Proceedings of ASME Turbo Expo 2011
GT2011
June 6-10, 2011, Vancouver, British Columbia, Canada

GT2011-45710

APPLICATION OF FORECASTING METHODOLOGIES TO PREDICT GAS TURBINE BEHAVIOR OVER TIME

A. Cavarzere, M. Venturini
Dipartimento di Ingegneria, Università degli Studi di Ferrara
Via G. Saragat, 1 - 44122 Ferrara, Italy

ABSTRACT

The growing need to increase the competitiveness of industrial systems continuously requires a reduction of maintenance costs, without compromising safe plant operation. Therefore, forecasting the future behavior of a system allows planning maintenance actions and saving costs, because unexpected stops can be avoided.

In this paper, four different methodologies are applied to predict gas turbine behavior over time: Linear and Non Linear Regression, One Parameter Double Exponential Smoothing, Baesyan Forecasting Method and Kalman Filter. The four methodologies are used to provide a prediction of the time when a performance limit will be exceeded in the future, as a function of the current trend of the considered parameter. The application considers different scenarios which may be representative of the trend over time of some significant parameters for gas turbines. Moreover, the Baesyan Forecasting Method, which allows the detection of discontinuities in time series, is also tested for predicting system behavior after two consecutive trends.

The results presented in this paper aim to select the most suitable methodology that allows both trending and forecasting as a function of data trend over time, in order to predict time evolution of gas turbine characteristic parameters and to provide an estimate of the occurrence of a failure.

NOMENCLATURE

A_1, A_2, A_3	matrices in KM
E	error
h	shift parameter in BFM
H	Bayes factor
m_1,m_2	space dimensions
n	number
RMSE	Root Mean Square Error
S	smoothed statistics in OPDES
t	time
U	optional control input in KM
w	process noise in KM
X	unknown state
Y	non-dimensional state parameter
Z	measurement in KM
α	smoothing constant in OPDES
β	model parameter in Regression Method
ε	error term in Regression Method
σ^2	variance of simulated data trends
λ	uncertainty limit in BFM
υ	measurement variance in BFM
ω_1,ω_2	variances in value and gradient in BFM
ξ	measurement noise in KM
Δt	time frame of the trend
ΔY	variation of Y over the time frame Δt

Subscripts and Superscripts

av	average
grad	gradient
meas	measurement
meth	methodology
min	minimum
p	prediction
t	time
T	trend
u	measurement uncertainty
v	value
Δt	time frame of the trend
[2]	double

Acronyms

BFM	Bayesian Forecasting Method
DLM	Dynamic Linear Model
KM	Kalman Method
OPDES	One Parameter Double Exponential Smoothing
SLRM	Simple Linear Regression Method
SNLRM	Simple Non Linear Regression Method

INTRODUCTION

The growing need to increase the competitiveness of industrial systems continuously requires a reduction of maintenance costs, without compromising safe plant operation. Therefore, forecasting the future behavior of a system allows planning maintenance actions and saving costs, because unexpected stops can be avoided. More generally, the prediction of future events can provide key information to the decision-making process, especially in today's ever competitive energy market.

Nowadays, most gas turbines in operation are equipped with monitoring and/or diagnostic tools [1-5]. In fact, it is common practice that the most significant operational parameters (e.g. overall gas turbine efficiency or exhaust gas temperature) are always monitored, and in many cases raw measurements are also processed by diagnostic tools to provide a further insight into the engine health state. For instance, Gas Path Analysis techniques [6-10] use gas turbine field measurements to determine the actual values of the parameters which are indices of the gas turbine health state, such as efficiencies, characteristic flow passage areas and pressure drops along the gas path. Such indices, usually called *health indices*, allow both the faulty component to be localized and the malfunction to be identified and quantified.

Once the trend over time of raw measurements or health indices is available, linear trending is one of the prognostic methods commonly applied in field operations, as reported for instance in [11,12]. However, three issues have to be considered. First, it is reasonable to assume that failure evolution over time is linear only in a short term period, since the failure rate is not constant, as it likely tends to occur during initial operation. For instance, the failure rate in case of fouling roughly follows an exponential law [13]. Second, the methodology to be used to reproduce the available data has to be selected, in accordance with the actual time evolution of the considered parameter. Finally, the parameter considered for trending should be rendered independent of machine operating point. In fact, otherwise, the prediction would also account for load variation, as may happen for measured variables, such as fuel mass flow rate or vibration levels. To do this, a procedure called *normalization* was presented in [11], where the measured values of thermodynamic quantities are normalized with respect to the respective expected values calculated in the same boundary conditions and actual working point. Health indices estimated through Gas Path Analysis techniques are instead independent of machine load and ambient conditions [7,8].

In literature, several prognostic methodologies for gas turbines have been reported. A good overview is offered by Roemer *et al.* in [14]. In the same area, Lipowsky *et al.* [15] present a statistical method called Bayesian Forecasting, which will also be applied in this paper. Zaluski *et al.* [16] develop a data mining methodology, while Bryg *et al.* [17] apply logistic regression to aircraft engine takeoff data. Finally, Puggina and Venturini [18] develop a methodology, which allows the prediction of future availability, starting from data trends collected in the past.

In this context, this paper stems from the work produced by Li and Nilkitsaranont [12] for gas turbine data trending and forecasting over *one* trend, by assuming different failure rate patterns over a given time frame. Moreover, this paper also continues the work by Lipowsky et al. [15] for forecasting system behavior after *two* consecutive trends, as for instance in the case of performance deterioration due to a gradual progressive failure followed by performance recovery due to a maintenance action. Such use of this methodology was not investigated in [15] and, so, it represents a new application carried out in this paper.

Four different methodologies are tested:

- Simple Linear and Non Linear Regression (SLRM and SNLRM);
- One Parameter Double Exponential Smoothing (OPDES);
- Baesyan Forecasting Method (BFM);
- Kalman Filter Method (KM).

The four models are used to provide a prediction of the time when a performance limit will be exceeded in the future (e.g. maximum decrease of efficiency, or highest turbine exhaust temperature), as a function of the current trend of the considered parameter. The application considers different scenarios which may be representative of the trend over time of several significant gas turbine parameters. The attention is focused on *gradual* deteriorations (e.g. compressor fouling or turbine erosion), since they can usually be tracked [18,19] and, therefore, their time evolution can be *predicted*. The influence of the presence of measurement uncertainty is also considered in the simulated scenarios. To this aim, a simulation model, with an easy-to-use graphical interface, is set up, to allow both the simulation of data trends and their prediction by means of the considered methodologies.

The Baesyan Forecasting Method, which allows the detection of discontinuities in time series, which may be associated with system failure followed by performance recovery, is also used for predicting system behavior after two consecutive trends.

Thus, the results presented in this paper are aimed to select the most suitable methodology that allows both trending and forecasting. As a consequence, maintenance planning may be optimized, as a function of the current machine health state, in order to (i) predict the time evolution of gas turbine state parameters and (ii) provide an estimate of the time point when the considered state parameter falls below a given threshold value.

FORECASTING METHODOLOGIES

To predict the future reliability of a system, it is common practice to make use of statistical methods to support system performance monitoring. The main features of some statistical techniques established in literature for this purpose are briefly reported below, while a thorough description, which is beyond the scope of this paper, can be found in specialized texts, which are referenced within each section.

Regression Method. Time series regression models relate the dependent variable X to polynomial functions of time. The pth-order polynomial trend model is given by Eq. (1) derived from [20]:

$$X = X_T + \varepsilon = \beta_0 + \beta_1 t + \beta_2 t^2 + ... + \beta_p t^p + \varepsilon \quad (1)$$

The error term ε represents random fluctuations that cause the X values to deviate from the average level X_T. Two trend expressions are considered in this paper: linear trend ($X_T=\beta_0+\beta_1 t$) and quadratic trend ($X_T=\beta_0+\beta_1 t+\beta_2 t^2$), which lead to the Simple Linear Regression Method (SLRM) and Simple Non Linear Regression Method (SNLRM), respectively. Least square point estimates of the β parameters in the trend models can be obtained by using regression techniques. To do this, the error term ε is assumed to satisfy the constant variance, independence and normality assumptions. For a complete description of regression method fundamentals, the reader is addressed to [20,21], while samples of its application to gas turbine data can be found in [11,12].

One Parameter Double Exponential Smoothing. In ongoing forecasting systems, forecasts of a time series are made each period for succeeding periods. Hence the forecasting equation and the estimates of the time series parameters need to be updated at the end of each period to account for the most recent observations.

Exponential smoothing is a forecasting method that weights the observed time series values unequally. In fact, more recent observations are weighted more heavily than more remote observations, by means of smoothing constants. Exponential smoothing has been found to be most effective when the parameters describing the time series change slowly over time, as may happen to the variation of gas turbine health parameters due to gradual failures. Unfortunately, exponential smoothing methods are not based on any formal statistical theory, but are rather intuitive methods that may produce adequate forecasts in some applications.

One Parameter Double Exponential Smoothing (OPDES) is an exponential smoothing method for handling a time series that displays a slowly changing linear trend, i.e. $X_T=\beta_0+\beta_1 t$. It is supposed that parameters β_0 and β_1 were determined at time (t-1), and a new observation is available at time t, to update the estimates of β_0 and β_1. To do this, single and double smoothed statistics (S_T and $S_T^{[2]}$) can be computed as follows [20]:

$$\begin{aligned} S_T &= \alpha X_T + (1-\alpha) S_{T-1} \\ S_T^{[2]} &= \alpha S_T + (1-\alpha) S_{T-1}^{[2]} \end{aligned} \quad (2)$$

Both equations employ the same smoothing constant α, which is defined to be between 0 and 1. The first equation smoothes the original time series observations, while the second one smoothes the S_T values obtained by means of the first equation.

For a discussion about the capabilities and the limitations of this method, the reader is once again addressed to [20,21].

Bayesian Forecasting Method. The Bayesian Forecasting Method (BFM) applied in this paper is derived from the paper authored by Lipowsky et al. [15]. The idea of Bayesian forecasting is based on Bayes' theorem for calculating conditional probabilities. In combination with Dynamic Linear Models (DLM), which break down the chronological sequence of the observed parameters into mathematical components (value, gradient, etc.), BFM can be used to calculate probability density functions prior to the next observation.

Starting from the information available at time (t-1), the guess values for the value, the gradient and the measurement at time t can be calculated as follows [15]:

$$\begin{aligned} (X_v)_t &= (X_v)_{t-1} + \Delta t \cdot (X_{grad})_t + \omega_{1,t-1} \\ (X_{grad})_t &= (X_{grad})_{t-1} + \omega_{2,t-1} \\ (X_{meas})_t &= (X_v)_t + \nu_t \end{aligned} \quad (3)$$

where the parameters ω_1, ω_2 and υ are the variances in value, gradient and measurement, respectively, which can be used to adjust the smoothing level. Equation (3) can be extended, so that, given the state at time (t-1), the BFM provides the predictions for any number of steps ahead, starting from the next time t. The conditional probability is accounted for by means of Bayes Factors H, which are the ratios of two probability density functions for the next observation: one is obtained by means of the current model and one is obtained by means of an alternative model, of which the mean value is shifted by h. Bayes Factors allow the identification of uncertainty limits for outlier detection, as defined in Eq. (4):

$$\lambda = \frac{\ln(H_{\min})}{h} + \frac{h}{2} \quad (4)$$

where h is the shift parameter.

Moreover, in the same paper [15], the logic for outlier detection is extended to change detection of parameter X, by means of cumulative Bayes Factors. In fact, as the method of Bayesian Forecasting provides guesses for both the value of the observed process and its dynamics (gradient, curvature, etc.), BFM is well suited to perform a prognosis of parameter X as an application of the forecast for a given number of steps ahead.

For the analytical description of this method, the reader is addressed to [22], while its application and validation for gas turbine prognostics can be found in the same paper [15]. Paper [23] also reports an application of Bayesian approach for fatigue life prediction.

Kalman Method. The method based on Kalman filter (KM) addresses the problem of estimating the state X in an m_1-dimensional space of a discrete-time controlled process that is governed by the linear stochastic difference equation [24]

$$X_t = A_1 X_{t-1} + A_2 U_{t-1} + w_{t-1} \quad (5)$$

with a measurement Z, defined in an m_2-dimensional space:

$$Z_t = A_3 X_t + \xi_t \quad (6)$$

The random variables w_t and ξ_t represent the process and measurement noise, respectively. They are assumed to be independent of each other, white, and with normal probability distributions. The m_1xm_2 matrix A_1 in the difference Eq. (5) relates the state at the previous time step (*t*-1) to the state at the current time step *t*, in the absence of either a driving function or process noise, while the m_1x1 matrix A_2 relates the optional control input *U* to the state *X*. Instead, the m_1xm_2 matrix A_3 in the measurement Eq. (6) relates the state to the measurement Z_t. It should be noted that matrices A_1 and A_3 generally change with each time step or measurement.

Therefore, the Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. Time update equations are responsible for projecting forward in time the current state, while error covariance estimates can be used to obtain the *a priori* estimates for the next time step. The measurement update equations are responsible for the feedback, i.e. for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations. Indeed, the final estimation algorithm resembles that of a predictor-corrector algorithm for solving numerical problems.

The approach of BFM is similar to the KM applied to time series, since both methods comprise predictor and corrector elements. In fact, the value of the next measurement is predicted and subsequently corrected when the next measurement is actually available.

Many specialized books deal with Kalman Filter, as for instance [25-27]. A fundamental application of the KM to gas turbine time series is reported by Provost in [28]. Other cornerstones in the field of gas turbine diagnostics through Kalman Filter based approaches are the papers authored by Volponi [29,30] and Doel [31].

SIMULATED DATA TRENDS

Let us assume that the considered state parameter, identified as *Y*(*t*), varies over time due to gradual gas turbine performance deterioration, as for instance compressor fouling. Let us also assume that data recording frequency is constant.

Equation (7) shows that the evolution over the time frame Δt of parameter *Y*(*t*) depends on the data trend and measurement uncertainty:

$$Y(t) = Y_T(t) + Y_u \qquad t = 0,1,2,...,\Delta t \qquad (7)$$

The trend values $Y_T(t)$ are simulated in this paper through four different curves (linear, quadratic, exponential and logarithmic). The assumption of linear trend means that the failure rate is constant and may approximate several types of degradation, mainly in a short-term period. However, the evaluation of different trends is made according to the fact that (i) performance degradation can occur in different forms, since the degradation rate is specific to the considered engine and installation and (ii) it is often found that the failure rate is not constant. For instance, in the case of fouling, the failure rate roughly follows an exponential law [13]. Moreover, the evaluation of different trends and of different $Y_T(t)$ variations also allows the behavior of different operational parameters, which may decrease slightly or considerably over a given time frame Δt, to be simulated.

The contribution of measurement uncertainty is given by Y_u, which is a random number taken from a Gaussian distribution with a zero mean and a variance σ_u^2, which can be varied to account for different instrumentation categories. In this paper, a variance of 1.0 % with respect to the actual value was imposed, to simulate a field instrumentation category. Moreover, the influence of halved measurement uncertainty (i.e. σ_u^2 = 0.5 %) was also evaluated for some of the analyzed cases. For each considered case, several simulations are carried out, to account for different data sets. In the following, the results will always refer (unless otherwise indicated) to the *average* of the errors occurred in the simulated data sets, in order to smooth the influence of the random numbers used to simulate the presence of measurement uncertainty.

Since the aim of this paper is to compare different forecasting methodologies, different minimum allowable values Y_{min} are considered. Reaching Y_{min} means that a shop visit is mandatory, since a limitation in the engine operation (e.g. a maximum power loss [5] or reaching turbine exhaust temperature limit [32]) has occurred.

Three *gradual* degradation scenarios, summarized in Tab. 1, are considered, in order to cover different rates of performance loss. Without any loss of generality, the first value of the trend is equal to 1, as the starting reference condition of the considered non-dimensional state parameter *Y* (i.e. Y_T = 1 @ t = 1). In fact, the state parameter *Y* is assumed non-dimensional, as in the case of gas turbine health indices estimated through Gas Path Analysis techniques. Moreover, the time frame is considered equal to 180 days in all cases.

Therefore, the difference among the scenarios is the relative change of parameter *Y* with respect to the initial *unit* value, i.e. (-1 %, -3 %, -5 %) and the respective minimum value Y_{min}

Table 1 – Scenarios for data trends (Y_T = 1 @ t = 1)

Scenario	$\Delta Y_T @ \Delta t$	Δt [days]	Y_{min}	Trend
#1	-1 %	180	0.98	Linear, Quadratic, Exponential, Logarithmic
#2	-3 %	180	0.95	Linear, Quadratic, Exponential, Logarithmic
#3	-5 %	180	0.92	Linear, Quadratic, Exponential, Logarithmic

(0.98, 0.95, 0.92). For instance, these values are in agreement with Schneider *et al.* [5], who reported experimental values of compressor performance degradation in terms of compressor efficiency drop and power output loss due to fouling, over a period of approximately seven months.

The values of parameter Y are reported in Fig. 1 for the four simulated trends for scenario #1. The same figure also illustrates Y_T values extrapolated after time Δt by means of two different procedures:

1. the line after time Δt is the linear interpolation line of Y_T values. Therefore, the interpolating and extrapolating lines in the case of Y_T linear trend are clearly superimposed. Moreover, it can be observed that this assumption moves the time point when $Y_T = Y_{min}$ further forward (e.g. Y_T logarithmic trend) or backward (e.g. Y_T quadratic trend) with respect to that of the linear trend;
2. the line after time Δt is the line passing through the last two Y_T values. Thus, also in this case, the interpolating and extrapolating lines in the case of Y_T linear trend are superimposed. However, unlike the previous situation, the time point when $Y_T = Y_{min}$ is always much further forward for all Y_T trends (quadratic, exponential and logarithmic) with respect to that of the linear trend.

The reason for the use of the linear interpolating line to extrapolate the data trend after time Δt is due to the fact that the actual trend is not known in practical applications and so a linear interpolation is the easiest and most likely solution. For instance, this same solution was also adopted in [12]. Instead, the rationale for the second procedure is that BFM prognosis for a second order DLM (i.e. only value and gradient are calculated) is essentially a continuation of the gradient calculated at the last observed cycle. Thus, the prognosis is linear [15]. Therefore, this procedure for extrapolation is applied to all methodologies, to make the results comparable.

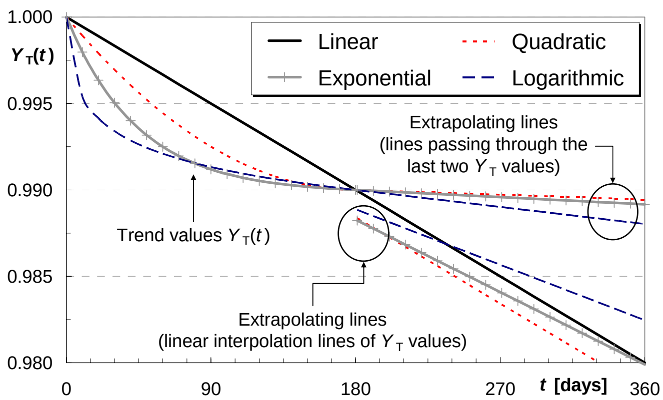

Figure 1 – Data trend for scenario #1 and extrapolating lines (linear interpolation lines of Y_T values or line passing through the last two Y_T values)

Indices for methodology assessment. In this paper, two indices are adopted to assess the reliability of the different methodologies.

The first index is the Root Mean Square Error *RMSE*, expressed in Eq. (8):

$$RMSE = \sqrt{\frac{1}{n_{\Delta t}} \sum_{i=1}^{n_{\Delta t}} \left(\frac{Y_T(t_i) - Y_{meth}(t_i)}{Y_T(t_i)} \right)^2} \tag{8}$$

The *RMSE* represents the error made by the methodology, which estimates Y_{meth}, with respect to the expected value Y_T, which does not account for measurement uncertainty. The sum is performed on all time points t_i from 1 to $n_{\Delta t}$. In this paper, $n_{\Delta t} = 180$, i.e. one Y value is available per day. Therefore, the RMSE can be used to evaluate the *trending* capability of each methodology.

The second index used for methodology reliability assessment is the prediction error E_p, expressed in Eq. (9):

$$E_p = | t^* - t^*_{meth} | \tag{9}$$

This value is obtained as the absolute value of the difference between two time points and so it is expressed in days. The time point t^*_{meth} represents the moment when the parameter Y_{meth} achieves the minimum allowable value Y_{min} and so it depends on the considered methodology and on data measurement uncertainty. Instead, the time point t^* represents the moment when the parameter Y actually achieves the minimum allowable value Y_{min} and is calculated by evaluating Y_T after time Δt up to the moment when $Y_T(t^*) = Y_{min}$. Therefore, this time value only depends on the considered type of trend, while it does not account for data measurement uncertainty. This time value t^* can be calculated by means of two different lines after time Δt, as previously discussed:

1. linear interpolation line of Y_T values;
2. line passing through the last two Y_T values.

Therefore, the prediction error E_p depends on the considered methodology and also on data measurement uncertainty. Therefore, this index can be used to evaluate the *forecasting* capability of each methodology.

The simulation model. All the forecasting methodologies have been implemented in the Matlab® environment. The simulation model is provided with a graphical user interface, which can be seen in Fig. 2.

The model allows the user to easily handle the data (either simulated or experimental) and to perform both trending and forecasting. The model also estimates the prediction error over a user-defined time frame.

Model inputs are the data trends. The model internal parameters depend on the considered methodology and are reported in Tab. 2, together with the respective values used for the simulations carried out in this paper. The values of model internal parameters were adjusted to achieve an optimal tuning of each forecasting methodology.

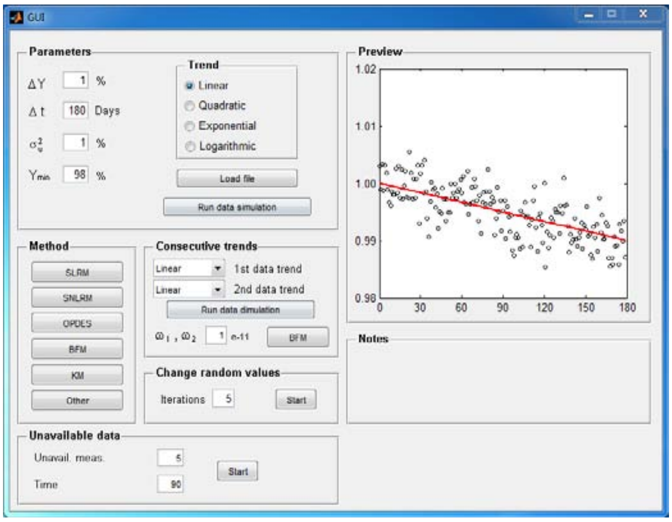

Figure 2 – Graphical user interface of the simulation model

Table 2 – Simulation parameters

Parameter	Description	Value
$Y(t)$	Data trends	Simulated, as in Tabs. 1 and 3
Y_{min}	Minimum value	
α	Smoothing constant (OPDES)	[0.01;0.30]
ω_1, ω_2	Variances in value and gradient (BFM)	$\omega_1 = \omega_2 = 10^{-11}$
h	Shift parameter (BFM)	[0.03;0.15]

RESULTS AND DISCUSSION

The comparison of the capability of the different methodologies to predict gas turbine behavior is reported in the following for the four types of simulated trends. First, trending capability *within* the time frame Δt is addressed. Second, the forecasting reliability *after* the time frame Δt is assessed. Finally, the forecast after two consecutive trends is also investigated through the BFM approach.

Trending capability. Two sample situations are reported in the following to illustrate trending capability evaluation procedure. Figure 3 shows the deviation between Y_T values (linear trend) and Y_{meth} values (obtained by using the SLRM methodology), which in this case is only due to the presence of measurement uncertainty. Instead, Figure 4 shows the deviation between Y_T values (quadratic trend) and Y_{meth} values (estimated through BFM), which accounts for both the presence of measurement uncertainty and the modeling approach. In this case, Y_{meth} values estimated through BFM oscillate at the first time steps, while they seem to reproduce Y_T values correctly after t = 10 days. Figure 4 also shows the uncertainty limits λ for outlier detection, calculated according to Eq. (4).

The complete results for the two scenarios #1 and #3 and all the considered trends for Y_T (i.e. linear, quadratic, exponential and logarithmic) are reported in Fig. 5, as a function of the methodology used for trending. It can be seen that SLRM,

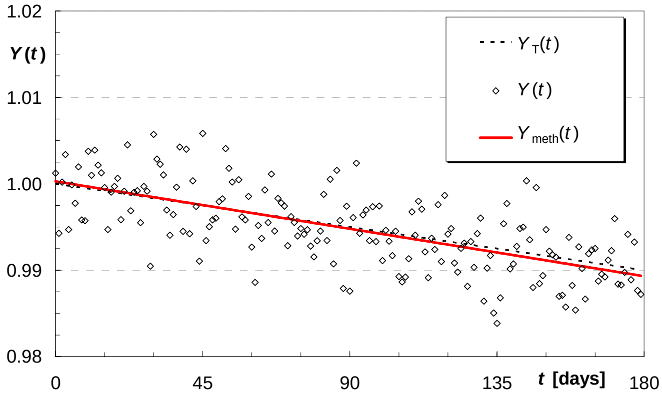

Figure 3 – Y_T linear trend and Y_{meth} values obtained through SLRM for scenario #1

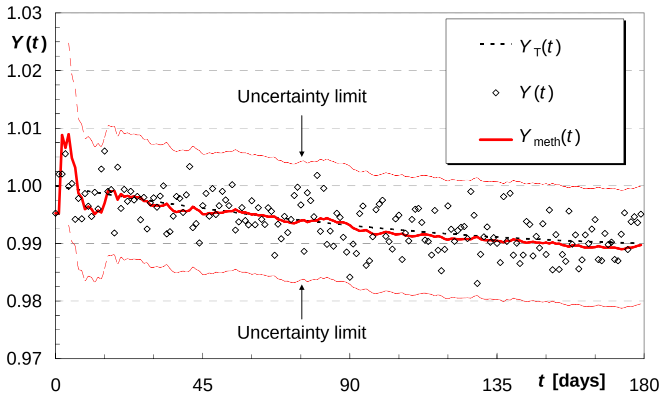

Figure 4 – Y_T quadratic trend and Y_{meth} values obtained through BFM for scenario #1 (uncertainty limits obtained according to Eq. (4))

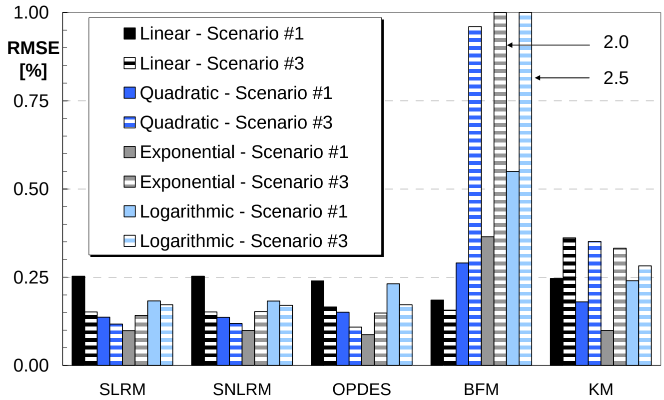

Figure 5 – RMSE values for scenarios #1 and #3 (σ_u^2 = 1.0 %)

SNLRM, OPDES and KM allow RMSE values, which are almost comparable and usually very low (i.e. less than 0.4 %).

Instead, with the exception of the linear trend for Y_T, RMSE values for BFM can be high (i.e. up to 2.5 %). Therefore, the analysis carried out in this section highlights that all methodologies offer good trending capability, with the exception of BFM, mainly because of high errors at the first time points. The reason for this behavior can be observed in Fig. 4, where the values estimated through the BFM at the first time points considerably deviate from the expected value Y_T. However, this only affects BFM trending capability, while, as shown in the next section, the reliability of BFM forecast will prove to be comparable to that of the other methodologies.

One more comment can be made about the influence of the magnitude of the degradation. The RMSE values for KM and, above all, BFM considerably increase by passing from scenario #1 to #3, as this latter method reproduces trends characterized by high gradients with difficulty. Otherwise, the RMSE variation which can be observed by passing from scenario #1 to #3 for the other methodologies depends on the considered trend and, in any case, RMSE values usually remain very low. As a general conclusion, BFM and KM are better suited to modeling parameter trends which change slightly with time.

Forecasting capability. The results of the analysis of the forecasting capability of the four considered methodologies after the time frame Δt is reported in Figs. 6 and 7 in the case of linear trend for Y_T. It can be observed that the prediction error:

- depends on the procedure used to forecast system behavior. However, the use of the linear interpolation of Y_T values is preferable in almost all cases, as shown in Fig. 6. In fact, according to Fig. 7. the line passing through the last two Y_T values provides lower E_p values in three cases only, i.e. in case of scenario #3 for SLRM, OPDES and BFM methodologies;
- decreases by passing from a slowly degrading (scenario #1) to a fast degrading (scenario #3) machine, since in the latter case the forecast is made for a lower number of future time steps and, as a consequence, the influence of the methodology and extrapolation procedure reduces;
- does not heavily depend on the selected methodology, as the order of magnitude of E_p is almost the same for all methodologies;
- in contrast with the results obtained in the previous section to evaluate trending capability, BFM proves as reliable as the other methodologies. This can be explained by considering that the prediction smoothes the influence of trending errors at first time points. Moreover, it should be noted that the E_p values for BFM in Fig. 7 are the *actual* prediction errors of this methodology, as discussed while presenting methodology characteristics.

The forecast capability of SLRM is evaluated in Fig. 8 for all the four Y_T trends. In this case, it can be concluded that the type of trend slightly affects the prediction error, if the interpolation line on the Y_T values is used. This is a remarkable result, since it confirms the reliability of this procedure, which is often adopted in practice.

Finally, it can be observed that the prediction errors may be considered acceptable for scenarios #2 and #3 and in the case of Y_T linear trend (Figs. 6 and 7). In fact, E_p values are lower than 15 days over a time frame Δt of 180 days. In contrast, for scenario #1, prediction errors can be very high, i.e. up to 58 days. Instead, if different trends are considered, as in Fig. 8 for the SLRM, prediction errors can be considerably lower.

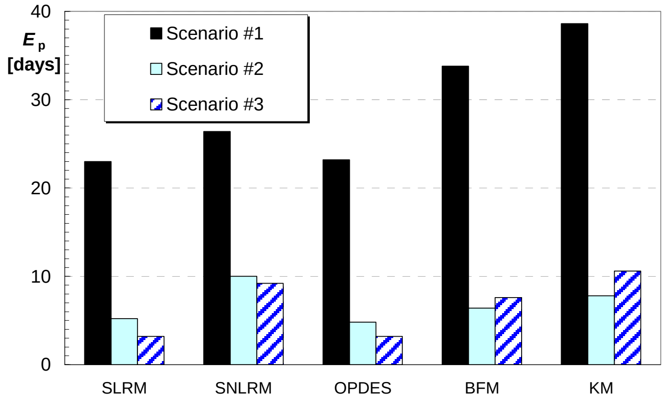

Figure 6 – Prediction errors by means of linear interpolation line of Y_T values for scenarios #1, #2 and #3 (Y_T linear trend)

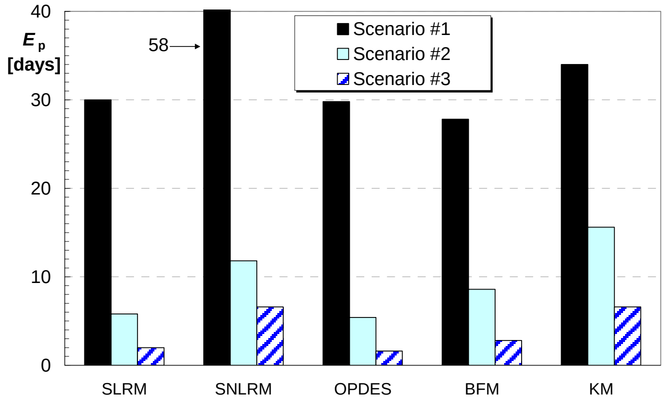

Figure 7 – Prediction errors by means of the line passing through the last two Y_T values for scenarios #1, #2 and #3 (Y_T linear trend)

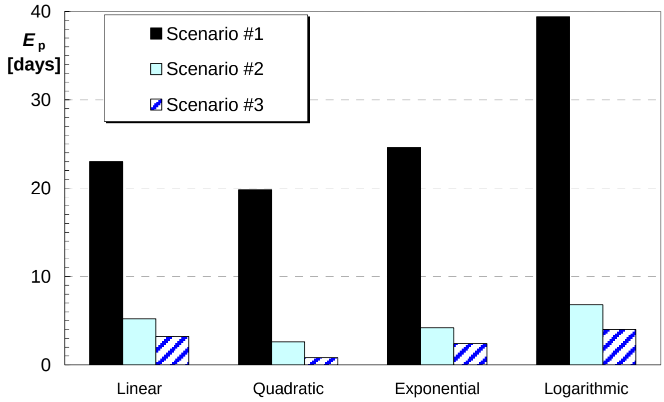

Figure 8 – Prediction errors by means of linear interpolation line of Y_T values obtained through SLRM for scenarios #1, #2 and #3

Influence of measurement uncertainty. In this section, reduced measurement uncertainty (i.e. σ_u^2 equal to 0.5 % instead of 1.0 %) is evaluated for scenarios #1 and #3. The results, reported in Figs. 9 and 10 for assessing trending and forecasting capability respectively, quantify the magnitude of the decrease in the errors, by direct comparison with the results presented in Figs. 5 and 6.

As regards RMSE values in Fig. 9, it can be observed that, when measurement uncertainty is halved, the RMSE values for SLRM, SNLRM and OPDES are also halved, while the reduction of measurement uncertainty is less effective or even negligible in the case BFM or KM are used.

As regards the prediction error in Fig. 10, the reduction of measurement uncertainty leads to an almost proportional reduction of E_p, with the exception of KM, which is less sensitive to σ_u^2 variation. Finally, Fig. 10 confirms that scenario #3 allows lower prediction errors, as previously discussed.

BFM capability to forecast system behavior after two consecutive trends. The results presented so far deal with trending and forecasting one trend only. However, it is common in practice that *two* or more consecutive trends for the considered parameter *Y* are available.

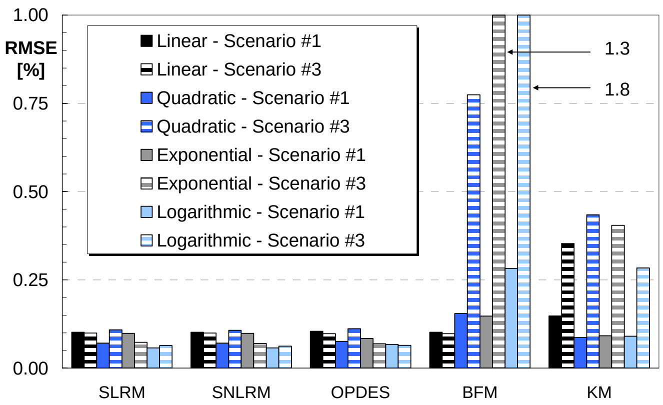

Figure 9 – RMSE values for scenarios #1 and #3 ($\sigma_u^2 = 0.5$ %)

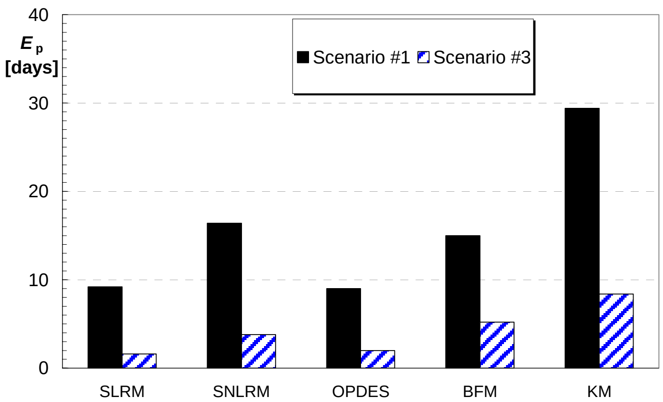

Figure 10 – Prediction errors by means of linear interpolation line of Y_T values for scenarios #1, #2 and #3 (Y_T linear trend; $\sigma_u^2 = 0.5$ %)

This situation may be due to performance deterioration followed by a maintenance action which recovers machine performance. Due to their modeling approach, the Regression Method and One Parameter Double Exponential Smoothing are useless, since they would interpolate *both* trends and, as a consequence, the effect of performance recovery on *Y* would not be accounted by these methodologies. In contrast, the Bayesian Forecasting Method is well suited for prognostics, since it provides estimates for both the value of the observed process and its dynamics (gradient, curvature, etc.). In fact, it inherently allows forecasting, by means of the recursive estimation of parameter *Y* at the next time step, as reported in Eq. (3). Moreover, it was previously verified that BFM also proves a reliable methodology to forecast system behavior when only *one* trend is considered.

The considered scenarios are summarized in Tab. 3. Two different variations of parameter *Y* are imposed (-1% or -3%) over the same time frame (180 days). Similar values of gradual deterioration for several consecutive trends were also considered by Puggina and Venturini in [18]. Moreover, it has to be noted that these values are consistent with the values for which BFM detection capability was tested in [15], i.e. a change height up to $5\sigma_u^2$. In fact, in this paper, the considered step changes are equal to $1\sigma_u^2$ or $3\sigma_u^2$, for scenario #4 or #5 respectively. Moreover, it has to be underlined that, for each scenario considered in this section, ten simulations are carried out (as reported in Tab. 4), to account for different data sets simulated through different combinations of random numbers. Finally, it has to be remarked that all trends Y_T are assumed to start from the unit reference condition, i.e. $Y_T = 1$, both at $t = 1$ and at $t = \Delta t+1$.

Figures 11 and 12 sketch the situations corresponding to scenarios #4a and #5a, respectively, to simulate two different values of the step change. An exponential trend is imposed for the first 180 days, followed by a linear trend for additional 180 days. BFM estimation of the trend is also reported, together with the two uncertainty limits.

Figure 11 shows that the step change at $t = (\Delta t+1)$ is reproduced by BFM with a noticeable delay, which may be responsible for the high RMSE values highlighted in the previous sections. This delay is due to the fact that the imposed step change is of the same order of magnitude as measurement uncertainty and therefore it can be recognized by BFM with difficulty.

Table 3 – Scenarios for data trend forecasting (total time frame: $2\Delta t = 360$ days; $\sigma_u^2 = 1.0$ %)

Scenario	ΔY	Y_{min}	Trend
#4	-1 %	0.98	a) Exponential + Linear
			b) Exponential + Quadratic
#5	-3 %	0.95	a) Exponential + Linear
			b) Exponential + Quadratic

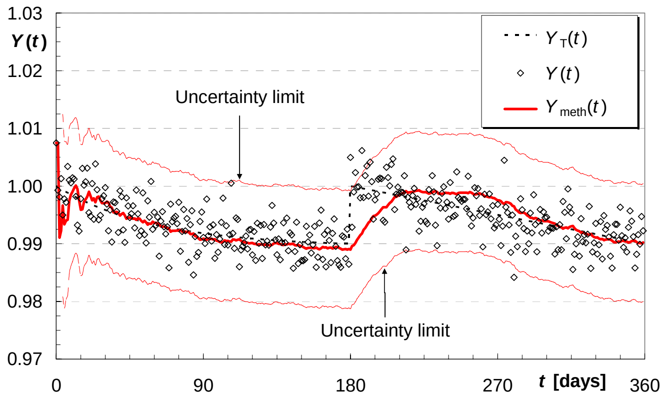

Figure 11 – Y_{meth} values and uncertainty limits obtained through BFM for scenario #4a

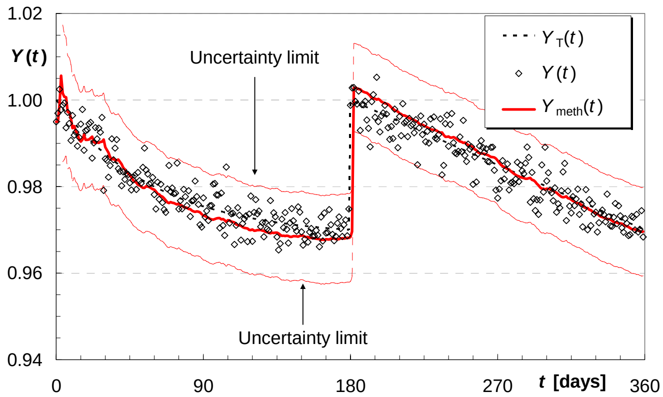

Figure 12 – Y_{meth} values and uncertainty limits obtained through BFM for scenario #5a

Table 4 – Prediction errors after two consecutive trends

Scenario	#4a	#4b	#5a	#5b
$(E_p)_1$	6	491	15	326
$(E_p)_2$	43	518	8	309
$(E_p)_3$	42	198	2	306
$(E_p)_4$	35	485	2	321
$(E_p)_5$	28	564	2	285
$(E_p)_6$	6	423	5	314
$(E_p)_7$	1	539	18	290
$(E_p)_8$	20	494	11	287
$(E_p)_9$	17	409	14	323
$(E_p)_{10}$	9	471	18	264
$(E_p)_{av}$	**21**	**459**	**10**	**303**

However, in spite of this, the average prediction error on ten data sets is 21 days for scenario #4a, as reported in Tab. 4. This represents a very good estimation, since the total time frame for scenarios #4 and #5 is 360 days, as both trends are made up of 180 days. This good result is also confirmed in scenario #5a (characterized by a faster degradation rate), where the average prediction error is 10 days. In fact, in the case of a more remarkable step change (i.e. 3 %), the BFM can track the change with a delay of a few days only, as clearly shown in Fig. 12. Otherwise, scenarios #4b and #5b, i.e. those ones of which the second trend is quadratic, are affected by very large prediction errors (i.e. larger than 300 days).

Therefore, it can be stated that, in the case the second trend is linear, BFM actually has the capability to track the step change-over time of the considered parameter and improves the prediction reliability in the case of a scenario with a faster degradation rate. In contrast, if the second trend decreases with a quadratic law, the BFM prediction error can be very high.

CONCLUSIONS

In this paper, several scenarios for gas turbine parameter trends, representative of time evolution of *gradual* deteriorations, were investigated, in order to identify the most suitable methodology which allows both trending and forecasting. Four different methodologies were applied, i.e. Regression, One Parameter Double Exponential Smoothing, Baesyan Forecasting Method and Kalman Filter.

As regards trending capability, Regression, One Parameter Double Exponential Smoothing and Kalman Filter allowed very low RMSE values (i.e. less than 0.4 %). Otherwise, the Baesyan Forecasting Method usually proved less reliable (i.e. errors up to 2.5 %). The analysis of the influence of the degradation scenario showed that the Baesyan Forecasting Method and the Kalman Filter are better suited to estimate parameter trends which change slightly with time.

As regards forecasting capability, it was proved that the prediction error (i) depends on the procedure used to forecast system behavior, (ii) decreases by passing from a slowly degrading to a fast degrading machine and (iii) does not heavily depend on the selected methodology. However, in contrast with the results obtained from the analyses on trending capability, the Baesyan Forecasting Method proved as reliable as the other methodologies.

The reduction of measurement uncertainty leads to an almost proportional reduction of errors both for trending and forecasting, with the exception of the Kalman Filter and the Baesyan Forecasting Method, which are less sensitive to this reduction.

Finally, the Baesyan Forecasting Method was also applied for predicting system behavior after two consecutive trends. It was observed that, when the second trend is linear, this methodology actually has the capability to track a step change over time of the considered parameter and improves the prediction reliability in the case of a scenario with a faster degradation rate. In contrast, if the second trend decreases with a quadratic law, the prediction error can be very high.

Future developments will deal with the assessment of the capability of the Baesyan Forecasting Method as a diagnostic tool for fault detection, both in the case that some data are not available during the considered time frame and in the case that this method is applied to an on-line monitoring system, in which the measurements are continuously recorded.

REFERENCES

[1] Bettocchi R., Pinelli M., Spina P. R., Venturini M., Sebastianelli S., 2001, "A System for Health State Determination of Natural Gas Compression Gas Turbines", *ASME Paper 2001-GT-223.*

[2] Jaw, L. C., 2005, "Recent Advancements in Aircraft Engine Health Management (EHM) Technologies and Recommendations for the Next Step", *ASME Paper GT2005-68625*.

[3] Therkorn, D., 2005, "Remote Monitoring and Diagnostic for Combined-Cycle Power Plants", *ASME Paper GT2005-68710*.

[4] Davison, C., Drummond, C., 2009, "Application of Cost Matrices and Cost Curves to Enhance Diagnostic Health Management Metrics for Gas Turbine Performance"*, ASME Paper GT2009-59630*.

[5] Schneider, E., Demirciogiu, S., Franco, S., Therkorn, D., 2009, "Analysis of compressor on-line washing to optimize gas turbine power plant performance"*, ASME Paper GT2009-59356*.

[6] Stamatis, A., Mathioudakis, K., Papailiou, K.D., 1990, "Adaptive Simulation of Gas Turbine Performance", *ASME J. Eng. Gas Turbines Power*, **112**, pp. 168-175.

[7] Bettocchi, R., Spina, P. R., 1999, "Diagnosis of Gas Turbine Operating Conditions by Means of the Inverse Cycle Calculation", *ASME Paper 99-GT-185*.

[8] Pinelli, M., Venturini, M., 2002, "Application of Methodologies to Evaluate the Health State of Gas Turbines in a Cogenerative Combined Cycle Power Plant", *ASME Paper GT-2002-30248*.

[9] Doel, D. L., 2003, "Development of baselines, influence coefficients and statistical inputs for gas path analysis"*, Von Karman Institute Lecture Series 2003-01 "Gas Turbine Monitoring & Fault Diagnosis"*, Jan 13-17.

[10] Li, Y. G., 2004, "Gas Turbine Diagnosis Using a Fault Isolation Enhanced GPA", *ASME Paper GT2004-53571*.

[11] Pinelli, M., Venturini, M., 2001, "Operating State Historical Data Analysis to Support Gas Turbine Malfunction Detection", *ASME IMECE2001/AES-23665.*

[12] Li, Y.G., Nilkitsaranont, P., 2009, "Gas turbine performance prognostic for condition-based maintenance"*, Applied energy*, **86**, pp. 2152 - 2161.

[13] Meher-Homji, C. B., Chaker, M., Bromley, A.F., 2009, "The fouling of axial flow compressors – Causes, effects, susceptibility end sensitivity", *ASME Paper GT2009-59239*.

[14] Roemer, M. J., Byington, C. S., Kacprzynski, G. J., Vachtsevanos, G., 2006, "An Overview of Selected Prognostic Technologies with Application to Engine Health Management"*, ASME Paper GT2006-90677*.

[15] Lipowsky, H., Staudacher, S., Bauer, M., Schmidt, K. J., 2009, "Application of Bayesian Forecasting to Change Detection and Prognosis of Gas Turbine Performance"*, ASME Paper GT2009-59447*.

[16] Zaluski, M., Letourneau, S., Bird, J., Yang, C., 2010, "Developing Data Mining-Based Prognostic Models for CF-18 Aircraft"*, ASME Paper GT2010-22944*.

[17] Bryg, D. J., Mink, G., Jaw, L. C., 2008, "Combining Lead Functions and Logistic Regression for Predicting Failures on an Aircraft Engine"*, ASME Paper GT2008-50118*.

[18] Puggina, N., Venturini, M., 2011, "Development of a Statistical Methodology for Gas Turbine Prognostics"*, ASME Paper GT2011-45708*.

[19] Borguet, S., Leonard, O., 2008, "A generalized Likelihood Ratio Test for Adaptive Gas Turbine Health Monitoring", *ASME Paper GT2008-50117*.

[20] Bowerman, B. L., O'Connell, R. T., 1993, "*Forecasting and Time Series - An Applied Approach*", Duxbury Classic Series.

[21] Hastie, T., Tibshirani, R., Friedman, J., 2001, "*The elements of Statistical Learning – Data Mining, Inference, Prediction*", Springer-Verlag.

[22] West, M., Harrison, J., 1999, "*Bayesian forecasting and dynamic models*", Springer.

[23] Kim, N. H., Pattabhiraman, S., Houck, L. A., 2010, "Bayesian Approach for Fatigue Life Prediction from Field Data"*, ASME Paper GT2010-23780*.

[24] Welch, G., Bishop, G., 2001, "An Introduction to the Kalman Filter"*, Proc. ACM SIGGRAPH 2001 Conference*, August 12-17, Los Angeles, CA, USA.

[25] Haykin, S. S., 2001, "*Kalman Filtering and Neural Networks*", John Wiley & Sons, Inc..

[26] Zarchan, P., Musoff, H., 2005, "*Fundamentals of Kalman Filtering: A Practical Approach*", American Institute of Aeronautics and Astronautics.

[27] Grewal, M. S., Andrews, A. P., 2008, "*Kalman Filtering Theory and Practice Using MATLAB*", John Wiley & Sons, Inc..

[28] Provost, M.-J., 2003, "Kalman Filtering Applied to Time Series Analysis"*, Proc. IEE Seminar on Aircraft Airborne Condition Monitoring*, May 14, Gloucester, UK.

[29] Volponi, A. J., Urban, L. A., 1992, "Mathematical Methods of Relative Engine Performance Diagnostics", SAE Trans., **101**; *Journal of Aerospace*, Technical Paper 922048.

[30] Volponi, A. J., DePold, H., Ganguli, R., and Daguang, C., 2000, "The Use of Kalman Filter and Neural Network Methodologies in Gas Turbine Performance Diagnostics: A Comparative Study", *ASME Paper 00-GT-547*.

[31] Doel, D. L., 1994, "TEMPER - A Gas Path Analysis Tool for Commercial Jet Engines", *ASME Paper 92-GT-315*.

[32] Muller, M., Staudacher, S., Friedl, W. H., Kohler, R., Weisschuh, M., 2010, "Probabilistic Engine Maintenance Modeling for Varying Environmental and Operating Conditions"*, ASME Paper GT2010-22548*.

