

# An Adaptive Model-Based Framework for Prognostics of Gas Path Faults in Aircraft Gas Turbine Engines

Ogechukwu Alozie<sup>1</sup>, Yi-Guang Li<sup>2</sup>, Xin Wu<sup>3</sup>, Xingchao Shong<sup>4</sup>, Wencheng Ren<sup>5</sup>

<sup>1,2</sup>*Department of Power and Propulsion, Cranfield University, MK43 0AL, Bedfordshire, UK*  
*o.alozie@cranfield.ac.uk*  
*i.y.li@cranfield.ac.uk*

<sup>3,4,5</sup>*Shenyang Engine Design and Research Institute (SEDRI), Aircraft Engine Corporation of China (AECC)*  
*jiasuzi136@sina.com*

## ABSTRACT

This paper presents an adaptive framework for prognostics in civil aero gas turbine engines, which incorporates both performance and degradation models, to predict the remaining useful life of the engine components that fail predominantly by gradual deterioration over time. Sparse information about the engine configuration is used to adapt a performance model which serves as a baseline for implementing optimum sensor selection, operating data correction, fault isolation, noise reduction and component health diagnostics using nonlinear Gas Path Analysis (GPA). Degradation models which describe the progression of faults until failure are then applied to the diagnosed component health indices from previous run-to-failure cases. These models constitute a training library from which fitness evaluation to the current test case is done. The final remaining useful life (RUL) prediction is obtained as a weighted sum of individually-evaluated RULs for each training case. This approach is validated using dataset generated by the Commercial Modular Aero-Propulsion System Simulation (CMAPSS) software, which comprises both training and testing instances of run-to-failure sensor data for a turbofan engine, some of which are obtained at different operating conditions and for multiple fault modes. The results demonstrate the capability of improved prognostics of faults in aircraft engine turbomachinery using models of system behaviour, with continuous health monitoring data.

## 1. INTRODUCTION

Prognostics and Health Management (PHM) as a field of specialization in engineering encompasses techniques employed to maximize the service life of various systems and equipment. It achieves this through continuous monitoring of

key system parameters and operating conditions, assessing the current health state from these measurements and making predictions about the future health or the time-to-failure of critical components within the system. It therefore offers great potential for improved availability, reliability, optimum performance and ensured safety of the system to which it is applied, thereby reducing the chances of unforeseen downtime during operation (SAE 2013).

Prognostic methods in general can be classified into 4 main types: (1) Knowledge-based methods, which employs the experience of domain experts in generating rule sets (DePold & Gass 1999; Biagetti & Sciubba 2004), (2) Data-driven approaches that apply statistical and machine learning algorithms to reveal underlying patterns in large CM data (Barad et al. 2012; Mosallam et al. 2016; Li & Nilkitsaranont 2009), (3) Physics-based models that provide a mathematical relationship between system operating conditions and time to failure (Cubillo et al. 2016), and (4) hybrid methods that combine the benefits of two or more of the above mentioned types (Saha & Goebel 2011; Baraldi et al. 2013).

Despite the recognized benefits and the large resource of proposed methods on the subject, there are some difficulties associated with the deployment of PHM systems in real-life industrial applications. The ease of adapting proposed prognostic methods to well-established and existing diagnostic schemes, and the availability of relevant run-to-failure data for verification and validation of various prediction methods have been identified as two key limiting factors, especially in the energy and aviation industry where gas turbines play a key role (Sikorska et al. 2011; Saxena et al. 2008).

To tackle the issue of validation data availability, the NASA Prognostics Centre of Excellence (PCoE) has provided a data repository for various engineering systems that would foster research in the field of prognostics. Most of the datasets comprise run-to-failure sensor data for training and testing

Ogechukwu Alozie et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

purposes in systems such as trebuchet, turbofan engine, MOSFET, transformer, Li-ion battery, etc. This approach has yielded significant returns in the number of prognostic methods that have been published based on these dataset. Some methods of worthy mention developed for the turbofan engine application include similarity-based approach (Wang et al. 2008), Kalman filter ensembles of Radial Basis Function networks (Peel & Gold 2008), recurrent neural networks (Heimes 2008), general path model using a Bayesian belief theorem (Coble & Hines 2014), and logistic regression of a state-space model (Yu 2017). A comprehensive review of over 70 of these methods based on their performance is provided in (Ramasso & Saxena 2014), while a classification based on the information used is provided in (Coble & Hines 2008).

While some of these methods have employed a form of sensor fusion and modelling in describing the degradation pattern, none has provided an investigation into the use of an engine performance model for prognostics. This could be attributed a number of reasons, not limited to the following:

- the nature of the data provided comprises sensor readings for training and testing the algorithm, with little or no engine performance specification,
- restricted access to the CMAPSS model and software used to generate the data, and
- the rules of the PHM Challenge may have inferred the desire for a data-driven solution that could be readily applied to other case studies (Ramasso & Saxena 2014).

On the other hand, some other proposed methods which have incorporated engine modelling however lack validation using externally-sourced data of the scale available in the PCoE turbofan data repository.

This paper proposes a solution to the identified issues above by combining a validated diagnostic routine with the available historical condition monitoring data to perform prognostics of gas path faults in the engine. The approach is classified as model-based for two reasons;

1. A performance model that describes the behaviour of the gas turbine components based on the thermodynamics of the working fluid is used to provide information about the configuration and operation of the real engine.
2. The degradation model that describes the progression of fault over time is used to determine the fitness of previous failure data to the diagnosed component health parameters and perform RUL prediction on test cases.

A key benefit of this approach is that it accounts for the peculiar behavior of the system to which it is applied – the gas turbine engine in this case. This is unlike a purely data-driven method that focuses only on the acquired sensor data, in isolation from the system. The major contribution of this paper, thus lies in the fusion of both adapted engine performance and degradation models with historical run-to-failure data from the same or similar engines in a fleet via a fitness evaluation function, to predict the remaining useful life of the currently deteriorating engine. This adaptive fusion improves prediction accuracy in a real-life scenario where no two engines, even of the same type and configuration, perform in the same way due to different manufacturing tolerances, installation variances and operating profiles. This method is also not limited to the gas turbine, but can easily be adapted to a different system, provided a performance model which describes its behavior is available.

In the following Sections, the proposed methodology is described in details, a CMAPSS engine case study along with the underlying assumptions is provided, and the results and the performance metrics of the prediction algorithm are presented. Finally, the implication of the study and areas for further research are provided in the conclusions.

## 2. METHODOLOGY

Figure 1 shows a block diagram of the proposed adaptive model-based prognostic framework, with the most vital element being the engine performance model produced in the modeling phase. The techniques employed for each process in the framework is described in detail below.

### 2.1. Engine Modelling and Adaptation

A gas path performance model can provide useful insight into the behavior of gas turbine components and its overall output in terms of efficiencies, thrust, fuel consumption, etc. Such a model is therefore considered a true representation of the real engine if it accurately predicts the performance output as provided in manufacturer specifications or from actual measured data. Figure 2 below shows a schematic for a typical turbofan engine with some of the installed sensors provided in the dataset.

Due to the proprietary nature of component performance maps, as well as individual differences between similar engine configurations arising from manufacturing or installation tolerances, the true component parameters at any operating condition are truly unknown. It is therefore necessary to carry out a performance adaptation to match the model output to the engine measured data.

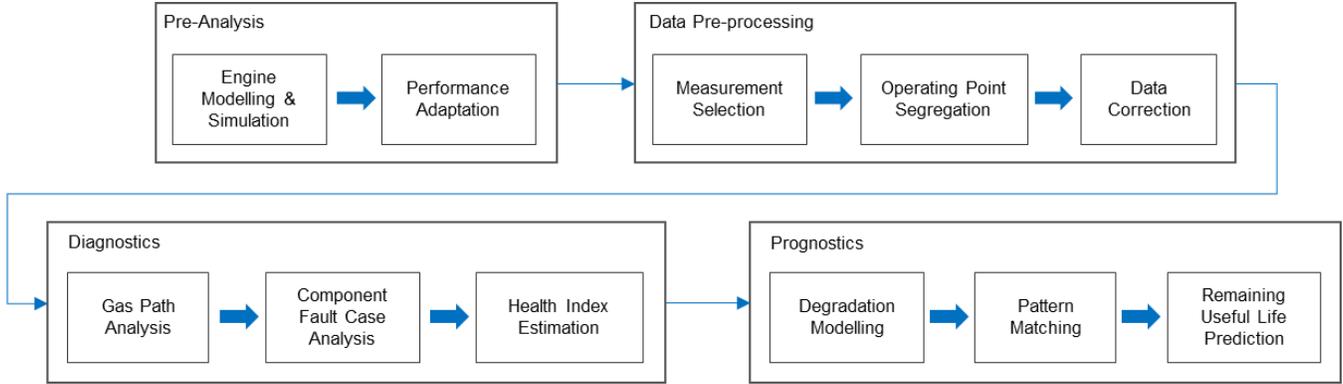


Figure 1: Adaptive model-based prognostics methodology.

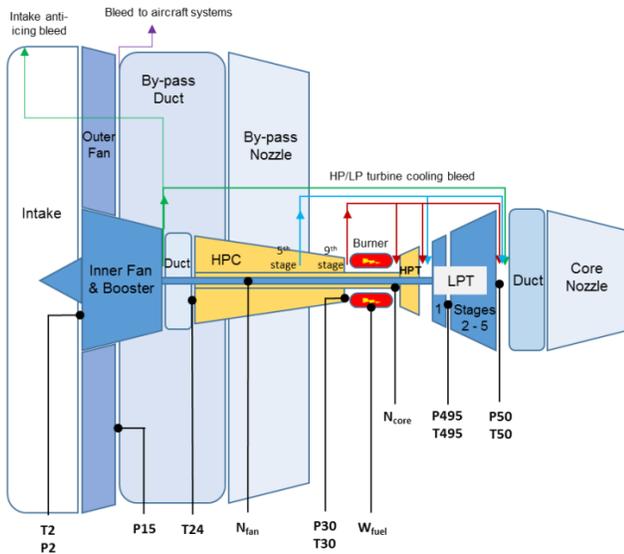


Figure 2: Turbofan engine gas path model schematic showing bleed and sensor locations.

A nonlinear form of the adaptation coefficient matrix (ACM) approach, which describes changes in sensor measurements  $\Delta z$  as a function of corresponding changes in component parameters  $\Delta x$  as shown in Eqs. (1) to (3), neglecting higher-order terms, was adopted. Here, an iterative procedure was applied to the linear ACM approximation in Eq. (3) until a predefined convergence criteria is achieved, thus accounting for the nonlinearity in engine behavior over large changes in measurements (Li et al., 2006).

$$z = z_0 + \frac{\partial z}{\partial x} \Delta x + HOTS \quad (1)$$

$$\Delta z = G \cdot \Delta x \quad (2)$$

$$\Delta x = G^{-1} \cdot \Delta z \quad (3)$$

Where  $G$  is the adaptation coefficient matrix,  $\Delta z$  is the deviation of the model-simulated measurement from the real

engine,  $\Delta x$  is the corresponding change in component parameter need to match the model to the real engine, and *HOTS* represent higher-order terms.

## 2.2. Sensor Selection

Although measured engine parameters from on-board sensors can provide crucial information on the state of the engine components and the presence of incipient faults, not all sensors might be relevant for a given fault case search. A three-step sensor selection process was therefore adopted to determine the optimum sensor subset required for isolation and quantification of all possible component(s) fault. This subset was chosen based on the criteria of maximum and unique visibility of engine health, while providing some redundancy to take into account the possible case of biased or faulty sensors.

First, a sensor sensitivity analysis, where a unit degradation was implanted in each component health parameter to obtain the deviation in each available sensor, was performed. The sensitivity norm which describes the overall sensitivity of each sensor can then be expressed using Eq. (4) below (Jasmani et al., 2010).

$$\|s_i\| = \sqrt{\sum_{j=1}^N (s_{i,j})^2} \quad (4)$$

Where  $s_{i,j}$  is the sensitivity coefficient,  $i = 1, 2, \dots, M$  is the sensor index and  $j = 1, 2, \dots, N$  the health parameter index.

A correlation analysis was also carried out to reveal the sensors with identical fault signatures. The correlation matrix, whose elements are obtained by the multiplying the matrix of normalized absolute sensitivity  $n_{i,j}$ , expressed in Eq. (5), by its transpose, was used to quantify this phenomenon (Chen et al. 2015).

$$n_{i,j} = s_{i,j} / \|s_i\| \quad (5)$$

When two or more sensors were correlated to one another, only the most sensitive was selected for further investigation.

Finally, a classification of the sensors according to their associated components, either by proximity/location in the engine or direct mathematical relationship was carried out. This enabled the selection of sensors associated with more than one component since they can provide a wider coverage for component fault detection.

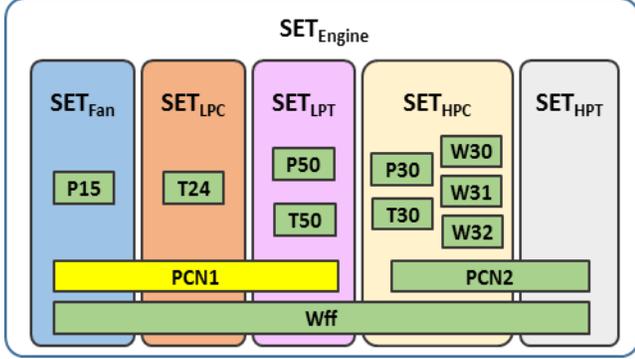


Figure 3: Venn diagram showing component-related sensor subsets. Fan speed (PCN1) is the engine handle parameter.

### 2.3. Data Correction

Sensor data from engine condition monitoring is seldom obtained at a fixed operating condition. The changing properties of air with ambient and flight conditions therefore makes it difficult to compare the sensor measurements acquired during the different phases of flight and in different flight cycles to their respective clean engine values. A data correction technique which refers each sensor data  $z_{B1}$  obtained at conditions  $Y_b$  to a pre-defined set of baseline conditions and power setting at  $Y_a$  as shown in Figure 4 and using Eq. (6) was adopted for this study (Li et al., 2002).

$$\Delta z_{A1A} = \Delta z_{B1B} - (\Delta z_{AB} - \Delta z_{A1B1}) \quad (6)$$

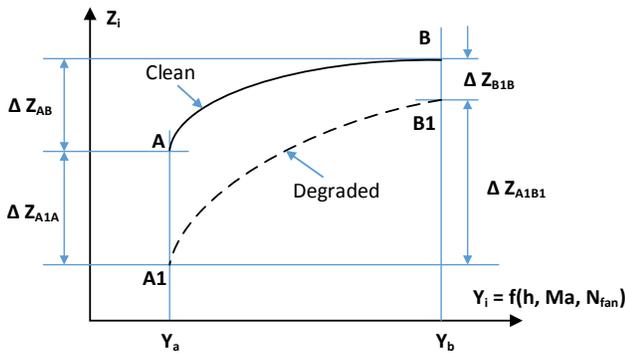


Figure 4: Data correction schematic.

### 2.4. Fault Quantification

To identify the faulty component(s) and quantify the level of degradation present using the selected sensor set, the nonlinear form of the Gas Path Analysis (GPA) with component fault cases was used. The GPA is based on

solving a linear system of equations which relate changes in each component health parameter to corresponding changes in sensors  $\Delta x/\Delta z$ , through an influence coefficient matrix. To account for the nonlinear relationship between  $\Delta x$  and  $\Delta z$ , the system of equations is solved iteratively until the convergence criteria, which is the accuracy of the fault prediction, is attained. A schematic of the nonlinear GPA process is presented in Figure 5.

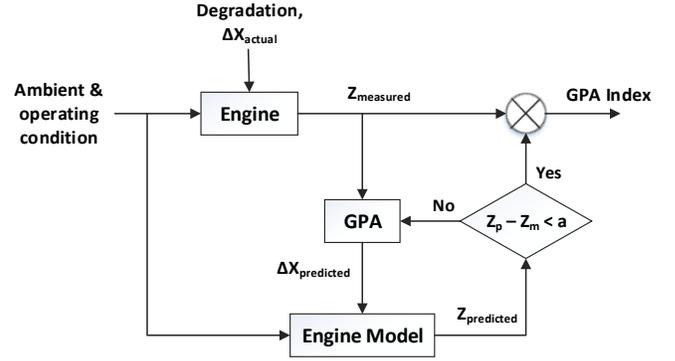


Figure 5: Flowchart for the non-linear GPA procedure.

In a multiple component fault case (CFC) analysis, a GPA index is calculated according to Eq. (7) and assigned to each combination of components evaluated, referred to as a fault case. The fault case with the highest GPA index therefore reveals the most accurate prediction of faulty components from the search space (Li et al., 2009).

$$GPA\ Index = \frac{1}{1 + \varepsilon} \quad (7)$$

Where  $\varepsilon$  is a measure of the difference between the measured and GPA-predicted deviations in the measurement parameters as expressed below.

$$\varepsilon = \frac{1}{N} \sum_{i=1}^N \left| \frac{\Delta z_{measured}}{z_{measured}} - \frac{\Delta z_{predicted}}{z_{predicted}} \right| \quad (8)$$

### 2.5. Health Index Simulation

Based on the GPA predictions of component(s) fault in terms of flow capacity and efficiency loss, a single health index (HI) that is representative of the overall engine health may be desirable. According to Saxena et al. (2008), this health index could represent deviations of fan, LPC or HPC surge margin by up to 15% or exhaust gas temperature deviations up to a 2% limit.

$$HI_t = f(X_t^{eff}, X_t^{fc}) \quad (7)$$

Where  $X_t^{eff}$  and  $X_t^{fc}$  represent the efficiency and flow capacity health parameters respectively over time.

To obtain either of the above HIs, the GPA-predicted fault progression is implanted back into the engine model and simulated at the specified operating condition. The obtained health index is then normalized as shown in Equation 9, using the average value at the end-of-life, to a range of 0 to 1; where 0 denotes failure and 1 signifies relatively clean/healthy engine condition.

$$HI_{norm} = \frac{HI_t - \overline{HI}_n}{HI_0 - \overline{HI}_n} \quad (9)$$

Where  $HI$  is the health index, and  $t = 0, 1, 2, \dots, n$  signifies the data time series indices for an engine unit, with 0 as initial and  $n$  as the end-of-life index.  $\overline{HI}_n$  is therefore the average health index at end-of-life for all the units in a dataset.

## 2.6. Degradation Modelling

The underlying degradation mechanism determines the form of the HI trend for the faulty component(s) or system. According to Saxena et al. (2008), the prevailing mechanism can be represented by a generalized exponential wear degradation model of the form in Eq. (10) below.

$$HI_t = c - a \cdot \exp(t^b) \quad (10)$$

Where  $c$  is the initial health index,  $a$  is the health index decay rate,  $b$  is the exponential time-scaling parameter, and  $t$  is the time in cycles.

This wear model was used to fit the HI time series and the model parameters were obtained using the non-linear least-squares regression approach, which minimized the errors between the model and the data.

A second model of the form in Eq. (11) below, was also chosen to provide an alternative perspective to the trend analysis, based on the assumption of a linear time exponent.

$$HI_t = c - a \cdot \exp(b \cdot t) \quad (11)$$

The model coefficients for each training unit was used to create a degradation model library which describes the various possible gradual degradation patterns that a given engine unit might experience. In a real-world application, this pattern library would be updated when new run-to-failure data from the engine becomes available.

## 2.7. RUL Prediction

The process for the RUL prediction of an on-wing engine is shown in Figure 6 below. The methods described in sections 2.1 to 2.7 transform the acquired multi-dimensional sensor data into individual engine component health indices which are aggregated, trended and used for prediction to the pre-determined health index threshold. This threshold which signifies the component's end-of-life could be determined statistically from previous engine operation, as the health index or combination of health indices values which lead to a maximum unacceptable reduction in engine performance

that impact engine life, safe operation and operating (mission fuel burn) costs. This threshold could therefore be the maximum acceptable loss in component flow capacity and efficiency arising from recoverable degradation such as fouling, or irrecoverable degradation like erosion, corrosion, blade tip rubs or seal clearance damage. It could also be in line with the engine certification requirement, where the consumption of inter-turbine temperature (ITT) or exhaust gas temperature (EGT) margin signifies the end-of-life or determines time-to-failure. For this case study investigation, the threshold measure of minimum permissible component health parameter indices was adopted.

Trends of health indices from previous run-to-failure cases of the same engine or similar engines in the fleet form a degradation pattern library to which the current engine degradation scenario is compared. Curve-fitting of the exponential degradation model to the test engine HI series produces the model parameters as coefficients and the measures of statistical fitness are used to evaluate the trends in the pattern library which best describes the history of the test case for prediction.

The best fitting models from the pattern library is selected based on two criteria:

1. The training models with initial HI values in the range of those for the test data are pre-selected,
2. Two statistical goodness-of-fit parameters – root mean squared error (RMSE) and the coefficient of determination ( $R^2$ ) – which are expressed in Eqs. (12) and (13) below, are used for the final selection.

$$RMSE_i = \sqrt{\frac{\sum_{t=1}^n (\widehat{HI}_t - HI_t)^2}{n}} \quad (12)$$

$$R_i^2 = 1 - \frac{\sum_{t=1}^n (\widehat{HI}_t - HI_t)^2}{\sum_{t=1}^n (\overline{HI}_t - HI_t)^2} = 1 - n \cdot \frac{RMSE_i^2}{SS_{total}} \quad (13)$$

Where  $t = 1, 2, \dots, n$  is the time in cycles,  $HI_t$  is the computed test unit health index (observation),  $\widehat{HI}_t$  is the predicted HI by the fitting model  $i$  from the training library,  $\overline{HI}_t$  is the mean of  $HI_t$  and  $SS_{total}$  is the total sum of squares (proportional to the variance of the data).

The final RUL of the test unit is therefore calculated as the weighted average of the RUL values obtained by subtracting the current time of the test unit from the end-of-life of each selected train unit as shown in Eq. (14). The weight is evaluated using an inverse of the RMSE in Eq. (15).

$$RUL_j = \sum_{i=1}^n w_i (eol_i - t_{p,j}) \quad (14)$$

$$w_i = 1 / \left( \frac{\sum_{k=1}^T e_{k_i}^2}{\sum_{i=1}^n \sum_{k=1}^T e_k^2} \right) \quad (15)$$

Where  $i = 1, 2, \dots, n$  represents the selected training models, and  $k = 1, 2, \dots, T$  represents the time step in each data series.  $w$  is the corresponding weighting factor based on the

deviation at each time step  $e_k$ , and  $eol$  is the final cycle.  $t_{p,j}$  is the current/prediction time for the test unit  $j$ .

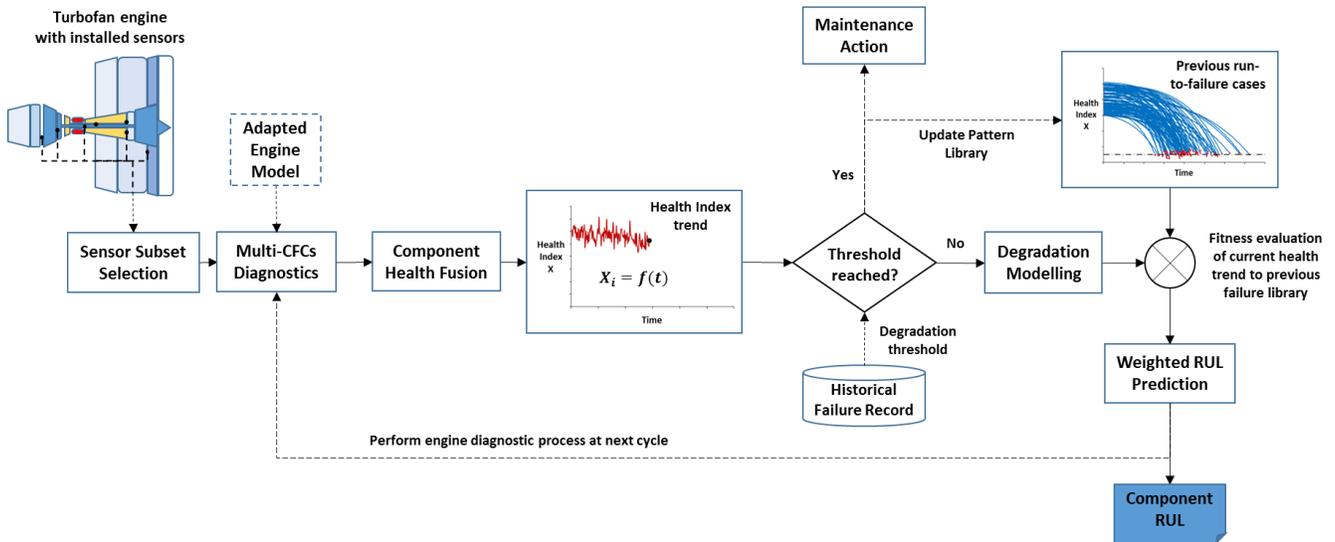


Figure 6: Flow chart for RUL Prediction

### 3. CASE STUDY

The case study being investigated is a 2-shaft high-bypass ratio turbofan engine, propelling a civil airline aircraft. Simulated data for 21 sensors from multiple engines of similar configuration, belonging to a fleet for example, is available as a multivariate time series of engine operation in cycles, where one cycle may refer to a certain number of flights or flying hours. The data is grouped into 100 training and 100 test sets, and indicate different levels of initial deterioration and different trends of gradually increasing degradation over time (Saxena & Goebel 2008a).

In the training dataset, the degradation progresses until a threshold is reached, where the engine is deemed inoperable or the affected component has failed. In the test set, sensor data is available until some point, assumed as the current time prior to engine component failure. Both training and test cases comprise data obtained at different operating conditions and for different fault modes, contaminated with random sensor noise. When the failure threshold is reached, the trend of the current engine HI is added to the pattern library such that the robustness of the RUL prediction of the system improves over time.

#### 3.1. Model Adaptation

Thus, the thermodynamic model of the CMAPSS engine was first built using the Cranfield University gas turbine performance simulation and diagnostics software, PYTHIA. This model was then adapted to match the first set of sensor

data in the FD001 training dataset – assumed as the clean engine output – using information available in (Decastro et al. 2008; Frederick et al. 2007) as initial component parameter specifications. Table 2 below shows the values of some of the target measurement parameters after design point adaptation.

Table 2: Design point adaptation results.

Measurement Parameter	Symbol	Target	Unit
Total temperature at Fan inlet	T2	288.15	K
Total temperature at LPC exit	T24	356.57	K
Total temperature at HPC exit	T30	883.17	K
Total temperature at LPT exit	T50	778.11	K
Total pressure at Fan inlet	P2	0.9948	atm
Total pressure in bypass-duct	P15	1.4705	atm
Total pressure at HPC exit	P30	37.722	atm
Total pressure at LPT exit	P50	1.2933	atm
Inlet air mass flow	W1	1246.6	kg/s
Combustor fuel flow	Wff	3.1201	kg/s
HPT coolant bleed	W31	17.717	kg/s
LPT coolant bleed	W32	10.623	kg/s
LP shaft speed	PCN1	2388	rpm
HP shaft speed	PCN3	9046.2	rpm
Corrected HP shaft speed	CN3	8138.6	rpm

### 3.2. Optimum Sensor Selection

The sensitivity analysis revealed the most sensitive sensors to a given health parameter fault as shown in Figure 7. The overall sensitivity of each sensor is summarized in the sensitivity norm in Table 3, where fuel flow is seen to have the highest value, hence the most sensitive parameter.

In the correlation matrix in Table 4, a relatively high element value denotes a correlation between the two corresponding sensors. For example, it can be seen that HPC exit mass flow W30 is highly correlated to the bleed flows W31 and W32.

Thus, the latter were discarded from further analysis without influencing the fault detection capability.

A total of 7 sensors – T30, T50, P15, P30, P50, PCN3 and Wff – were selected as the optimum subset based on the above criteria. These correspond to sensors 3, 4, 6, 7, 9, 10 and 12 respectively as denoted in the dataset., and share some similarity to those obtained using other methods in Xu et al. (2014), Ramasso (2014) and Wang et al. (2008).

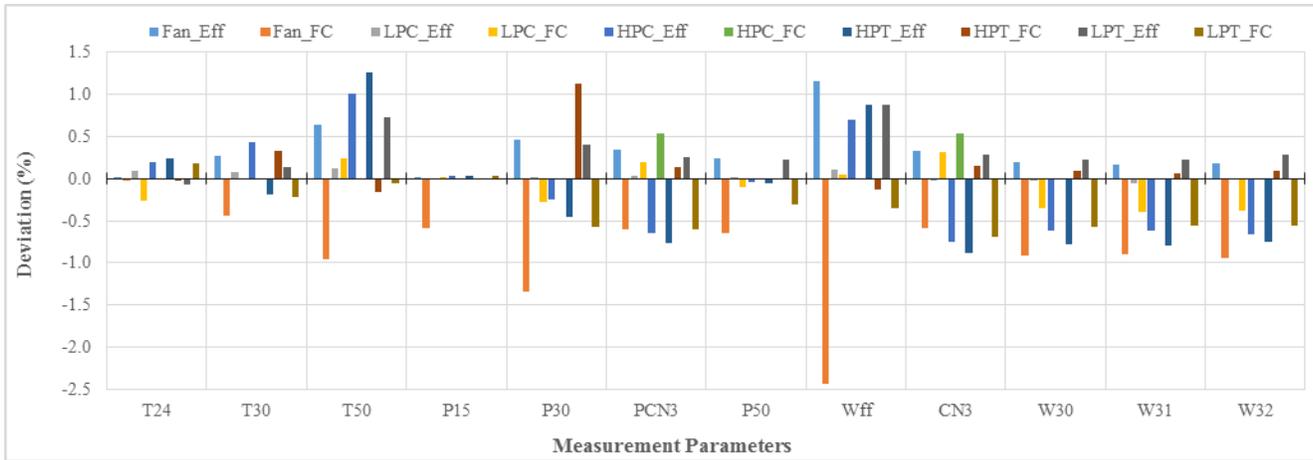


Figure 7: Measurement Sensitivity

Table 3: Sensitivity Norm

Sensor	T24	T30	T50	P15	P30	PCN3	P50	Wff	CN3	W30	W31	W32
<b>Sensitivity Norm</b>	0.455	0.820	2.137	0.589	2.038	1.504	0.794	3.081	1.664	1.546	1.544	1.571

Table 4: Correlation Matrix

Sensor	T24	T30	T50	P15	P30	PCN3	P50	Wff	CN3	W30	W31	W32
<b>T24</b>	1.00	0.01	0.43	0.11	-0.23	-0.68	-0.13	0.21	-0.75	-0.45	-0.44	-0.44
<b>T30</b>	0.01	1.00	0.49	0.53	0.75	0.35	0.67	0.67	0.31	0.41	0.39	0.40
<b>T50</b>	0.43	0.49	1.00	0.52	0.20	-0.18	0.48	0.85	-0.23	-0.16	-0.17	-0.14
<b>P15</b>	0.11	0.53	0.52	1.00	0.61	0.32	0.77	0.82	0.28	0.50	0.50	0.51
<b>P30</b>	-0.23	0.75	0.20	0.61	1.00	0.66	0.80	0.59	0.63	0.78	0.76	0.78
<b>PCN3</b>	-0.68	0.35	-0.18	0.32	0.66	1.00	0.64	0.26	0.99	0.85	0.83	0.84
<b>P50</b>	-0.13	0.67	0.48	0.77	0.80	0.64	1.00	0.85	0.60	0.79	0.78	0.80
<b>Wff</b>	0.21	0.67	0.85	0.82	0.59	0.26	0.85	1.00	0.20	0.36	0.35	0.38
<b>CN3</b>	-0.75	0.31	-0.23	0.28	0.63	0.99	0.60	0.20	1.00	0.83	0.82	0.82
<b>W30</b>	-0.45	0.41	-0.16	0.50	0.78	0.85	0.79	0.36	0.83	1.00	1.00	1.00
<b>W31</b>	-0.44	0.39	-0.17	0.50	0.76	0.83	0.78	0.35	0.82	1.00	1.00	1.00
<b>W32</b>	-0.44	0.40	-0.14	0.51	0.78	0.84	0.80	0.38	0.82	1.00	1.00	1.00
<b>Correlation Norm</b>	1.71	1.92	1.68	2.03	2.38	2.40	2.51	2.10	2.36	2.53	2.50	2.53

### 3.3. GPA Diagnostics

Given 5 degradable components and a maximum of two simultaneously degrading components, a total of 125 possible fault combinations would have been analyzed if the actual faulty components were unknown. For this study however, only 3 fault cases were considered based on the disclosed potential faulty components – Fan and HPC.

The average GPA indices for each fault case when applied to the FD001 dataset is shown in Table 5. Case 2 is seen to have the highest value as expected, since it was stated as the site of the implanted faults (Saxena & Goebel 2008b), while Case 1 has a very low index value since it gives a wrong prediction when diagnosed. This validates the GPA method as a reliable means of isolating and predicting component faults in a multi-component diagnostic analysis.

Table 5: Average GPA Indices for FD001 Unit 1.

Fault Case	Components	Linear GPA Index	Non-linear GPA Index
1	Fan only	0.332	0.363
2	HPC only	0.842	0.843
3	Fan + HPC	0.835	0.835

### 3.4. Predicted Health Trends

Figures 8 (a) and (b) show the run-to-failure plot for the GPA-predicted HPC efficiency values relative to the initial values of the reference engine unit, which has a starting value of 1.0, before and after applying the degradation model to all the training units. Despite the random noise implanted in the sensors, a gradual trend of performance loss is apparent in the health parameter over time in Figure 7a. It is also clear that while each case follows a different path from healthy engine to failure, they terminate at points normally distributed about a mean failure threshold. Applying the degradation model to each data trend generates model parameters which define the smoothed trends in Figure 7b, where the mean failure threshold was calculated as 0.9825. A similar data trend and model was also obtained for the relative HPC flow capacity index.

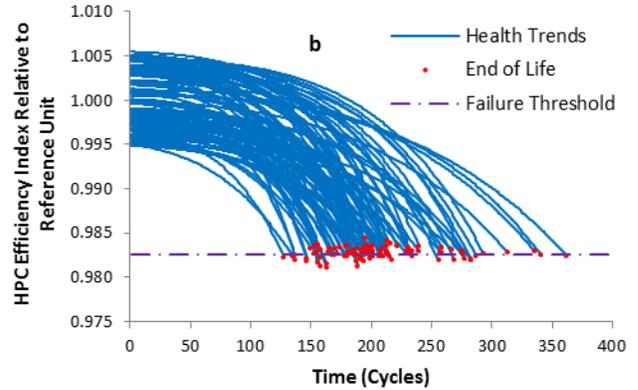
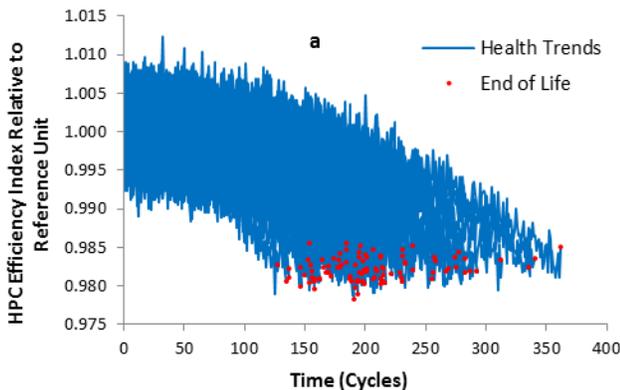


Figure 8: Relative HPC efficiency index run-to-failure plots of (a) diagnosed and (b) modelled trends for the train dataset.

A single health index was obtained from the normalized average of the relative HPC efficiency and flow capacity index degradation. This combined HI end-of-life value for all the training cases exhibited a skewed normal distribution about the mean value as shown in Figure 9, with a standard deviation of 0.058.

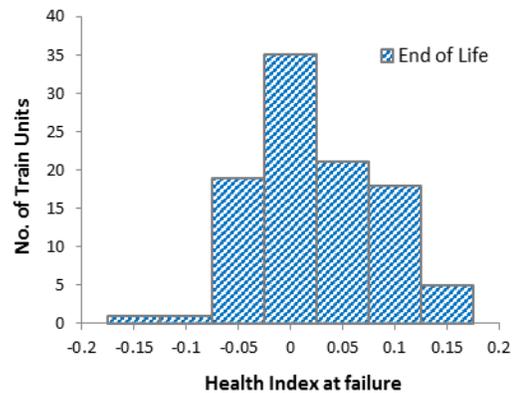


Figure 9: Distribution of the normalized combined health index at end-of-life for FD001 training run-to-failure data.

### 3.5. RUL Prediction

The final RUL for each test case was calculated as a weighted sum of the RUL predicted by the best matching set of cases in the training model library as described in Section 5. Figures 10 to 12 show some plots of test instances and their corresponding training models for three most popular types of prediction encountered – long-, mid- and short-term prediction cases respectively, where  $t_P$  is the time at which prediction is done and  $t_{EOL}$  is the predicted end-of-life from the weighted summation of the identified training cases.

It can be seen that more training data are required for fitness evaluation and RUL prediction of a test unit that has only operated for a short period of time (Figure 10) compared to one that has run for longer. This is because the limited

amount of data produces a trend which is not fully-formed, thus a larger number of training models might be considered as a match at the initial stages. The trend is however more visible for test units with longer time series data, resulting in fewer, more precise train models. Hence, test unit 25 with only 48 cycles of operation recorded 12 matching train models, test unit 51 had 7 matches, while for test unit 20 with 184 spent cycles and only 15 cycles left to failure, the number of training models used for prediction was reduced to 6. Intuitively, the data available for unit 20 was sufficient to fit a degradation model without need for the training models.

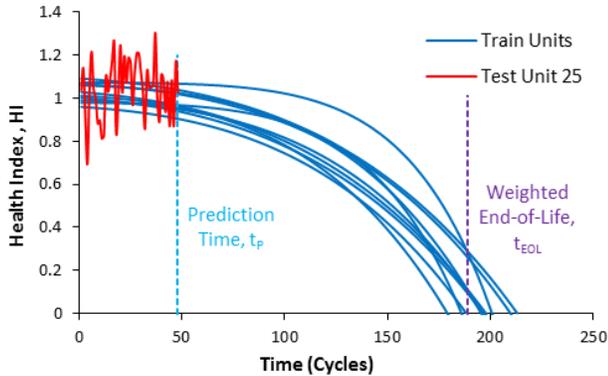


Figure 10: Training and test HI data series for unit 25.

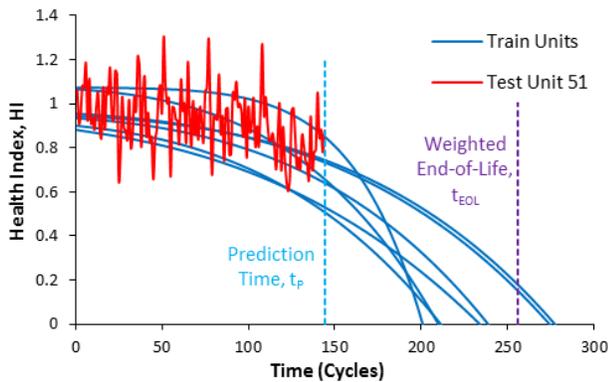


Figure 11: Training and test HI data series for unit 51.

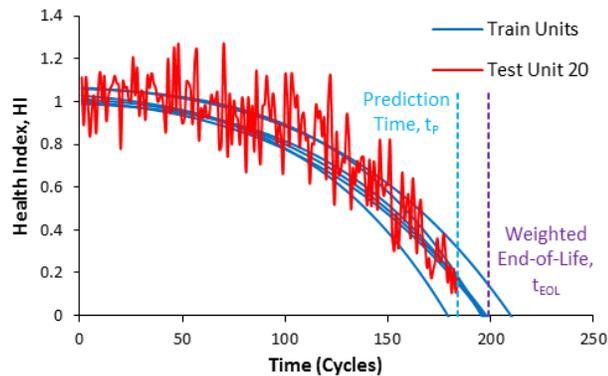


Figure 12: Training and test HI data series for unit 20.

A comparison between the relative accuracy of the predicted RUL using the degradation progression model types 1 and 2 is provided in the histogram in Figure 13. For both models, more than 85% of the predictions were within an error of 20 cycles above or below the true RUL. It can therefore be inferred that a prior knowledge of the degradation mechanism parameters may not necessarily improve the prediction accuracy, as either model type when used consistently for both the training and test cases would produce similar outcomes.

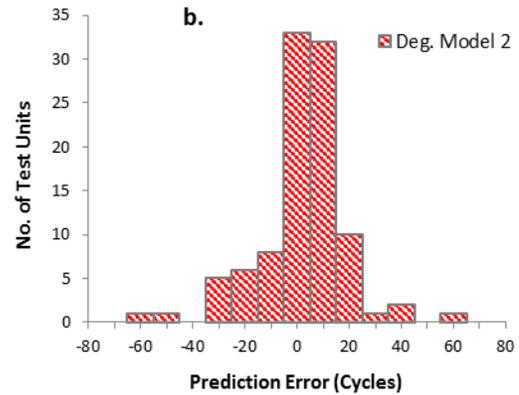
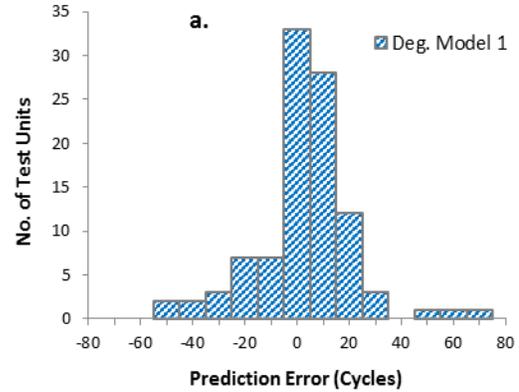


Figure 13: Distribution of RUL prediction errors for test cases using degradation models (a) 1 and (b) 2.

### 3.6. Prognostics Metrics

Using the inter-quartile range of the number of test data series for each unit compared to its overall life, it was possible to classify each case into short-, mid- and long-term prediction. Hence, in a short-term prediction, the number of data series would fall in the 4th quartile of the overall engine life, 1st quartile for long-term prediction and anywhere in-between was regarded as mid-term. It can be seen in Table 6 that the model-based approach is above 78% accurate across board for all prediction types. The reduced accuracy for the short-term prediction could be attributed to the large uncertainty arising from implanted noise, which was found to be of the same magnitude as the predicted RUL in some instances.

Table 6: Prediction accuracy according to prediction horizon for degradation model 1.

Prediction Type	Number of cases	Prediction Accuracy (%)
Short-term	4	78.3
Mid-term	64	86.5
Long-term	32	83.6

A quantitative analysis of the predicted RUL accuracy compared with the true RUL using various error metrics is provided in Table 7. These metrics are relevant because they make it possible to compare various prognostics techniques, based on the datasets to which they are applied. The PHM score, which is based on a scoring algorithm developed principally for the PHM 2008 Challenge competition, to

penalize late predictions more severely than early predictions is shown in Eq. 16.

$$S_{PHM} = \sum_{i=1}^N S_i \quad (16)$$

$$\text{where } S_i = \begin{cases} e^{-d_i/10} - 1, & d_i < 0 \\ e^{d_i/13} - 1, & d_i \geq 0 \end{cases}$$

$$d_i = \text{Estimated RUL} - \text{True RUL}$$

A final score of 1193 and 1355 was achieved for the FD001 test case, both using degradation models 1 and 2 respectively. These scores represent a measure of the RUL prediction errors, hence a higher value infers less precise or inaccurate predictions.

Table 7: Overall prediction scoring metrics of the test dataset for each degradation model.

Deg. Model	MSE	MAE	MAPE (%)	MSPE (%)	SD	R <sub>p</sub>	FPR (%)	FNR (%)	Accuracy (%)	PHM Score
1	335.94	11.97	17.94	2.57	18.33	0.912	52.0	46.0	82.8	1193
2	279.82	10.68	15.98	2.37	16.73	0.981	53.0	46.0	86.6	1355

#### 4. CONCLUSION

An adaptive model-based prognostic method for predicting the remaining useful life of similar degrading gas turbine engines was proposed and validated using the CMAPSS prognostics dataset. This approach comprised distinct methods for optimum sensor selection, fault isolation and quantification, and health index estimation. The technique of matching the health index data for a test unit by statistical goodness of fit parameters to dynamic degradation models from a training library of previous run-to-failure cases was shown to provide accurate predictions of RUL, without need for further pruning of the results. The presence of random sensor and process noise was mitigated by applying an outlier exclusion algorithm to the normalized HI data.

The approach showed capability for short-, mid- and long-term RUL predictions, even in the presence of random noise. An average prediction accuracy of over 80% was achieved using the default degradation model. The accuracy only change slightly when a different model was used. Hence, the choice of fitting model, though important from the perspective of obtaining the degradation mechanism's parameters, might not necessarily translate to significant changes in RUL prediction accuracy provided that the chosen model can fit the data properly. This is worth considering especially in real-life scenarios where multiple failure modes are in effect at any given time.

Further work based on this study would involve providing more robust predictions by quantifying the uncertainty

contributions from the various processes involved up to the final prediction step. Overall accuracy could also be improved by employing original engine performance information to build the model, using actual engine component maps during performance adaptation, and taking a methodical approach to random noise reduction, such as through non-linear state estimation filters.

#### NOMENCLATURE

<i>CFC</i>	Component Fault Case
<i>CM</i>	Condition Monitoring
<i>CMAPSS</i>	Commercial Modular Aero-Propulsion System Simulation
<i>CN</i>	relative corrected shaft speed
<i>FNR</i>	false negative rate
<i>FPR</i>	false positive rate
<i>G</i>	adaptation coefficient matrix
<i>h</i>	altitude
<i>H</i>	influence coefficient matrix
<i>HI</i>	health index
<i>HP</i>	high pressure
<i>LP</i>	low pressure
<i>Ma</i>	Mach number
<i>MAE</i>	mean absolute error
<i>MAPE</i>	mean absolute percent error
<i>MSE</i>	mean squared error
<i>MSPE</i>	mean squared percent error
<i>P</i>	pressure
<i>PCN</i>	relative physical shaft speed

<i>RMSE</i>	root mean squared error
<i>R<sub>p</sub></i>	Pearson's correlation
<i>RUL</i>	remaining useful life
<i>s</i>	sensitivity coefficient
<i>SD</i>	standard deviation
<i>t</i>	time in cycles
<i>T</i>	temperature
<i>W</i>	air mass flow rate
<i>W<sub>ff</sub></i>	fuel flow rate
<i>X</i>	component parameter
<i>Z</i>	measurement parameter

## REFERENCES

- Barad, S.G. et al., 2012. Neural network approach for a combined performance and mechanical health monitoring of a gas turbine engine. *Mechanical Systems and Signal Processing*, 27, pp.729–742. Available at: <http://www.sciencedirect.com/science/article/pii/S0888327011003797>.
- Baraldi, P. et al., 2013. Model-based and data-driven prognostics under different available information. *Probabilistic Engineering Mechanics*, 32, pp.66–79. Available at: <http://dx.doi.org/10.1016/j.probenmech.2013.01.003>.
- Biagetti, T. & Sciubba, E., 2004. Automatic diagnostics and prognostics of energy conversion processes via knowledge-based systems. *Energy*, 29(12–15 SPEC. ISS.), pp.2553–2572.
- Chen, M., Quan Hu, L. & Tang, H., 2015. An Approach for Optimal Measurements Selection on Gas Turbine Engine Fault Diagnosis. *Journal of Engineering for Gas Turbines and Power*, 137(7), p.71203. Available at: <http://gasturbinespower.asmedigitalcollection.asme.org/article.aspx?doi=10.1115/1.4029171>.
- Coble, J. & Hines, J.W., 2014. Incorporating prior belief in the general path model: A comparison of information sources. *Nuclear Engineering and Technology*, 46(6), pp.773–782. Available at: <http://dx.doi.org/10.5516/NET.04.2014.722>.
- Coble, J.B. & Hines, J.W., 2008. Prognostic algorithm categorization with PHM challenge application. *2008 International Conference on Prognostics and Health Management, PHM 2008*.
- Cubillo, A., Perinpanayagam, S. & Esperon-miguez, M., 2016. A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery. , 8(8), pp.1–21.
- Decastro, J.A., Litt, J.S. & Frederick, D.K., 2008. A Modular Aero-Propulsion System Simulation of a Large Commercial Aircraft Engine. , (September).
- DePold, H. & Gass, F.D., 1999. The Application of Expert Systems and Neural Networks to Gas Turbine Prognostics and Diagnostics. *Journal of Engineering for Gas Turbines and Power*, 121, p.607.
- Frederick, D.K., Decastro, J.A. & Litt, J.S., 2007. User's Guide for the Commercial Modular Aero-Propulsion System Simulation ( C-MAPSS ) October 2007. , (October).
- Heimes, F.O., 2008. Recurrent Neural Networks for Remaining Useful Life Estimation. *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pp.1–6. Available at: <http://ieeexplore.ieee.org.etechconricyct.idm.oclc.org/stamp/stamp.jsp?tp=&arnumber=4711422>.
- Jasmani, M.S., Li, Y.-G. & Ariffin, Z., 2010. Measurement Selection for Multi-Component Gas Path Diagnostics using Analytical Approach and Measurement Subset Concept. In *Proceedings of ASME Turbo Expo 2010: Power for Land, Sea and Air*. Glasgow, UK.
- Li, Y.G. & Nilkitsaranont, P., 2009. Gas turbine performance prognostic for condition-based maintenance. *Applied Energy*, 86(10), pp.2152–2161. Available at: <http://dx.doi.org/10.1016/j.apenergy.2009.02.011>.
- Mosallam, A., Medjaher, K. & Zerhouni, N., 2016. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, pp.1037–1048. Available at: <http://dx.doi.org/10.1007/s10845-014-0933-4>.
- Peel, L. & Gold, I., 2008. Data Driven Prognostics using a Kalman Filter Ensemble of Neural Network Models. , pp.1–6.
- Ramasso, E., 2014. Investigating computational geometry for failure prognostics. , pp.1–18.
- Ramasso, E. & Saxena, A., 2014. Performance Benchmarking and Analysis of Prognostic Methods for CMAPSS Datasets. , pp.1–15.
- SAE International, 2013. *Prognostics for Gas Turbine Engines - AIR5871*, Available at: <https://doi.org/10.4271/AIR5871>.
- Saha, B. & Goebel, K., 2011. Model Adaptation for Prognostics in a Particle Filtering Framework. , pp.1–10.
- Saxena, A. et al., 2008. Damage Propagation Modeling for Aircraft Engine Prognostics. *Proceedings of IEEE International Conference on Prognostics and Health Management*, pp.1–9.
- Saxena, A. & Goebel, K., 2008a. C-MPASS data set. *NASA Ames Prognostics Data Repository*.
- Saxena, A. & Goebel, K., 2008b. *Turbofan Engine Degradation Simulation Data Set*, Moffett Field, CA.
- Sikorska, J.Z., Hodkiewicz, M. & Ma, L., 2011. Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*, 25(5), pp.1803–1836.
- Wang, T. et al., 2008. A Similarity-Based Prognostics Approach for Engineered Systems. In *International Conference on Prognostics and Health Management*.
- Xu, J., Wang, Y. & Xu, L., 2014. PHM-oriented integrated

fusion prognostics for aircraft engines based on sensor data. *IEEE Sensors Journal*, 14(4).

Yu, J., 2017. Aircraft engine health prognostics based on logistic regression with penalization regularization and state-space-based degradation framework. *Aerospace Science and Technology*, 68, pp.345–361. Available at: <http://dx.doi.org/10.1016/j.ast.2017.05.030>.