Proceedings Paper:

Fu, R., Harrison, R.F., King, S. et al. (1 more author) (2016) Lean burn combustion monitoring strategy based on data modelling. In: 2016 IEEE Aerospace Conference. 2016 IEEE Aerospace Conference, 05/03/2016-12/03/2016, Big Sky, MT, USA. Institute of Electrical and Electronics Engineers . ISBN 9781467376761

https://doi.org/10.1109/AERO.2016.7500876

Lean Burn Combustion Monitoring Strategy Based on Data Modelling

Ruowei Fu
Department of Automatic Control and Systems Engineering
The University of Sheffield
Mappin Street
Sheffield, S1 3JD, UK
+44 (0) 114 222 5682
Ruowei.Fu@sheffield.ac.uk

Robert F. Harrison
Department of Automatic Control and Systems Engineering
The University of Sheffield
Mappin Street
Sheffield, S1 3JD, UK
+44 (0) 114 222 5139
r.f.harrison@sheffield.ac.uk

Steve King
Rolls-Royce Controls and Data Services Limited
Sinfin D-11, PO Box 31,
Derby, DE24 8BJ, UK
+44(0) 1332 777 129
steve.king@controlsdata.com

Andrew R. Mills
Department of Automatic Control and Systems Engineering
The University of Sheffield
Mappin Street
Sheffield, S1 3JD, UK
+44 (0) 114 222 5634
a.r.mills@sheffield.ac.uk

Abstract—New designs of gas turbine lean burn combustors are under development to deliver lower emissions. To identify deterioration of combustion performance and engine health due to the increased complexity in these lean burn fuel system, one solution is through monitoring variation in Turbine Gas Temperature (TGT) profile. In this work, a data-driven monitoring strategy is designed and a prediction model for TGT associated with other crucial parameters is constructed. Due to limitations on sensing techniques and constraints on weight, only a limited number of TGT measurements downstream of combustion system are feasible in production engine, this along with gas swirling effects through the turbine, reduces the magnitude of temperature anomaly caused by an incipient fault. The model must meet EHM requirements on accuracy and sensitivity of the TGT monitoring model, be robust to influence of environmental changes. To accommodate these requirements, an adaptive model structure is proposed. A data-driven modelling framework with complexity control strategies for both a linear and a non-linear model are developed. The risk of overfitting is controlled by hyper-parameter optimization and cross-validation. The models are trained using data collected from combustor rig tests and test bed experiments. The fault mode behaviour is validated by augmenting the rig data with computational models of fault behaviour. Results show that with suitably selected range of data, and the application of the presented modelling framework, that a linear in parameter model provides an effective monitoring solution for lean burn systems. The adaptive modelling framework presented is also applicable to general data modelling tasks.

TABLE OF CONTENTS

1. INTRODUCTION

Aircraft equipment health management (EHM), especially the practices of gas turbine engine monitoring have evolved in form and complexity for several decades [1]–[3]. The EHM techniques, which are responsible for assessing and tracking the asset's state of health, have significantly reduced costs by preventing or delaying maintenance of large civil engines, as well as identifying potentially costly technical problems. Ultimately the implementation of an EHM system increases asset availability and readiness while supporting safe operation [4].

The EHM system uses a range of on-board sensors to record key technical parameters such as pressure, temperature, shaft speed and vibration level during engine operation. For large civil gas turbine engines an Aircraft Condition Monitoring System (ACMS) is used to acquire a series of snapshot records at key steady-state points in flight for EHM. The ACMS reports are routinely transmitted from aircraft to ground, using a digital data-link system called Aircraft Communication Addressing and Reporting System (ACARS) via a VHF radio or satellite link whilst the aircraft is in-flight. Raw data measurements are transmitted directly or after on-board signal processing, as ACMS data from each flight, which is then fed into the monitoring models to produce health indices corresponding to critical engine components. The health indices are trended, and subtle changes in indices can be detected. Integrated with historical fault signature library and expert knowledge, potential faults will be diagnosed, and inference on the most likely physical cause of a particular fault, as well as recommendation on how urgently the inspection and maintenance needs to be carried out, will be made. Today, health indices are typically generated from a model standardised across a fleet of aircraft, which is not sensitive to the effects of variability between units.

There is a need to continue development of EHM systems so they may accommodate new product subsystems. Lean burn systems provide environmental benefits, but entail increased complexity and novel potential failure modes. A larger number of valves in fuel systems and more complicated fuel scheduling and splitting strategy makes fuel and combustion systems susceptible to potential faults in demanding operating conditions. Therefore, the operation of lean burn system requires closer monitoring to mitigate abnormality propagation and to prevent operational disruption. As the combustion and fuel system constitutes a critical module of the lean burn engine, in this work we focused on developing monitoring capability for faults occurring in these components.

Physical faults result in changes in observable engine parameters, such as pressure and temperature, which are measured along gas path. In these circumstances a data based monitoring model is more suitable than a traditional physics-based model, without the difficulties of precisely formulating the combustion and heat transfer process physics through the turbine. Essentially the monitoring problem can be characterized as comparing the estimated value of key operating variables with their measurement to distinguish fault from normality. For combustor faults the temperature profile in the combustor exit plane and downstream has been identified as a direct indication.

Various data modelling methodologies and techniques have been applied in modern EHM systems on gas path measurements to monitor, detect, predict and trend degradation, fault and failures of the engine and components. Conventional data modelling and estimation methods, e.g. Kalman filtering, polynomial curve fitting, and multivariate statistical analysis algorithms have been considered for detecting and isolating component fault [6]–[8]. Machine learning techniques such as Logistic Regression (LR), neural network (NN), Support Vector Machine (SVM), Hidden Markov Model (HMM), as well as combinations of these are more often adopted to not only discriminate fault from normal, but also diagnose possible defects [5], [9]–[12]. Despite widespread use in academic publications, application issues concerning how to select appropriate data set and model, as well as the impact of data and model selection on performance are not systematically addressed in the literature.

Data based machine learning methods are generic in nature, but the EHM system design must be tailored to each application. Considering that over a thousand engines from a fleet need to be monitored in real time, the modelling and monitoring need to be computationally efficient and compact. Parsimonious models are expected to be trained autonomously to ease the deployment difficulty. Sample size for acquiring a satisfactory model is generally proportional to the model structure complexity. Considering that only snapshots of measurements at certain operational conditions are available in service, to minimise the interval between initial flights and a well-established monitoring model, a successful monitoring system demands that the model structure be optimized.

Knowledge about the detailed physics of how the gas heat transfer process occurs within the combustor and through the turbine is limited and difficult to model, thus the physics based model predicting temperature distribution changes does not exist. Tarassenko et al. proposed a novelty detection approach, learning a linear and a neural network predictor of normality from the normal class data [5]. We reassess necessity for NN with lean burn data, and also note his approach to controlling complexity through NN structure (varying number of hidden units) rather than regularisation, which can be a challenge. Compared to the previous industrial study, further evaluation of model structure is recommended. Allegorico's work [12] focused on using machine learning classification algorithms to discriminate distorted exhaust gas temperature pattern. This strategy is not suitable for our monitoring strategy, because currently only the normal data is available, and no historical faulty patterns exists to support classification training. Monitoring capability is expected to be initialized soon after the new lean burn engine has been operating for a few flights. In addition, the convergence issue for nonlinear models such as neural network and logistic regression and systematic guidelines for optimizing model structure and complexity is not fully addressed in prior art studies.

Work concerning the design of monitoring strategy for a lean burn combustion system is presented in this paper. The features of Turbine Gas Temperature (TGT) profile and its monitoring approach is proposed in Section 2. In Section 3 the top-level design procedures, as well as strategies accounting for enhancing model efficiency and robustness, resolving application difficulties are described. Section 4 provides a general description on the modelling framework and selection of model complexity control strategy. Results are discussed in Section 5 and the efficacy of implementing the modelling technique on lean burn combustion data is demonstrated. Section 6 provides a summary of the lean burn EHM strategy.

2. TGT PROFILE MONITORING

TGT Profile

Engines with fully annular combustors have a uniform exit temperature during normal operation. Temperature near combustor exit deviating from nominal values is a direct indication of combustion anomalies, for example fuel injector blockage or valve stiction. However, due to limitations on sensing techniques due to extreme high temperatures, only a limited number of temperature measurements downstream of combustion system are feasible in production engine. A sensor apparatus, consisting of multiple thermocouple probes, is arranged at the Low Pressure Turbine (LPT) Nozzle Guide Vanes (NGVs), circumferentially around LPT inlet. The multiple TGT harness system gives either a quadrant or individual measurements of the circumferential temperature distribution of the turbine gas flow, i.e. the TGT profile.

TGT profile provides an indirect observation of the combustion process, but the estimate of the combustor temperature is compromised by the gas mixing as it flows through the rotating turbine. The associated turbulence effect makes the prediction of heat transfer and temperature distribution at TGT position even harder. The "swirl effect" caused by turbines depends on shaft speed and would cause small nonlinear distortion to temperature signature. Hence added model complexity is needed to meet EHM requirements on accuracy and sensitivity of the TGT monitoring model.

Operating Conditions

It is known that measures such as inlet temperature (T30) and pressure (P30), shaft speed (NH) and altitude define the flight condition at which the engine is operating, and thus the TGT profile. As a consequence these measures are included as inputs of the monitoring model in order to let the monitoring model be robust to operating conditions variations.

EHM system aims to detect faults that emerge from deterioration mechanisms and develop slowly along time, thus there is no advantage to demand continuous measurements for monitoring in this application. A monitoring strategy, which continuously measures system health over the whole flight envelope, inevitably requires consideration of dynamics and nonlinearity, which complicates the model. Monitoring an engine's status in a periodic manner is sufficient, and is suited to today's data transfer protocols.

The three most opportune conditions for detecting combustor related faults are identified, namely during take-off, climb and cruise. These conditions correspond to specific points in the fuel staging strategy of lean burn operation. The best operating point for detection is when the deviation in temperature profile due to fuel distribution tolerances on parts such as fuel splitting valve is at a minimum. This occurs when there is a high fuel flow through either the pilot or mains. The maximum pilot fuel flow occurs during take-off where the pilot is operating in isolation (100% pilot split). The maximum fuel flow through the mains occurs during climb.

Discussions about whether to adopt a unified model for the three specific conditions, or three individual models, are elaborated in the following section.

Lean Burn Monitoring Strategy

TGT profile in normal condition provides a signature of temperature distribution, associated with operating condition at that time. When a fault develops in one of the combustion chambers, there is a local effect on the temperature profile; only a small number of adjacent thermocouples are significantly affected whilst the rest of the profile is largely unchanged [5]. The monitoring strategy is to construct a predictive model based on normal data, making use of the spatial correlations inherent in the TGT profile, combined with context parameters such as shaft speed, fuel flow etc. to indicate operating conditions.

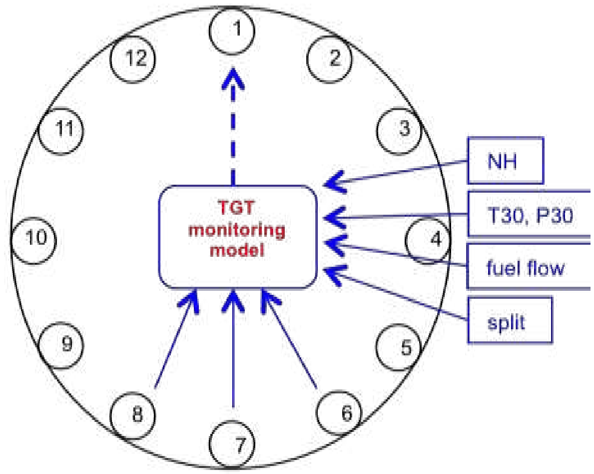

Figure 1 Lean Burn TGT profile monitoring strategy

As shown in Figure 1, the model of normality is trained by learning the function f_n between the target thermocouple reading (T_n) and the temperature values from the three thermocouples in opposite direction, together with shaft speed (NH), inlet temperature and pressure (T^{in}, P^{in}), fuel flow (m_f) and pilot/mains fuel split ratio (s),

$$\boldsymbol{\hat{T}_n = f_n(T_{n+i}, T_{n+i+1}, T_{n+i+2}, NH, T^{in}, P^{in}, m_f, s)}, \quad (1)$$

where T_{n+i} to T_{n+i+2} represent value from opposite thermocouples. Multiple regression models (one per thermocouple) are obtained by training with data collected under normal conditions, providing complete coverage of temperature profile prediction around annulus.

In normal condition, the estimation of model obtained from a finite sample contain an error, thus each model will give a prediction error defined as:

$$e_n = T_n - \overline{T}_n. \quad (2)$$

The statistical distribution threshold $[e_{LL}, e_{UL}]$ for prediction error in normal condition with certain confidence is calculated to characterize normality – it is shown that prediction error follows the Hotelling's T^2 distribution. After learning, new test data can be fed into the model. Prediction error of test data will be compared with the distribution, to determine whether the new prediction is significantly deviated from intervals of trained error. Off-nominal state of TGT profile is thus detected and alert will be issued.

3. Design Procedures

There is much more to data modelling than selecting an algorithm and applying it to the data. The art of modelling from data mixes the insight and design choice of the model designer with the information contained in the observations. Typically, there are many analyses and practical issues to be addressed to help achieving a satisfactory model, and a modelling process is better represented as an iteration of several steps, through refinements at each iteration, see Figure 2. After the monitoring problem is formulated, general procedures of design and developing the monitoring model are highlighted in this section. They include data generation and selection, model selection, training and validation, as well as deployment. The model structure decisions and data selections made in the first few steps may be readjusted according to results from model validation and verification. The iteration of adaptation ensure that the obtained model achieve expected performance. The following are all the issues we need to consider and shall address in our design.

Figure 2 General design procedures for data modelling

Data and Model Requirements

In application, the normal profile signature and monitoring model ought to be derived from data from a sufficient number of flights prior to or early after entering into service, and then continue to use the model to monitor the engine with this trained model, throughout its operation whilst it is deemed to be valid. The size of training set should be as small as possible to reduce the period during which the TGT monitoring module being inactive, whilst delivering the required certainty over a sufficient operating range.

The monitoring model needs to accommodate a wide range of engine operation. External conditions (e.g. aircraft load, ambient day condition, flight route and distance, take-off thrust chosen by operator etc.) might vary significantly so the model should be robust to those external factors, but sensitive to internal changes due to faults.

Synthetic Data Generation

In the current programme, the design and development of a new lean burn engine and its EHM system is ongoing in parallel. At the moment there exists no ability to acquire genuine data from real flights, or from a fleet of assets. Generating synthetic data for EHM model training and verifying the model capability through engine life is recommended. Augmented rig data could serve as a type of synthetic data. In addition, one task is to generate synthetic faulty data based on a combination of modelling, expert knowledge and (fault-free) engine testbed data. Given that the current available rig test data does not relate to failure cases, fault seeded synthetic data is generated based on knowledge of engine operation and the propagation mechanism of fault cases, as well as observations of faulty data from other lean burn platform.

At the current development stage, data from only one single test bed engine is available. Another important aspect is to produce profiles under various conditions to characterise model capability over variations and changes due to uncontrollable factors (engine to engine variation, environmental changes, component deterioration etc.). Historical in-service flight records are able to provide knowledge on the magnitude of all kinds of variations.

Data Selection

In this work, the TGT monitoring model is essentially a static regression model, so only steady state data should be selected for training and testing.

Three steady states at which monitoring is going to be carried out have been identified. There is a choice to be made: either a unified ‘global’ model, or three specific ‘local’ models targeting ‘take-off’, ‘climb’ and ‘cruise’ respectively, ought to be constructed. The coverage of data used for training is a significant factor in this decision.

There is a trade between model local accuracy and robustness over a full operational range. With a local model, a precise description of behaviour is achievable with a model of a given structure. However, since in deployment a model is likely to experience test data beyond training region, the model has to extrapolate into unknown regions. A global model captures the general trend in the whole range and might not be as flexible to fit every particular signature, but robust to the influence of variation across operating range.

In this paper, models using both applicable ranges are constructed through selecting local and global datasets, and their performance are compared to support the decision making on training data range selection.

Model Selection

Without knowing the exact underlying function of the data, the selection of an appropriate model structure is a central issue of data modelling. The selection of model type, e.g. linear/nonlinear, parametric/non-parametric, is based on the data characteristics and constraints on implementation. There is an open question we addressed in this paper, which determines whether it is sufficient to use a linear-in-the-parameter (LIP) model, or a neural network model. Comparing model performance on data will give answer to selecting the best model structure; the general principle followed is to always explore a simple, linear model in the first instance.

Initial data visualization showed weak non-linearity in the steady state data. Therefore, a multivariate polynomial regression model is chosen as a starting point for the TGT profile modelling problem. To verify whether the precision of prediction can be improved, if taking into account the nonlinear distortion of TGT signature due to the very small change of swirl angle caused by increasing speed, a multi-layer perceptron (MLP) type of neural network broadly based on a legacy study of an aero-derivative industrial engine [5] is used for comparison. As a representative nonlinear model, a neural network can approximate any nonlinear continuous function to any extent of accuracy.

As for the practical implementation restrictions, using a linear-in-parameter model provides certain advantages. Techniques for design, training, and analysing a linear model are well established. There exists an analytical solution of the linear-in-parameter model given selected order and terms, thus no concerns over the convergence or local optima, though exhaustive search or iterative methods are needed for model hyper-parameter selection.

Strategies for determining appropriated hyper-parameters for both polynomial and MLP model are elaborated in the framework of modelling.

Model Training

Regardless of model formation, the regression model training is essentially an optimization problem of minimizing the following error function from a finite set of data:

$$E(W) = \frac{1}{2}\|f(X,W) - y\|^2, \tag{3}$$

where regression model $y = f(X, W) + e$ relates the target value of output y to a function of inputs X and the unknown parameter vector W.

In the case of regularized neural network, a regularized error function is used:

$$\widetilde{E}(W) = E(W) + \frac{\rho}{2}W^TW, \tag{4}$$

where ρ represents the regularization coefficient, which is a hyper-parameter to be decided.

Model Verification and Validation

Model validation evaluates model performance such as whether the model fulfils its design specification, as well as whether it fits the usage scenarios. To guarantee users' confidence of the monitoring results, the model coverage, convergence, sensitivity and accuracy are of concern. For assessing the model's performance, e.g. whether it is overfitting, k-fold cross validation are actually embedded in the training process to cycle through all k subsets of the data without demanding extra data for validation. Though full verification and validation activities are not presented here due to the emphasis and length of this paper, coverage and robustness to operating range variation are applied in the presented analysis.

In our EHM system configuration TGT profile model is only one of several monitoring subsystems. Its prediction error is trended over time and diagnosed together with information from other subsystems to assess engine health. The use of the model output as a trended parameter allows an additional layer of robustness to be incorporated and for the prediction error to be compared against similar engines across the fleet.

Deployment

Once reaching a satisfactory model, the model structure and its hyper-parameters such as model order and terms for polynomial model, or number of hidden units and regularization coefficient for MLP model are fixed for the target fleet of engine. Only unknown parameters (regression coefficients of polynomial or weights of MLP) are to be trained individually based on data from each engine.

3. MODELLING FRAMEWORK

Multivariate Polynomial Model

Firstly a multivariate polynomial regression model such as

$$f(X,W) = \sum_{i_1\cdots i_D} w_{i_1\cdots i_D} x_1^{i_1} \cdots x_D^{i_D} \tag{5}$$

is employed, in which $x_1, \cdots, x_D$ are regression variables, $w_{i_1\cdots i_D}$ is the unknown coefficient and $i_1, \cdots, i_D$ are non-negative integers. It is able to provide enough degrees of freedom to fit with polynomial terms. Obtaining its coefficients is linear with respect to the objective function, which can be solved through a least-squares error problem.

Multi-layer Perceptron Model

The two-layer perceptron model with $(D + 1)$ dimensional inputs $x_0, \cdots, x_D$ (an additional constant input $x_0 = 1$), and M hidden units is with the following overall network function [13]:

$$f(X,W) = \sigma\left(\sum_{j=0}^{M} w_{kj}^{(2)} h\left(\sum_{i=0}^{D} w_{ji}^{(1)} x_i\right)\right), \tag{6}$$

where $h(\cdot)$ and $\sigma(\cdot)$ are activation functions of the 1st and 2nd layer, respectively, and all weights w_{kj}, w_{ji} are grouped together to form the weight vector W.

Mechanisms of Complexity Control

An overly flexible model is at risk of overfitting and having burdensome training data requirements. For overfitting, an over complex model may display perfect regression on training data, but produce large value of error with test data. For large models with many inputs and parameters, large, multiple parameter datasets are required; these impose a burden on communications and computing infrastructures. A compact and simple model is therefore a preferred choice, naturally mitigating overfitting issues. When more flexible model behaviour is required, a regularization method with cross validation is recommended. The regularised methods are applicable to both a polynomial and an MLP model.

In this design framework, for all models, the approach is to optimize hyper parameters in some way to control the model flexibility. The main advantage of linear-in-parameter model is that the weight optimization step has unique analytical solution so reduces difficulty slightly. A significant feature of the model complexity control approach is that it does not confine the complexity of model at the very beginning of modelling. The mechanism of model complexity control is embedded in the training through regularization and optimization of model ‘hyper-parameters’. Generally, this mechanism is applicable to all other data modelling tasks and regardless of the type of regression model; it is able to achieve a most appropriate model by this flexible mechanism.

For the polynomial model, complexity is controlled in two ways: cross validation and model order selection. The process of training is essentially based on solving a quadratic error minimisation problem for the parameters; its candidate hyper-parameters are from a discrete set (i.e. model order must be positive integers). Selecting the appropriate hyper-parameters is performed by repeating the process of solving a least square problem with respect to a range of model orders. Firstly the maximum order of polynomial terms is restricted (e.g. maximum order is no greater than five), then starting from the first-order, each model order is executed up to the maximum order. The optimal order for the model is selected if the performance indicator (PI), i.e. the normalized sum of squared error (NSSE), reaches its minimum with the corresponding order. In addition, as a designer’s judgment, if a lower ordered model is good enough and compared with which a higher ordered model does not provide a significant improvement in NSSE, we may still choose the lower ordered model.

For a NN model the situation slightly differs. For each choice of its hyper-parameters (e.g. number of hidden units), model training requires the solution of a non-convex optimization, of which only local optimum may obtained. If one NN produces smaller training error than another does, it is hard to say that the former one has better generalization than the latter one, because there are possibilities that the latter model only find a not-so-good local minimum. Strategy adopted in a regularized NN is to fix the number of hidden units a relatively large value so the network will overfit, then to use one dimensional regularization parameter optimization. This gives a bigger model than necessary but makes the development easier by avoiding optimization of hidden units number. Training the regularized neural network with the addition of a regularization term to the cost function prevents the weights from being too large. The strategy is also known as weight decay. The effective model flexibility and smoothness is decided by the regularization parameter ρ, which can be optimized within a heuristic range. This strategy enables automating the modelling process easily rather than requiring manual intervention.

The modelling framework for the polynomial model and the regularized MLP model is illustrated in Figure 3. Major steps for generating a model consist of: (1) Model initialization: set maximum model order, or a large enough hidden unit number and range of regularization parameter; (2) K-fold cross validation, calculate performance indicator – NSSE, loop with increased hyper-parameter; (3) Select the optimal value of model order or regularization parameter, retrain with the complete dataset, save the model and corresponding training data; and (4) Pass model to execute unit, executing model with new test data.

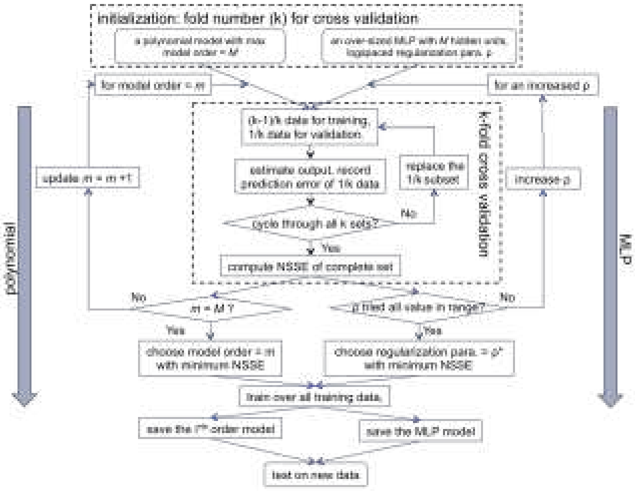

Figure 3 Modelling framework: training, cross validation and hyper-parameter optimization

4. RESULTS

TGT profile monitoring capability is described in this section based on data from a lean burn engine demonstrator and blocked burner test. A synthetic fault is added upon normal data. The results use T550, sensors fitted to rig test close to TGT. They are at slightly different position but the principle is the same. Data at steady state are extracted from the engine test data.

Data Characteristics

In normal condition the temperature distribution is reflected by measurements from multiple thermocouples, which constitutes a unique pattern for each engine. The pattern of eight T550 recordings divided by their mean value at three critical operating conditions are demonstrated in Figure 4, with take off shown in black, climb in blue and cruise in green. It shows that gas temperature is relatively uniformly distributed in normal condition, with about $\pm 3\%$ section-to-section variations compared with its mean value. However, a fault in combustion process changes the uniform pattern, as the two red lines in Figure 4 portrays the faulty pattern caused by 5% and 15% partial blockage at take off, respectively. The degree of pattern deviation depends on the severity of the fault. For an incipient fault, e.g. 5% blockage, the changed pattern lies on the margin of the normal temperature band (ring area between the dashed lines), but not exceed it, thus fault of this magnitude is not distinct. The prediction model developed is targeted at providing higher accuracy and differentiation than a simple profile radar plot can provide.

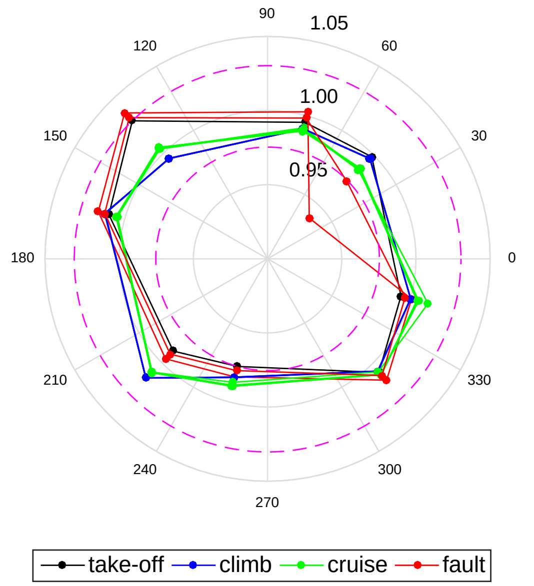

Figure 4 Radar plots of normalised T550 profiles

Due to the difficulty in visualizing the linearity of high dimensional data, we explored non-linearity in data by fitting a linear and a nonlinear model and analysing goodness of fit.

Fault Synthesis

In the lean burn engine development tests only fault-free data is obtained from a test rig. To derive an understanding on monitoring performance of the model, i.e. how model prediction error shift when there is an abnormality, multiple sources are combined together to generate synthetic faulty data for testing.

Computational fluid dynamic (CFD) simulation of combustor exit temperature profile T40 in normal and with various fault levels is available. A heuristic concerning T40 to TGT mapping is adopted, based on turbine experts' experience, in order to allow faulty T40 to be projected onto T550 measurement plane. In addition, data from another similar industrial lean burn platform with seeded incipient faults (5% and 10% blockage in fuel injector) is also analysed and referenced to support the faulty data synthesis.

At the moment the new lean burn engine have very little normal data to train fault detection algorithms, and no data from real fault. Researchers are left to rely on data from simulations and prototype rigs, as well as experience from previous similar systems, to generate synthetic data. Considering only limited sensors are implemented to capture the change in temperature distribution, we simplify the symptoms of the combustion fault, and synthetic data are generated with magnitude changes in one of the TGT measurements. Tests with the synthetic faulty data make sure that the models are able to differentiate anomaly from normality as long as the pattern of the temperature profile changes.

Choice of Model Applicable Range

Data from a small range consisting of a single steady state of shaft speed with variations from independent runs, as well as from a large range covering several steady states are selected for training the model, respectively. The spread of NH is extracted from the steady state regions of the complete engine running profile, as shown in Figure 5. In the figure the "range A" training data set with shaft speed around 78%, represents one of the three operating conditions. This is opposed to "range B", which consists of data from five steady conditions. A 'local' model is trained based on data from range A and a global model on range B data. The performances of the two models are compared in the following section.

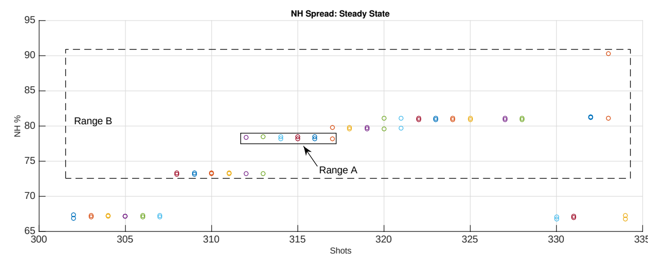

Figure 5 Spread of shaft speed at steady states (derived from engine running profile)

Polynomial and RNN Model Selection and Training

In total 4 monitoring models, i.e. a polynomial model and a RNN model on range A and B, are trained. In this work 200 points are randomly selected as training samples from both ranges, for both model types. In range B equal number of samples are selected from a group of steady states. In this subsection we focused on demonstrating the efficacy of complexity control framework and comparing the prediction accuracy of the two model types.

For the polynomial model, the maximum order of polynomial terms (including cross terms) is set to be five. Although slightly impacted by the randomly selected training data, a first or a second order model consistently provides the lowest NSSE in both range A and B, as illustrated in Figure 6. In this application model order higher than two could lead to a severe overfitting problem.

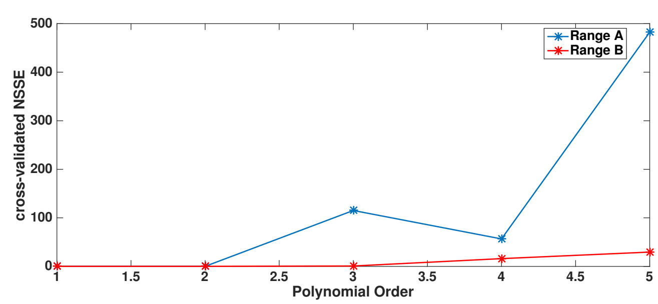

Figure 6 Polynomial model order selection

In this study the MLP model is obtained by using the scaled conjugate gradient (SCG) algorithm for optimization. Without loss of generality a linear activation function is chosen for output layer. Netlab toolbox is used in the core algorithms [14], [15]. Hidden units number is chosen to be 100 to ensure that an over-sized network is established. Regularization parameter is trialled with 100 logarithmically spaced values in the interval $[10^{-5}, 10^{2}]$. The optimal value for regularization parameter ρ is determined, by which the normalized sum of squared error is minimized, as shown in Figure 7.

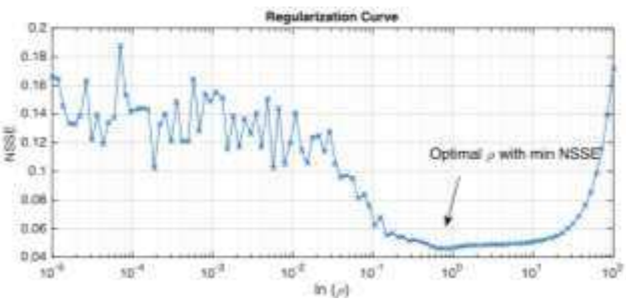

Figure 7 Regularization curve

As shown in Figure 8 and Figure 9, on either range A or B, polynomial and RNN model with the same sample size provide comparable prediction accuracy. Both models are able to predict targeting temperature accurate to less than 2 Kelvin in normal condition. The models are able to detect anomalies (deliberately induced synthetic faults) with the magnitude of 20 Kelvin deviation of T550, which represents 5% blockage of a fuel injector.

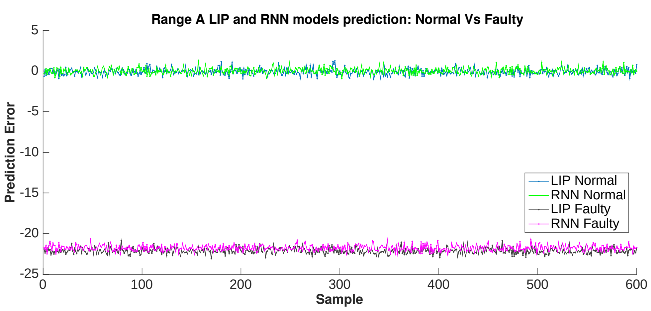

Figure 8 Prediction errors in normal and faulty condition (range A)

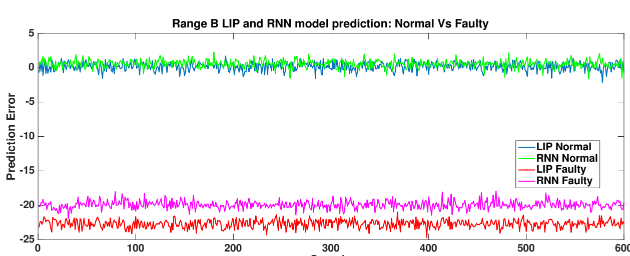

Figure 9 Prediction errors in normal and faulty condition (range B)

The result that a low order polynomial model achieves similar precision to RNN model may be attributed to the linearity of steady state data in the whole above-idle region.

Results for Inside Range Test Data

The two regression models have been constructed, it is important to confirm the goodness of fit of the model and the statistical significance of the estimated parameters. Commonly used checks of goodness of fit include the coefficient of multiple determination (i.e. R^2) and residual analysis.

With 600 normal testing data points chosen within the range of the training data, the residuals appear approximately normally distributed around zero, indicating that both the polynomial and RNN models on range A and range B describe the TGT profile well. The acceptable error margin derived from the histogram of residuals (as shown in Figure 10) in normal condition is ± 2K, except that a few prediction errors of RNN model on range B is larger than that due to outliers in test data. Robustness to outlier can be improved with simple model structure selection.

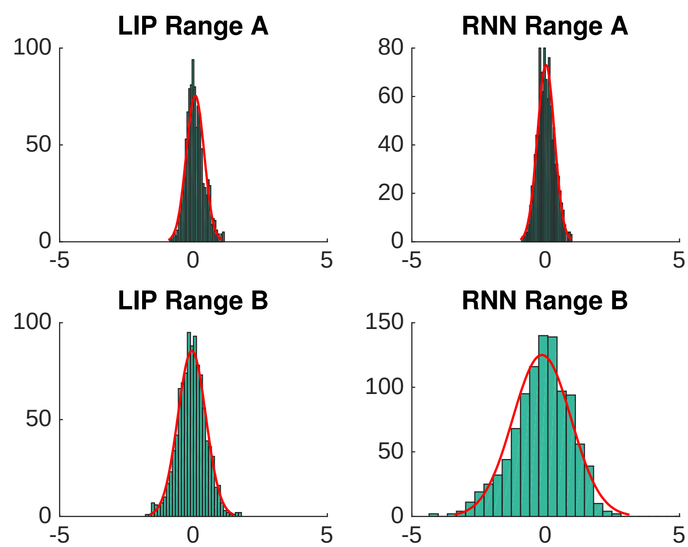

Figure 10 Distribution of prediction error in normal condition

The R^2 indicate the proportionate amount of variation in the response variable explained by the independent variables in the regression model. The R^2 of the polynomial and RNN model achieved in prediction are listed in Table 1. With a value closer to 1, R^2 indicating that a greater proportion of variance is accounted for by the model.

Table 1. R^2 Achieved by the Models

Polynomial (range A)	RNN (range A)	Polynomial (range B)	RNN (range B)
0.9574	0.9656	1.0000	1.0000

Results from Outside Range Testing Data

Both polynomial and RNN model display good prediction when test data is within the modelling range, i.e. the operating condition of test data has been covered by training set. However there is concern that data from real operations may be driven by external disturbances away from the specified steady state encountered in training stage, i.e. the flight profile and environmental conditions may not result in the engine passing through the specified operating points contained in training set. In application the model must be resilient to training – testing condition offset, which is very common in real flights caused by disturbances. To analyse the influence of training data range selection and the model's response on untrained states, performances of a local model associated with range A and a global model trained over range B are compared here. Steady states represented by NH speed in the two ranges for training and a state for testing are shown in Table 2.

Table 2. Shaft Speed (NH) as A Percentage of Maximum Speed for Range A and B Training Data and Test Data

Range A Training (% NH)	Range B Training (%NH)	Test Data (%NH)
$78.3 \pm 0.1\%$	$73.7 \pm 0.1\%$ $78.3 \pm 0.1\%$ $79.7 \pm 0.1\%$ $81.0 \pm 0.1\%$ $90.1 \pm 0.1\%$	$79.7 \pm 0.1\%$

All models are tested for their accuracy with a set of test data selected to have an approximately 2% greater NH value than range A. Test data from this power condition essentially interpolate range B models and extrapolate for range A models.

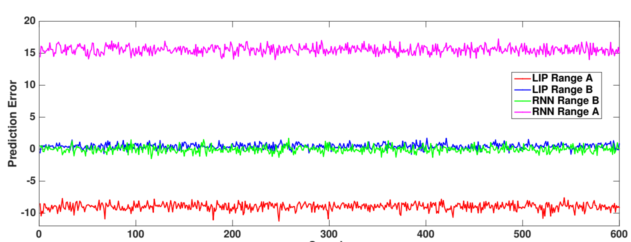

Figure 11 Prediction error of LIP and RNN model with out-of-range data

As indicated in Figure 11, prediction error of range A models are biased from zero, which could be easily confused with the effect of a blockage. Monitoring results of range A model might be vulnerable to minor changes in operating condition, as the simulation shows that extrapolating a model with a small applicable range is riskier than interpolating one with a wider range. While range B models provide a non-biased error even when the testing condition does not appear in training data but only within the range B interval.

The global model is robust over a larger range. A local model trained by data around a single set point display strong sensitivity to state mismatch. In other words, the model over a small range is only locally applicable in the sense that a positive result of anomaly detection is credible only if the test data is from a similar, comparable condition as training.

5. Conclusion

This paper demonstrates a robust and generic modelling strategy for detection of lean burn combustion abnormalities via TGT profile monitoring. Two predictive models are presented with an emphasis on model complexity control framework utilized in the TGT profile monitoring. Results demonstrated polynomial and RNN model are able to predict turbine gas temperature effectively, thus a range of combustor related faults in gas turbine engine can be detected by identifying temperature deviations from its nominal value in the annular profile of the TGT distribution. Based on modelling with rig data, the polynomial model on a large operating range is recommended, though it needs to be validated by real flight data.

Given the generic nature of the detection method, deployment on other platforms or targeting other engine parameters is also feasible, but will require issues to be revisited to address specific engine requirements and architecture. Its advantage in capturing complex behaviour may not be fully reflected in this application, but a RNN model can be used on other circumstances to deal other data modelling problem with strong nonlinearity.

To enhance the performance of future lean burn combustion monitoring models, uncertainties such as measurement error should be taken into account. Further studies concerning monitoring performance over application uncertainties are needed.

Acknowledgements

This work was funded by Innovate UK under the HITEA-II (Highly Innovative Technology Enablers for Aerospace, phase 2) project. The authors would like to thank Rolls-Royce plc for their support and guidance in this work.

References

[1] I. Y. Tumer and A. Bajwa, "A survey of aircraft engine health monitoring systems," Proc. AIAA, 1999.

[2] L. C. Jaw, "Recent Advancements in Aircraft Engine Health Management (EHM) Technologies and Recommendations for the Next Step," 2005, pp. 1–13.

[3] A. J. Volponi, "Gas Turbine Engine Health

Management: Past, Present, and Future Trends," J. Eng. Gas Turbines Power, vol. 136, no. 5, p. 051201, Jan. 2014.

[4] A. Volponi and B. Wood, "Aircraft Propulsion Health Management," in System Health Management: With Aerospace Applications, John Wiley & Sons Ltd, 2011, pp. 389 – 403.

[5] L. Tarassenko, A. Nairac, N. Townsend, I. Buxton, and P. Cowley, "Novelty detection for the identification of abnormalities," Int. J. Syst. Sci., vol. 31, no. 11, pp. 1427–1439, Nov. 2000.

[6] S. Borguet and O. Léonard, "Coupling principal component analysis and Kalman filtering algorithms for on-line aircraft engine diagnostics," Control Eng. Pract., vol. 17, no. 4, pp. 494–502, Apr. 2009.

[7] A. Kumar, A. Banerjee, A. Srivastava, and N. Goel, "Prediction of exhaust gas temperature in GTE by multivariate regression analysis and anomaly detection," in Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on, 2014.

[8] C. Kong, "Review on Advanced Health Monitoring Methods for Aero Gas Turbines using Model Based Methods and Artificial Intelligent Methods," Int. J. Aeronaut. Sp. Sci., vol. 15, no. 2, pp. 123–137, Jun. 2014.

[9] S. Menon, O. Uluyol, and K. Kim, "Incipient fault detection and diagnosis in turbine engines using hidden markov models," ASME Turbo Expo, Pap. No. GT2003-38589, pp. 493–500, 2003.

[10] R. A. Martin, M. Schwabacher, N. Oza, and A. Srivastava, "Comparison of unsupervised anomaly detection methods for system health management using space shuttle main engine data," in Proceedings of the Joint Army Navy NASA Air Force Conference on Propulsion, 2007.

[11] S. Lee, T. Roh, and D. Choi, "Defect diagnostics of SUAV gas turbine engine using hybrid SVM-artificial neural network method," J. Mech. Sci. Technol., vol. 23, pp. 559–568, 2009.

[12] C. Allegorico and V. Mantini, "A Data-Driven Approach for on-line Gas Turbine Combustion Monitoring using Classification Models," in European Conference of the Prognostics and Health Management Society, 2014, pp. 1–9.

[13] C. Bishop, Pattern Recognition and Machine Learning. 2006.

[14] I. Nabney and C. Bishop, "Netlab Neural Network Software." [Online]. Available: Http://www.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/.

[15] I. Nabney, Netlab: Algorithm for Pattern recognition. Springer, 2001.

BIOGRAPHY

Ruowei Fu received a B.S. in Industrial Automation in 2006, and a Ph.D. in Control Science and Engineering in 2012, both from Zhejiang University, Hangzhou, China. She has been a research associate in Control and Monitoring Systems Technology Centre sponsored by Rolls Royce, the University of Sheffield, working on developing intelligent adaptive EHM system for new generation of lean burn large civil gas turbine engine. Her research interests include data-driven modelling, machine learning, engineering system fault monitoring and diagnosis.

Robert F. Harrison received the B.Sc. and Ph.D. degrees in sound and vibration engineering in 1979 and 1983, respectively, from the University of Southampton, England. He is currently Professor of Computational Data Modelling in the Department of Automatic Control & Systems Engineering at the University of Sheffield, England. His main research interests are in the development and the application of machine learning and data modelling solutions to problems of practical interest in fields ranging from clinical and bio-medicine, communications, and transport. He is author of more than 80 journal articles and 150 conference papers in these and other areas.

Steve King is a Rolls-Royce Engineering Associate Fellow and EHM Specialist working within the Controls and Data Services part of the Rolls-Royce Group focusing mainly in the application of data mining and advanced analytical techniques in support of in-service issues. Steve holds a degree in Mathematics and Computer Science and a PhD in the application of expert systems for vibration analysis. In addition to being a Chartered Engineer, he is a fellow member of both the Institution of Engineering and Technology and the Institute of Mathematics and its Applications. He is also a Visiting Professor at Cranfield University.

Andrew R Mills, CEng, PhD, is a Senior Researcher and Research Programme Manager at the University of Sheffield. He has worked in the defence industry on aerospace and automotive applications before his current post within the University Technology Centre supported by Rolls-Royce. Research interests are in a broad range of aspects related to control and health management systems including: signal processing algorithms, instrumentation, advanced processing technologies and model based design.