracking [12] and is the backbone of the algorithm presented in this contribution. The newly developed diagnosis tool combines a Kalman filter, which relies on the assumption of a smooth variation of the health parameters, and a secondary system that monitors the residuals. This auxiliary component implements a *generalized likelihood ratio test* (GLRT) in order to detect and estimate a rapid fault. In that way, the performance of the Kalman filter is maintained under normal operation (long-time-scale deterioration), whereas the event detector acts as a fault compensator in the case of a sudden variation in the engine health. The enhancement in terms of accuracy and reactivity brought by this adaptive Kalman filter is illustrated for a variety of simulated fault cases that may be encountered on a commercial aircraft engine.

## 2 Description of the Method

The scope of this section is to provide the mathematical background of the adaptive algorithm. The generic diagnosis tool re-

<sup>1</sup>A domestic object damage is caused by an element of the engine (e.g., part of a blade) that breaks off and strikes a downstream flow path component.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 1, 2008; final manuscript received April 2, 2008; published online October 13, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

lying on a Kalman filter is briefly presented. Then, based on a model that can handle abrupt variations in the health parameters, the auxiliary component that performs the detection and estimation of the “jumps” is developed. Finally, the integration of this adaptive component with the Kalman filter is described.

**2.1 The Engine Performance Simulation Tool.** One of the master pieces of the gas path analysis approach is a simulation model of the engine. Considering a steady-state operation of the engine, these simulation tools are generally nonlinear aerothermodynamic models based on mass, energy, and momentum conservation laws applied to the engine flow path. Equation (1) represents such an engine model where  $\mathbf{u}_k$  are the variables defining the operating point of the engine (e.g., fuel flow, altitude, and Mach number),  $\mathbf{w}_k$  are the aforementioned health parameters,  $\mathbf{y}_k$  are the gas path measurements, and  $k$  is the discrete time index.

$$\mathbf{y}_k = \mathcal{G}(\mathbf{u}_k, \mathbf{w}_k) \quad (1)$$

Within the frame of turbine engine diagnostics, the quantity of interest is the difference between the actual engine condition and a guessed one,  $\hat{\mathbf{w}}_k^-$ . Assuming a linear relationship between the measurements and the parameters as well as fixed operating conditions, the performance model (Eq. (1)) is reformulated according to Eq. (2).

$$\mathbf{y}_k = \mathbf{G}_k \mathbf{w}_k \quad (2)$$

where

$$\mathbf{G}_k = \frac{\partial}{\partial \mathbf{w}_k} \mathcal{G}(\mathbf{u}_k, \mathbf{w}_k) \Big|_{\mathbf{w}_k = \hat{\mathbf{w}}_k^-} \quad (3)$$

is the Jacobian matrix of the engine model around the health condition  $\hat{\mathbf{w}}_k^-$ . With some abuse of notation,  $\mathbf{y}_k$  and  $\mathbf{w}_k$  are now deviations of the measurements and the health parameters from their values at the linearization point.

**2.2 An Improved Model for Adaptive Estimation.** The core of the diagnosis tool is a Kalman filter that is to be made more responsive to abrupt degradation. The derivation of the Kalman filter (see, for instance, Ref. [13] in the particular scope of engine performance monitoring) implies that the system is formulated as a state-space model. Within the frame of adaptive estimation, it is proposed in Ref. [11] to express the state-space model under the following form:

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mathbf{v}_k + \Delta_{\mathbf{w}} \delta_{\tau,k} \quad (4)$$

$$\mathbf{y}_k = \mathbf{G}_k \mathbf{w}_k + \epsilon_k \quad (5)$$

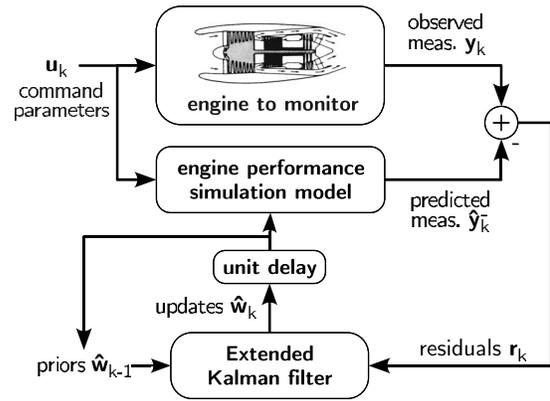
Classical assumptions about the different variables of the model are the following:

- The health parameters  $\mathbf{w}_k$  are considered as Gaussian random variables with initial condition  $\mathbf{w}_0 \in \mathcal{N}(\hat{\mathbf{w}}_0, \mathbf{P}_{\mathbf{w},0})$ .
- The random variable  $\mathbf{v}_k \in \mathcal{N}(0, \mathbf{Q}_k)$  is the so-called process noise that provides the required mobility to track gradual deterioration. It is assumed to be independent of  $\mathbf{w}_k$ .
- The random variable  $\epsilon_k \in \mathcal{N}(0, \mathbf{R}_k)$  is the measurement noise that represents sensor inaccuracies. It is assumed to be uncorrelated with  $\mathbf{v}_k$ .

The last term in Eq. (4) accounts for a possible jump in the health parameters:

- $\Delta_{\mathbf{w}}$  is a vector modeling the jump.
- $\tau$  is a positive integer that represents its time of occurrence.
- $\delta_{i,j}$  is the Kronecker delta operator.

Note that  $\Delta_{\mathbf{w}}$  and  $\tau$  are regarded here as *unknown* parameters and not as random variables, which means that no prior distribution is



**Fig. 1 Health parameter update mechanism using an extended Kalman filter**

attached to them.

The strategy of adaptive estimation comes from viewing Eqs. (4) and (5) according to the following two different hypotheses:

- $H_0$ : no jump up to now ( $\tau > k$ ).
- $H_1$ : a jump has occurred ( $\tau \leq k$ ).

Under assumption  $H_0$  (no jump), the Kalman filter provides an optimal estimation of the health parameters in the least-squares sense. Due to the mild nonlinearity of the system, the extended Kalman filter (EKF) is selected. The basic processing step of the EKF at time  $k$  is reported in algorithm 1 in a pseudocode style.

Line 1 is the prediction step: prior values of the health parameter distribution are generated by means of the transition model (Eq. (4)) under the no-jump assumption. Then, the data are acquired (line 2) and used in building the residuals  $\mathbf{r}_k$ , namely, the difference between the measurements and the model predictions (line 3). The Jacobian matrix is assessed on line 4 and subsequently used in the computation of the covariance matrix of the residuals  $\mathbf{P}_{\mathbf{y},k}$  (line 5) and of the Kalman gain  $\mathbf{K}_k$  (line 6). Loosely speaking, the gain weighs the uncertainty on the parameters versus the uncertainty on the measurements. Finally, the posterior distribution of the health parameters is generated at the correction step (line 7). To complete the description, Fig. 1 sketches the closed-loop structure of the generic diagnosis tool relying on the EKF. The engine performance model, actually embedded in the EKF, is represented outside of it in order to underline its importance in this *model-based* estimation process.

Algorithm 1 Basic step of the extended Kalman filter

- 1:  $\hat{\mathbf{w}}_k^- = \hat{\mathbf{w}}_{k-1}$  and  $\mathbf{P}_{\mathbf{w},k}^- = \mathbf{P}_{\mathbf{w},k-1} + \mathbf{Q}_k$
- 2: acquire  $\mathbf{u}_k$  and  $\mathbf{y}_k$
- 3:  $\mathbf{r}_k = \mathbf{y}_k - \mathcal{G}(\mathbf{u}_k, \hat{\mathbf{w}}_k^-)$
- 4: compute Jacobian matrix as per Eq. (3)
- 5:  $\mathbf{P}_{\mathbf{y},k} = \mathbf{G}_k \mathbf{P}_{\mathbf{w},k}^- \mathbf{G}_k^T + \mathbf{R}_k$
- 6:  $\mathbf{K}_k = \mathbf{P}_{\mathbf{w},k}^- \mathbf{G}_k^T \mathbf{P}_{\mathbf{y},k}^{-1}$
- 7:  $\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_k^- + \mathbf{K}_k \mathbf{r}_k$  and  $\mathbf{P}_{\mathbf{w},k} = (\mathbf{I} - \mathbf{K}_k \mathbf{G}_k) \mathbf{P}_{\mathbf{w},k}^-$

Under assumption  $H_1$ , the residuals  $\mathbf{r}_k$  and the estimated parameters  $\hat{\mathbf{w}}_k$  can be expressed as functions of the jump characteristics  $\tau$  and  $\Delta_{\mathbf{w}}$ . Given the first-order approximation on the measurement equation, the residuals  $\mathbf{r}_k$  can be seen as the sum of two terms:

$$\mathbf{r}_k = \mathbf{r}_{k,H_0} + \mathbf{H}_{k,\tau} \Delta_{\mathbf{w}} \quad (6)$$

where  $\mathbf{r}_{k,H_0}$  are the residuals in the no-jump case, distributed as  $\mathcal{N}(0, \mathbf{P}_{\mathbf{y},k})$  and the second term represents the influence of a jump

$\Delta_{\mathbf{w}}$  that has occurred at time  $\tau$  on the residuals at time  $k$ .

Similarly, we can split the estimated health parameters in a sum of two terms:

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k,H_0} + \mathbf{F}_{k,\tau} \Delta_{\mathbf{w}} \quad (7)$$

where  $\hat{\mathbf{w}}_{k,H_0}$  are the estimated values in the no-jump case and the second term represents the influence of a jump  $\Delta_{\mathbf{w}}$  that has occurred at time  $\tau$  on the estimated values at time  $k$ .

As will be shown in the next section, the detection and estimation of the jump largely rely on the matrices  $\mathbf{F}_{k,\tau}$  and  $\mathbf{H}_{k,\tau}$ . They can be computed from the system models (4) and (5) and the equations of the Kalman filter in Algorithm 1 (see the Appendix for a particular formulation implemented in this study or Ref. [11] for extensive details).

**2.3 Event Detection by Means of the Generalized Likelihood Ratio Test.** A common method to determine which hypothesis between  $H_0$  and  $H_1$  is true is to use the likelihood ratio test (LRT). In short, it is a statistical test in which a ratio is computed between the maximum probability of a result under two different hypotheses, so that a decision can be made between them based on the value of this ratio (see Ref. [14]). To apply the LRT, prior distributions for  $\tau$  and  $\Delta_{\mathbf{w}}$  have to be postulated. Moreover, the test does not directly yield estimates of those quantities. These two issues are clearly drawbacks in the present application. As a workaround, the GLRT, described in Ref. [14], is used. This procedure essentially consists in computing the maximum likelihood estimates  $\hat{\tau}$  and  $\hat{\Delta}_{\mathbf{w}}$ , assuming that  $H_1$  is true, and then in substituting these values into the usual LRT to decide whether a jump has happened.

Practically, the detection at time  $k$  of a hypothetical jump that occurred at time  $\tau \leq k$  relies on the log-likelihood ratio

$$l_{k,\tau} = 2 \log \frac{p(\mathbf{r}_1, \dots, \mathbf{r}_k | H_1, \hat{\Delta}_{\mathbf{w},k})}{p(\mathbf{r}_1, \dots, \mathbf{r}_k | H_0)} \quad (8)$$

where

- $p(\mathbf{r}_1, \dots, \mathbf{r}_k | H_1, \hat{\Delta}_{\mathbf{w},k})$  is the conditional probability of observing the residual history if hypothesis  $H_1$  is true,
- $p(\mathbf{r}_1, \dots, \mathbf{r}_k | H_0)$  is the conditional probability of observing the residual history if hypothesis  $H_0$  is true, and
- $\hat{\Delta}_{\mathbf{w},k}$  is the maximum likelihood estimate of the jump,

$$\hat{\Delta}_{\mathbf{w},k} = \arg \max_{\Delta_{\mathbf{w}}} \{p(\mathbf{r}_1, \dots, \mathbf{r}_k | H_1, \Delta_{\mathbf{w}})\} \quad (9)$$

Since all conditional densities in Eq. (8) are Gaussian, the log-likelihood ratio writes down the following after some algebraic steps:

$$l_{k,\tau} = \mathbf{d}_{k,\tau}^T \mathbf{C}_{k,\tau}^{-1} \mathbf{d}_{k,\tau} \quad (10)$$

where matrix  $\mathbf{C}_{k,\tau}$  is deterministic and vector  $\mathbf{d}_{k,\tau}$  is a linear combination of the residuals,

$$\mathbf{C}_{k,\tau} = \sum_{j=\tau}^k \mathbf{H}_{j,\tau}^T \mathbf{P}_{j,\tau}^{-1} \mathbf{H}_{j,\tau} \quad (11)$$

$$\mathbf{d}_{k,\tau} = \sum_{j=\tau}^k \mathbf{H}_{j,\tau}^T \mathbf{P}_{j,\tau}^{-1} \mathbf{r}_j \quad (12)$$

Equations (11) and (12) show that the likelihood ratio (Eq. (10)) actually implements a matched filter, i.e., a correlation test between the variations in the residuals and the signature of a jump, represented by  $\mathbf{H}_{k,\tau}$ .

At each time step  $k$ , the likelihood ratio  $l_{k,\tau}$  is computed for  $\tau = 1, \dots, k$  to perform the jump detection. Hence, the original implementation of the GLRT algorithm involves a growing bank

of matched filters. To keep the problem tractable, the jump detection is restricted to a sliding data window of width  $M$ . If the window is sufficiently wide to ensure detection of all important faults, this approximation does not lead to serious difficulties.

So, for  $k-M < \tau \leq k$ , the likelihood ratio  $l_{k,\tau}$  is computed. The value  $\hat{\tau}$  that maximizes  $l_{k,\tau}$  represents the most likely time at which a jump occurred during the last  $M$  time steps. The decision rule to choose between  $H_0$  and  $H_1$  is

$$l_{k,\hat{\tau}} \underset{H_0}{\overset{H_1}{\geq}} \eta \quad (13)$$

where  $\eta$  is a threshold value chosen to provide a reasonable trade-off between false alarms and detection delay.

If hypothesis  $H_1$  is verified at time step  $k$ , a jump has occurred at estimated time  $\hat{\tau}$ , and its maximum likelihood estimate is given by

$$\hat{\Delta}_{\mathbf{w},k} = \mathbf{C}_{k,\hat{\tau}}^{-1} \mathbf{d}_{k,\hat{\tau}} \quad (14)$$

The latter relation provides a least-squares estimate of the jump  $\Delta_{\mathbf{w}}$ , assuming that  $\tau$  is known and that no prior information is available about the value of the jump. In that case,  $\mathbf{C}_{k,\hat{\tau}}^{-1}$  is the error covariance of the estimate  $\hat{\Delta}_{\mathbf{w}}$ .

From a computational standpoint, it is interesting to note that recursive relations can be derived for the quantities involved in the GLRT when the detection scheme is limited to a buffer containing the latest  $M$  data points. These relations are provided in the Appendix.

#### 2.4 Practical Implementation of the Adaptive Algorithm.

Once a jump is detected by the GLRT algorithm, the maximum likelihood estimates  $\hat{\tau}$  and  $\hat{\Delta}_{\mathbf{w},k}$  can be used directly to increment the parameters estimated by the Kalman filter,

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k,KF} + \underbrace{(\mathbf{I} - \mathbf{F}_{k,\hat{\tau}}) \hat{\Delta}_{\mathbf{w},k}}_{\delta \hat{\mathbf{w}}_k} \quad (15)$$

where  $\hat{\mathbf{w}}_{k,KF}$  are the Kalman filter estimates,  $\mathbf{I} \hat{\Delta}_{\mathbf{w},k}$  is the contribution to the parameters of a jump that occurs at  $\hat{\tau}$ , and  $\mathbf{F}_{k,\hat{\tau}} \hat{\Delta}_{\mathbf{w},k}$  represents the response of the Kalman filter to the jump prior to its detection.

In some cases, the estimate of the jump may be quite inaccurate. To reflect this degradation in the quality of the estimate, caused by the jump, it is advised in Ref. [15] to increase the covariance matrix of the health parameters accordingly. This rise in parameter covariance results in an increased Kalman gain, i.e., an increased bandwidth. The filter can improve its response to the jump and hence compensate for inaccuracies in  $\hat{\tau}$  and  $\hat{\Delta}_{\mathbf{w},k}$ . The covariance modification is done through

$$\mathbf{P}_{\mathbf{w},k} = \mathbf{P}_{\mathbf{w},k,KF} + \underbrace{(\mathbf{I} - \mathbf{F}_{k,\hat{\tau}})^T \mathbf{C}_{k,\hat{\tau}}^{-1} (\mathbf{I} - \mathbf{F}_{k,\hat{\tau}})}_{\delta \mathbf{P}_{\mathbf{w},k}} \quad (16)$$

Engine performance monitoring generally suffers from negative redundancy; it means that the number of health parameters,  $n$ , outweighs the number of measurements,  $m$ . As a consequence, the system is only partially observable and matrix  $\mathbf{C}_{k,\tau}$  is singular. In that case, the pseudoinverse [16] usefully replaces the common inverse. The possible jump directions are then restricted to the observable subspace of the parameter space (see Ref. [17] for further details). Another possibility is to estimate the jump with a Bayesian approach. The introduction of prior knowledge about the jump focuses the search for a solution in a neighborhood of the a priori values. Note that this solution could also be used for fully observable systems in the case where some a priori information about the possible range of values for  $\Delta_{\mathbf{w}}$  is available.

In the implemented GLRT system, the buffer of residuals  $\mathbf{d}_{k,\tau}$  is reset immediately after each detection. This is done in order to

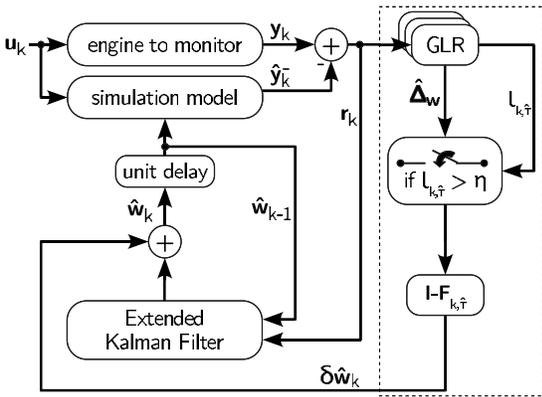


Fig. 2 Integration of the adaptive component with the Kalman filter

avoid possible instabilities due to multiple detections of the same jump, which would lead to “overcompensation.” With this modification, the GLRT system can detect several successive jumps.

Two parameters are available to tune the GLRT system: the width  $M$  of the sliding window and the threshold  $\eta$  in the hypothesis testing (Eq. (13)). The selection of  $M$  is dictated by a trade-off between accurate jump estimation and minimum detection delay. The former suggests choosing  $M$  large enough to minimize the uncertainty  $\mathbf{C}_{k,\tau}^{-1}$  in the estimated jump, while the latter advocates a small-sized buffer. On the other hand, the threshold  $\eta$  is directly related to the probability of false alarms  $P_F$  in jump detection,

$$P_F = \int_{\eta}^{\infty} p(l=L|H_0)dL \quad (17)$$

where  $p(l=L|H_0)$  is the probability density of  $l_{k,\tau}$  conditioned on  $H_0$ , which is a chi-squared density with  $n$  degrees of freedom (see Ref. [14] for a proof). Specifying an allowable false alarm rate (small, but nonzero) gives the threshold value.

Figure 2 depicts the integration of the adaptive component, which comprises all the elements in the dashed box, with the Kalman filter. For the sake of clarity, only the most relevant data streams are sketched in the diagram. Briefly explained, the adaptive component works as follows:

- The GLR box performs the GLRT in three stages:
  1. update of the quantities in the  $M$ -sized buffer according to the relations given in the Appendix
  2. computation of the likelihood ratio by means of Eq. (10)
  3. search for the maximum value of the likelihood ratio

The outputs of the GLR box are the estimated time of occurrence of the jump  $\hat{\tau}$  (not represented), the estimated jump  $\hat{\Delta}_{\mathbf{w},k}$ , and the maximum value of the likelihood ratio in the buffer  $l_{k,\hat{\tau}}$ .

- This value is compared with the threshold  $\eta$  in order to determine whether a jump has occurred.
- In case hypothesis  $H_1$  is true, the increment in the parameters  $\delta\hat{\mathbf{w}}_k$  due to the jump is computed and added to the Kalman filter estimate, as stated in Eq. (15). At the same time, the covariance matrix of the health parameters is also modified according to Eq. (16).

From a computational viewpoint, the aforementioned operations performed in the adaptive components are roughly equivalent to  $M+1$  runs of a linear Kalman filter (i.e., no call to the nonlinear engine model) per time step. Yet, the most demanding part of the whole adaptive algorithm lies in the evaluation of the Jacobian matrix  $\mathbf{G}_k$  and in the prediction of the measurements,

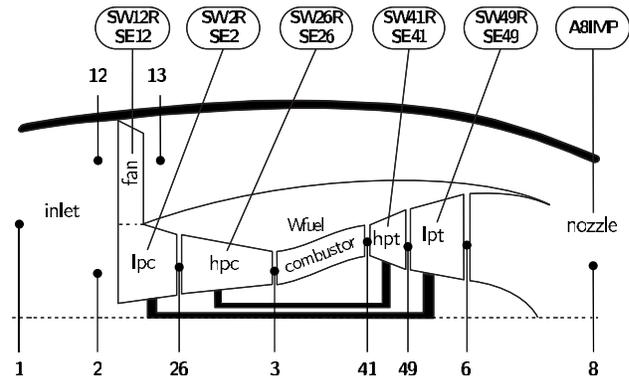


Fig. 3 Turbofan layout with station numbering and health parameter location

which involves, in the case of centered differences,  $2n+1$  calls to the nonlinear engine model. As a result, the overhead in CPU time needed by the adaptive algorithm is rather limited.

### 3 Application of the Method

**3.1 Engine Layout.** The application used as a test-case is a high bypass ratio, mixed-flow turbofan. The engine performance model has been developed within the frame of the OBIDICOTE<sup>2</sup> project and is detailed in Ref. [18]. A schematic of the engine is sketched in Fig. 3 where the location of the 11 health parameters and the station numbering are also indicated. One command variable, which is the fuel flow rate fed in the combustor, is considered here.

The sensor suite selected for diagnosing the engine condition is representative of the instrumentation available on-board contemporary turbofan engines and is detailed in Table 1 where the nominal accuracy of each sensor is also reported.

**3.2 Definition of the Test-Cases.** Simulated data have been generated to examine how the adaptive algorithm performs. Cruise conditions (Alt=10,668 m, Mach=0.8, and  $\Delta T_{ISA}=0$  K) are assumed. The flight sequence is 5000 s long, and the sampling rate is set to 2 Hz. Gaussian noise, whose magnitude is specified in Table 1, is added to the clean simulated measurements to make them representative of real data.

Engine wear due to normal operation is simulated by linearly drifting values of nearly all health parameters, starting from a healthy engine (all parameters at their nominal values) at  $t=0$  s and with the following degradation at the end of the sequence ( $t=5000$  s):  $-0.5\%$  on SW12R,  $-0.5\%$  on SE12,  $-0.4\%$  on SW2R,  $-0.5\%$  on SE2,  $-1.0\%$  on SW26R,  $-0.7\%$  on SE26,  $+0.4\%$  on SW41R,  $-0.8\%$  on SE41, and  $-0.5\%$  on SE49. To demonstrate the improvements achieved with the adaptive estimation algorithm, a number of fault cases, summarized in Table 2, are superimposed one at a time to the global performance deterioration. While far from being exhaustive, these cases are representative of typical accidental component faults that can be expected on present turbofans and are added as a step change to engine wear at  $t=2500$  s. This library of degradations (both distributed and localized) has been devised within the frame of the OBIDICOTE project too (see Ref. [19]) and has been used in a number of studies.

**3.3 Definition of a Figure of Merit.** The quality of the estimation performed by the generic diagnosis tool and the adaptive

<sup>2</sup>A Brite/Euram project for on-board identification, diagnosis, and control of turbofan engines.

**Table 1 Selected sensor suite (uncertainty is three times the standard deviation)**

Label	Uncertainty	Label	Uncertainty
$p_{13}^0$	$\pm 100$ Pa	$T_{13}^0$	$\pm 2$ K
$p_3^0$	$\pm 5000$ Pa	$T_3^0$	$\pm 2$ K
$N_{lp}$	$\pm 6$ rpm	$N_{hp}$	$\pm 12$ rpm
$T_6^0$	$\pm 2$ K		

one is assessed in terms of the maximum root mean square error (RMSE) over the last 50 s of the sequence (i.e., over 100 samples),

$$RMSE = \max \left( \sqrt{\frac{1}{100} \sum_{t=4950}^{5000} \left( \frac{\mathbf{w}_t - \hat{\mathbf{w}}_t}{\mathbf{w}^{hl}} \right)^2} \right) \quad (18)$$

where  $\mathbf{w}^{hl}$  are the nominal values of the health parameters.

Given the stochastic character of the measurement noise, each test-case has been run 20 times. The RMSEs reported in Table 3 are the average values over the 20 runs in order to guarantee that they are statistically representative. A test-case characterized by an averaged maximum RMSE below 0.25% is declared successful, which is indicated by a check mark. This threshold corresponds to roughly three times the standard deviation of the identified health parameters (i.e., the square root of the diagonal terms of the covariance matrix  $\mathbf{P}_{\mathbf{w},k}$ ).

**Table 2 List of abrupt component faults**

Label	Definition of the fault		Faulty component
<b>a</b>	-1% on SW12R	-0.5% on SE12	Fan
	-0.7% on SW2R	-0.4% on SE2	lpc
<b>b</b>	-1% on SE12		
<b>c</b>	-1% on SW26R	-0.7% on SE26	hpc
<b>d</b>	-1% on SE26		
<b>e</b>	-1% on SW26R		
<b>f</b>	+1% on SW41R		hpt
<b>g</b>	-1% on SW41R	-1% on SE41	
<b>h</b>	-1% on SE41		
<b>i</b>	-1% on SE49		lpt
<b>j</b>	-1% on SW49R	-0.4% on SE49	
<b>k</b>	-1% on SW49R		
<b>l</b>	+1% on SW49R	-0.6% on SE49	
<b>m</b>	+1% on A8IMP		Nozzle
<b>n</b>	-1% on A8IMP		

**Table 3 Comparison of the figure of merit obtained with the generic EKF and with the adaptive algorithm (EKF+GLR)**

Case	EKF		EKF+GLR	
Wear ( <b>w</b> )	0.09%	✓	0.10%	✓
<b>w+a</b>	0.50%	-	0.12%	✓
<b>w+b</b>	0.27%	-	0.09%	✓
<b>w+c</b>	0.20%	✓	0.11%	✓
<b>w+d</b>	0.18%	✓	0.13%	✓
<b>w+e</b>	0.32%	-	0.12%	✓
<b>w+f</b>	0.21%	✓	0.09%	✓
<b>w+g</b>	0.43%	-	0.11%	✓
<b>w+h</b>	0.23%	✓	0.10%	✓
<b>w+i</b>	0.20%	✓	0.09%	✓
<b>w+j</b>	0.53%	-	0.57%	-
<b>w+k</b>	0.45%	-	0.09%	✓
<b>w+l</b>	0.40%	-	0.09%	✓
<b>w+m</b>	0.16%	✓	0.11%	✓
<b>w+n</b>	0.34%	-	0.15%	✓

## 4 Results

Both the generic and the adaptive diagnosis tools have been run on the previously described test-cases. The tuning parameters of the adaptive tool have been set to the following values: a sliding window of width  $M=10$  time steps (i.e., 5 s) and a probability of false alarm  $P_F=10^{-4}\%$ . These settings were found satisfactory for the level of noise and the magnitude of the abrupt faults considered in the reported application.

Figure 4 depicts the tracking of engine wear with the extended Kalman filter. It can be seen that the identified values are in good agreement with the true ones, which is confirmed by a RMSE of 0.09%. This figure underlines the fact that the Kalman filter is able to filter the measurement noise and to give an accurate assessment of the level of deterioration of the gas turbine, provided that the evolution of the health parameters is “sufficiently” smooth.

The tracking of the engine condition with the EKF in the case where accidental fault **a** (fan+lpc) is superimposed to wear is sketched in Fig. 5. It can readily be seen that the Kalman filter is unable to follow the abrupt drift in the performance of the fan and the lpc. This fast variation is indeed not accounted for in the transition model of the health parameters that is blended in the Kalman filter. As a result, the fault is spread over multiple components such as the hpc (SW26R) and the lpt (SE49). Moreover, it can be noticed that several hundreds of seconds are required by the filter to (erroneously) recognize the abrupt fault.

When the adaptive algorithm is applied to the same fault case (wear+**a**), the identification of the health parameters is as shown in Fig. 6. The improvement with respect to the results discussed previously is obvious. First, the abrupt fault is localized on the fan and the lpc, and its magnitude is quite fairly assessed, the drop in SW2R being slightly overestimated. Second, the responsiveness for recognizing the fault has dramatically improved. The examination of Table 4 will show that the accidental event is detected and recognized a few seconds after it has happened. These two points greatly enhance the relevance of the results and provide a detailed insight on the temporal evolution of the engine condition.

Table 3 reports the figure of merit defined by Eq. (18) for the different fault-cases that have been considered in the present study. The first line is related to the case of engine wear (abbreviated **w** in the subsequent lines). For that case of long-time-scale deterioration, the performance of both diagnosis tools is essentially the same. Indeed, as long as the adaptive component does not issue any detection flag, the adaptive algorithm will be confounded with the generic one. This can be clearly seen from the schematic in Fig. 2.

Considering the cases involving the global drift in engine performance and one of the 14 component faults defined in Table 2, it can be seen that the EKF achieves a reasonable identification for a mere 6 cases out of 14. Furthermore, the figure of merit is a measure of the estimation accuracy at the end of the sequence but does not reflect the delay in fault recognition, which has been shown to be quite long from the analysis of Fig. 5.

On the other hand, the adaptive tool succeeds in solving all test-cases but case **w+j**. The figure of merit is roughly the same as for the case of pure engine wear (first line of the table), which hints at the capability of the adaptive component to estimate quite accurately the jump in engine health. The issue of timely recognition of the fault will be discussed in the next paragraph. The misdiagnosis of case **w+j** is due to a lack of observability of the health parameters of the turbines—more specifically the triplet SE41, SE49, and SW49R—with the sensor configuration used here. Figure 7 depicts the evolution of the health parameters of the hot section estimated with the adaptive algorithm. It can be noticed that the fault on SW49R is largely underestimated, while the efficiency drop (SE49) is simply not detected. On top of that, a false alarm on SE41 is issued, which leads to a wrong interpretation of the actual health condition of the engine.

To conclude the analysis of the result, it is worth taking a look

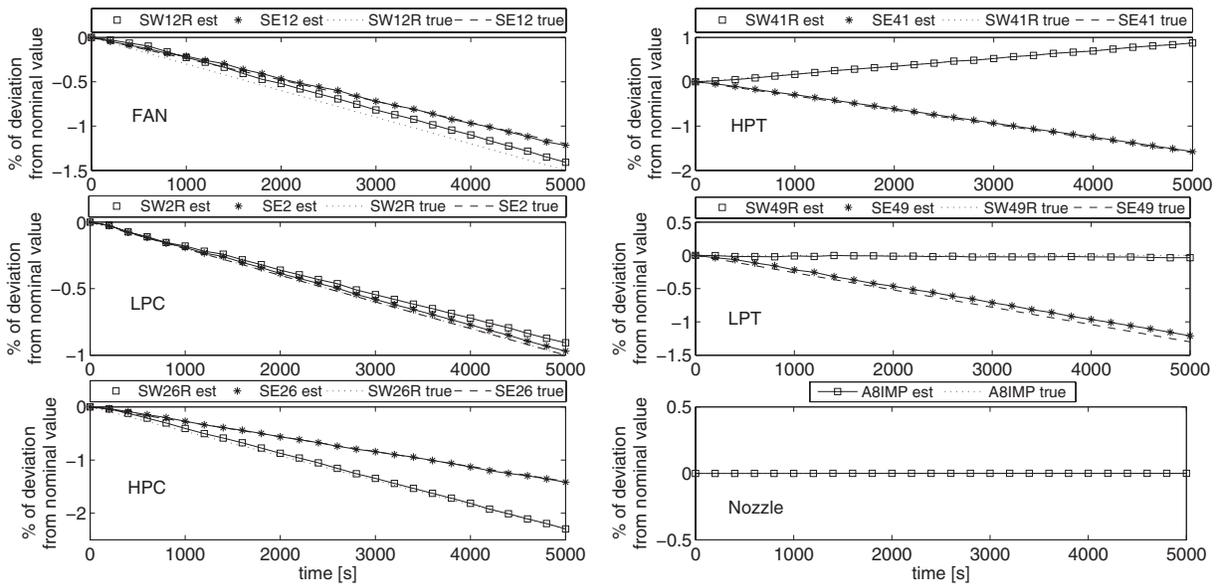


Fig. 4 Tracking of engine wear with the generic diagnosis tool. In the legends, “est” stands for estimated.

at the delay needed by the adaptive algorithm to recognize the fault. Recall that the adaptive component gives access to the time of detection of the jump,  $k_{det}$ , defined as the time at which the maximum likelihood ratio in the buffer exceeds the threshold, and to the estimated time of occurrence of the jump,  $\hat{\tau}$ , which is the time corresponding to the maximum value of the likelihood ratio in the buffer.

The middle column of Table 4 lists the delay between the occurrence of the fault  $\tau_{true}$  and its detection  $k_{det}$  by the GLRT algorithm expressed in number of data samples, which is a more relevant quantity than the detection time itself. It can be noticed that the delay between the true time of occurrence of the fault (which is  $\tau_{true}=2500.5$  s) and the time of detection amounts to a couple of samples. Most of the faults are detected four time steps after they appear. For fault cases  $\mathbf{w}+\mathbf{b}$  and  $\mathbf{w}+\mathbf{j}$ , the delay is slightly longer. This is due to the fact that the signature of these faults on the measurements is somewhat weaker for the other cases. On the contrary, the delay for the hpt faults is a bit shorter since those

faults cause a deeper variation in the measurements. As soon as the jump is detected, it is also estimated and subsequently added to the Kalman filter estimate in order to compensate for the rapid drift in the engine health. The estimation of the jump relies on a residual averaged over a few time steps, which provides some attenuation of the measurement noise. Looking at the right column of Table 4, one can notice that the time of occurrence of the jump is fairly well assessed. This probably constitutes a piece of information of secondary matter, but which might be helpful in some cases.

Given the fast detection of the jump, only a fraction of the buffer is actually used. One can thus think about shortening the width of the sliding window to spare memory space and computational time. Yet, the rapid detection of the jump is due to two causes. First, the abrupt fault has been introduced as a pure step (i.e., the full fault occurs in one time step); most rapid deterioration happen, however, over a few seconds or so rather than instantly. Second, we consider a static system so that a modification

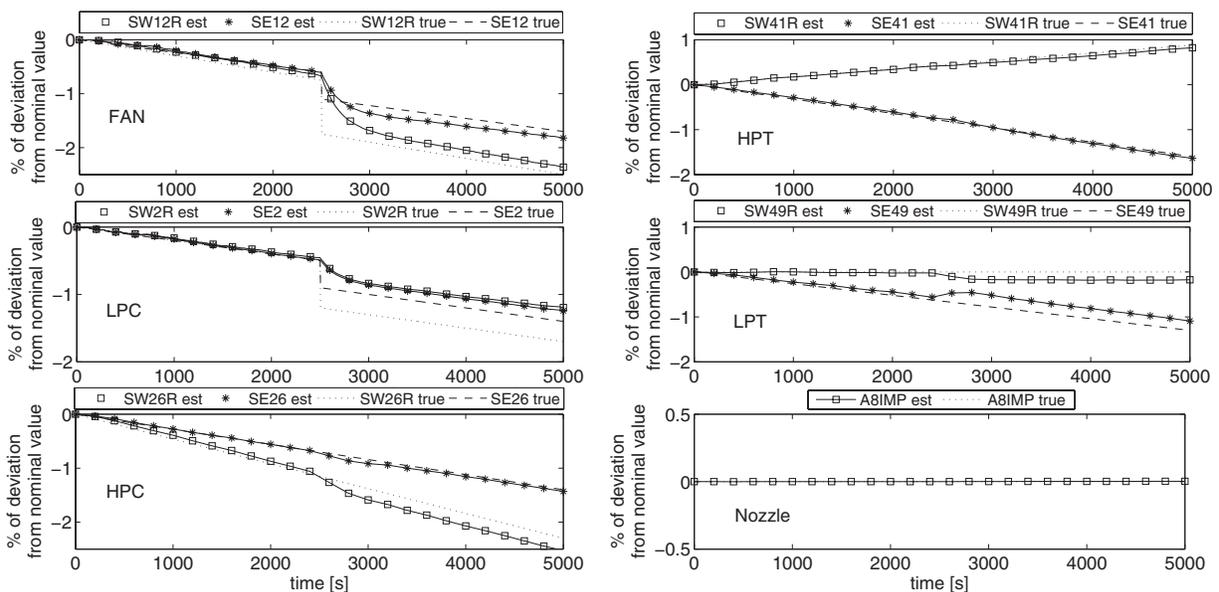


Fig. 5 Tracking of engine wear+fault a with the generic diagnosis tool. In the legends, “est” stands for estimated.

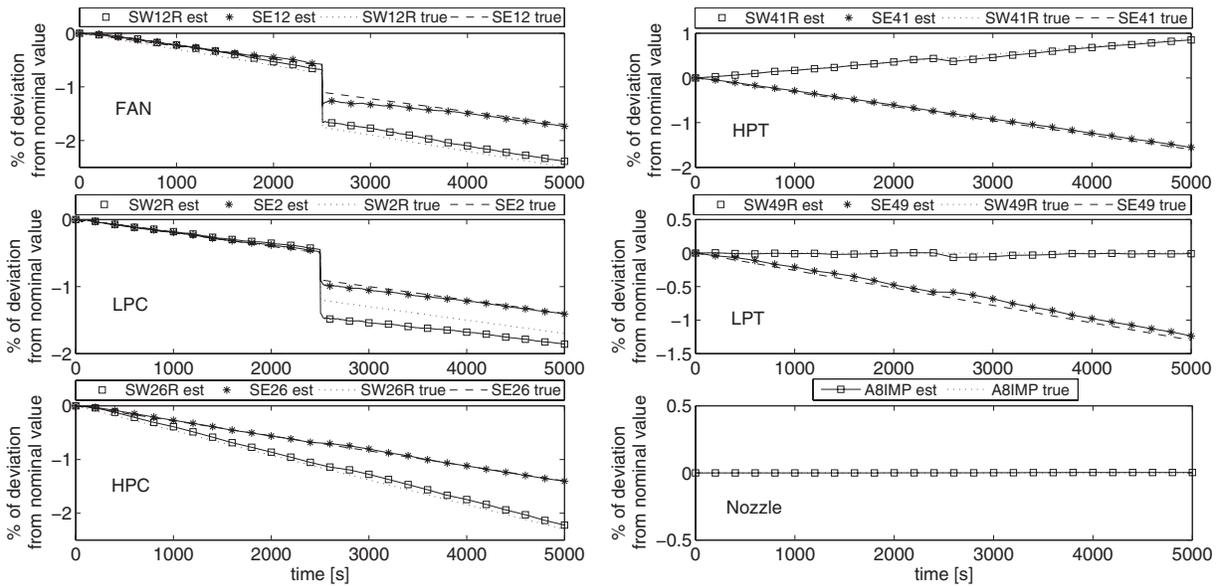


Fig. 6 Tracking of engine wear+fault a with the adaptive algorithm. In the legends, “est” stands for estimated.

in the state of the system instantaneously shows up in the measurements. The length of the buffer should be carefully selected with these two parameters in mind.

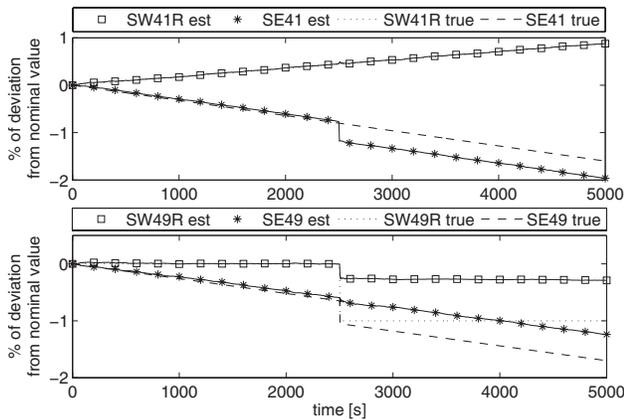


Fig. 7 Tracking of engine wear+fault j with the adaptive algorithm. In the legend, “est” stands for estimated.

Table 4 Detection delay  $k_{det} - \tau_{true}$  and estimated time of jump  $\hat{\tau}_{k_{det}}$ . The true time of jump is  $\tau_{true} = 2500.5$  s.

Case	$k_{det} - \tau_{true}$	$\hat{\tau}_{k_{det}}$
w+a	4	2500.5
w+b	9	2501.5
w+c	4	2500.5
w+d	4	2500.5
w+e	4	2500.5
w+f	2	2500.5
w+g	2	2500.5
w+h	2	2500.5
w+i	4	2500.5
w+j	6	2501
w+k	4	2500.5
w+l	4	2500.5
w+m	4	2500.5
w+n	4	2500.5

## 5 Discussion

The enhancements in fault recognition and estimation accuracy brought by the adaptive algorithm have been pointed out in the previous section. To complete the analysis of the results, additional issues that may lead to further developments of the proposed methodology are discussed in the following.

The first point concerns the estimation of the jump  $\Delta_w$  in the health parameters. In the present implementation of the adaptive component,  $\Delta_w$  is assumed to be totally free. It would be interesting to replace this unconstrained estimate with a selection of the most likely fault in a bank of predefined events, as in Ref. [5]. The modifications to the GLRT that need to be carried out in that case are described in Ref. [11]. However, this solution requires a great expertise for the definition of the fault library. In turn, it may be restricted to engine manufacturers. Moreover this implementation requires additional fault recognition logic for handling novel fault modes.

A second interesting development of the adaptive diagnosis tool consists in extending the fault coverage to system (and possibly sensor) malfunctions such as stuck bleed valves or mistuned variable stator vanes. These faults typically occur in an abrupt manner so that they could be kept constant in the Kalman filter (by assuming perfect knowledge about them, i.e., a tiny covariance) and be estimated only by the adaptive component in case a jump in the health condition is detected. Sensor biases could be dealt with in a similar fashion.

Information fusion comes to mind as a third axis for future research. Indeed, the adaptive component allows detection and estimation of a rapid deterioration. Once a jump is detected, another diagnosis tool could isolate the fault so that its result is fused with the fault estimated by the GLRT. Information fusion between model-based and data-driven diagnosis tools has already been investigated by the authors in previous publications [20,21] and has shown potential improvement in the assessment of the engine condition.

## 6 Conclusion

In this contribution, an adaptive algorithm for engine health monitoring has been developed. It combines a Kalman filter, which provides an accurate estimation of the health condition for long-time-scale deterioration (such as engine wear), and an adaptive component, which monitors the residuals and looks for abrupt

changes in the health condition. The adaptive component implements a generalized likelihood ratio test that detects and estimates rapid variations in the engine condition.

The advantages of the proposed approach have been pointed out. First, the performance of the Kalman filter is not degraded for slow evolutions of the engine condition. Second, the adaptive component acts as a fault compensator for the Kalman filter and can handle a variety of fault events. Third, the adaptive component is easily integrated with the generic diagnosis tool, while the increase in computational burden remains limited (about 15% in the present implementation). The improvement in estimation accuracy and fault recognition achieved by the adaptive diagnosis tool with respect to the generic one has been illustrated for a series of typical component fault cases that may occur on a turbofan engine.

## Nomenclature

$\hat{\phantom{x}}$	= estimated value
A8IMP	= nozzle exit area (nominal value: 1.4147 m <sup>2</sup> )
hpc	= high pressure compressor
hpt	= high pressure turbine
$k$	= discrete time index
lpc	= low pressure compressor
lpt	= low pressure turbine
$N$	= rotational speed
$p_i^0$	= total pressure at station $i$
SE $_i$	= efficiency factor of the component whose entry is located at section $i$ (nominal value: 1.0)
SW $_i$	= flow capacity factor of the component whose entry is located at section $i$ (nominal value: 1.0)
$T_i^0$	= total temperature at station $i$
$\mathbf{u}_k$	= actual command parameters
$\mathbf{w}_k$	= actual but unknown health parameters
$\mathbf{y}_k$	= observed measurements
$\Delta\mathbf{w}$	= unknown abrupt fault
$\boldsymbol{\epsilon}_k$	= measurement noise vector
$\mathbf{v}_k$	= process noise vector
$\tau$	= unknown time of occurrence of the abrupt fault
$\mathcal{N}(\mathbf{m}, \mathbf{R})$	= Gaussian probability density function with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$

## Appendix: Recursive Relations for Computation of the GLRT

Jump detection by means of the GLRT requires that the likelihood ratio  $l_{k,\tau}$  (Eq. (10)) be evaluated for all  $k-M < \tau \leq k$ . The necessary quantities for this calculation can be computed from the following equations:

$$\forall \tau \in k-M, k[$$

$$\mathbf{C}_{k,\tau} = \mathbf{C}_{k-1,\tau} + \mathbf{H}_{k,\tau}^T \mathbf{P}_{k,\tau}^{-1} \mathbf{H}_{k,\tau} \quad (\text{A1})$$

$$\mathbf{d}_{k,\tau} = \mathbf{d}_{k-1,\tau} + \mathbf{H}_{k,\tau}^T \mathbf{P}_{k,\tau}^{-1} \mathbf{r}_k \quad (\text{A2})$$

$$\mathbf{H}_{k,\tau} = \mathbf{G}_k (\mathbf{I} - \mathbf{F}_{k-1,\tau}) \quad (\text{A3})$$

$$\mathbf{F}_{k,\tau} = \mathbf{F}_{k-1,\tau} + \mathbf{K}_k \mathbf{H}_{k,\tau} \quad (\text{A4})$$

where  $\mathbf{I}$  is the identity matrix but also the transition matrix of the health parameters in the application at hand.

For the most recent element in the buffer ( $\tau=k$ ), the quantities are initialized according to

$$\mathbf{C}_{k,k} = \mathbf{G}_k^T \mathbf{P}_{y,k}^{-1} \mathbf{G}_k \quad (\text{A5})$$

$$\mathbf{d}_{k,k} = \mathbf{G}_k^T \mathbf{P}_{y,k}^{-1} \mathbf{r}_k \quad (\text{A6})$$

$$\mathbf{F}_{k,k} = \mathbf{K}_k \mathbf{G}_k \quad (\text{A7})$$

## References

- [1] Volponi, A. J., 2003, "Foundation of Gas Path Analysis (Part I and II)," *Gas Turbine Condition Monitoring and Fault Diagnosis*, von Karman Institute Lecture Series No. 1, von Karman Institute, Rhode-Saint-Genèse, Belgium.
- [2] Li, Y. G., 2002, "Performance-Analysis-Based Gas Turbine Diagnostics: A Review," *Proc. Inst. Mech. Eng., Part A*, **216**(5), pp. 363–377.
- [3] Kalman, R. E., 1960, "A New Approach to Linear Filtering and Prediction Problems," *ASME J. Basic Eng.*, **82**, pp. 35–44.
- [4] Provost, M. J., 1994, "The Use of Optimal Estimation Techniques in the Analysis of Gas Turbines," Ph.D. thesis, Cranfield University.
- [5] Volponi, A. J., 2003, "Extending Gas Path Analysis Coverage for Other Fault Conditions," *Gas Turbine Condition Monitoring and Fault Diagnosis*, von Karman Institute Lecture Series No. 1, von Karman Institute, Rhode-Saint-Genèse, Belgium.
- [6] Mehra, R. K., 1972, "Approaches to Adaptive Filtering," *IEEE Trans. Autom. Control*, **17**(5), pp. 693–698.
- [7] Mehra, R. K., 1970, "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Trans. Autom. Control*, **15**(2), pp. 175–184.
- [8] Odelson, B., Rajamani, M., and Rawlings, J., 2006, "A New Autocovariance Least-Squares Method for Estimating Noise Covariances," *Automatica*, **42**, pp. 303–308.
- [9] Jazwinski, A. H., 1969, "Adaptive Filtering," *Automatica*, **5**, pp. 475–485.
- [10] Dewallef, P., Léonard, O., and Borguet, S., 2006, "An Adaptive Estimation Algorithm for Aircraft Engine Performance Monitoring," University of Liège, Technical Report No. 06-02.
- [11] Willsky, A., and Jones, H., 1974, "A Generalized Likelihood Ratio Approach to State Estimation in Linear Systems Subject to Abrupt Changes," *Proceedings of the IEEE Conference on Decision and Control*, Phoenix, AZ, November, pp. 846–853.
- [12] Bensalah, F., and Chaumette, F., 1994, "Real-Time Visual Tracking Using Abrupt Changes Detection," INRIA, Technical Report No. 2425.
- [13] Dewallef, P., 2005, "Application of the Kalman Filter to Health Monitoring of Gas Turbine Engines: A Sequential Approach to Robust Diagnosis," Ph.D. thesis, University of Liège.
- [14] van Trees, H. L., 1968, *Detection, Estimation and Modulation Theory*, Wiley, New York.
- [15] Willsky, A., and Jones, H., 1976, "A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems," *IEEE Trans. Autom. Control*, **21**(1), pp. 108–112.
- [16] Strang, G., 1988, *Linear Algebra and its Applications*, 3rd ed., Brooks-Cole, Pacific Grove, CA.
- [17] Borguet, S., and Léonard, O., 2008, "A Study on Sensor Selection for Efficient Jet Engine Health Monitoring," *Proceedings of the 12th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu, HI, February, Paper No. ISROMAC-2008-20072.
- [18] Stamatis, A., Mathioudakis, K., Ruiz, J., and Curnock, B., 2001, "Real-Time Engine Model Implementation for Adaptive Control and Performance Monitoring of Large Civil Turbofans," *ASME Paper No. 2001-GT-0362*.
- [19] Curnock, B., 2000, "Obidicote Project—Work Package 4: Steady-State Test Cases," Rolls-Royce PLC, Technical Report No. DNS62433.
- [20] Dewallef, P., Romessis, C., Léonard, O., and Mathioudakis, K., 2006, "Combining Classification Techniques With Kalman Filters for Aircraft Engine Diagnostics," *ASME J. Eng. Gas Turbines Power*, **128**(2), pp. 281–287.
- [21] Borguet, S., and Léonard, O., 2007, "Coupling Principal Component Analysis and Kalman Filter Algorithms for On-Line Aircraft Engine Diagnostics," *Proceedings of the 18th International Symposium on Air Breathing Engines*, Beijing, China, Sept., Paper No. ISABE-2007-1275.