CRANFIELD UNIVERSITY

Faisal Khan

Vehicle Level Health Assessment through Integrated Operational Scalable Prognostic Reasoners

School of Aerospace, Transport and Manufacturing
IVHM Centre

PhD Thesis
Academic Year: 2015 - 2016

Supervisors: Prof Antonios Tsourdos
Dr. Tarapong Sreenuch

Nov 2015

CRANFIELD UNIVERSITY

School of Aerospace, Transport and Manufacturing
IVHM Centre

PhD Thesis

Academic Year 2015 - 2016

Faisal M Khan

Vehicle Level Health Assessment through Integrated Operational Scalable Prognostic Reasoners

Supervisors: Prof Antonios Tsourdos
Dr. Tarapong Sreenuch

Nov 2015

Abstract

Today's aircraft are very complex in design and need constant monitoring of the systems to establish the overall health status. Integrated Vehicle Health Management (IVHM) is a major component in a new future asset management paradigm where a conscious effort is made to shift asset maintenance from a scheduled based approach to a more proactive and predictive approach. Its goal is to maximize asset operational availability while minimising downtime and the logistics footprint through monitoring deterioration of component conditions. IVHM involves data processing which comprehensively consists of capturing data related to assets, monitoring parameters, assessing current or future health conditions through prognostics and diagnostics engine and providing recommended maintenance actions.

The data driven prognostics methods usually use a large amount of data to learn the degradation pattern (nominal model) and predict the future health. Usually the data which is run-to-failure used are accelerated data produced in lab environments, which is hardly the case in real life. Therefore, the nominal model is far from the present condition of the vehicle, hence the predictions will not be very accurate. The prediction model will try to follow the nominal models which mean more errors in the prediction, this is a major drawback of the data driven techniques.

This research primarily presents the two novel techniques of adaptive data driven prognostics to capture the vehicle operational scalability degradation. Secondary the degradation information has been used as a Health index and in the Vehicle Level Reasoning System (VLRS). Novel VLRS are also presented in this research study. The research described here proposes a condition adaptive prognostics reasoning along with VLRS.

Acknowledgements

To have this thesis completed is a significant step in my life and career, and it could not have been achieved without receiving precious support from several people. Therefore, I would like to express sincere gratitude to my supervisor Professor Antonios Tsourdos and co-supervisor Dr Tarapong Sreenuch for all their help and guidance carrying out the work presented in this thesis. I would also like to extend my gratitude to Professor Ian K Jennions for providing specialist knowledge in the IVHM field and valuable feedback. I would also thank Dr. Frederic Coulon for his help in collaboration in the reviews.

I would like to acknowledge the full sponsorship and feedback for this project provided by the Cranfield University IVHM Centre and its industrial partners: Boeing, Rolls-Royce, Meggitt, BAE Systems, Thales, Alstom, UK Ministry of Defence and all other IVHM partners.

I would also like to thank my colleagues in the IVHM centre for all kinds of support and assistance. In particular, I would like to thank Omer Eker for the technical discussions on all sorts of interesting engineering topics which indirectly helped with this thesis completion.

Finally, I would like to thank all my family members who have supported me at every stage.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	**Artificial Intelligence**
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Networks
ARMA	Auto-Regressive Moving-Average
BPNN	Back Propagation Neural Networks
CBM	Condition-Based Modelling
CDF	Cumulative Density Function
CM	Condition Monitoring
C-MAPSS	Commercial Modular Aero-Propulsion System Simulation
CPNN	Confidence Prediction Neural Networks
CRA	Cumulative Relative Accuracy
DBN	Dynamic Bayesian Networks
DDM	Data-Driven Modelling
DP	Damage Prognosis
DoD	Department of Defence (US)
DWNN	Dynamic Wavelet Neural Networks
EKF	Extended Kalman Filter
EoL	End-of-Life
eUAV	Electric Unmanned Aerial Vehicle
FFNN	Feed Forward Neural Networks
FMECA	Failure Mode, Effects, and Criticality Analysis
GM	Gray Model
GPR	Gaussian Process Regression
HHMM	Hierarchical Hidden Markov Models
HMM	Hidden Markov Models
HPC	High Pressure Compressor
HSMM	Hidden Semi Markov Models
HUMS	Health and Usage Monitoring Systems
IEEE	Institute of Electrical and Electronics Engineers
IGBT	Insulated Gate Bipolar Transistor
ISHM	Integrated System Health Management
IVHM	Integrated Vehicle Health Management
KBM	Knowledge-Based Modelling
KF	Kalman Filters
kNN	k-Nearest Neighbour
LSSVR	Least Squares Support Vector Regression
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error

MLE	Maximum-Likelihood Estimation
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
MTBF	Mean Time Between Failures
MTTR	Mean Time To Restore
NASA	National Aeronautics and Space Administration
NDE	Non-Destructive Evaluation
NI	National Instruments
NN	Neural Networks
OSA-CBM	Open Systems Architecture for CBM
PbM	Physics-Based Modelling
PDF	Probability Density Function
PEEK	Poly-ether-ether-ketone
PF	Particle Filters
PHM	Prognostics and Health Management
PoF	Physics of Failure
RBF	Radial Basis Function
RCM	Reliability Centred Maintenance
RMSE	Root Mean Squared Error
RUL	Remaining Useful Life
SBP	Similarity Based Prognostics
SHM	Structural Health Monitoring
TDNN	Time Delay Neural Networks
CMAPPS	Commercial Modular Aero-Propulsion System Simulation
VLRS	Vehicle Level Reasoning System

Chapter 1

This chapter briefly describes the basics of Integrated Vehicle Health Management (IVHM), a capability that enables a number of maintenance philosophies emphasising prognostics and reasoning, which are the most attractive research topics in this area. The research problem, aims, objectives of this study are outlined and the PhD contribution is presented. Finally the publications that have been made during this research are also listed and the architecture of the thesis is provided.

Introduction

IVHM is a latest technology in maintenance field, enabling disciplines with an integrated structure. Maintenance approaches like Condition Based Maintenance (CBM) or Reliability Centred Maintenance (RCM) are aided using IVHM. Prognostics and diagnostics are combined into the framework comprising the monitoring of sensory data and predicting the future health and diagnosing the faults of the system. IVHM technology has potential applications in many fields such as aerospace, military systems, electronics, machinery, energy, and manufacturing. In IVHM, real-time sensory data obtained from the equipment is analysed continuously to detect and forecast the health states and to plan maintenance based on the forecasted health.

Prognostics is challenging and the fundamental technology within IVHM, where it requires identification of the current health level and extrapolating it to a predefined failure threshold, concluded with the estimation of remaining useful life (RUL). The output of prognostics (i.e. RUL) is the duration between the current time and the time at which the forecasted health level reaches to a predefined threshold. Benefits of the prognostics motivate researchers and the industry to achieve reduced costs, increased safety and availability via better maintenance planning. In contrast with traditional maintenance philosophies, the IVHM approach enables modelling and tracking of individual equipment deterioration leading to a maintenance action only when it is necessary rather than performing scheduled maintenance. Note that, Prognostics and Health Management (PHM) is a relevant technology to IVHM where slight differences may appear which are reported in the literature. IVHM endeavours bringing a business model within the integrated scheme which is missing in the PHM. However, this research coverage involves both PHM and IVHM.

1.1 Research Problem Definition

There are two major types of prognostic reasoning approaches, data-driven and model-based. Both approaches are by no means without their own limitations. The model-based approach on the other hand has its limitations in terms of derivation of a physical degradation model which can be infeasible for the case of complex degradation mechanisms which makes a model based approach limited to simple systems. On the other hand a data-driven approach requires enormous amount of data to train the system. Data driven approach uses the data to create a nominal model of the system. However if the condition changes and the true system data is far from the nominal model, which is often learnt from run-to-failure accelerated aging laboratory tests, then the data driven approach will try to stick with the nominal model regardless as what the present conditions might be of the system. Until now, there are very few prognostic examples addressing both of these shortcomings reported in the literature. Primarily this problem will be addressed in this thesis.

Therefore, a prognostic approach that allows a complex degradation profile to be accurately modelled and is able to adapt to a wide range of operating profiles, which is likely to be the case in practice, would be of particular interest.

In this thesis, we explore how a complex degradation profile, where a traceable physical model might not be feasible, is modelled and adapted to a variation of operating profiles. Secondly, in complex systems it's very difficult to find out the indicators which are affecting the overall health of the system and what certain data actually means in context of the present mission.

Presently, in state of art systems there are many sensors, these systems take the input of the sensor and by expert knowledge the threshold gets setup whenever the value of the sensor reaches the threshold, when it does it generates false warmings which causes confusion to vehicle operators. The secondary problem which this research is trying to solve is, how to achieve the complex system reasoning in terms of overall health of the system and which components are effecting or contributing towards the overall health in the mission context? Furthermore how the prognostic framework can aid the Vehicle Level Reasoning System (VLRS) to improve the vehicle level reasoning results which could improve the safety of the vehicle.

1.2 Research Aims & Objectives

The primary research aim is to develop a proof-of-concept demonstration of a novel prognostic framework capable of providing accurate prediction and potentially utilising multiple fault precursors, a performance improvement in comparison to conventional data-driven approaches, over a wide range of operation uncertainties based on degradation patterns. The secondary aim is to demonstrate how this framework can aid the vehicle level reasoning system.

The core objectives of this research are the following:

1. Review and identify key modelling and prediction aspects of state-of-the-art model-based and data-driven prognostic approaches.

2. Identify key technical requirements to a predictive power of a prognostic reasoner if to be deployed in a real-world operation uncertain environment.

3. Design a generic prognostic framework (or problem formulation) that can be operationally scaled to usage and imperfect manufacturing uncertainties and yet can be simply constructed using standard data-driven and parameter estimation building blocks.
4. The conceptual design of VLRS by providing prognostics engine output as a Health Index to VLRS

5. Proof-of-concept demonstration and performance evaluation of the developed framework using adaptive/scalable machine learning and online estimation models based on aerospace related prognostic case studies.

1.3 Contributions

The intellectual contributions of this research are outlined below:

1. Develop a prototype design of vehicle level reasoning system and a literature survey on the adaptive, scalable prognostics and VLRS.
2. The development of generic prognostic framework that can be adaptable to the operating conditions and operationally scaled to usage, which are simple to be constructed using standard data-driven and parameter estimation building blocks.
3. Systematic way to contextualise sensor data in terms of operation and safety which will enable relevant alerts and alarms to be generated in a timely manner.

4. Adaptive reasoning approach based on probabilistic contextual models that is able to estimate usage condition (and/or component specific) factors and accurately predict future health
5. A non-linear functional mapping from multiple probabilistic fault precursors to uniform monotonic cumulative condition index function which is component life or safety margin correlated.

1.4 List of Publications

A list of publications that contributes to the literature regarding this research is listed below:

Journal papers:

1. **F. Khan**, T. Sreenuch, O. F. Eker, and A. Tsourdos " Adaptive Degradation Prognostic Reasoning by Particle Filter with Neural Network Degradation Model for Turbofan Jet Engine", Journal of Aerospace Information Systems, AIAA, (Submitted)

2. T. Sreenuch, **F. Khan**, and J. Li. "Particle Filter with Operational-Scalable Takagi–Sugeno Fuzzy Degradation Model for Filter-Clogging Prognosis", Journal of Aerospace Information Systems, AIAA, Vol. 12, No. 5 (2015), pp. 398-412. doi: 10.2514/1.I010385.

3. **F. Khan**, O. F. Eker, T. Sreenuch and A. Tsourdos, "Multi-Domain Modelling and Simulation of an Aircraft System for Advanced Vehicle-Level Reasoning Research and Development," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 5, no. 4, pp. 86-96, 2014.

Conference Proceedings:

1. **F. Khan,** O. F. Eker, T. Sreenuch, A. Tsourdos, "*Health Index and Behaviour Based Vehicle Level Reasoning System*", IEEE PHM, Conference of the Prognostics and Health Management, Austin, Texas, USA, 21-25 June 2015.

2. **F. Khan, ,** O. F. Eker, I. K. Jennions, A. Tsourdos,., "*Prognostics of Crack Propagation in Structures using Time Delay Neural Network*", IEEE PHM, Conference of the Prognostics and Health Management, Austin, Texas,, USA, 22-26 June 2015.
3. **F. Khan,**, T. Sreenuch, A. Tsourdos, W. A. Orfali, "*A Simulation-based Health Monitoring System Test-bed for an Electrical Power Distribution System*", IEEE PHM, Conference of the Prognostics and Health Management, Austin, Texas, USA, 22-26 June 2015.

4. O. F. Eker, **F. Khan,** Z. Skaf, F. Camci, I. K. Jennions, "*Collection of a benchmark Dataset for Prognostic Modelling*", The Twelfth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, from sensors through diagnostics and prognostics to maintenance, Oxford, UK, 09-11 June 2015.

5. **F. Khan**, I. Jennions and T. Sreenuch, "*Integration Issues for Vehicle Level Distributed Diagnostic Reasoners*," SAE Technical Paper 2013-01-2294, 2013, doi:10.4271/2013-01-2294, 2013.

1.5 Thesis Layout

Organisation of the thesis is as follows:

Chapter 2: provides a detailed literature review of maintenance schemes by IVHM. Secondly a diagnostics and prognostics reasoning literature reviews is presented and thirdly the VLRS and their types are discussed. It also provides the detail literature review of maintenance schemes of IVHM. Finally The VLRS and their types are also discussed.

Chapter 3: There is a brief discussion about VLRS, next their design is studied which is followed by a presented novel VLRS design. It also shows the need of a health index calculation and adaptive data driven degradation models.

Chapter 4: An operational scalable data driven technique is presented along with a brief literature review and industrial need for such system. The design and methodology and implementation of the scalable model are also presented along with data-set information and results.

Chapter 5: The adaptive data driven techniques for complex systems are mainly focused in this chapter, the design, methodology, implementation and results are also presented. The Particle Filter and RBF (Radial Bases Function) Neural Network design are also discussed.

Chapter 6: A conclusion of the research study, it also concludes the outcomes of the research as well as summarises the research presented in this thesis and the potential future work that can be done on and with this research study.

Chapter 2

Literature Review

The primary aim of this chapter is to provide a detailed literature review regarding IVHM, reasoning approaches and prognostics along with a review of maintenance strategies. The prognostics and diagnostics approaches are categorised and discussed in detail. Furthermore, an analysis of the strengths and weaknesses of the approaches has also been conducted. However the chapters 3, 4 and 5 also provide brief literature reviews closely related to the subjects discussed.

2.1 Integrated Vehicle Health Management (IVHM)

An Integrated Vehicle Health Management (IVHM) system consistently evaluates the health of the vehicle and is also responsible for communicating and coordinating with various sub-systems within the vehicle/system.

The main aims of the IVHM system are:

- Increase the safety of the aircraft
- Diagnose the faults within the system
- Suggest the maintenance to perform
- Detect the degradation of any parts / component
- Predict the failure/remaining useful life of the components

The Integrated Vehicle Health Management (IVHM) system plays a vital role in the management of the aircraft operation.

Generally, the IVHM is used in the aerospace sector, although some authors have proposed that IVHM functions can be found in other vehicle types including also helicopters, land vehicles and maritime systems and the definition and scope of IVHM has been defined differently, please see Table 1 for details.

TABLE 2-1 IVHM DEFINITIONS
Author (date) definition of IVHM
NASA (1992) "The capability to efficiently perform checkout, testing, and monitoring of space transportation vehicles, subsystems, and components before, during, and after operation.'…'must support fault-tolerant response including system/subsystem reconfiguration to prevent catastrophic failure; and IVHM must support the planning and scheduling of post-operational maintenance."
Aaseng (2001) "All the activities that are performed to understand the state of the vehicle and its components, to restore the vehicle to nominal system status when malfunctions occur, and to minimize safety risks and mission impacts that result from system failures"
Baroth et al. (2001) "an 'effort to coordinate, integrate and apply advanced software, sensors, and design technologies to increase the level of intelligence, autonomy and health state determination and response of future vehicles"
Roemer et al. (2001) "integrates component, subsystem and system level health monitoring strategies, consisting of anomaly/diagnostic/prognostic technologies, with an integrated modeling architecture that addresses failure mode mitigation and life cycle costs"
Price et al. (2003) "an example of an intelligence sensing system. The purpose of such system is to detect and measure certain quantities, and to use the information and knowledge obtained from the measured data, and any prior knowledge, to make intelligent, forward-looking decisions and initiate actions"

Paris et al. (2005)
"the process of assessing, preserving, and restoring system functionality across flight and ground systems"

Jakovljevic et al. (2006)
"ensures the reliable capture of the "health status" of the overall aerospace system and helps to prevent its degradation or failure by providing reliable information about problems and faults"

Karsai et al. (2006)
"its goal is to provide better ways for operating and maintaining aerospace vehicles using techniques, such as condition monitoring, anomaly detection, fault isolation, and managing the vehicle operations in the case of faults"

There is also literature that describes applications of IVHM to non-vehicle systems, such as power generation plants , production machines, drilling rigs, industrial process plants, or electronic equipment (IGBT (Insulated-Gate Bipolar Transistor), power supplies) [1]. NASA first conceptualised IVHM in a report that described "IVHM as the highest priority technology for present and future NASA space transportation systems". However, their concept was established in the early 1970s [1]. Most of the literature on IVHM has been published since the late 1990s with conferences proceedings being the most common. The 'IEEE Aerospace Conference' being prominent among these. These articles cover a range of aspects of IVHM, with approximately 35% describing potential impacts or cost benefit analyses; 15% discussing design approaches and about 25% focusing on examples of IVHM systems in use or under development [1]. Other topics are related to technology evolution and integration.

Most contributions to the literature on IVHM come from prime contractors or government agencies such as NASA, the Boeing Company or the US DoD. There have been rather few contributions from academics, and those that do exist have originated in the US and UK research centres', such as the Applied Research Laboratory at Pennsylvania State University or the Intelligent Control Systems Laboratory at GIT (Georgia Institute of Technology) [2], [3].

There are multiple definitions of IVHM within the context of vehicle health management. Table 2.1 provides a tabulation of some of the important IVHM interpretations within the literature.

IVHM enables many disciplines with an integrated framework. CBM, Health and Usage Monitoring Systems (HUMS), and RCM (Reliability Centred Maintenance) are some of the maintenance strategies offered under IVHM where diagnostics and prognostics were considered under the analytics category. IVHM builds the background of this thesis along with the relevant technology, PHM.

The following sections present the maintenance strategies including CBM and its sub-disciplines which provide a basis for this research.

2.2 Maintenance Strategies Overview

Maintenance philosophies are classified into two categories, these are:

1. Reactive Maintenance (unplanned)
    - Corrective Maintenance
2. Proactive Maintenance (pre-planned)
    - Preventative Maintenance
    - Predictive Maintenance

2.2.1 Reactive Maintenance

Since World War II the most remarkable changes have been done in last sixty years; this gives the historical prospect of maintenance [4]. Up until then the only option available was corrective maintenance where the equipment could only be fixed or replaced. It is less risky, cost effective, and failure consequences are not fatal, therefore, corrective maintenance is still in use for simple mechanisms such as light bulbs or a basic pipeline.

Mechanisation and automation steps have grown since the mid-nineties due to the extreme intolerance of downtime and the significant increase in cost of labour. Development of proactive maintenance is a result of improved machinery of lighter construction which ran at a higher speed provoking wear out more quickly. Sub-discipline of proactive maintenance is the preventative maintenance in which the maintenance tasks are performed periodically. Periods are fixed intervals determined by using historical data (e.g. MTBF: Mean-Time-Between-Failures) and without any input from the individual equipment itself. Service is required or not regardless the equipment is serviced on a routine schedule which increases the cost. However, both approaches, reactive and blindly, have financial and safety implications associated with them. Few examples of preventative maintenance are routine inspection rounds and lubrication, bi-monthly bearing replacements, or maintenance inspections and overhauls on aircraft systems.

In the late seventy's, questions arose regarding the effectiveness of conducting preventative maintenance due to a common concern about 'over-maintaining' which led to the development of predictive maintenance. Major features that distinguish predictive maintenance from preventative maintenance are adaptively determined scheduling of maintenance actions. On the contrary, predictive maintenance is limited to those applications where the cost and consequences are critical and technically feasible [5]. Predictive maintenance is classified as two: Condition-based Maintenance (CBM) and Reliability-Centred Maintenance (RCM). RCM performs two tasks: first, analyse and categorise failure modes (e.g. FMEA) and second, assess the impact of maintenance schedules on system reliability [6]. RCM is based on manual inspections and basic data trending. CBM is discussed in further detail later in this chapter.

2.2.2 Proactive Maintenance

The Proactive maintenance is a preventive maintenance strategy for maintaining the reliability of systems or machines. The purpose of proactive maintenance is to manage machine failure that can be anticipated and setup the maintenance routine before they occur to reduce the downtime. From the 1980's, systems became more complex in nature resulting in a more competitive marketplace and the intolerance of increased downtimes. The daily loss of revenue due to downtime is £320,000 for a Boeing 747 aircraft [7] is an example from the 21st century. Moreover, risk analysis and environmental safety issues have become predominant. New concepts have emerged such as condition monitoring and expert systems. The Institute of Asset Management has been established in the UK in mid-90's which has received significant attention from most organisations. In the 2000's, terms such as prognostics, IVHM, and integrated system health monitoring (ISHM) have emerged and taken place in literature gradually thus far.

Nowadays, in engineering practices, maintenance activities have become predominantly intuitive and are based on experts' or personnel's experience that are familiar with the equipment. However, it is becoming exceedingly difficult to gain experience due to an ageing engineering workforce and improved asset reliability. Additionally, when dealing with complex equipment, human decision making is not always adequately reliable due to the multitude of interrelating failure modes [8].

Current systems are becoming more complex and expensive which lead to an increase in competition drive and as a result industrial and military areas are highly concerned about the availability and reliability of the systems. For many industries it is vital to maximise the cost of system availability and reliability and minimise failure and downtime. Today's sophisticated sensor technology enables engineers to track degradation processes and empowers for prognostic reasoning of equipment being monitored [9].

2.2.3 Condition-Based Maintenance

Condition-Based Maintenance (CBM) is a predictive maintenance strategy, whereby the maintenance tasks are performed when the need arises. The necessity concept is determined by assessing the health condition of the equipment continuously and extrapolating it to a predefined failure threshold [10], [11].

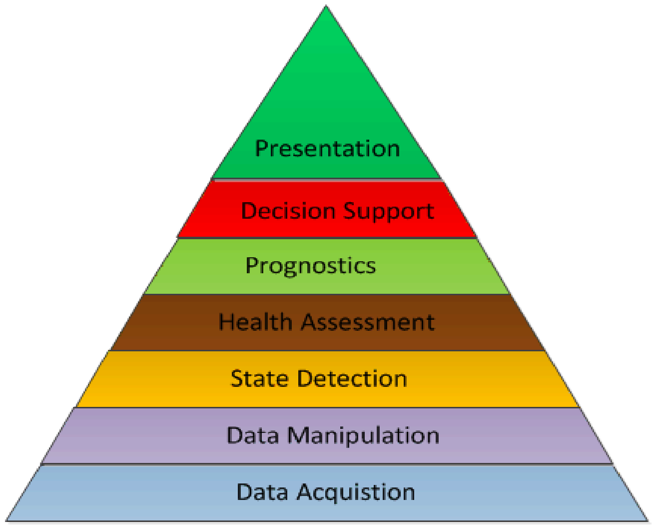


Figure 2.1 The OSA-CBM architecture

Figure 2.1 depicts the hierarchical steps of standardised Open Systems Architecture for Condition-Based Maintenance (OSA-CBM). OSA-CBM, a layered approach, describes a standardised information delivery system in between its functional blocks. First layer is acquisition of data which gets data from sensor. Second layer is where the data signal is processed (e.g. feature extraction). Third layer is for detect the state of the system and checks abnormality. Forth layer checks where different faults types are detected and isolated for comparisons (e.g. FMECA failure mode analysis). Fifth layer handles the degradation level of the system, where fourth layer, provides an input to a prognostic block in order to be able to predict the remaining useful life of the asset. The top two layers are responsible for the intelligent decisions for maintenance activity by means of the prognostic results and instrumentation, respectively.

Figure 2.2 shows an example of degradation in health level of an asset. P-F (Physics of Failure) interval represents the time interval between potential failure which is identified by health indicators, and an eventual functional failure. With CBM, it's necessary that the P-F interval is long enough to enable corrective maintenance action to be taken [12].

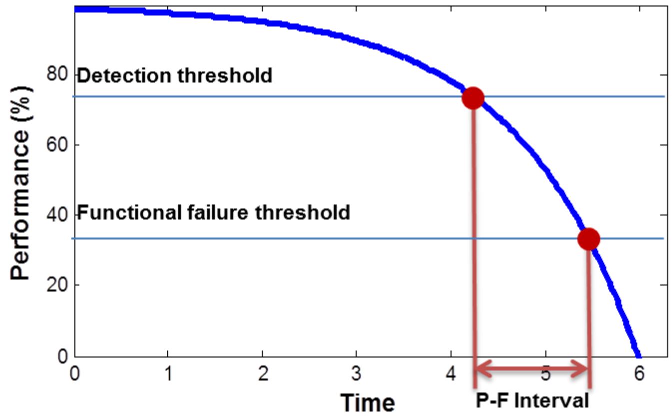


Figure 2.2. P-F curve of an equipment

Maintenance preparations are achieved on the system while it's up and running which greatly impacts the reduction of operation and support cost. The inventory cost will be reduced, in addition to the reduced down time, as more time will be available for obtaining required parts. Furthermore, better maintenance preparation will increase the efficiency in logistics and supply chain. Eventually, the life cycle cost of the equipment will be reduced, as they are used until the end of their lives.

2.2.3.1 Diagnostics and Prognostics in CBM

Two major disciplines of CBM are diagnostics and prognostics. In literature, there is a minor disagreement that prognostics is related to and highly dependent upon diagnostics [8].

Diagnostics is a relatively mature area compared to prognostics as it involves detecting and reporting abnormalities in signals as well as identifying the fault type, and the quantification of current health status of an asset.

Once the abnormality has been detected CBM with diagnostics targets to stop and schedule a maintenance task for the system, otherwise the system continues to operate. Unscheduled maintenance should be performed once degradation is detected to prevent the failure consequences. Rather than executing actual maintenance it is very common to spend more time on maintenance preparation due to the lack of resources. Ideally, in prognostics, maintenance preparation could be performed when the system is up and running, since the time-to-failure is known early enough.

Therefore, in CBM only the actual maintenance duration becomes the major contributor of the downtime which is way less than the fault diagnostic approach. For example if prognostics can present a warning of a failure of an asset before 10 flight hours, re-test and installation steps can be pre-planned, yielding in saving of maintainer time and significant reduction in its variability [13]. The comparison of diagnostics and prognostics in CBM is illustrated in Figure 2.3.

In general, incipient failures follow a progressive degradation path [14]. Once it has reached the sever point detection of failure progression is more valuable compared to the detection of failure. Moreover, it is a prerequisite for prognostics [15]. In other words, prognostics utilise the health severity or health status information transmitted from the diagnostics base. [13] states that prognostics for avionics is essential as the increasing of the number of complex systems comprising of electro-mechanical components in current and future aircrafts and a possible shortage of technicians capable of servicing them.

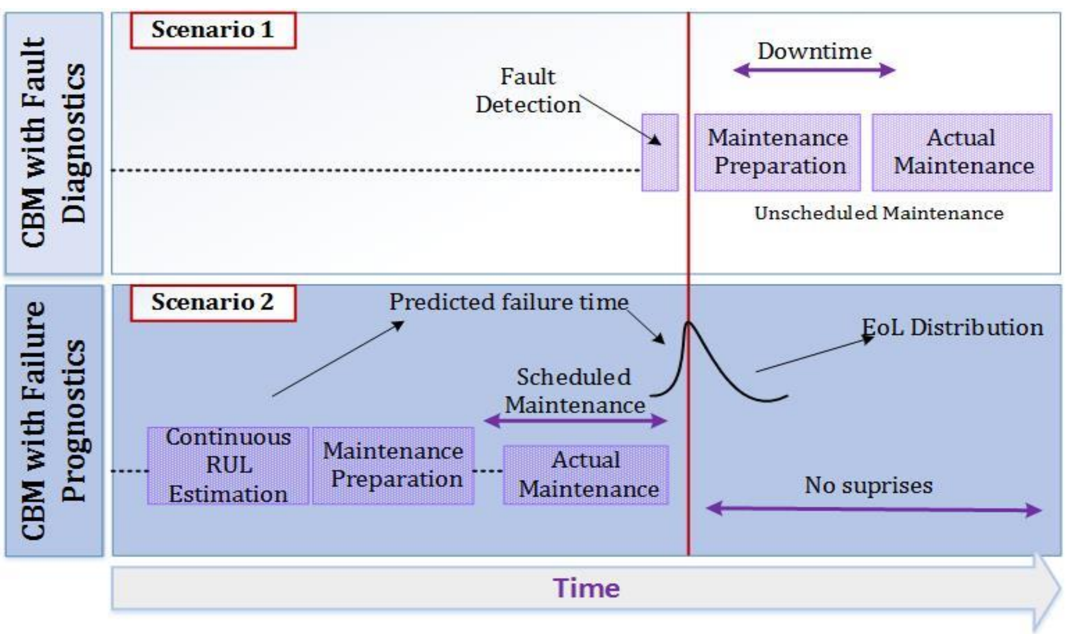


Figure 2.3. Fault Diagnostics vs. Failure Prognostics in CBM

Figure **2.4** displays two phases of prognostics. To access the current health condition is the aim of the first stage of prognostics. In literature, the terms used for describing this stage are severity detection of system, health assessment, and degradation identification. Usually, Bayesian filtering and/or pattern recognition techniques such as classification or clustering are employed in the health assessment part. The goal of the second phase, also known as true prognostics, is to predict the failure time by forecasting the degradation trend leading to the estimation of remaining useful life (RUL). The terms used to describe this stage are time series analysis, extrapolation, propagation, trending, projection and tracking.

Prognostics indicate forecasting of the system's/component's future health level by propagating the current health level until a failure threshold and therefore, enables an ability to provide an estimate of the remaining useful life (RUL). Prognostics are considered to be one of the most challenging and key enabling technologies among the CBM steps [16], [17], [18].

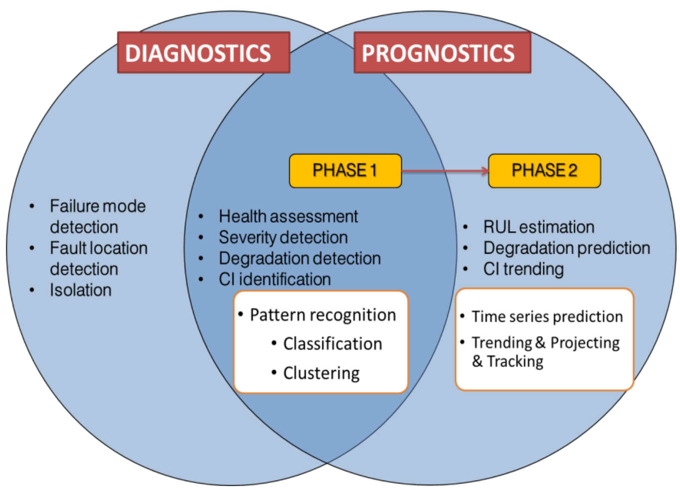


Figure 2.4. Prognostic and diagnostic phases [19]

2.2.3.2 Benefits of Prognostics with CBM

The CBM approach has significant advantages on reducing the support and operating costs and leading to a more effective planning and operational decision making. An unexpected one-day stoppage in machinery industry may cost up to £160,000 [17]. Another example from the return on investment for companies is the investment of £9,500 on monitoring the condition of systems which prevents £315,000 of maintenance costs per year [6].

In another example, FAA's BRITE radar was maintained either with pre-arranged (proactive) or unscheduled (reactive) maintenance. Pre-arranged maintenance decisions were taken reasonably before the potential failure utilising prognostics by the monitoring of degradation in the radar. Unscheduled maintenance took seventeen hours more than the pre-arranged maintenance in mean time to restore (MTTR) which was fifteen times higher of downtime than that of the pre-arranged maintenance in comparison [13]. A detailed review of prognostic approaches is presented in the following section.

2.3 Prognostics and Reasoning Approaches

There are several different approach / methods for prognostics/ reasoning. There are three major methods or approaches for the reasoning. The Figure 2.5 shows the hierarchical of major techniques of the reasoning, prognostics and diagnostics.

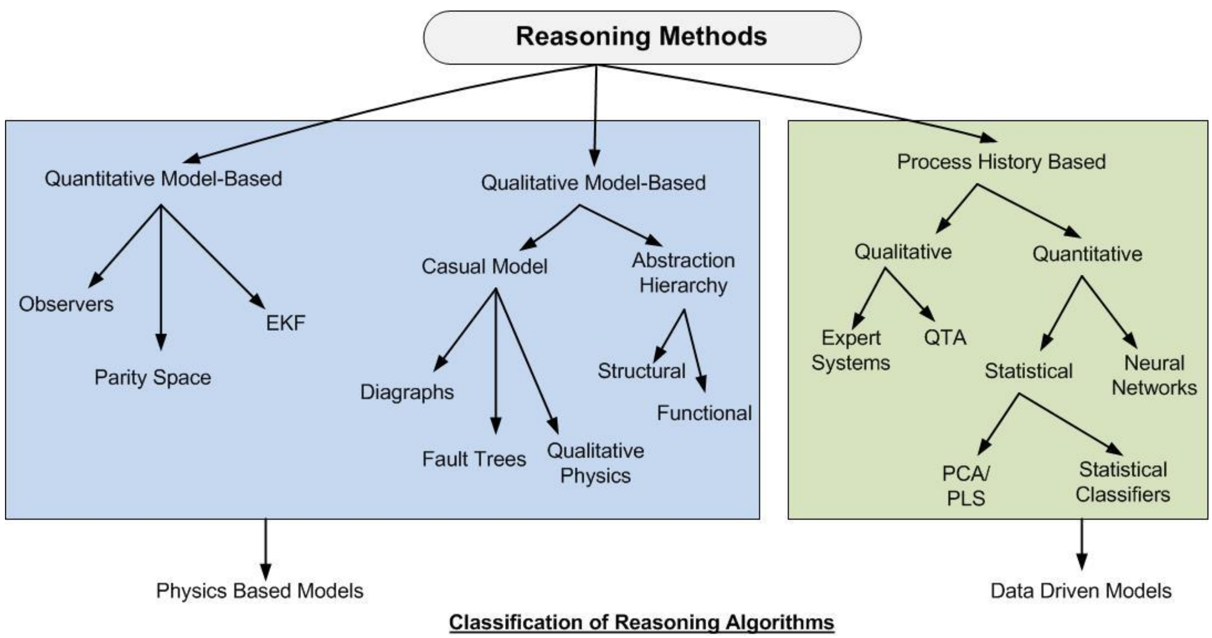


Figure 2.5. Classification of Diagnostics/Reasoning Algorithms

In general, prognostic models can be categorised into main three classes, these are:

1. Data-driven models
2. Physics-based models
3. Knowledge-based models

The first three categories are illustrated in Figure 2.6; this Figure presents the hierarchy of prognostic models based on the range of applicability, cost, and accuracy where knowledge-based models, being the most cost effective, find themselves a maximum applicability range in systems/components, albeit the accuracy of these models is less than the high accurate and costly physics-based models. Data-driven models fit in the middle of these models mentioned.

Several literature surveys covering the prognostic models have been presented by [6], [9], [16], [17], [20], [21], [22], [23], [24], [25]. This literature review builds on the surveys referred in this section. In addition, current prognostic applications which are mentioned in the literature are further discussed. In the following sections, a literature review of prognostic approaches within these categories is presented.

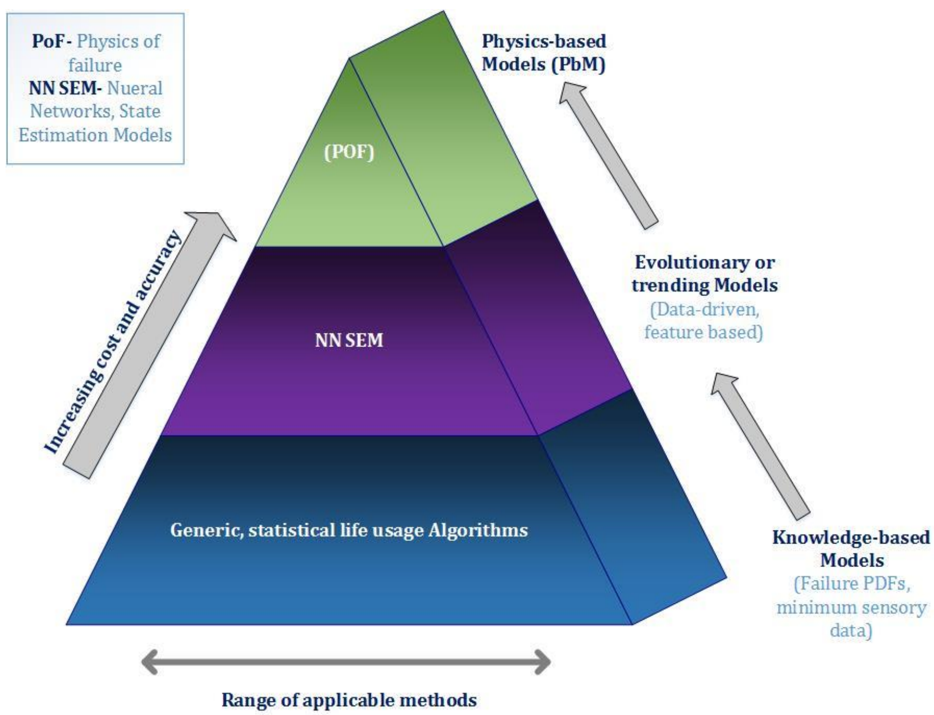


Figure 2.6. Prognostic models hierarchy [21]

2.3.1 Data-Driven Models

Data-driven models (DDM) utilise collected condition monitoring data and/or historical event data instead of constructing a model based on system physics or human expertise. DDMs attempt to track the degradation of an asset using extrapolation or projection techniques (e.g. regression, exponential smoothing, and neural networks) or match similar patterns in the history of relevant samples to infer RUL [20]. They also rely on the past patterns of deterioration to forecast future degradation. Usually system or loading inputs are not involved in data driven prognostic modelling. The assumption for models in this category is that the future system inputs or operational profile remains constant or consistent with the past data. Since data-driven prognostics have no elaborate information (e.g. physical information) related to the asset or system, it is considered to be a black-box operation [26]. Data-driven models are divided into two categories: Statistical models and Artificial Intelligence-Based (i.e. machine learning) models.

2.3.2 Model based / Physics base Models

Physics-based Models (PbM), also known as 'Model-based Prognostics or Model-based Approaches', generally involve defining the physics of the equipment and the failure mechanism. The author prefers to use the term 'physics-based models' rather than 'model-based prognostics' since the most data-driven approaches use models as well. This way of categorisation gives a better ability to distinguish physics-based and data-driven models [27].

In PbMs, mathematical models of failure are usually employed which is directly tied to health degradation. In order to provide knowledge rich prognostics output; PbMs are attempted to combine defect growth formulas, system specific mechanistic knowledge and condition monitoring data. These models assume that an accurate mathematical model for component degradation can be constructed from first principles. Residuals, the outcomes of consistency checks between sensor measurements and mathematical model outputs, are utilised as features of health condition in PbM approaches. Thresholds to detect the presence of faults are determined by using statistical techniques. In addition, model parameters are identified using empirical data obtained from specifically designed experiments [20]. Physics-Based Models are implemented in three different ways [8]; firstly, dynamic ordinary or partial differential equations that can be solved with approximation approaches (e.g. Lagrangian or Hamiltonian dynamics), secondly, state-space methods (i.e. no differential equations), thirdly, simulation methods. In [28] employed a physical stochastic model on gears. They calibrated the parameters for physical stochastic prognostics & diagnostics using system level features extracted from test specimens. In [29] they developed a fault detection and prediction algorithm for flight actuators which applies parametric identification and physical modelling techniques. In [30] and [31] they applied physics-based approaches to prognostics which have involved deriving the explicit relationship between condition variables and the current lifetime and failure lifetime via mechanistic modelling. Both of them applied their model for energy processors and bearings by employing vibration sensor measurements respectively. A general method for tracking the progress of a hidden damage process was proposed by [32]. The proposed model is applicable for a given situation where a slowly evolving damage process is connected to a fast, directly observable dynamic system. [33] Fused diagnostic information and physics of failure modelling is applied for helicopter gear prognostics. A hierarchical modelling approach proposed by [34] is used for system simulation to determine remaining useful life.

A physics-of-failure approach reinforced with Kalman filters were used to track the dynamics of the frequency of accelerometer sensor signals in tensioned steel band by [35]. [36] Used a Kalman Filter with an associated interacting multiple model to perform tracking of sensor-level test-failure probability vectors for prognostics. Assumptions for Kalman Filters are that the system exhibits a linear process and the noise in the system follows Gaussian distribution. Extended Kalman Filters and Unscented Kalman Filters are some of the extensions to the traditional Kalman Filters in which the system is not bound by the linear process. [37] Presented an Extended Kalman Filter approach for estimation of Lithium-ion battery life. Particle filters are a generic type of Bayesian tracking method used with physics laws (i.e. in the form of differential equations) in which the model is not bound by the assumption of linearity in the system and Gaussian noise. Instead of using deterministic probability distributions, significant numbers of particles are employed, representing the health state of the system distribution. A number of examples are available in prognostic modelling literature for particle filters [18], [38], [39]. Detailed discussion of particle filters is given in section 4.7 and 5.6.

Physics-based models are considered to be more accurate if an accurate mathematical model representing the degradation process is fitted in the model thoroughly [20]. And the requirement concept on the data is significantly less, compared to the data-driven models. However, PbMs are usually component or system specific models which mean usually they cannot be applied to other types of components or systems in which the physics of the failure mechanism is different. Another disadvantage is that the PbMs are costly compared to other approaches whereas they are the most suitable approach for cost-justified applications where accuracy outweighs most other factors [22].

The Physics based models and Data driven models are compared in table 2-2. The major advantages and disadvantages are listed. The advantages and disadvantages are mainly compared in the view of the prognostic and diagnostic reasoning.

TABLE 2-2. COMPARISON OF THE BENEFITS OF PROGNOSTIC AND DIAGNOSTIC APPROACHES

	Advantages	**Disadvantages**
Physics-Based Models	Accurate compared to other approaches (if a good representative of mathematical model is available) Higher precision Requires less data compared to other approaches Suitable for creation in design phase	Difficult to create a model especially for complex systems Sensitive to the design and material properties Sufficient component information and a good insight of the failure mechanism is required High cost of implementation Component or system specific
Data-Driven Models	Easy to conduct & simplicity in implementation Flexible and adaptable	Need data representing the failure progression, which is often not possible to obtain

	Suitable to all levels (component, system) More robust to changes in material or design compared to physics based Low cost	Computational complexity may be high Difficulty in determining of the failure thresholds
Knowledge-Based Models	Simple and easy to understand No model is required Wide application area and lower cost Ability of dealing with incomplete, noisy or imprecise input information	Not always easy to obtain domain knowledge and extract rules Handling of new situations which are not stored in knowledge base is limited Computational difficulty increases dramatically as the number of rules increases Limited capability of learning No confidence limits are provided

2.3.3 Knowledge-Based Models

It is usually difficult to obtain an accurate mathematical model in real-world applications which limits the use of physics-based prognostic models. Due to the absence of a complex model, systems tend to be maintained with simpler models such as knowledge-based models (KbM). Knowledge or experience-based prognostic approaches are the simplest way of performing prognostics where the statistical historical failure information of systems is utilised for predicting the RUL [40]. The use of knowledge-based models is automated representation of how a human domain expert solves a problem [20]. Expert systems and fuzzy logic are two generic examples of these models.

Disadvantages of knowledge based systems can be listed as:

- ❖ Hard to obtain domain knowledge and extract rules

- ❖ Handling of new situations which are not stored in knowledge based models is limited
- ❖ Computational difficulty increases dramatically as the number of rules increases (i.e. combinational explosion problem)
- ❖ No confidence limits are supplied

The detailed literature review shows that all the approaches have their own advantages and disadvantages, therefore it would not be correct to imply that one approach is better than the other. It really depends on what model you would like to create and how complex the system is, whether the previous data is available for a modelled system etc.

2.3.4 Expert Systems

The Expert systems were one of the earliest systems in the reasoning and prognostics field. Expert systems have been used since the 1960s, and are considered as an artificial intelligence (AI) program that represent domain expert knowledge in solving a problem related to a particular domain. In expert systems, knowledge of domain experts is stored in the knowledge base where the extracted rules are applied into the failure situations by the maintainer. Knowledge-based rules are generated from collections of real experiments. Basic IF-THEN statement rules are often based on heuristic facts acquired by experts over a number of years [8]. Outputs of expert systems are singular rather than a distribution of RUL usually.

Expert systems have traditionally been used in failure diagnostics cases and it is starting to be implemented in prognostics applications as well. [41] Developed an online expert system called CASSANDRA, which was built to monitor the condition of industrial equipment with the intent of fault prognostics.

In [42] an expert system called PROMISE (Prognostics and Intelligent Monitoring Expert System) was presented which carries out both diagnostic and prognostic duties and provides solutions to system maintenance in plants. However no RUL information was provided with their proposed method.

In [43] the authors have developed an expert system based framework called FDPM (Failure Detection and Predictive Maintenance) which consists of several expert-system-related databases and components. It was applied on a power distribution system component for predicting maintenance demands.

2.3.5 Fuzzy Logic

Similar to expert systems, fuzzy logic is a problem solving mechanism providing a robust mathematical framework to deal with non-statistical uncertainty and real world imprecision. A fuzzy system consists of a knowledge base; fuzzy rule, and the implementation algorithms for applying the logic. Fuzzy logic has a wide application area from simple small components to large workstations. Unlike expert systems, the fuzzy logic system has the ability to model system behaviours in continuum mathematics of fuzzy sets rather than with traditional discrete values. Fuzzy logic systems are usually incorporated with other methodologies such as neural networks (NN) or expert systems.

In [44] a fuzzy expert system called 'Alarm Filtering' and Diagnostic System (AFDS) are proposed which provide clean alarm pictures and system wide failure information during abnormal states. It also provides alarm prognosis to notify the operator of process abnormalities.

In [45] a fuzzy logic process is presented in which the input data is mapped into fuzzy variables (i.e. fuzzification) using membership functions and de-mapping the fuzzy variables processed into numerically precise outputs (i.e. defuzzification). This methodology has been used widely in control applications such as in [46].

In [47] the author has proposed a dynamic fuzzy system for real-time condition monitoring and incident prevention. However, the RUL was not calculated whereas the applicability of fuzzy logic into prognostics was demonstrated.

A comparison of a fuzzy logic model and neural networks is conducted by [48] for predicting the life of boiler tubes. Results show that neural network performed better where the applicability of NNs was favourable compared to the fuzzy logic model.

Unlike [48] work, fuzzy logic is usually integrated in RUL calculations as an auxiliary method for the primary method to enhance prediction results. Fuzzy logic has an ability of dealing with incomplete or imprecise input information with the use of linguistic variables such as 'low', 'very low', which provides an intuitive way of reasoning and the representation of the failure health level. On the other hand; having no memory, limited capability of learning, difficulties of determining good fuzzy rules and membership functions are some of the disadvantages of fuzzy logic.

The next section will discuss the vehicle level reasoning system, their design and the development has been achieved and reported in the literature review.

2.4 VLRS (Vehicle Level Reasoning System)

The Vehicle Level Reasoning System (VLRS) is a system within the IVHM which is responsible for processing the input from several sub-systems and units to make logical decisions and predictions about the vehicle health.

Generally, the reasoning system is an AI based application, where computational functions have to generate results from available knowledge/Data using logical techniques of deduction, diagnosing and prediction or other forms of reasoning [12].

According to NASA: the reasoning system is: a system to detect, diagnose, predict, and mitigate adverse events during the flight of an aircraft.

Most aircraft sub-systems detect for simple threshold exceedance and report them to a CMC (central maintenance computer), the vehicle level reasoning system (VLRS) that operates on these evidences tends to be reactive rather than proactive. [12]

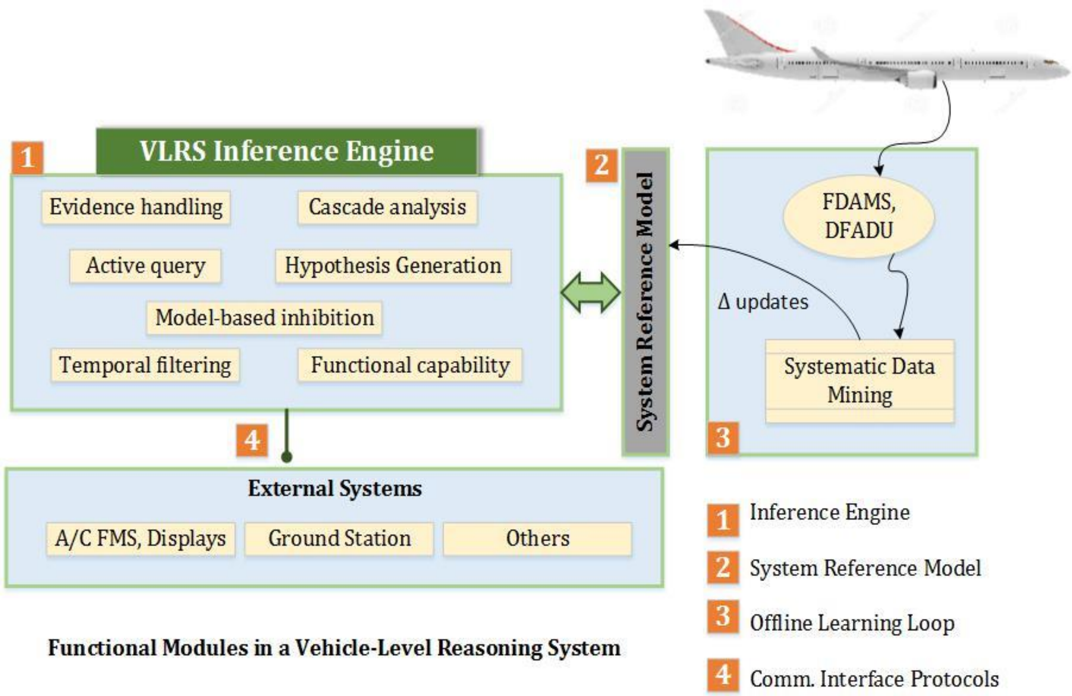


Figure 2.7. Basic Vehicle Level Reasoning System [12]

Figure 2.7 shows the basic VLRS which has been presented by Ashok N Srivastava et al. at NASA Ames Research Centre [12]. This VLRS diagram has been divided into four parts.

1. Inference Engine
2. System Reference Model
3. Data mining and learning Loop
4. Communication Interfaces

2.4.1 Reasoners / Inference Engine

This module (showed in Figure 2.8) considers health evidence generated from all components, sub-systems and systems within the vehicle to produce the current diagnostic state or predict the future evolution of a fault.

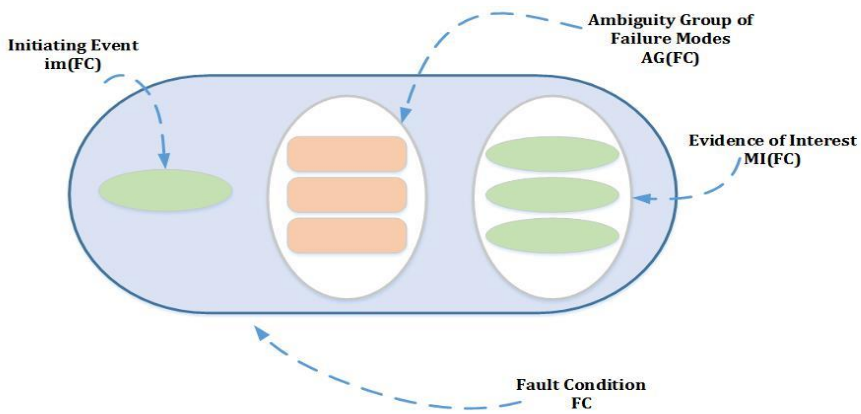


Figure 2.8 Inference Engine [12]

In this process, it produces a most plausible explanation for all the symptoms provided by various sources; creates new hypotheses to track multiple faults and deletes hypotheses that may have weak or no evidence support.

2.4.2 System Reference Model

The necessary relationships for the inference process are typically separated in a static system reference model. This partitioning allows the same inference engine software code to be reused in multiple vehicles and minimise certification and qualification costs for deploying the VLRS on-board an aircraft. The system reference model, an aircraft loadable software module, describes the relationship between evidence generated at the component and/or sub-system level and failure modes that can be mapped to specific maintenance or correction action [12].

2.4.3 Data Mining and learning Loop

Fleet modelling, data mining and knowledge discovery methods working with historical data can detect anomalies and precursors to critical failure modes. Discovering new patterns and updating old relationships in the system reference model can continually improve aircraft safety to a higher level. Information from this learning loop, resulting in a f'-change in the reference model, enables the VLRS to provide an accurate health assessment of the component, sub-system, or system and support condition-based equipment maintenance and replacement.

2.4.4 Communication Interfaces

By design, the VLRS takes a system-wide view of the adverse event detection problem. While the input interfaces defines how the VLRS receives health information from various member components, the output interface defines how it communicates its outputs to the flight crew (displays), ground maintainer (ground station) or a flight management system for automatic fault accommodation.

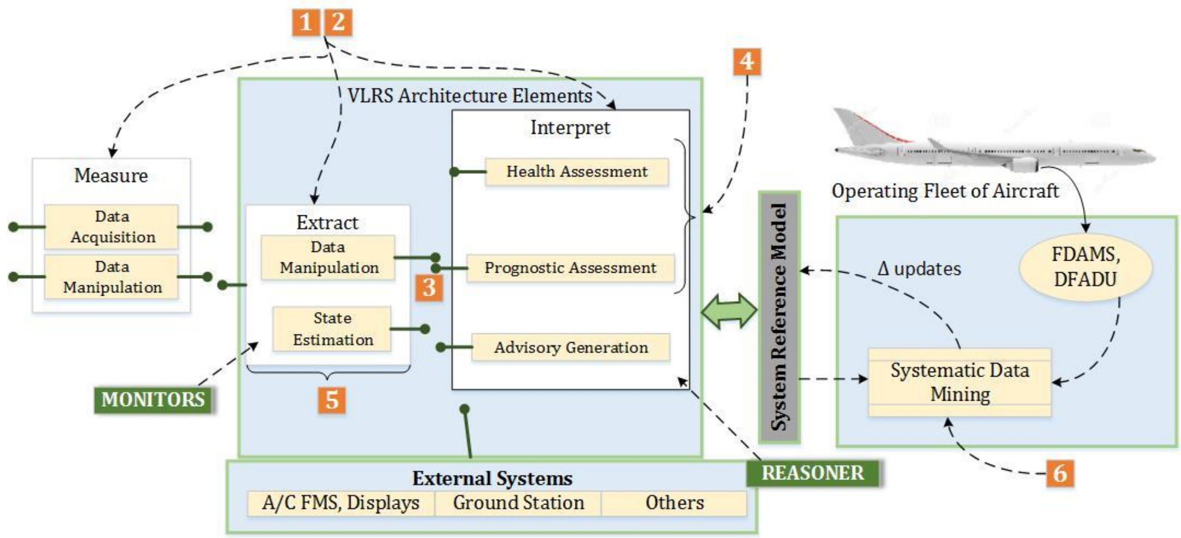


Figure 2.9. Honeywell Data Driven VLRS+ Architecture

The Figure 2.9 demonstrates the architecture of the data driven VLRS which is developed by Honeywell. There are many different approaches to designing the VLRS. These various approaches will be discussed in detail in this thesis.

2.4.5 VLRS background

In the early development of complex aerospace systems, the idea of built in testing became an important method to verify the integrity of on-board systems. The main idea is that it allows a machine to test itself and in some cases perform diagnostics to determine the source of a problem. The hope was that these systems could help improve the reliability and increase mission assurance. In many cases, the built-in test capability was implemented in hardware and helped a technician determine the state of health of the system. Early systems have little software and relied primarily on the combination of human analysis of the results of the built-in test. As a result, such systems were not designed to operate in real time or during a mission [12]. In contrast, modern aerospace systems are fundamentally different. These systems are comprised of a combination of hardware and software operating on embedded systems. Because of these highly complex systems, the need to develop techniques that can analyse their state of health in real-time with little to no human intervention is essential [12].

NASA is amongst the first organisations to contribute towards VLRS. Their VLRS was mainly for the space rovers and space vehicles of which the maintenance environment is not suitable.

NASA launched a software LV2 (Livingstone Version 2) in 2004. This software has been uploaded to the EO-1 satellite to test and analyse errors in the spacecraft system [40]. Traditionally this troubleshooting is achieved from the ground station. Many tests have been conducted to detect and diagnose the simulated failures in the instruments and system of the satellite.

The Livingstone version 2 application provides the opportunity to recover from errors to protect the system, and continue to achieve mission goals. On this mission, LV2 also monitored another software application. LV2 detected the error, made a diagnosis, and sent its analysis to mission control.

The LV2 has a model of the spacecraft and it compares the model with the current behaviour of the spacecraft. On the detection of different behaviour from the model, the LV2 will search for the faults and provide the controller a list of suggestions of possible faults.

The remote agent reasoning system is also related to the same league. This system enables autonomous planning and execution of many tasks on-board the spacecraft [12]. With this capability, only general directions are commanded from ground controllers on earth.

The need for vehicle level reasoning in aircraft will grow in the next generation because of the increasing complexity of aircraft and the higher reliance on automation. Civil aviation is also demanding greater reliability and safety from its vehicles.

A reasoning system also reduces maintenance costs [49], as the human troubleshooting decreases [1]and chances of replacing wrong parts are reduced. As well as the minimisation of the overall maintenance time.

2.5 Research Gap

In the detailed literature review the advantages and disadvantages of the data driven and model based approaches were shown in this chapter. The need of VLRS also has been highlighted. However the following are the research gaps which this thesis has contributed towards and focused on.

- In the data driven approach the major problem is if the operating conditions are different than the nominal model, in this case the prediction errors increases and the precision of the RUL results are worse. The prognostic reasoner will try to stick with the nominal model regardless of the operational conditions.
- When the sensor data is collected from any component what does the information really mean? In a traditional system the expert person sets the threshold with upper and lower bounds on the sensor values and when certain values reach the system it generates the associated warning to the user. The problem with this approach is the 1) Sensor values are mostly contaminated with huge amount of noise which results in false alarms. 2) The values do not indicate any information regarding the health of the system or components.
- The reasoning of the system is usually achieved on the sub-system level, and no overall vehicle health is defined. Each sub-system reasoner achieves the results by checking their own sensor values, and the other system health's are ignored. However most of the systems rely on each other on the vehicle level point of view, which results in 'No Fault Found' problem or 'False Fault' indication. To develop a VLRS system which can overcome these limitations require all the prescribed component/system sensor information. Intensive

amounts of sensor data cause the problem that most of the algorithms have limitations and too many components' data will cause convergence problems. There is a need of VLRS which can take sub-system results (which are already reported to the operator) and other indications which can relate the faults.

This study presents two different techniques for the Adaptive Data Driven Techniques which overcomes these major issues with the data driven technique, the presented adaptive data driven technique is scalable to the operational conditions and is adaptive their usage (see chapter 4 and 5).

This study also provides the context aware prognostics reasoning which identifies what the sensor value means to the operator in the context of the system health without identifying any threshold values for warming generation (see chapter 5).

The study also provides the vehicle level reasoning without having issues of convergence and it provides extra information of each system (see chapter 3).

Chapter 3

Health Index and Behaviour Based Vehicle Level Reasoning System

This chapter will discuss the VLRS and presented design. This VLRS takes the input from the adaptive degradations models which are presented in chapter 4 and 5, therefore this chapter also highlights the importance of the adaptive and scalable prognostics models. Finally, this chapter also discusses the novel design of VLRS and their industrial use.

Introduction

A Vehicle Level Reasoning System (VLRS) helps with improving the safety of the aircraft. These systems comprise of various sub-system diagnostics/detection units that monitor related components or/and sub-systems for functional status and relay back the operational status to the entities of interest. Therefore, a main purpose of the VLRS, is to deduce the overall operational health of the aircraft rather than each individual component, see Figure 3.1.

To comprehend the overall health, the VLRS takes into account the individual health state of the various components within the aircraft. These distinct components often employ specific sensory inputs / outputs and are subject to a constant inference mechanism to establish the functional health. The nature of the component governs the technique used by the inference mechanism. A real-time embedded computer system incorporates this inference mechanism by employing a suitable deduction technique. These deduction techniques are domain specific and are often implemented as an algorithm within the inference mechanism. However, several complex components require advanced artificial intelligence (AI) techniques and machine learning algorithms to deduce their health status [50].

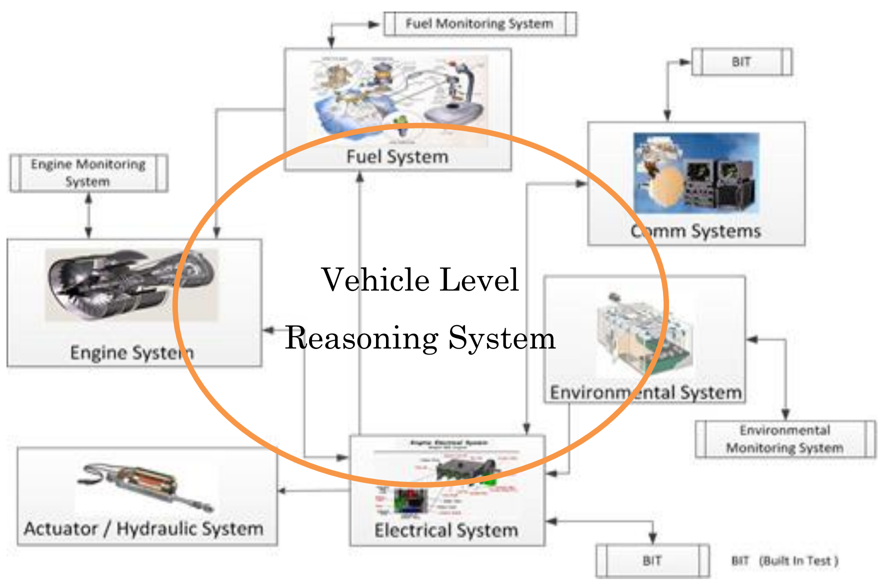


Figure 3.1. VLRS Overview

To establish the overall vehicle health, the VLRS coordinates with the component level inference mechanisms to [51]. Furthermore, this liaison aid in the identification of fault domains and the execution of a rectification strategy.

3.1 Background

The fault finding can be achieved in different ways, basic abnormal change in sensor values (threshold detection) or other types of detection, or fault diagnostics which could be achieved manually, by rule-based systems, model based techniques, mathematical or other learning.

Fault isolation, detection and diagnosis in engineering systems have been widely used in commercial industry over the past few decades.

Historically, troubleshooting is a major component of the maintenance strategy, mainly for mechanical equipment of many kinds. In the traditional detection system it monitors the sensor thresholds when it reaches the fixed thresholds it triggers a warning to the ground based system. Almost all complex systems, especially aerospace systems are in need of a precise diagnostics system. The precise diagnostics system will be cost effective as it will reduce the troubleshooting activity and speed up the repair and decrease the down time of the equipment [52]. This has an impact on the further development for the diagnostic and detection reasoner, introducing several different techniques such as Model Based Reasoning (MBR) or data driven methods which has been discussed earlier in this thesis.

3.1.1 VLSR Industry Key Players

Modular diagnostic systems / reasoners were the initial starting point for the aircraft industry. Where each vital sub-system (such as the engine, electrical power unit, and hydraulic system) had their own reasoning/diagnostic system and these reasoning systems provided the health status to the flight deck. In case of emergency, the pilot received a warning message related to the specific failure. However, this approach had many drawbacks due the systems being inter-dependable. Failure in one system could cause or trigger failure in another. For instance, aircraft generator is dependent on the engine and the engine is dependent on the fuel system. If fuel system fails it would have catastrophic results as this failure will affect the engine and the engine will affect the power generation.

These systems are inter-dependable but their diagnostics and reasoning systems are not. There is no knowledge sharing mechanisms between Modular diagnostic systems and their diagnostics/reasoning systems. A VLRS has an advantage because it is capable of foreseeing and predicting one system failure effect on another system which is interdependent. There are many challenges in the development of VLRS which will be discussed later in this chapter.

There are very few organisations who are working on VLRS as it is a fairly new concept in fault diagnostics and reasoning. Some of the industries and organisations are working on VLRS such as NASA, BAE Systems, Boeing, Lockheed Martin and Honeywell.

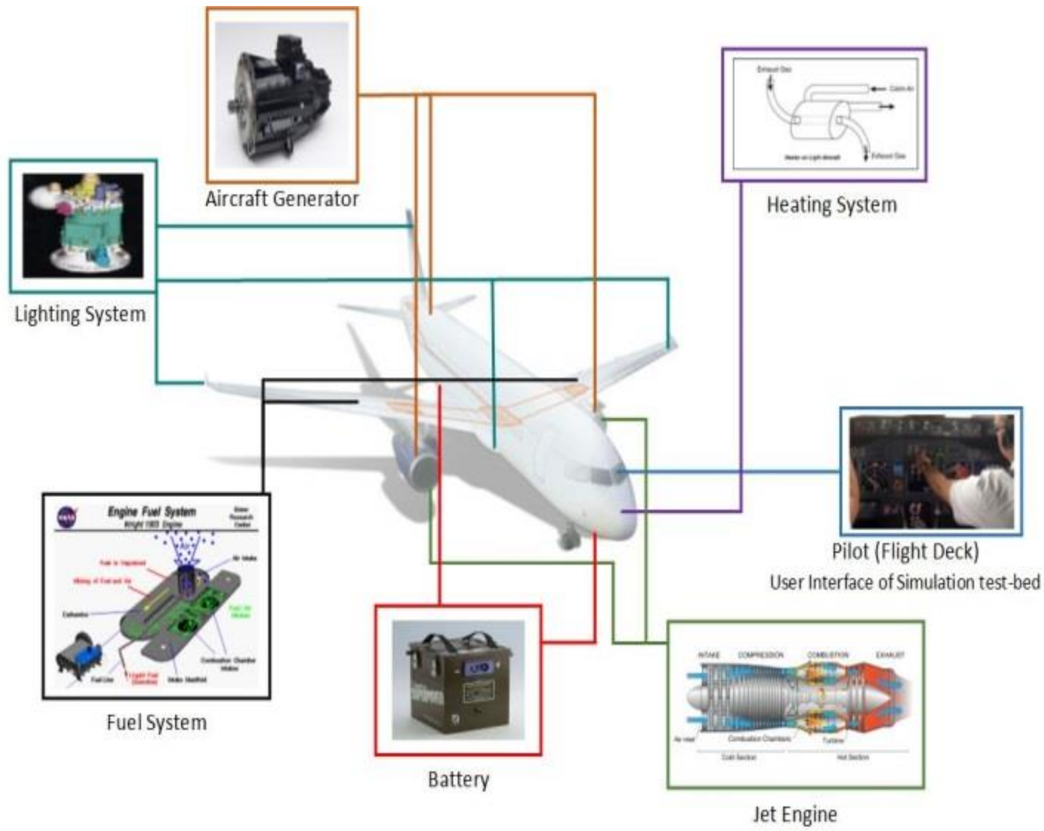


Figure 3.2. Aircraft basic Sub-System

The overview of an aircraft major systems such as the fuel system, electrical system, jet engine etc. is shown in figure 3.2.

Boeing and Honeywell are the designers of Aircraft Diagnostic and Maintenance System (ADMS) [53], [54]. Boeing first incorporated on-board health monitoring into the 747-400 in the late 1980s in the Central Maintenance Computer (CMC) to collect fault data from various system components and perform diagnostics [55]. Please see table 3.1 for list of reasoners and their use.

TABLE 3-1. LIST OF DIAGNOSTIC REASONERS

Intelligent Reasoner	Type	Known Applications	Company
CMC	Fault propagation modelling	Boeing 777; Primus Epic (business jets, helicopters)	Honeywell International
TEAMS Toolset	Multi-signal dependency modelling (advanced form modelling)	Consult Company	Qualtech Systems Inc.
eXpress Design Toolset	Dependency modelling (similar to fault propagation modelling)	Consult Company	DSI international
Livingstone	Artificial intelligence based reasoner (mixture of functional and parametric modelling)	DEEP Space One Spacecraft ; Earth observing one (EO-1) satellite	NASA Ames Research Centre
BEAM	Artificial intelligence based reasoner (mixture of functional and parametric modelling)	NASA Deep Space Missions (Voyager, Galileo, Megellan, Cassini and Extreme Ultraviolet explorer	NASA Jet Propulsion Laboratory

At the moment there are hardly any aircraft that actually have VLRS capability built in, they are mainly implemented at a research level. Those built in VLRS that are there, are based on pattern recognition diagnostics and basic detection.

There is very little information out there about VLRS implementation in military and civil aircraft, as the literature review shows us.

The Eurofighter, F22 and the Joint Strike Fighter are when it comes to defense aircraft the best developing health management system examples [56]. The VLRS of the JSF has advanced AI capabilities, although the JSF is still in the developmental stage. The AI algorithms are achieved by using mathematical techniques. We see less of the defence developments discussed in the literature due to the complex and sensitive nature of the subject.

In current VLRS development arguably the best example of the VLRS would be the developing example of program F35 (Joint strike Fighter), it has a highly integrated VLRS system for health management [56].

One of the difficulties in developing a health management system from the beginning has been in sifting out the objectives and requirements of the user to an acceptable level that contains buy in from the expected and diverse user groups.

3.2 Methodology

Customarily, the reasoning system is an AI based software and hardware application or it is a combination of hardware and software whose computational purpose it is to achieve conclusions from available information using rational techniques of deduction, prediction and diagnosing or other styles of reasoning [57]. Nonetheless, if the VLRS is constructed in a comparable manner as the system level diagnostics, the main difficulty would be algorithm merging. The convergence of the algorithm becomes more difficult because of the growth of the input of the diagnostics system. This will be a computational issue, because the VLRS has to oversee far too many factors and systems. The VLRS which is presented in this academic work is somewhat different in nature. The following part will discuss the VLRS architecture.

3.2.1 Proposed VLRS

VLRS can be implemented in different ways. Generally, it's linked to the major part of the sub-system's sensors or components, thereafter, the sensory signal will be analysed and processed by the diagnostic engine.

In order to drive the vehicle level reasoning in the presented approach, the diagnostics algorithm has to take thousands of the sensory information from many sub-systems. Most of the diagnostics algorithm such as Bayesian belief networks increase the complexity of the diagnostics reasoned rather than converging to produce effective diagnostics results.

However, this study exhibits a novel method for vehicle level reasoning. The presented technique is to consider the following:

1. Condition/health index of each sub-system of the vehicle.
2. Physical behaviour and the responses of the vehicle.
3. Sub-system diagnostics results along with probability where possible.

As presented in Figure 3.3 on the basis of the engineering principals' vehicle level reasoning engine will also assign weights to each input.

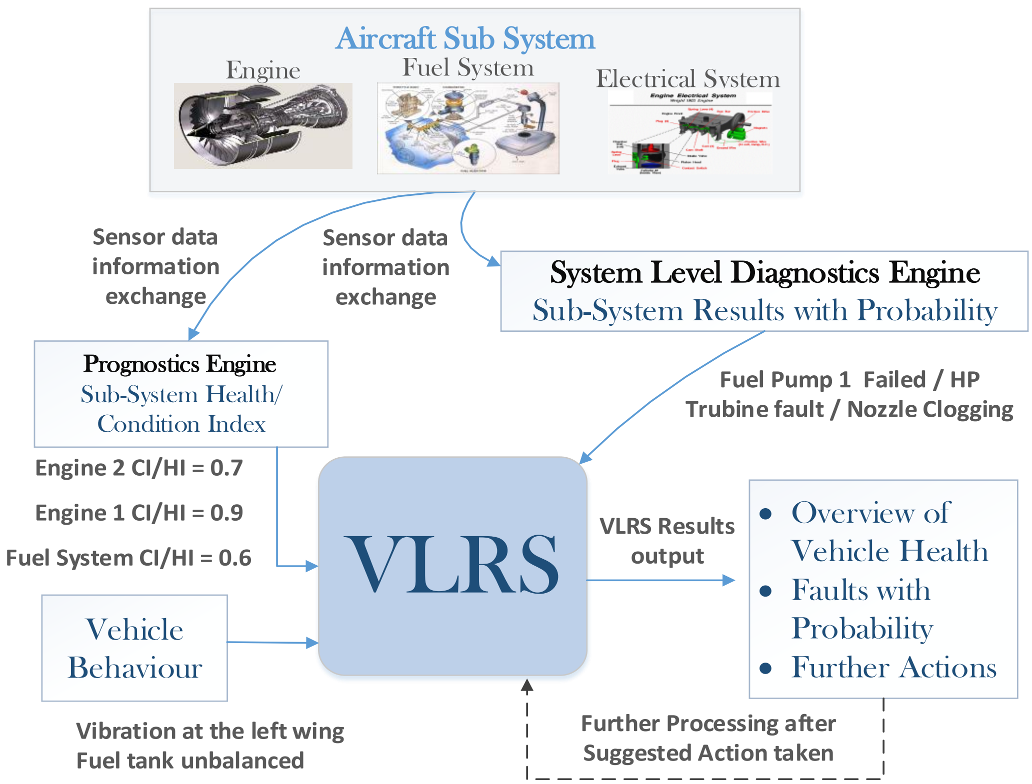


Figure 3.3. Presented VLRS with Inputs.

Figure 3.3 displays the inputs of VLRS, which are condition indicator (CI)/ health index (HI), sub-system results and vehicle behaviour of the vehicle. The details of the CI, sub-system results and behaviours are discussed later in this chapter.

3.2.2 Health Index of Sub System

The term, health index (HI), is described by a discrete number representing the actual health of the system. For example, 1 indicates the best health condition and 0 signifies the worst health condition. A mid-range value such as 0.5 represents half-life, see Figure 3.4. Following equation can be used to determine the virtual health index of the system:

$$HI = \frac{Current\ RUL}{EoL} \tag{3.1}$$

The formulised health index at a specific time point as the rate of current RUL (remaining useful life) to the end-of-life (EoL) of the system.

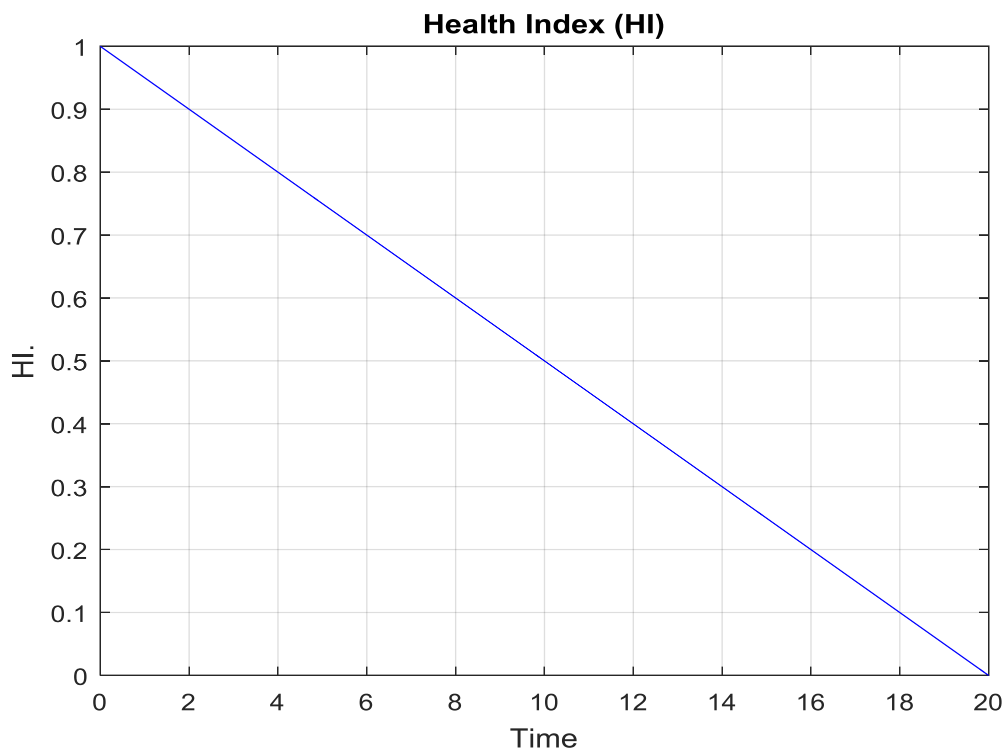


Figure 3.4. Initial HI (Health Index) for Component

The Figure 3.4 illustrates the linear HI calculated from the initial state of the system as the HI is calculated from the equation 3.1. However, when the sensor readings are received the HI calculation will be dynamic according to the sensor data received.

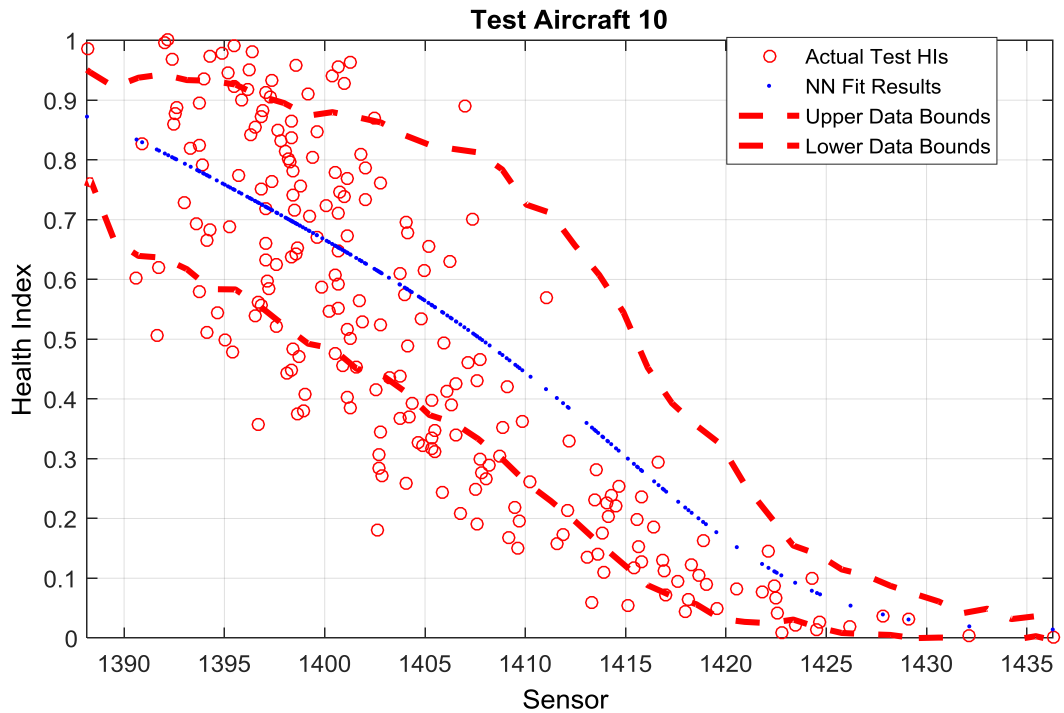


Figure 3.5. Degradation of the of health index of component /sub-system

Figure 3.5 displays the degradation of a sub-system; as the sub-system health degrades with time, the health index will decrease. The Y-axis is the HI, the X-axis is the sensor reading. The red outbound also provides the uncertainty of the calculation.

The HI has been fed to the VLRS and taken into account, as probability shows the components in a worse health index/condition are more likely to have a problem. Because this calculation will also be providing the RUL of the component, it is part of the prognostic engine. To prevent a catastrophic failure this warns the operator before the component/system completely crashes. The main focus of this chapter is VLRS rather than the calculation of HI in detail; therefore, further information on calculating the health index is provided in chapter 5 [58].

3.2.3 Vehicle Behaviour

The performance of the vehicle is one of the most important signs for the aircraft pilot. There are scenarios where the sub-system diagnostic engine does not supply any fault warning, or in the case of No-Fault-Found the pilot has to depend partly on the way the aircraft performance and other indicators.

In most cases, with an aircraft, it would be difficult to link certain behaviour to the defect(s), mainly because each behaviour could possibly be linked to many different problems or faults.

In addition, the behaviours and the performance of the aircraft have to generally be seen by the pilots of said aircraft. Therefore it has to be fed into the behaviour engine by hand. Diagnosing faults by using the performance and behaviour of the aircraft has been done before in different fields [59]. Aircraft flight behaviour studies are also discussed at simulation level in [60].

The behaviour and performance of the aircraft, which provides different fault possibilities to the main VLRS is discussed in this chapter. The behaviour of the vehicle can provide certain fault information. For example, in the case study presented in section 3.2.4, the pilot was not given any fault warning.

Nonetheless, the performance of the aircraft provided very beneficial information which was not noticed by the pilot. It was very clear that the aircraft was unbalanced and that the fuel tank on the one wing was indeed heavier than the other tank, due to the leak on one side.

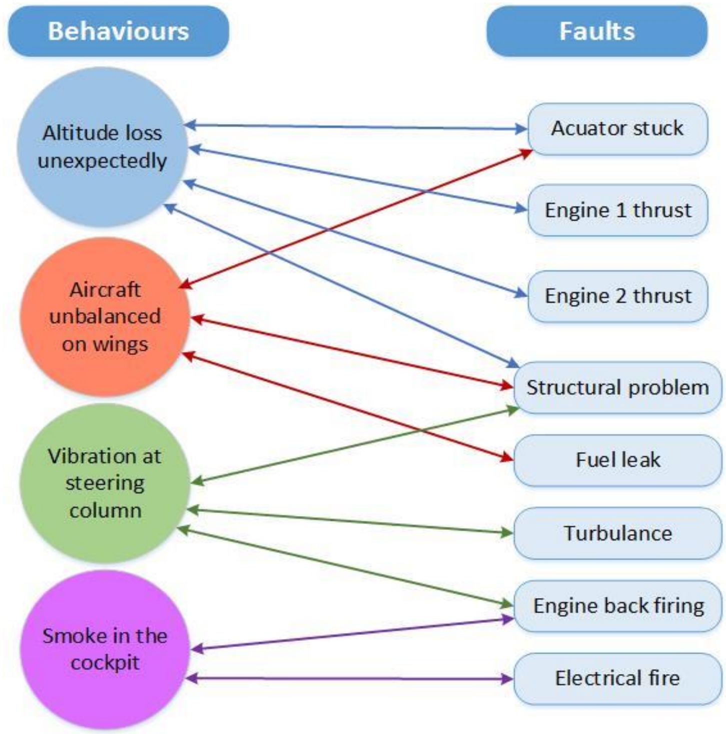


Figure 3.6. Behaviour relations to fault

Figure 3.6 describes the basic faults connected with the behaviours of the aircraft; however, this linkage regarding the behaviours is essential to be defined in detail by an expert. However, the vehicle faults and behaviour can be recorded and linked from the real accidents/incidents. Following is a real life incident case study which could be used for capturing the faults and their behaviour and other useful information.

3.2.4 Case study of a real incident

Aircraft details: Airbus A330, Manufactured in March 1999, France.

Fault Type: Fuel leak (at the entrance of the engine inlet pipe line).

Detection: No fault found (at the aircraft system) only warning has been issued for low oil temperature and high oil pressure.

Details:

Flight TS 236, took off from Toronto at 0:52 (UTC) on Friday August 24, 2001 (local time: 8:52 pm (ET) on Thursday August 23, 2001) bound for Lisbon.

At 05:16 UTC, in flight deck warnings have been issued for 'low oil temperature' and 'high oil pressure' at engine #2. There was no understandable connection between the oil temperature or pressure problem and fuel leak. The Captain Piché and First Officer DeJager assumed that the warning were false warnings and shared that opinion with their (MCC) Maintenance Control Centre, who advised them to observe the situation. At 05:36 UTC, the pilots received a warning of a left fuel tanks imbalance. The pilots followed a civil aviation procedure to resolve the imbalance by transferring fuel from the left wing tank to the right wing tank which was almost empty. At this point pilots were unaware regarding the real fault of the aircraft, which was a fuel leak in a line to the #2 engine. The fuel transfer from the left wing tank to the right wing triggered fuel to be wasted through the leak in the line to the #2 engine.

The damaged fuel line was leaking at a gallon per second approximately, which caused a higher than normal fuel flow through the fuel-oil heat exchanger (FOHE) causing the low temperature of the oil. The Portuguese Aviation Accidents Prevention and Investigation Department (GPIAA) investigated the accident together with Canadian and French establishments.

The investigation revealed the reason of the accident was a fuel leak in the #2 engine, caused by an incorrect part fitted in the hydraulics system. Air Transat maintenance staff had replaced the engine as part of regular maintenance, using a spare engine, lent by Rolls-Royce, from an older model. Despite the lead mechanic's concerns, Air Transat ordered the use of a part from a similar engine, an adaptation that did not maintain satisfactory clearance between the hydraulic lines and the fuel line. This lack of clearance caused the transfer of vibrations to the hydraulic lines to degrade the fuel line, causing the fracture which led to leak.

3.2.5 Sub-System Results

There are many sub-systems in every aircraft such as an engine, fuel system, hydraulics system, electrical system etc. shown in Figure 3.7. Latest civil aircrafts are equipped with a sub-system diagnostics/detection system, which reports faults and errors to the flight deck. The health status is displayed to the flight deck for further action to be taken by the pilot. Most of the systems are connected and dependent on each other, such as the electrical system is dependent on generators; generators are dependent on the engine; engines are dependent on the fuel system; fuel system is dependent on the electrical system. If one system breaks down in the aircraft it will affect the entire chain of systems. However, further problems will be reduced due to the backup in most of these systems. The linkages between the systems are still very important.

An advanced diagnostics system is not very commonly used in most of the civil aircrafts. Usually, they are equipped with a simple diagnostics / detection system where a range of thresholds has been fixed, hence, a warning to the pilot flight deck is generated when the sensor values are out of the pre-set threshold values.

3.3 VLRS Engine

The main purpose of a vehicle level reasoning system (VLRS) is to identify faults and failures at the aircraft level. The VLRS receives health information from sub-systems and fuses them to provide an overall health state of the aircraft. The aim of VLRS is not to replace the diagnostics system but merely to aid the pilot in providing extra beneficial information which would assist in the decision making.

The presented VLRS takes the fault report or health status from several sub-systems. The sub-systems, in this chapter, are supposed to be the fuel system, engine, and electrical system etc. as shown in Figure 3.7

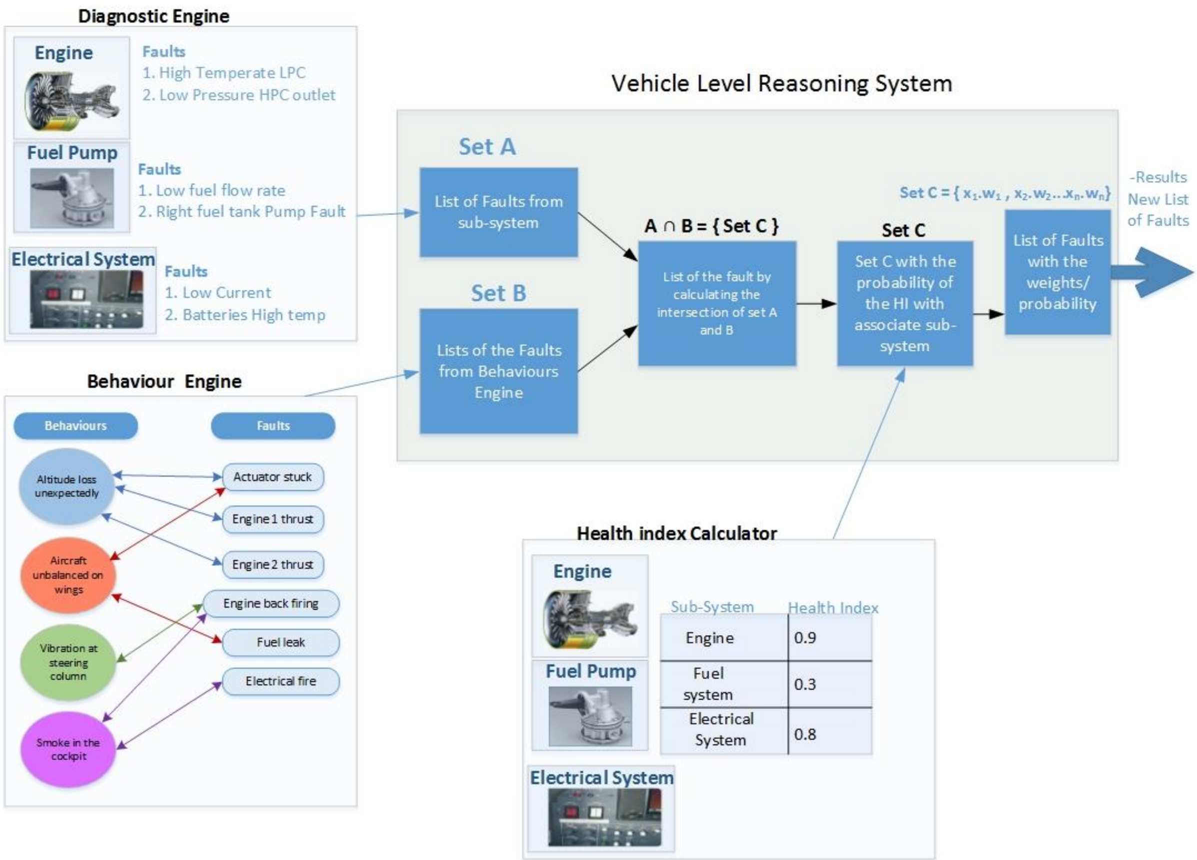


Figure 3.7. VLRS Engine

The faults are logged with time stamps in the demonstrated VLRS. In case of multiple break down, it would be helpful to identify the root cause of the failure. Commonly, the first failure would be the source of other failures.

The comprehensive design of the reasoning system has been demonstrated in Figure 3.7. The outcomes from the sub-system reasoner are stored into a list, let's say set A and the behavior of the aircraft has been inserted into the behavior engine which provides the fault list associated with the behavior which are stored in a list (set B). The intersection of Set A and B are stored and given weights according to the health index of the sub-system. Once the health indexes are given to their sub-system, the faults which have a higher weight are more likely to be the real fault.

There are occasions when there are faults in the aircraft which do not get detected or the wrong fault gets detected by the diagnostics/detection system. These occasions can be taken as case studies and as a starting point to see how the VLRS system performs.

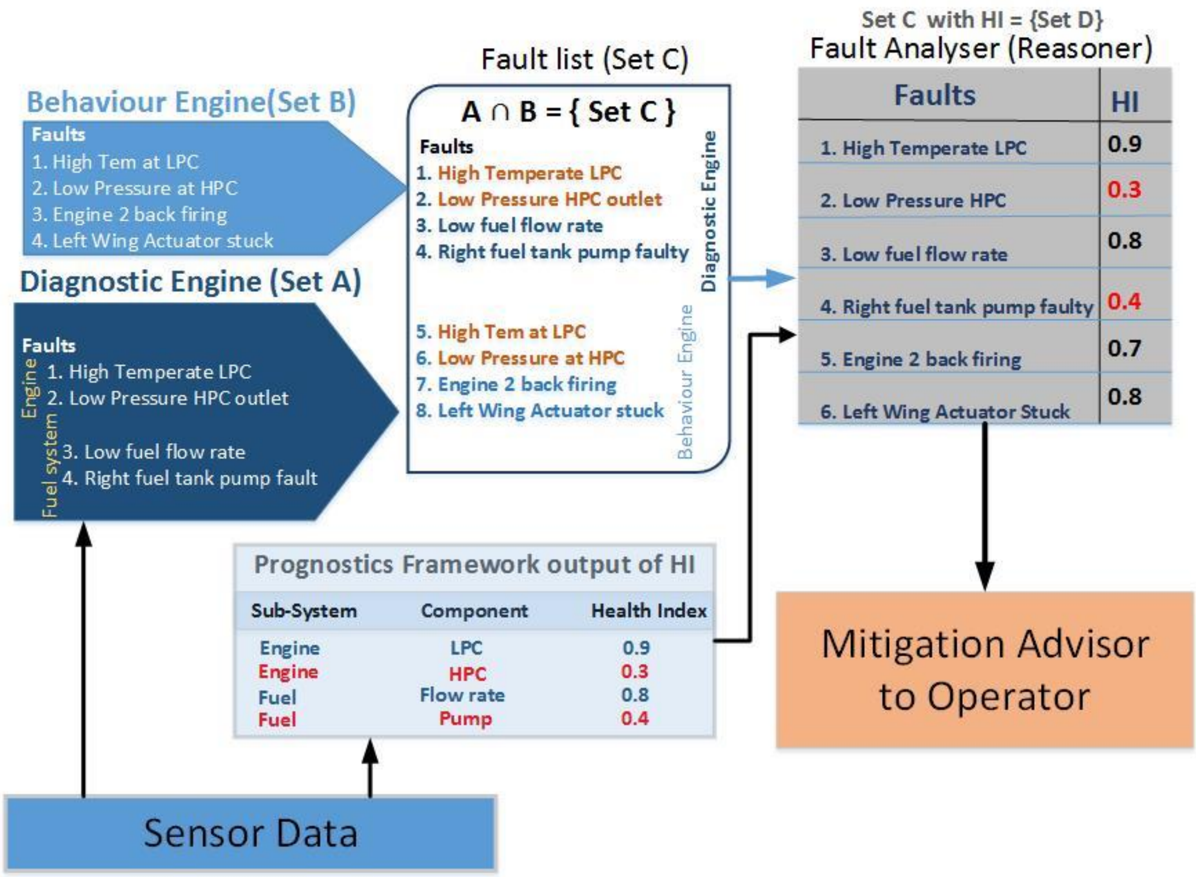


Figure 3.8. Reasoner of VLRS

Figure 3.8 demonstrate the reasoner of the presented VLRS. In this figure it shows how the results are fused from the sub-system diagnostics (Set A) and from the Behaviour engineer (Set B). The fusion is done in very simple way as simple addition of the both list. Which makes the new fault list (Set C), thereafter the HI are been marked on the component which are listed in the fault list. The HI is used in a similar manner like a probability of the fault occurrence. The final list after the fault along with the HI/probability will be fed to the flight deck and the mitigation advisor which would advise for further actions to take for the operator / pilot. However, the design of the mitigation unit is out of scope of this thesis.

3.4 Summary

This chapter presents the design of the VLRS and its use in aerospace applications. VLRS designs and its use in aerospace applications have been presented in this chapter. Some of the challenges to VLRS have also been discussed. The ambiguity in diagnosis is mainly due to the lack of sensing data. However, if more diagnostic related data is available from other sub-systems an improvement in reasoning accuracy can be made, particularly through using the health index, health status of other sub-systems and vehicle behavior of each sub-system. We proposed, in this chapter, that vehicle-level reasoning systems would significantly improve, if this information would be used to accomplish the diagnostics results. The presented case study of an actual accident demonstrates that if these VLRS like systems would have been equipped in the aircraft, the accident could have been avoided with the help of the VLRS.

The Health index (HI) has been used as a weight/probability; however optimisation of these weights could improve the results. This led to the problem how to calculate the HI and how to give probability of failure to each component/system. How the health index has been calculated of each component is discussed in chapter 5. However, it is very apparent that the adaptable and scalable degradation systems are in need of developing, which could provide many benefits, including it being used in the VLRS. The next chapters will discuss more how to develop the adaptable, scalable degradation prognostics system and HI calculation of each component.

Chapter 4

Operational Scalable Data Driven Technique for Prognostics

This chapter provides information regarding the essence of prognostics, problems of prognostics predictions, and the approach presented for a single component / system. The approach has been explained and the results are also provided.

4.1 The Essence of Prognostics

The main involvement of prognostics is the prediction of the present health and the calculation of the remaining useful life of the system or component. In the literature review (chapter two) different types of prognostics modelling such as Physical based modelling and data driven modelling have been discussed. However this section only will explain the present problems with the data driven prognostics.

To perform the data driven prognostics modelling, the first requirement would be to have run-to-failure data of the focused system. Generally, most of the run-to-failure data are collected on the accelerated test and in a laboratory environment which are very different in time scale as compared to the real environment. When these datasets are used to create models, those models are mainly made for the environment of the laboratory and for accelerated data. In a real environment the degradation pattern might be similar but the time scales are completely different.

The next section will discuss the issues and problems with the present data driven modelling and the solution provided along with the methodology and results.

4.2 Operational-Scalable Degradation Model for Filter Clogging Fault Prognosis

Prognostics in general involve a construction of a degradation model and then use it to estimate a current degradation state and to predict EoL of a component. The purpose of a model is to capture how degradation evolves over time. The current degradation state is estimated based on noisy measurement update and possible paths (determined by the model) from the previous degradation states. The model then uses the estimated degradation state to propagate the state until it reaches EoL at some pre-defined degradation threshold. The other operational parameters, like usage or component specific degradation coefficients, can also be taken into account in the state estimation and EoL prediction. The model itself forms a key part in prognostics. It acts as a physical constraint which is used to determine not all possible but the most likely degradation paths.

4.2.1 Problem with the Presented Prognostics

There are two commonly used prognostic approaches; one is 'data-driven' and the other is 'model-based'/"physics base model". In the data-driven approach, multiple training data sets are used to learn a model. The resulting model can be considered as a nominal model which on average represents the degradation paths used in training. Notable data-driven modelling approaches are Time-Delay Neural Network (TDNN), Hidden Markov Model (HMM), Gaussian Process (GP) and Fuzzy Inference System (FIS). The key advantage of these approaches is that a complex nonlinear degradation path can be easily captured in the model without having to know the exact physics of the system. However, the model structure itself will not naturally reflect the physics of the system despite modelling the degradation paths, which is somewhat expected.

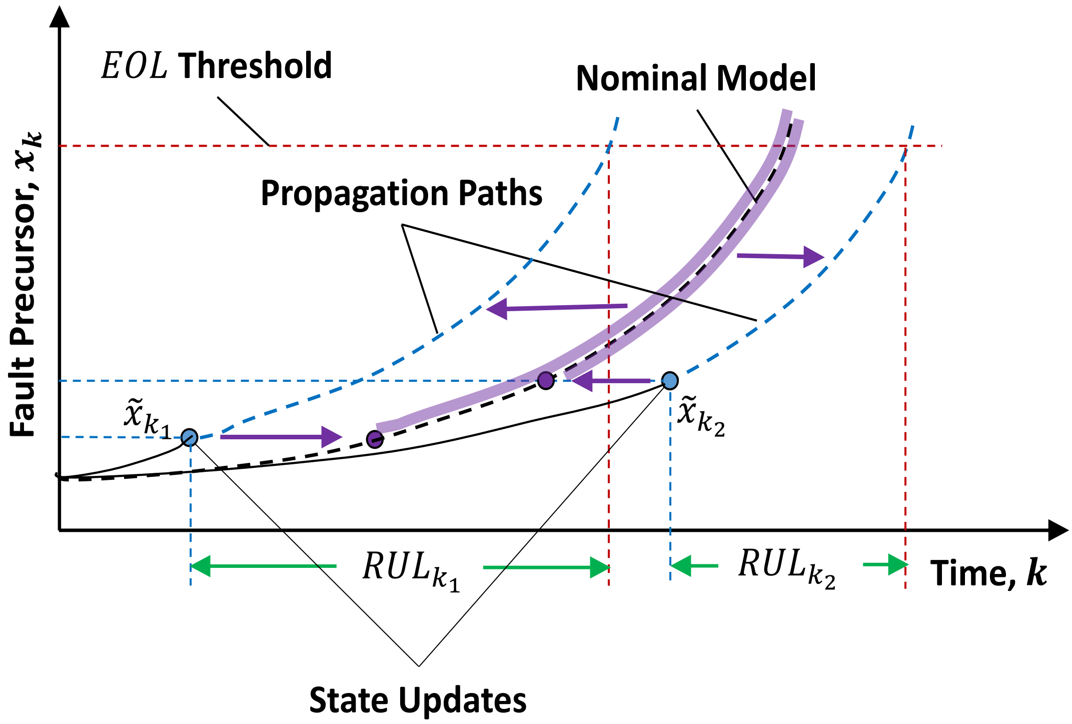


Figure 4.1. Data-Driven Prognostics Approach

Figure 4.1 conceptually describes how prognosis is carried out in the data-driven approach, where $k, x, \tilde{x}$, EoL and RUL are a discrete time, system degradation state (e.g. fault precursor), estimated degradation state, EoL threshold and remaining useful life, respectively. At each time step, a prediction is carried out by referring to the nominal model. $\tilde{x}$ is calculated from a measurement update and used as an initial condition for propagation of a degradation path. The degradation path is propagated until the predicted x reaches the EoL threshold. Actual RUL at time k simply equals to EOL-k. In the data-driven approach, the prognostic process assumes there are no changes in the system and operating conditions except the initial condition for degradation path propagations.

The prognostic algorithms developed using the data-driven approach will generally perform well if the runtime (or test) operating condition is similar to the conditions used to generate the training data. The overall prognostic accuracy will decrease and become more apparent as the difference in the operating condition from the trained condition increases. However, as the model is some kind of an average degradation path, if trained operating conditions are significantly varied, then the prognostic performance will also not be good, in particular for the outliers, even though the operating condition is similar to one of the trained conditions. In the data-driven approach, physical properties of a system are in general not explicitly parameterised in the degradation model. Consequently, the model will not be able to adapt when there are changes in operating conditions or system physical properties, and hence a decrement in prognostic performance when operating away from the trained condition.

In the model-based approach, a model is in a generic sense referring to a degradation model which can be mathematically formulated using physical knowledge of a system. In contrast to the data-driven approach, the model is in a form of mathematical equations where the component's physical properties can be explicitly parameterised in the model. The usage or system aging factors can be directly or indirectly related to the model parameters. Notable techniques related to the model based approach are Physics of Failure (PoF) [61], Particle Filter (PF) and Kalman Filter (KF). PoF is a modelling approach, but PF and KF are not; they are stochastic estimators. Figure 4.2. Model-Based Prognostics Approach) conceptually describes how prognosis is carried out in a model-based approach. Addition nomenclatures θ, $\tilde{\theta}$ and f (·, ·) are a model parameter (e.g. usage or system aging factors), estimated model parameter and degradation equation, respectively. From measurement updates, true values of degradation related variables are estimated. $\tilde{x}$ and $\tilde{\theta}$ estimates are continuously updated using PF or KF whenever measurements become available. They are the values that are most likely to generate a prediction x consistent with the measurements. In contrast to the data-driven approach, the model dynamics used in propagating a degradation path will vary depending on the estimated $\tilde{\theta}$. This way the real degradation path can be more accurately predicted, which otherwise will either be over or under estimated as illustrated in Figure 4.2.

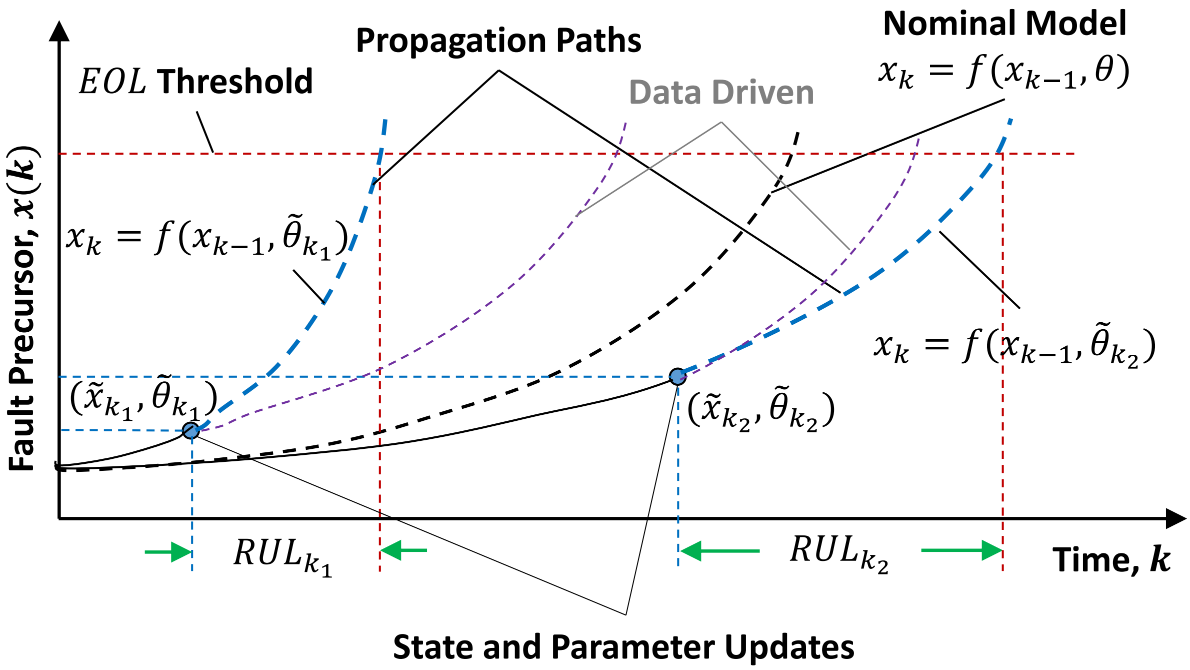


Figure 4.2. Model-Based Prognostics Approach

A physical model and PF (or KF) are always used in combination with model based prognostics. This way the degradation model is able to adapt according to the real usage or component specific aging conditions, and hence good prognostic accuracies can be expected over a wide range of operating conditions. However, mathematical formulation in itself can be a disadvantage in the model based approach. Physics of component degradation may not be known. A complex nonlinear degradation path can be difficult to formulate mathematically. Therefore, if a derived model does not sufficiently capture the degradation physics of a component, then a poor prognostic performance is likely to be expected.

4.3 Operational Scalable Prognostics Model

In terms of operating condition, the model-based approach has certain advantages over the data-driven approach. However, in many cases, degradation physics of a system is not always known and hence a model cannot be formulated. On the other hand, a degradation model can be physically well defined but has too many unknown parameters that need to be assigned. This makes the model-based approach difficult to apply in these situations. In this thesis, we propose that a data-driven model can be used in combination with a model-based estimator to address this shortcoming. It is important to recognise that multiple training data sets of a system will have similar degradation patterns. Hence, in contrast to how data-driven models are conventionally constructed, the focus should be on extracting a generalised degradation pattern rather than trying to learn a model that overall fits the data. The model should be learned in such a way that there are explicit parameters (or inputs) that can scale the model to fit data differed in degradation condition. By using the scaling parameters, a data-driven model constructed in this way will be able to adapt when there are variations in usage condition or component physical properties.

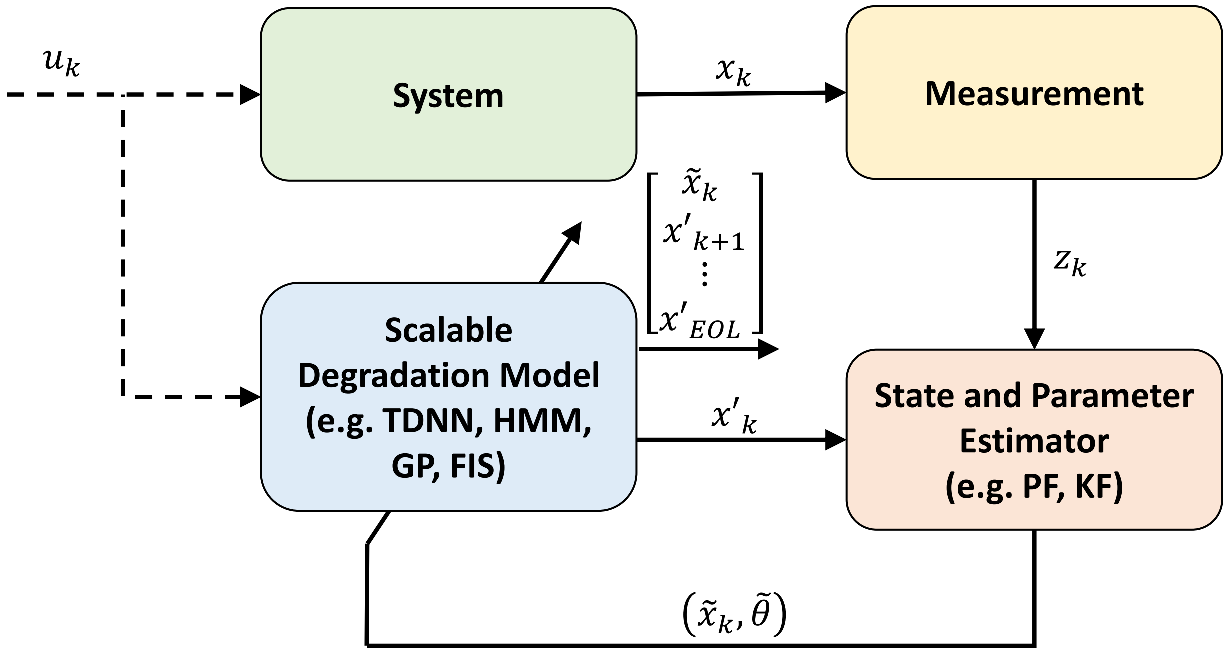


Figure 4.3. Operational-Scalable Data-Driven Prognostics Block Diagram

Figure 4.3 depicts a prognostics architecture where a data-driven degradation model is used in combination with a model-based estimator (e.g. PF or KF). Additional nomenclatures u, z and x' are a system input, noisy measurement update and model state prediction, respectively. In this prognostic approach, a data-driven model is parameterised (or inputted) by a scaling parameter θ, e.g. usage or system aging factors. The model predicts a current degradation state x'_k by taking previous estimated degradation states $\tilde{x}_{k-1}$,..., $\tilde{x}_{k-N}$ (and other system input information u_k) as input. An estimator determines the most likely degradation state $\tilde{x}_k$ and scaling parameter $\tilde{\theta}_k$ from the predicted state x'_k and measurement z_k. $\tilde{x}_k$ and $\tilde{\theta}_k$ are then fed back to update the predicted current state and model scaling parameter. Changes in θ reflect variations in operating condition and system specific properties. The degradation path $[\tilde{x}_k, x'_{k+1}, \dots, x'_{EOL}]'$ is then further propagated until the predicted state x' reaches the EOL threshold. In this approach, the degradation paths are propagated using the updated scaling parameter estimated during runtime instead of using a nominal value as in a conventional data-driven prognostic approach. This way a data-driven based propagation path will be able to adapt to match real current usage and system variations.

4.3.1 Filter Clogging Problem

Filtration is the removal of suspended particles from a fluid performed by a filter medium. It is an important process in many engineering systems where purification of fluid is required. Failures due to clogging are the most common for a filter. The consequences of filter failure are more pronounced in safety critical systems such as aircraft fuel and engine systems. In aircraft, fuel filter clogging is commonly caused by fuel contamination and debris from degraded fuel pumps. If a filter at the engine fuel inlet becomes clogged, it can cause an aircraft to return to the ground or divert and a remote possibility of engine shutdown [62], [63]. Moreover, pressure build up can consequently cause leakage (or damage) in other parts of the fuel system. Minimising unexpected downtime is key for an aircraft operating in a commercial environment. Hence, there is a need for a capability to monitor and accurately predict degradation in a critical system, or a component like filter to be developed and matured. This will enable maintenance to be planned and performed before clogging reaches a level allowed for safe operation, therefore minimising unexpected downtime and associated cost.

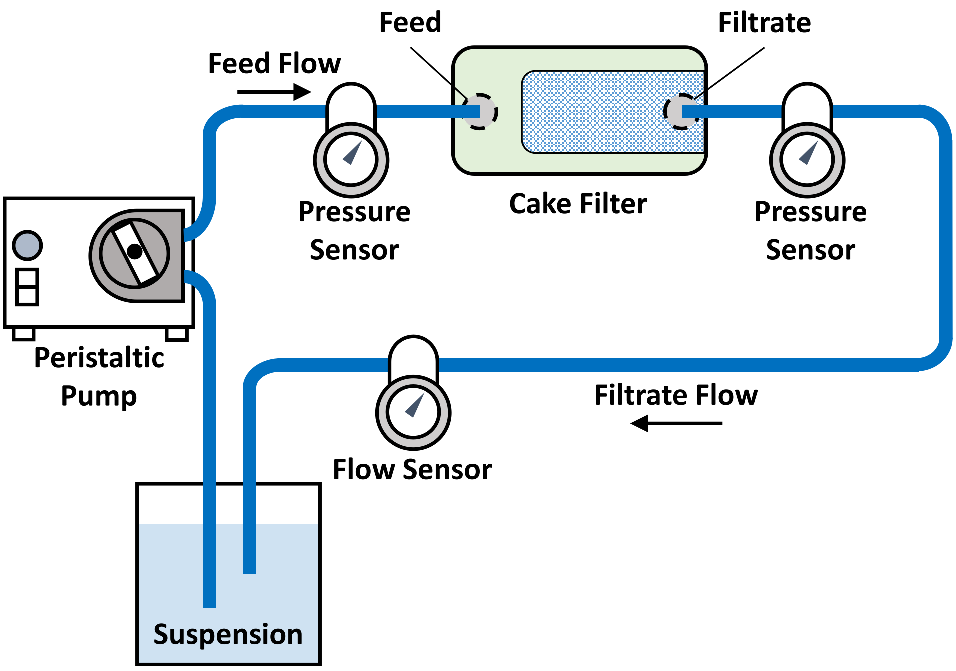


Figure 4.4. Filter Clogging Experiment Setup

In our study on the filter clogging the experimental data is similar to the one used in [64].

An overview of the experiment setup where the data was collected from is shown in Figure 4.4. The experiment was designed to create filter clogging using continuous cyclical flow. The filter used in the experiment is Baldwin's fuel filter BF7725 model, whose pore size is 125 μm. The filters were tested using different solid-ratio suspension (slurry) composed of Polyetheretherketone (PEEK) particles and water. The particle size is 61 μm on average with 24 μm standard deviation. During each test, the peristaltic pump was kept at constant speed at around 211 revolutions per minute (RPM). However, there was no control on the actual pump speed or flow rate, and therefore in addition to solid ratio, variations in the operating conditions can occur during the test and as well between the tests. Two pressure sensors and one flow sensor were used for data collection. It has been shown that the pressure drop ΔP across the filter is a good precursor for detection of incipient clogging. In this setup, the data collections were stopped when the pressure drop had reached about 35 psi.

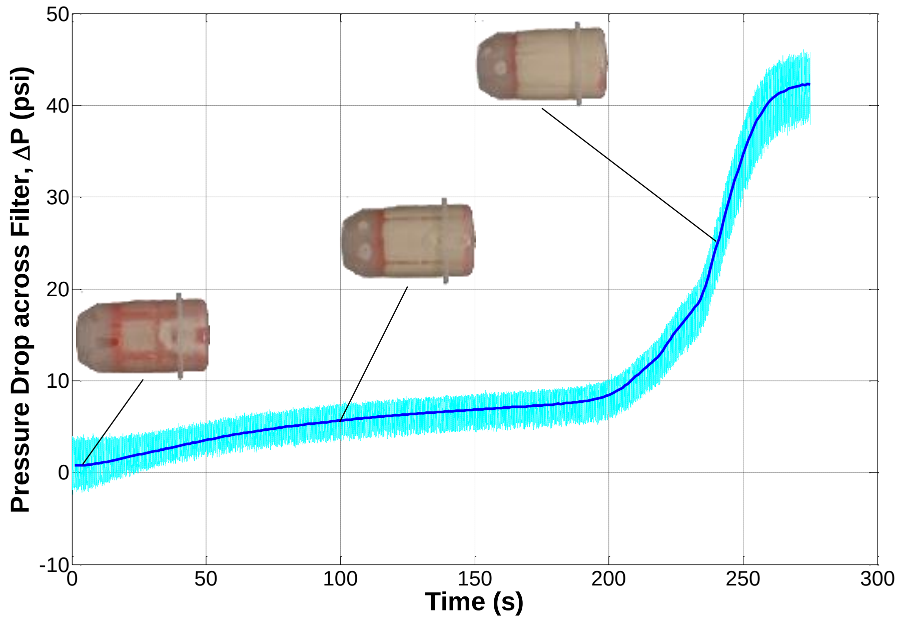


Figure 4.5. Raw and Filtered Filter-Clogging Data

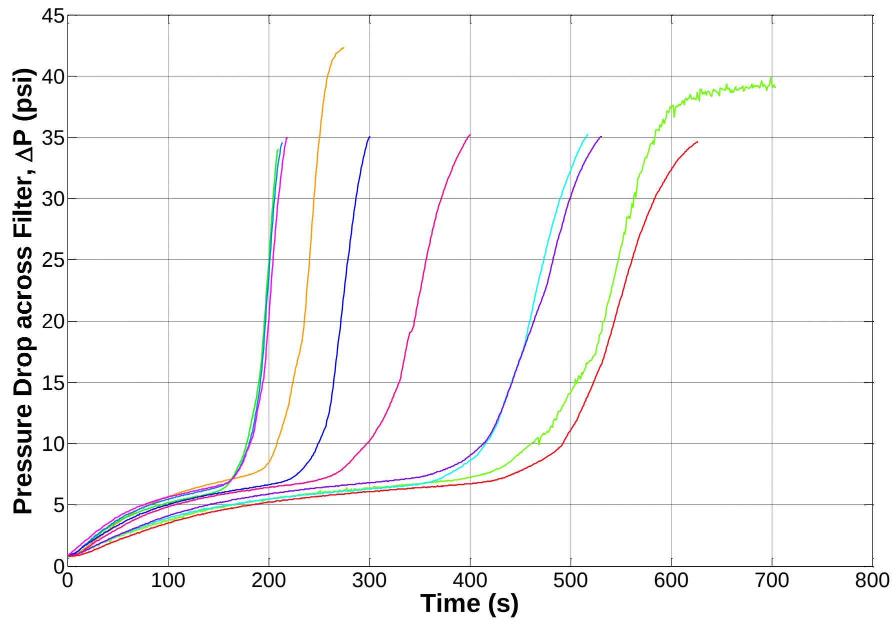


Figure 4.6. Filter Clogging Data Profiles

Figure 4.5 and 4.6 show examples of run-to-failure data collected during the experiment. In this study, raw pressure data are filtered using a low-pass Chebyshev Type I Infinite Impulse Response (IIR) filter [65] of order 8 and down sampled from 1,000 Hz to 1 Hz, see Figure 4.5. The filtered data in Figure 4.6 shows a similar degradation pattern for all the 10 samples. In this particular setup, three clogging stages can be observed; 1) almost zero ΔP, 2) steady increase in ΔP and 3) rapid increase leading to saturation in ΔP, see Figure 4.5. It is worth comparing the filter's pore size (125 μm) with the particle size (61±24 μm). From the distribution, 2% of the particles will be greater than the pore size. In the beginning of filtration (stage 1), only a small proportion of solid particles are retained on the filter's mesh surface. The solid suspension can pass through the filter almost without restriction, hence a very small build-up in ΔP. More large particles are retained as a more solid suspension is filtered and eventually providing filtering action for the smaller size particle. As time goes by the thickness of the cake increases, as more particles are filtered. This results in a steady increase of the pressure resistance across the filter; stage 2. As more solid suspension is filtered, more particles build up from the filtrate end, and the cake has increased till it touches the walls of the filter, which decreases the filtrations area dramatically; stage 3 clogging. This results in a rapid increase in ΔP as the filter surface area decreases. Recall that the pump was kept at a constant RPM, but not the flow rate. With the same RPM, the pump is unable to provide the same flow rate as the flow becomes too restricted; thus saturation in ΔP as the flow rate decreases.

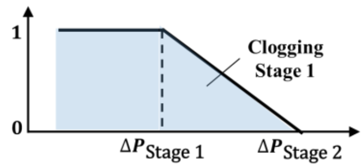
4.4 Operation-Scalable Takagi-Sugeno Fuzzy Model

4.4.1 Multiple Linear Models

Takagi-Sugeno (TS) fuzzy model approach provides a useful and uniform framework for system identification and modelling of non-linear systems [66], [67]. TS fuzzy models are weighted combinations of multiple linear models, and can be used as a universal approximator with good interpolation and extrapolation characteristics. TS models can either be derived from input-output data via system identification or from mathematical models of nonlinear systems. In this chapter, a TS fuzzy model is used to illustrate the concept proposed in (section 4.3 Operational Scalable Prognostics Model). The model structure is transparent in terms of its linear dynamics. This is useful in providing an intuitive example of how a data-driven model can be scaled to match a variation in system operating conditions.

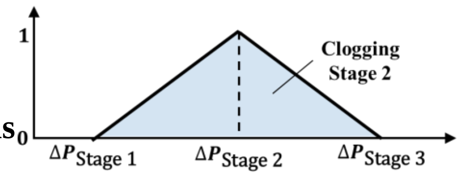
In our case, a TS fuzzy model is constructed as a data-driven model for the filter clogging process shown in Figure 4.7. The dynamics of ΔP is system identified by fitting a grayscale TS model to the time series clogging data. Following the clogging stages described in section 4.3, a TS fuzzy filter clogging model can be described by a set of N (=3) fuzzy rules as follows:

Rule 1: IF ΔP is



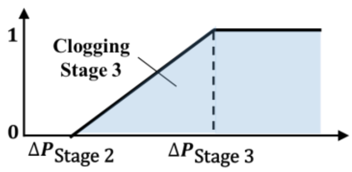
THEN $\Delta \dot{P} = A_1 \Delta P + B_1$.

Rule 2: IF ΔP is



THEN $\Delta \dot{P} = A_2 \Delta P + B_2$.

Rule 3: IF ΔP is



THEN $\Delta \dot{P} = A_3 \Delta P + B_3$.

Figure 4.7. Takagi-Sugeno Fuzzy Rules

A_i and B_i are local model parameters. For simplicity, the trapezoidal and triangular membership functions $\mu(\cdot)$ are used in this chapter. These membership functions are defined by scaling between 0 and 1 how representative the local clogging models are for a certain value of ΔP. However, filter clogging is a continuous dynamic process, smoothly transitioning from one clogging stage to another clogging stage. In a TS fuzzy model, this fact is reflected by the defuzzification process in which $\Delta \dot{P}$ can be calculated as:

$$\Delta \dot{P} = f(\Delta P)$$

$$= \sum_{i=1}^{N} w_i(\Delta P)(A_i \Delta P + B_i). \quad (4.1)$$

The normalised weighting functions $w_i(\cdot)$ are given by the fuzzy inference

$$w_i(\Delta P) = \frac{\mu_i(\Delta P)}{\sum_{j=1}^{N} \mu_j(\Delta P)}. \quad (4.2)$$

It is worth noting that the number of fuzzy rules (i.e. local models) need not be limited to 3. N can be set without having to consider the physics of a system.

Intuitively, the higher number of N the more accurate the model is. However, this also means more difficulty in learning the model; more parameters need to be identified.

4.5 Rate Variation and Scaling

From Figure 4.8, it can be observed that all the 13 samples have a very similar degradation (or clogging) pattern. The only difference is in the clogging rate. The time that ΔP reaches a certain value, let say 15 psi, happens earlier or later than one another. This is due to variations in operating condition and filter physical properties.

Table 4-1. Examples of Linear Time-Invariant System.

Differential Equations	Integral Solutions
$\dot{x} = B$	$x(t) = Bt + x(0)$
$\dot{x} = Ax$	$x(t) = x(0)e^{At}$
$\dot{x} = Ax + B$	$x(t) = \frac{B}{A}(1 - e^{At}) + x(0)e^{At}$

The question is 'How can the variation rate be intuitively captured in the model?' Let's consider linear time-invariant (LTI) differential equations and their integral solutions shown in table 4.1. These equations are, in fact, in the form of local TS fuzzy models described in section 4.1. The first two can be the case where A or B are very small. Suppose the integral solutions are scaled in time by a constant θ, i.e. $Bt \rightarrow B(\theta t)$ or $e^{At} \rightarrow e^{A(\theta t)}$. In our case, this means clogging will reach a certain level earlier or later depending on the value of θ. Note $B(\theta t)$ and $e^{A(\theta t)}$ can be rewritten as $(B\theta)t$ and $e^{(A\theta)t}$, respectively. From table 4.1, by replacing A with θA and B with θB, the differential equations become θB, θAx and $\theta(Ax + B)$; hence multiplying the rate equation is equivalent to time-scaling the integral solution. Conceptually, if $\dot{x}(t) = Ax(t) + B$ is a nominal degradation model, then individual data samples can be closely fitted with $\dot{x} = \theta(Ax + B)$ by choosing the right value of θ. The θ values will be different for different data samples.

Following this line of reasoning, the TS fuzzy model defined in (4.1) can be rewritten to include an operational-scalable parameter as

$$\begin{aligned}\Delta\dot{P} &= f(\Delta P, \theta) \\ &= \sum_{i=1}^{N} w_i(\Delta P)\gamma_i(\theta)(A_i\Delta P + B_i).\end{aligned} \quad (4.3)$$

The local scaling functions $\gamma_i(\cdot)$ can be given by

$$\gamma_i(\theta) = C_i\theta \quad (4.4)$$

Where θ and C_i are a global scaling parameters and constants for a specific local model, respectively. C_i allows the rates to be scaled differently for different local models. The nominal degradation path will have $\theta = 1$. $\theta > 1$ will represent the data samples that have faster clogging rates. The time that ΔP increases to a certain level will be earlier than the nominal model. Conversely, $\theta < 1$ will be for the data samples that have a slower clogging rate. However, (Eq. 4.3) is a nonlinear system where multiple local linear models are scaled, weighted and combined. In this case, multiplying the rate will approximately be scaling the time of the integral solution, but it will not be an exact equivalent of rate and time as for a LTI system.

4.6 System Identification

Sections 4.4 describe the structure of the TS fuzzy degradation model. However, for the model to closely approximate the clogging process, the model parameter values need to be tuned (or system identified) from the collected data samples. The aim is to minimise the prediction errors between the experimental clogging data and simulated data from the TS fuzzy model. In this chapter, where there is no ambiguity, the terms 'optimise' and 'minimise' are used interchangeably.

In this study, Mean Square Error (MSE) is used as a performance measure for system identification of the TS clogging model defined in (Eq.4.3). For a given data sample, MSE can be evaluated using

$$MSE = \frac{1}{N_D}\sum_{k=1}^{N_D}(\Delta P'_k - \Delta P_k)^2, \quad (4.5)$$

Where ΔP, $\Delta P'$ and N_D are a train data (in this case pressure drop across filter) sample, predicted value calculated from the model and number of data points in the sample, respectively. However, our model is aimed to generalise a degradation pattern from multiple data sets. Hence, the measure has to combine the MSE values evaluated against multiple train data. In this chapter, Euclidean norm L_2 is used as a combined cost function, it can be defined by

$$Cost = \left(\sum_{j=1}^{N_S} MSE_j \right)^{1/2}, \quad (4.6)$$

Where N_S is a number of data sets used in learning the model. To learn the model, we minimise $Cost$, i.e. overall minimising MSE_j. In this case, the optimising parameters for the nominal TS fuzzy degradation model are the values characterising the membership functions $\Delta P_{\text{Stage } i}$, the local model parameters A_i, B_i and the scaling constants C_i. By expanding (Eq. 4.3), it can be realised that C_iA_i and C_iB_i are essentially 2 constants. Substituting C_iA_i with A_i and C_iB_i with B_i, (Eq.4.3) can be rewritten with mathematically equivalent as

$$\Delta \dot{P} = \theta \sum_{i=1}^{N} w_i(\Delta P)(A_i \Delta P + B_i). \quad (4.7)$$

Therefore there are in total $3 \times N(= 3) = 9$ local model parameters instead of 12. In addition to θ_j there are also the free parameters that scale the nominal model to fit N_D training data.

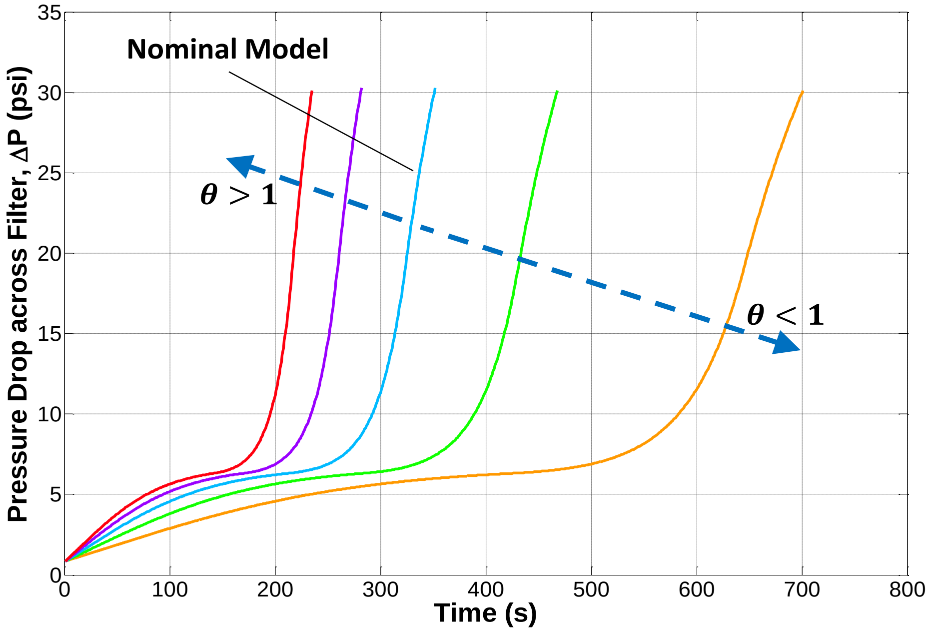


Figure 4.8. Scalable Filter Clogging Profiles

In this research, a global optimisation algorithm, namely Scatter Search [68], is used to find $9 + N_S$ optimal values for $\Delta P_{\text{Stage}\,i}$, A_i, B_i and θ_j. The search is implemented using the `GlobalSearch` function of the MATLAB's Global Optimisation Toolbox. In this chapter, we used 6 of 10 data samples for training the model and the other 4 for prognostic testing. The sampling period T is 1 s. The optimised model parameters, where $Cost$ is minimal, are $\Delta P_{\text{Stage}\,i} = \{-3.6322, 6.3099, 18.9145\}$, $A_i = \{0.0, -0.0155, -0.0168\}$ and $B_i = \{0.0, 0.1044, 0.8134\}$. Figure 4.8 shows variation in time response as the nominal model is scaled by the parameter θ from 0.5 to 1.5. The initial condition $\Delta P(0)$ is selected as 0.8 for this simulation. For pressure drop across filter ΔP at 15 psi, the corresponding time scales are 0.6688, 0.8025, 1.0, 1.3312 and 1.9936 for the scale parameter θ values of 0.5, 0.75, 1.0, 1.25 and 1.15, respectively. It can be shown that the scaling parameter θ is correlated with the resulting time scale, where $1.12113\theta - 0.1121$ is the linear regressed output time scale at $\Delta P = 15$ psi.

4.7 Particle Filter

4.7.1 Probabilistic Model

In order to predict EOL of a system, the system's current state of degradation and unknown operational parameter have to be continuously estimated from the measurement updates. In this case, the fault precursor ΔP and the scaling parameter θ are the variables that are required to be estimated. Kalman and Particle Filters are commonly used for estimating the degradation state and unknown model parameter. In this chapter, PF is used as it is more applicable to general non-linear systems. The estimation process of PF is based on the discrete stochastic state transition and measurement equations:

$$\Delta P_k = f(\Delta P_{k-1}, \theta_{k-1}) + \mathcal{N}(0, \sigma_{\Delta P}^2) \quad (4.8)$$

$$\Delta \tilde{P}_k = \Delta P_k + \mathcal{N}(0, \sigma_{\Delta \tilde{P}}^2) \quad (4.9)$$

Where $\Delta \tilde{P}$,$\mathcal{N}(\cdot,\cdot)$, $\sigma_{\Delta P}^2$ and $\sigma_{\Delta \tilde{P}}^2$ are noisy measurement pressure, Gaussian random function, process uncertainty variance and measurement noise variance, respectively. In PF framework, (4.8) and (4.9) are equivalently formulated in terms of probability density functions as:

$$\Delta P_k \sim p(\Delta P_k | \Delta P_{k-1}, \theta_{k-1}) \quad (4.10)$$

$$\Delta\tilde{P}_k \sim p\left(\Delta\tilde{P}_k \middle| \Delta P_k, \theta_k\right) \quad (4.11)$$

where $p(\cdot)$ is a probability density function and $\sim$ means 'sample from'. $f(\Delta P_{k-1}, \theta_{k-1})$ is deterministic and known through optimisation as described in section 4.3. The stochastic parts $\sigma^2_{\Delta P}$ and $\sigma^2_{\Delta\tilde{P}}$ of (Eq. 4.8) and (Eq.4.9) are unknown and need to be evaluated as required by the PF probabilistic framework. In this chapter, Random-Walk Metropolis-Hastings (MH) algorithm [69, 70, 71, 72] is used as a Markov Chain Monte Carlo (MCMC) sampling method to estimate probability distributions of $\sigma^2_{\Delta P}$ and $\sigma^2_{\Delta\tilde{P}}$. The pseudo code of a generic MH algorithm is listed in algorithm 1, where $x_i = \left(\sigma^i_{\Delta\tilde{P}}, \sigma^i_{\Delta P}\right)$, $q(\cdot \mid \cdot), \pi(\cdot)$, $\mathcal{U}(\cdot,\cdot)$ and $\alpha(\cdot \mid \cdot)$ are sample at i^{th} iteration, Gaussian proposal distribution, full joint density, uniform random function and acceptance probability, respectively. A proposal (or candidate) sample x' is computed using

$$\sigma'_{\Delta P} = \sigma^{i-1}_{\Delta P} + \mathcal{N}(0, \sigma^2_x) \quad (4.12)$$

$$\sigma'_{\Delta\tilde{P}} = \sigma^{i-1}_{\Delta\tilde{P}} + \mathcal{N}(0, \sigma^2_x) \quad (4.13)$$

Where σ^2_x is a variance (small positive number) of random walk step in the dimensions $\sigma_{\Delta P}$ and $\sigma_{\Delta\tilde{P}}$. The proposal distribution $q(x'|x_{i-1}) = \mathcal{N}\left(x_{i-1}, \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{bmatrix}\right)$ randomly perturbs the current sample of the chain, and then either accepts or rejects the candidate x' depending on its acceptance probability. In this case, we set the random walk variances to be equalled (i.e. σ_x) for both dimensions, however they can also be fine-tuned to different ones depending on the true values of $\sigma_{\Delta P}$ and $\sigma_{\Delta\tilde{P}}$.

Algorithm 1: Metropolis-Hastings [72]

Initialise $x_0 = \left(\sigma^0_{\Delta\tilde{P}}, \sigma^0_{\Delta P}\right)$:

Output: $\left\{x_i = \left(\sigma^i_{\Delta P}, \sigma^i_{\Delta\tilde{P}}\right)\right\}^N_{i=N_B}$

for $i = 1$ to N **do**

Step 1 (Propose)

$x' \sim q(x'|x_{i-1})$

Step 2 (Acceptance Probability)

$$\alpha(x'|x_{i-1}) = \min\left\{1, \frac{q(x_{i-1}|x')\pi(x')}{q(x'|x_{i-1})\pi(x_{i-1})}\right\}$$

Step 3 (Accept)

$u \sim \mathcal{U}(0,1)$

if $u < \alpha$ **then**

$x_j \leftarrow x'$

else

$x_j \leftarrow x_{j-1}$

end if

end for

The MH acceptance function is designed to tend to visit higher probability areas given by the ratio $\frac{q(x_{i-1}|x')\pi(x')}{q(x'|x_{i-1})\pi(x_{i-1})}$ In the case of random walk, $q(\cdot\,|\,\cdot)$ is symmetric (i.e. $q(x_{i-1}|x') = q(x'|x_{i-1})$), hence the acceptance function is simplified to be

$$\alpha(x'|x_{i-1}) = \min\left\{1, \frac{\pi(x')}{\pi(x_{i-1})}\right\} \quad (4.14)$$

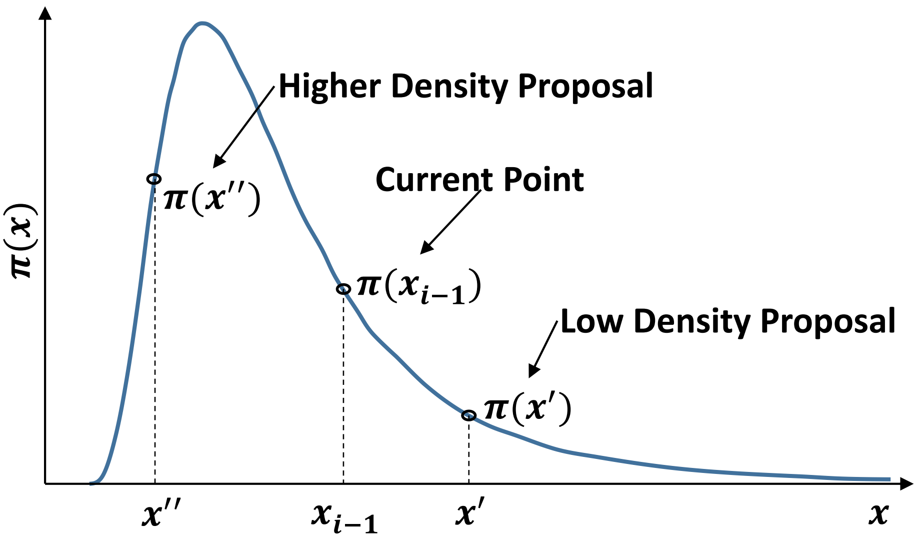


Figure 4.9. Illustration of Full Joint Density Function

The acceptance becomes proportional to how likely each of the current state and the proposed state are under the full joint density (see Figure 4.9), which can be defined as a maximum likelihood function

$$\pi(x) = \prod_{j=1}^{N_S} \prod_{k=1}^{N_D^j} \frac{1}{\sqrt{2\pi}\sigma_{\Delta\tilde{P}}} \exp\left(-\frac{(\Delta\tilde{P}_k^j - \Delta P_k^j)^2}{2\sigma_{\Delta\tilde{P}}^2}\right) \quad (4.15)$$

where $\Delta\tilde{P}_k^j$ and ΔP_k^j are obtained from the test data set and computed using (4.8), respectively In order for the MH algorithm not to get stuck at one sample, the space is ensured to be explored by randomly sampling the acceptance probability u from a uniform distribution $\mathcal{U}(0,1)$. The proposed sample x' is accepted if $u < \alpha$; otherwise we reject it. This way the distribution of x is ensured to converge to the true distribution that we are interested in.

We used 9 data samples as in section 4.4 and their associated optimised scaling parameters θ_j for the evaluation of the full joint density function $\pi(\cdot)$. σ_x, N and N_B were set to 0.001, 1,000,000 and 500,000 burn-in time, respectively. The burn-in time N_B is required to allow the Markov chain to approach the equilibrium; thus the representative draws from the true distribution. In this run, the means of the distributions $\sigma_{\Delta P}$ and $\sigma_{\Delta\tilde{P}}$ are respectively 0.0168 and 0.0998, where 0.0017 and 3.7115×10^{-4} are the corresponding standard deviations.

4.7.2 State and Parameter Estimation

PF uses a statistical method called Bayesian inference, in which measurements are used to estimate and update the pressure (state) variable ΔP and model parameter θ in a form of probability density function (pdf). This chapter focuses on how a data-driven approach can be formulated to handle large variations in the data. Therefore, to simplify we use the simplest form of the particle filter, named Sequential Important Resampling (SIR) [73, 71], to demonstrate the concept in this chapter.

Algorithm 2: The pseudocode of algorithm SIR Particle Filter is provided below [73] .

Inputs: $\{(\Delta P_{k-1}^i, \theta_{k-1}^i)\}_{i=1}^{N_P}$ and $\Delta\tilde{P}_k$

Outputs: $\{(\Delta P_k^i, \theta_k^i), w_k^i\}_{i=1}^{N_P}$

Step 1 (Update)

for $i = 1$ to N_P **do**

$\theta_k^i \sim \mathcal{N}(\theta_k^i, \sigma_\theta)$ or $\mathcal{N}(m(\theta_k^i), h^2 V_k)$

$\Delta P_k^i \sim p(\Delta P_k | \Delta P_{k-1}^i, \theta_k^i)$

end for

Step 2 (Resampling)

for $i = 1$ to N_P **do**

$w_k^i \leftarrow L(\Delta\tilde{P}_k | \Delta P_k^i, \theta_k^i)$

end for

$W \leftarrow \sum_{i=1}^{N_P} w_k^i$

for $i = 1$ to N_P **do**

$w_k^i \leftarrow w_k^i / W$

end for

$\{(\Delta P_k^i, \theta_k^i)\}_{i=1}^{N_P} \leftarrow$

Resample $\left(\{(\Delta P_k^i, \theta_k^i), w_k^i\}_{i=1}^{N_P}\right)$

The pseudo code of a generic SIR particle filter algorithm is listed in algorithm 2 and graphically illustrated in Figure 4.10.

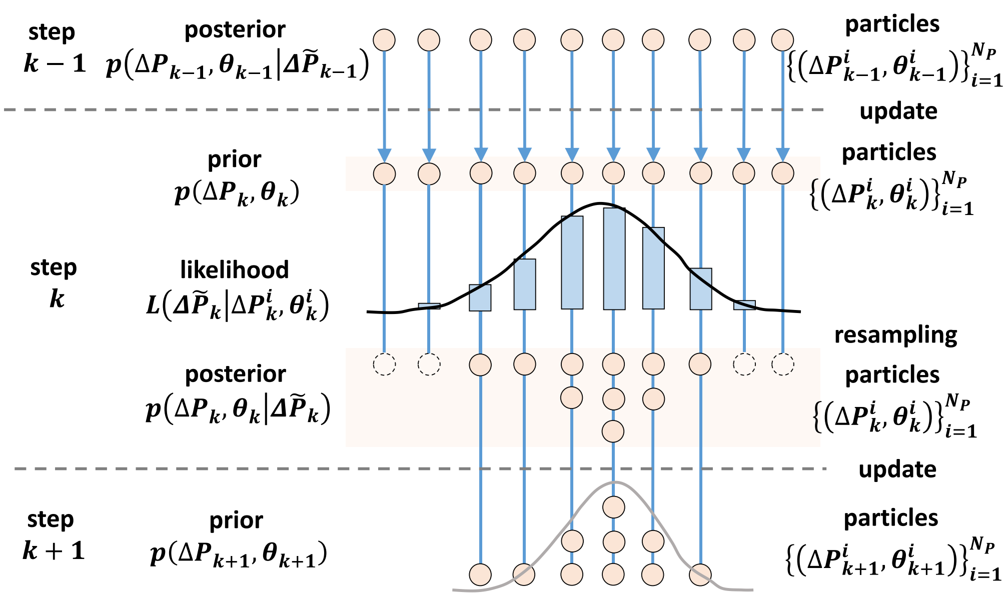


Figure 4.10. Illustration of Particle Filtering Process [71, 39]

In PF, pdf is not explicitly defined, but instead N_P number of samples$\{(\Delta P^i, \theta^i)\}_{i=1}^{N_P}$, so called particles, are used as an approximation of the pdf. A prior probability information of $(\Delta P_{k-1}, \theta_{k-1})$ and a measurement update $\Delta \tilde{P}_k$ are the algorithm inputs. To initialise the algorithm, the particles $(\Delta P_0^i, \theta_0^i)$ are often sampled uniformly from the possible (or arbitrary) intervals of ΔP and θ. If N_P is sufficiently large, then $\{(\Delta P^i, \theta^i)\}_{i=1}^{N_P}$can be regarded as the representative draws of $(\Delta P, \theta)$, which effectively ensures consistent results between runs.

PF consists of two main steps: update and resampling. In the update step, the prior pdf $\{(\Delta P_{k-1}^i, \theta_{k-1}^i)\}_{i=1}^{N_P}$ is propagated forward to time k by some random processes. ΔP_k is evolved using (4.8) (or equivalently (4.10)), which represents the physics of filter clogging. However, some type of evolution needs to be defined for the parameter θ_k. The typical solutions are to use either a random walk [74, 18] or kernel smoothing function [75, 76, 73]. The random walk can be defined by

$$\theta_k = \theta_{k-1} + \mathcal{N}(0, \sigma_\theta^2), \quad (4.16)$$

Where σ_θ defines a random walk step size. σ_θ determines the rate and estimation performance of the parameter θ_k. A large σ_θ will give fast convergence but high fluctuations, whereas a small value of $\sigma_\theta$will produce a smoother (but slower) convergence of θ_k. In kernel smoothing, the evolution of θ_k on the other hand takes into account of the probability information of θ_k. The parameter $\theta_k^i$is indirectly sampled through

$$\theta_k^i \sim \mathcal{N}\left(m(\theta_k^i), h^2 V_k\right), \quad (4.17)$$

Where

$$m_k^i = a\theta_k^i + (1-a)\bar{\theta}_k \text{ and} \quad (4.18)$$

$$V_k = h^2 \frac{1}{N_p} \sum_{i=1}^{N_P} (\theta_k^i - \bar{\theta}_k)^2 \quad (4.19)$$

With $\bar{\theta}_k$ is the mean value of θ_k^i and $a = \sqrt{1-h^2}$ where $h > 0$ is the smoothing parameter. Similarly to the random walk method, a larger a will give faster convergence but high fluctuations, whereas a small value of a will give a smoother (but slower) convergence. It is common practice to use a around 0.98 or higher [73]. The kernel smoothing method provides an adaptive way to evolve the parameter θ_k.

In the resampling step, the likelihood of the particles $\left\{\left(\Delta P_{k-1}^i, \theta_{k-1}^i\right)\right\}_{i=1}^{N_P}$ are evaluated using

$$L\left(\Delta\tilde{P}_k \middle| \Delta P_k^i, \theta_k^i\right) = \frac{1}{\sqrt{2\pi}\sigma_{\Delta\tilde{P}}} \exp\left(-\frac{\left(\Delta\tilde{P}_k - \Delta P_k^i\right)^2}{2\sigma_{\Delta\tilde{P}}^2}\right) \quad (4.20)$$

.

In equation (4.20) quantitatively determines how likely a measurement $\Delta\tilde{P}_k$ is produced by a particle $\left(\Delta P_k^i, \theta_k^i\right)$. The particles are weighted in which where their weights w_k^i are proportionally (or equal) to the computed likelihood values. This is in a way proposing a pdf for $(\Delta P_k, \theta_k)$. The weights are then normalised and used to systematically resample (or commonly known as roulette wheel) the particles, i.e. $\left\{\left(\Delta P_k^i, \theta_k^i\right)\right\}_{i=1}^{N_P} \leftarrow \text{Resample}\left(\left\{\left(\Delta P_k^i, \theta_k^i\right), w_k^i\right\}_{i=1}^{N_P}\right)$ in algorithm 2. Intuitively, a particle that has a higher weight will have a higher probability of being duplicated, and vice versa (see Figure 10). The resampled particles (posterior) are used to estimate the state $\Delta\bar{P}_k$ and model parameter $\bar{\theta}_k$, which can be by either taking the mean or median of the particles, and then form the prior pdf for the next filtering iteration.

4.7.3 End of Life Prediction

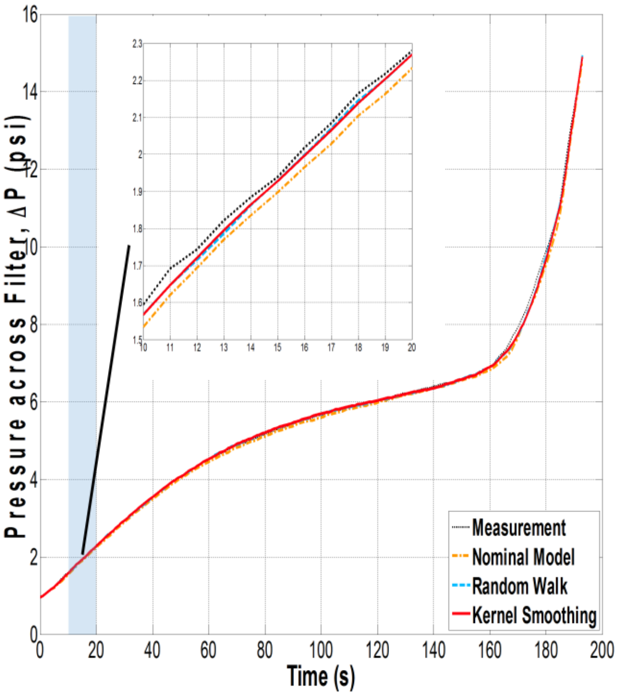
At a given time k, the future state of ΔP of a particle $\left(\Delta P^i, \theta^i\right)$ can be predicted by propagating forward using (2.8) where ΔP_k^i and θ_k^i are an initial condition and fixed model parameter, respectively. To compute RUL, we then propagate each particle until ΔP reaches the ΔP_{EOL} threshold to obtain the probability distribution of predicted EOL'. The distribution of predicted RUL' can then be obtained by simply subtracting the pdf of EOL' with k. The estimated $\overline{RUL}$ can be calculated by either taking the mean or median of the distribution. Note that prediction requires hypothesising future operating conditions of the system as the clogging rate is dependent on the pump speed, change in the solid-ratio, and so on. In this chapter, these operating conditions are encapsulated into a single model parameter θ. In prediction, fixing θ to the current estimated value will mean the operating conditions are assumed to be unchanged in the future state. If the operational profile is to be varied, then the model parameter fixing θ has to be time-varying during the simulation (forward propagation) according to the profile and this will be based on a pre-defined mapping (or lookup table) between the model parameter and operational conditions.

4.8 Results

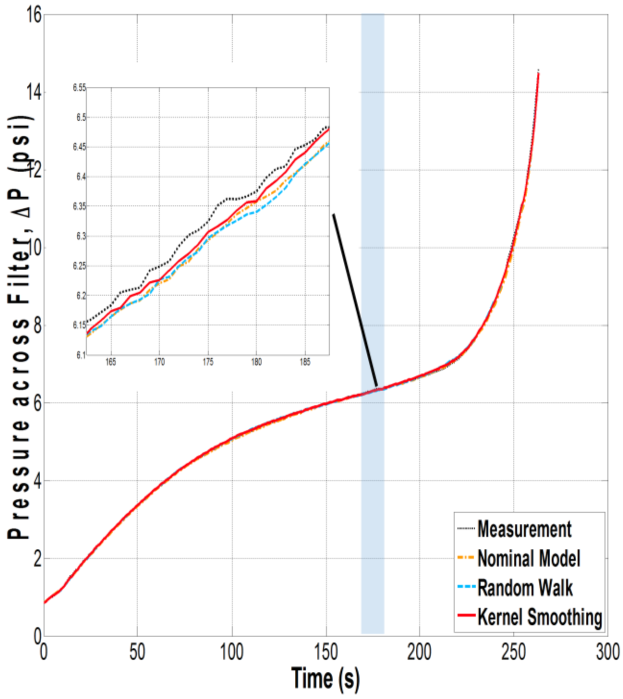
To demonstrate the approach, we tested the developed algorithm based on 4 test data samples, all of which have different operating conditions. In this chapter, we have implemented our algorithms in MATLAB. The measurement data is rendered point by point to simulate online estimation and prediction. Here, the end of life threshold ΔP_{EOL} was set at 15 psi. The PF parameters N_P, σ_θ and a were trial and error set at 1,000, 1.0×10^{-4} and 0.998, respectively. As discussed in section 4.7.2, the values of σ_θ and a will affect how the estimations converge.

However, how to optimally tune them is not the focus of this chapter. Further information can be found in [73, 76].

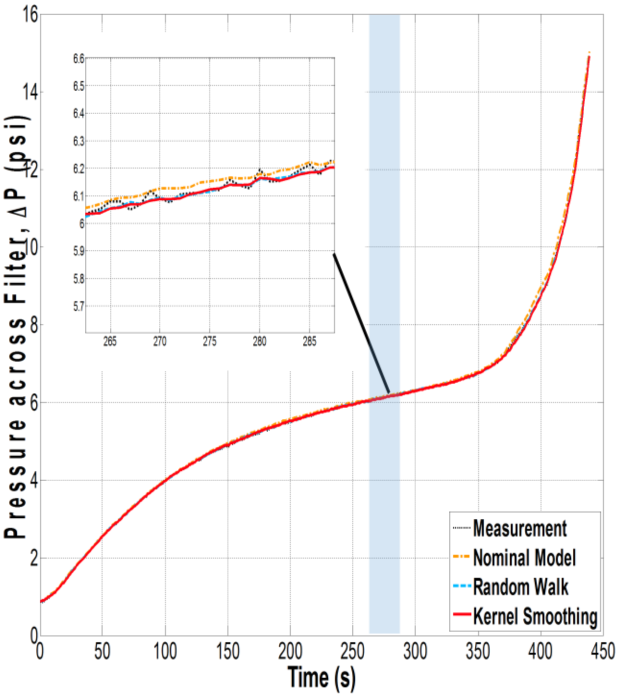
The resulting state estimation, parameter estimation and remaining useful life prediction are shown in Figures 4.11, 4.12 and 4.13, respectively. The results were generated using three different algorithm settings; 1) nominal TS fuzzy model + PF without parameter estimation, 2) scalable TS fuzzy model + PF with random walk parameter estimation and 3) scalable TS fuzzy model + PF with kernel smoothing parameter estimation. The results highlight the differences in prediction performance of the non-scalable and scalable data-driven TS fuzzy models. In overall, both non-scalable and scalable models are able to track the degradation (or clogging) profiles generated from different operating conditions, see Figure 4.11. With the scalable TS model, the online parameter estimation results show good convergences of the model parameter θ to its true values, see Figure 4.12. Recall that the operating conditions are encapsulated in the model parameter θ, hence significantly better accuracies in the RUL prediction are to be expected for the scalable TS fuzzy model, and, regardless of the operating conditions, this indeed reflects in the results shown in Figure 13. In contrast to the scalable model, the prediction errors increase as the operating conditions deviate away from the nominal condition.



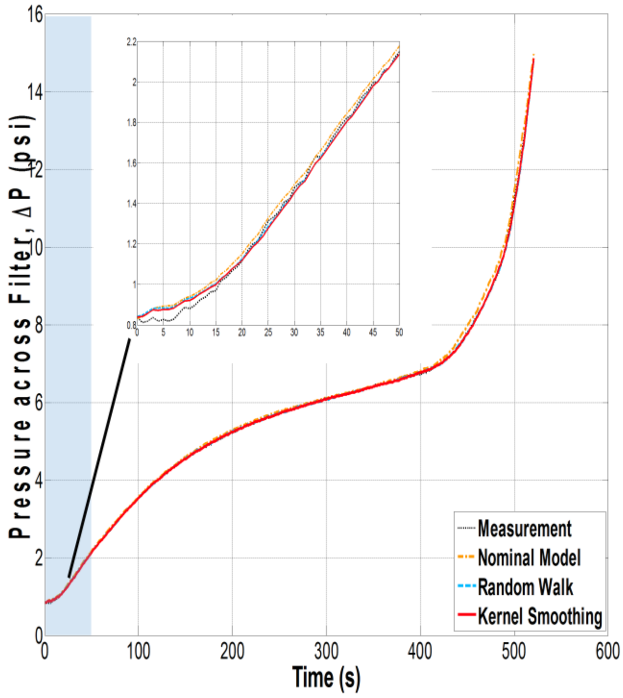
a) Test Sample #1



b) Test Sample #2

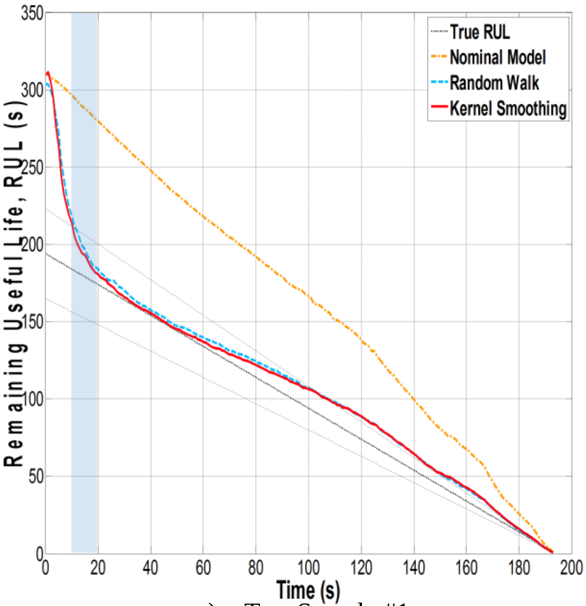


c) Test Sample #3

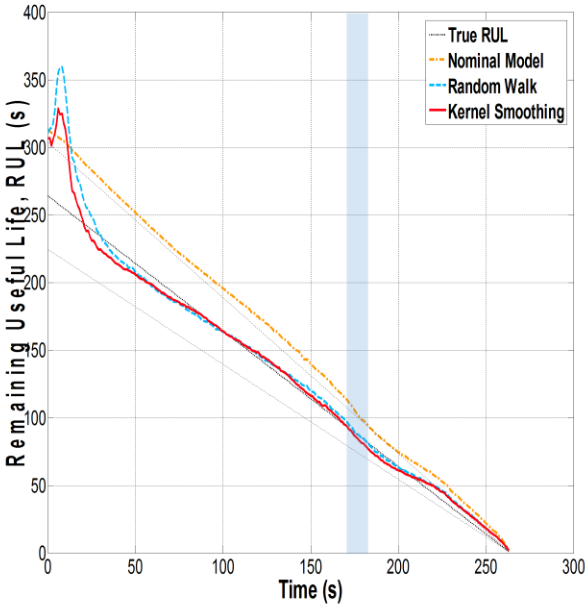


d) Test Sample #4

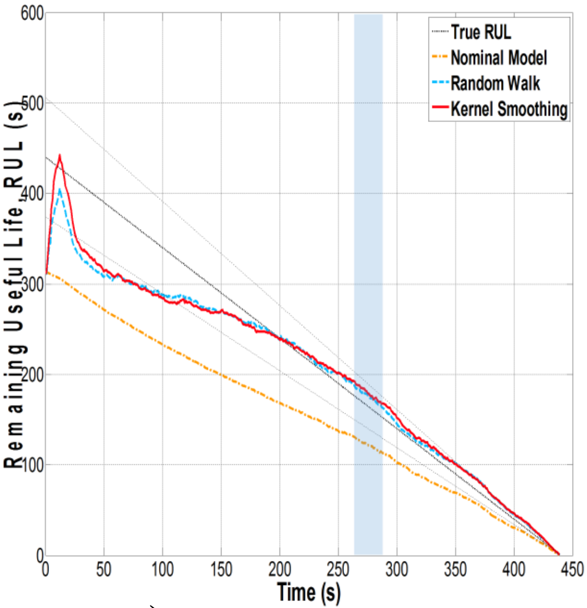
Figure 4.11: Estimated Pressure Drop Across Filter (State) ΔP.



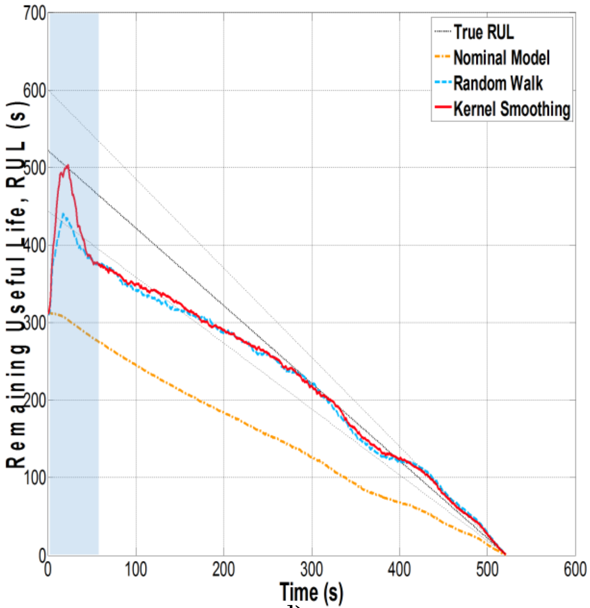
a) Test Sample #1



b) Test Sample #2



c) Test Sample #3



d) Test Sample #4

Figure 4.12 Estimated Remaining Useful Life (RUL).

One of the commonly used performance metrics for prognostics is Root Mean Square Error ($RMSE$) [77]. For a given RUL prediction profile, $RMSE$ can be offline evaluated using

$$RMSE = \sqrt{\frac{1}{N_P}\sum_{k=1}^{N_P}(\overline{RUL}'_k - RUL_k)^2} \qquad (4.21)$$

Where $\overline{RUL}'_k$ and RUL_k are predicted mean (or median) RUL (either mean or median) and true RUL at time k, respectively. Table 4.2 shows the RUL prediction errors measured in $RMSE$ for the 4 test samples. It can be seen that the RUL prediction results are significantly better in the case of online parameter estimations (i.e. using the scalable TS degradation model). The $RMSE$ ratio between nominal and scalable models is more significant if the operating condition θ of a test sample is more different from the nominal condition $\theta = 1.0$ ($RUL =$ 312.5 s). From the $RMSE$ results, the kernel smoothing method gives a better (but not significantly) RUL prediction than the random walk based parameter evolution. It can be seen that the kernel smoothing method has a better convergence in estimating the parameter θ (see Figure 4.12), hence this explains the better $RMSE$ results as observed in this experiment.

Table 4-2. Remaining Useful Life (RUL) Prediction Error for Test Samples

Particle Filter Prediction Models	Root-Mean-Square-Error (RMSE)			
	Test 1	Test 2	Test 3	Test 4
Nominal Model	76.242	29.155	73.846	125.992
Random Walk	36.464	19.010	21.840	50.151
Kernel Smoothing	34.526	13.262	21.238	44.266

To gain insight into the working mechanism of the proposed prognostic approach, we shall analyse the results in more detail. Let's consider the *light blue* shade areas of the estimated ΔP, θ and predicted RUL results for the test sample #1 shown in Figure 4.11 to 4.13. The estimated ΔP is able to closely track the measurement $\Delta\tilde{P}$, but somewhat underestimated as shown in the zoom-in of Figure 4.11 (a). This is due to the fact that the prior pdf of the parameter θ is being less than its true value (*black solid* line in Figure 4.12 (a). Therefore, the estimated ΔP is evolving at a lesser rate than the true clogging profile. For a similar ΔP, a particle that has a higher value of θ will more likely to be duplicated as it will produce ΔP at the next step with a lesser underestimation. This can be observed in an increment of the estimated θ towards its true value ~1.61. The closer the parameter θ is to its true value means the better the model represents the true clogging profile. This in a way provides a more accurate model for the EOL prediction, hence a decrement in the RUL prediction error as the time went on as seen in Figure 4.13a.

In Figure 4.11 (b) to Figure 4.13 (d) and Figure 4.12 (b) to Figure 4.12 (d), large fluctuations in the parameter estimation and RUL prediction can be observed at the beginning of the data renderings, between 0 and 50 s. This is caused by the modelling error and, as earlier described in the case of Figure 4.11a to 4.13a, the difference between the model parameter θ and its true value. To understand the former, let's consider the shade areas of the Figure 4.11d to 4.13d. From the zoom-in of Figure 4.11 (d), it can be seen that the ΔP estimation is varied from overestimate, close estimate to underestimate over the range of 0 to 50 s. In contrast to where the parameter is un-tuned, the modelling error is the variation in the scaled TS model from the real clogging profile; the model does not accurately represent the data in some parts of the profile. At the start, the across pressure ΔP is overestimated (see Figure 4.11d), and thus the model parameter has adapted itself towards ~0.6 to reduce the estimation error $\left|\Delta\tilde{P} - \Delta P\right|$. The model parameter θ is converging to its true value up to the point where the estimated ΔP closely tracks the measurement $\Delta\tilde{P}$. From > 20 s, the estimated ΔP changes from close estimate to underestimate, and therefore a reverse in the direction of the model parameter θ. Regardless of its true value, the particle filter will evolve θ in a way that the estimation error is minimised. This can be seen in Figure 4.11d and Figure 4.12d, where the model parameter θ fluctuates away from ~ 0.6 while the estimation error $|\Delta\tilde{P} - \Delta P|$ is decreasing. In a similar way, this explains the large alternations in the parameter estimation and RUL prediction as observed for the case of test samples #2 and #3.

Now let us consider the shade areas of the estimated ΔP, θ and predicted RUL results for the test samples #2 and #3 shown in Figure 4.11 - Figure 4.13. Figure 4.12b and Figure 4.12c show that the model parameter θ has converged to around it true values 1.19 and 0.7, respectively. Moreover, the estimated ΔP closely tracks the measurement $\Delta\tilde{P}$ for both samples #2 and #3, see the zoom-ins of Figure 4.11b and 4.11c. Therefore, the RUL predictions will be expected to close the real RULs as shown in Figure 4.13b and Figure 4.13c. For clarity, it is worth further discussing the difference between the test sample #1 and #2 in terms of how the model parameter θ converges in the shade areas shown in Figure 4.11 to Figure 4.13. To answer 'Why does the model parameter θ of the test sample #2 not vary in the similar way as the test sample #1?', notice the difference in the scale of the zoom-in graphs of Figure 4.11a (10×0.8) and 4.11b (25×0.45). The estimation error $|\Delta\tilde{P} - \Delta P|$ of the test sample #2 is in fact smaller than the test sample #1 (even though visibly larger). Furthermore, the rate of change in ΔP is also much smaller for the test sample #2 (~ 0.23 of the test sample #1); hence less variation in the likelihood due to the difference in θ for particles with a similar ΔP. For the shade areas, this explains (in relation to the test sample #1) a slower rate of change in the model parameter θ as observed in Figure 4.12b.

From the results, it is worth emphasising that our scalable TS fuzzy model is only a nonlinear approximator of the experimental data. It is not a physical model that the data were generated from, and hence modelling (or pattern) errors are inevitable. Recall that PF will only try to scale the TS fuzzy model in such a way that the estimation error $|\Delta\tilde{P} - \Delta P|$ is decreasing, it does not actually know what the true value of θ is. Thus, how good the parameter estimation and RUL

prediction are at a given time will depend on how much deviation the approximated degradation pattern has from the data at an interval around that particular time step. This consequently results in the small fluctuations in θ and prediction errors in RUL around their associated true values as observed in Figure 4.12 and Figure 4.13.

The summery of this chapter has been discussed in detail in the conclusion chapter. The next chapter will discuss the second technique of adaptive/scalable data driven techniques.

Chapter 5

Adaptive Degradation Prognostic Reasoning by Particle Filter with Neural Network Degradation Model for Turbofan Jet Engine

This chapter provides information regarding the essence of prognostics, problems of prognostic predictions, and the approach presented for single component. The approach has been explained and the results are also provided.

5.1 Introduction

This chapter presents the second adaptive data driven technique. This adaptive data driven technique is mainly useful where the sensor data does not show directly the degradation of the system or such system where several sensor data are available to track the degradation of the system. In the previous chapter, it has been discussed where the sensor data directly shows the degradation of the system. However, in this chapter the dataset that has been selected has several different sensors to track one complex system, for instance the jet engine has several sensors to monitor the whole jet engine, different sensors are monitoring different parts, such as temperature, pressure, fuel flow etc. This dataset also provides the health index of the part of the system which has been analysed. The health conditions are essential in terms of the mission context. The next section will discuss the motivation behind the philosophy of this technique.

5.2 Motivation

Generally, Prognostics involve a construction of a degradation model and then use it to estimate a current degradation state and to predict EoL of a component. The purpose of a model is to capture how degradation evolves over time. The current degradation state is estimated based on noisy measurement updates and possible paths (determined by the model) from the previous degradation states. The model then uses the estimated degradation rate to propagate the state until it reaches EoL at some pre-defined degradation threshold. The other operational parameters, like usage or component specific degradation coefficients, can also be taken into account in the state estimation and EoL prediction. The model itself forms a key part in prognostics. It acts as a physical constraint which is used to determine not all possible, but the most likely degradation paths.

Particle filters (PF) are commonly used Bayesian estimators for tracking the state of a degrading system. The common usage involves tracking the parameters of the state transition equation as well as tracking the main degradation indicator (i.e. health index). Health Monitoring (HM) data could be classify into two categories: direct HM data, and indirect HM data. Direct HM data indicates the health level of system directly (e.g. crack size, wear level) whereas indirect HM reflects the underlying system health partially or indirectly (e.g. sensor information, vibration, oil based monitoring). Wiener and Gamma processes, regression-based models, and Markovian-based models are based on direct HM data, while Stochastic Filtering-Based Models, Covariate-Based Hazard Models, Hidden Markov Models (HMMs) and, Hidden Semi Markov Models (HSMMs) would classify as the indirect HM kind. The sensory data also introduces noise as well as accuracy problems with it. Therefore PF are used to estimate the actual state of the system behaviour which is the actual response supposed to be measured.
On the other hand, for the cases where the actual health index of the system is not directly measured (i.e. indirect CM), the health index is approximated via transforming the sensory information. The sensory information does not directly

exhibit the health of the component/system, however it shows a degradation profile to be used in the transformation process [78].

The engine degradation simulation dataset doesn't provide health index information, however the dataset suppliers (NASA) provided the actual RUL of the system which could be treated as virtual HI information and system degradation information. The information they provided are the operational profiles and sensory data corresponding to some components. Therefore one has to establish a mechanism which transforms the sensory data into the health index to be used with particle filters.

To construct a prognostics degradation model for complex system like Turbofan jet engine is extremely difficult task. As it would need to measure and weight all the parameters which would be involved in degradations of the system. In literature there are control models of the Turbofan jet engine however, no prognostics degradation model has been found. Furthermore the NASA Turbofan jet engine dataset has been created from the simulation testbed. However, dynamic system state transition model can be approximated using regression or machine learning algorithms which are data driven techniques. The complexity of modelling the degradation behaviour is achieved by the learning algorithm which maps the sensory data into HI. However, the approximated data-driven nominal state transition model is not adaptable for different profiles of data. If the current conditions are far away from the nominal model then the errors will be higher in the prediction this is the major down side for the data driven technique in context of prognostics and they provide you with a model which will give you a dynamic model with fixed parameters. These limitation makes less adaptable to data driven technical in regression context.

The primary novelty of this research is, it provides the adaptable data driven technique which mainly captures different operational profiles. The secondary novelty of this work, is to transform the RUL information into HI as well as mapping the sensory data to HI values using a machine learning algorithm (i.e. NN) to be used as a measurement function within the particle filter tracking mechanism. In this way, the nominal mapping algorithm results are becoming adaptable by taking into account the particle filters which leads to more accurate prognostic results. The details of the methodology are elaborated in the next section.

5.3 Proposed Approach

This reasoning system designed and developed machine learning algorithms that can identify causal links between precursors and their associated system health indices in the form that will be relevant to the alert/alarm generation, operation and further predictive trend analysis. Meaningful information would be a quantitative value like probability, certainty or likelihood of an event or a combination of events will compromise the safety margin or cause a particular impact on operation.

5.3.1 Turbofan Engine Degradation Simulation Dataset

The NASA C-MAPSS turbofan data set has been used for this reasoning system, as the requirement of the dataset was it had to be a real or accurate dataset.
This dataset has four sets of data each, which is a combination of two failure modes and two operating conditions. Every dataset has at least hundred engine degradation simulations carried out using C-MAPSS, which are divided into training and test subsets [78]. Twenty one different sensor measurements as well as RUL values for test subsets are given. However, health indicators are not provided with the dataset.

The degradation in the Fan and HPC of the turbofan engine is simulated. The model that the dataset suppliers applied is exponential degradation shown in Eq. 5.1 where (d) is initial degradation, (A) is a scaling factor, (B(t)) time varying exponent, and (th_w) is upper wear threshold. The model is a generalised equation of common damage propagation models (e.g. Arrhenius, Coffin-Manson, and Eyring models) [78].

$$h(t) = 1 - d - Ae^{B(t)}/th_w \tag{5.1}$$

The dataset is eligible for a data-driven approach, as sufficient data and RUL values are available within dataset. Either statistical or machine learning data-driven models can be applied to predict the RUL of turbofan engines.

However, the physics-base modelling would not be appropriate for this sort of system as the Turbofan jet engines are very complex systems and many parameters have to be accounted for. Therefore the data driven approach would be more suitable for this dataset/ system.

5.3.2 NASA Turbofan Dataset

The dataset given in [78] consists of multivariate time series signals that are collected from the turbofan engine dynamic simulation process. The engine run-to-failure simulations were carried out using C-MAPSS. A Hundred engine's run-to-failure time series trajectories are considered in this study (dataset FD001) which can be considered to be forming a fleet of engines of the same type. The aircraft gas turbine engine's RUL is closely bound up with its conditions. To monitor the aircraft gas turbine engines conditions, several kinds of signals could be used, such as temperature, pressure speed and air ratio. In this simulation a total of 21 sensors were installed in the aircraft engine's different components (Fan, LPC, HPC, LPT, HPT, Combustor and Nozzle) to monitor the aircraft engine's health conditions.

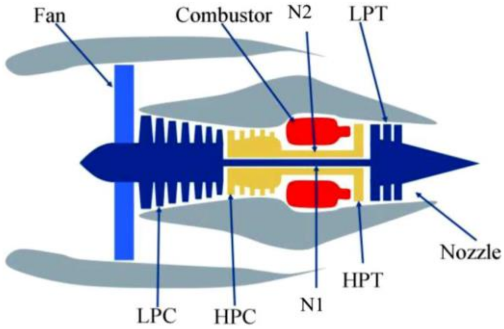
The 21 sensory signals, as detailed in the Table 5.1, where obtained from the above mentioned sensors. Among the 21 sensory signals, some signals contain little or no degradation information whereas the other shows a degradations trend. Most of the sensors data were contaminated with measurement noise.

TABLE 5-1.DESCRIPTION OF THE SENSOR SIGNALS FOR THE AIRCRAFT GAS TURBINE ENGINE DATASET [78]

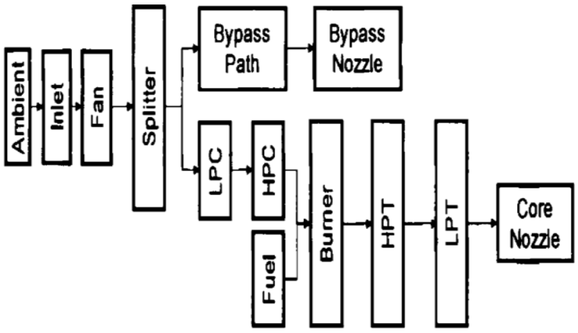
Inde	Symbol	Description	Units
1	T2	Total Temperature at fan	°R
2	T24	Total temperature at LPC	°R
3	T30	Total temperature at HPC	°R
4	T50	Total temperature LPT	°R
5	P2	Pressure at fan inlet	psia
6	P15	Total pressure in bypass-	psia
7	P30	Total pressure at HPC	psia
8	Nf	Physical fan speed	rpm
9	Nc	Physical core speed	rpm
10	Epr	Engine pressure ratio	--
11	Ps30	Static pressure at HPC	psia
12	Phi	Ratio of fuel flow to Ps30	psi
13	NRf	Corrected fan speed	rpm
14	NRc	Corrected core speed	rpm
15	BPR	Bypass ratio	--
16	farB	Burner fuel-air ratio	--
17	htBleed	Bleed enthalpy	--
18	Nf_dmd	Demanded fan speed	rpm
19	PCNfR_dm	Demanded corrected fan	rpm
20	W31	HPT coolant bleed	ibm/s
21	W32	LPT coolant bleed	ibm/s

°R **The Rankine temperature scale**
paisa **Pounds per square inch absolute**
rpm **Revolution per minute**
pps **Pulse per second**
psi **Pounds per square inch**
ibm/s **Pound mass per second**

To improve the RUL prediction accuracy and efficiency, important sensory signals must be carefully selected to characterize degradation behaviour for the aircraft gas turbine engine health prognostics. By observing the degradation behaviour of the 21 sensory signals, seven of them (2, 4, 7, 8, 11, 12, and 15) were selected in this study. The engine run-to-failure simulations were carried out using C-MAPSS simulation; they are a presentative simulation model of a modern commercial turbofan engine [79].



Simplified diagram for the aircraft gas turbine engine



A layout of modules and connections in the simulation

Figure 5.1. Gas turbine Engine dataset architecture [78]

C-MAPSS simulation have fourteen inputs and thirteen health parameters which allow the user to simulate the outcomes of degradation and faults in any of the five rotating parts of the engine i.e. Fan, LPC, HPC, HPT, and LPT. The time series data is recorded to characterize the evolution of the obscured health state of the engine.

The operability margins, for example temperature margins and stall, define the safe threshold operation region for the engine. The 6 flight conditions that have been simulated in this simulation, comprising arrange of values for 3 operational settings: Mach number (M: 0–0.84), altitude (Alt: 0–42Kft), and throttle solver angle (TRA: 20–100). Every engine degradation simulation starts with a different initial degradation state due to different degrees of initial wear and manufacturing variation in practice.

In the selected data set snapshots of engine performance parameters sat at sea level for each flight were used, i.e., Alt= 0, M=0, and TRA = 100). For more details on the engine run-to-failure simulation, the reader is referred to [80]. More information regarding the sensory signal screening is available from [78].

5.4 Reasoning and Prediction Thought Health Index Calculation

5.5 Health Index

Generally, the detection/reasoning system takes the input from the sensor and has some kind of threshold setup by the expert. The threshold usually has two limits the higher and the lower, each one is assigned with some kind of message or warning from the expert’s knowledge. However, until the threshold is reached the warnings are not displayed and the sensor values are just a data rather than information. This data doesn't provide any information regarding the health of the system. The HI calculation converts the raw sensor data into HI information. The HI term is defined by a discrete number representing the actual health of the system. For instance, 1 indicates the best health condition and 0 signifies the worst health condition. A mid-range value such as 0.5 represents half-life. The virtual health index of the system can be calculated in this system by using the following formula:

$$HI = \frac{Current\ RUL}{EoL} \tag{5.2}$$

In this chapter we formulized the health index at a specific time point as the rate of current RUL to the EoL of the system.

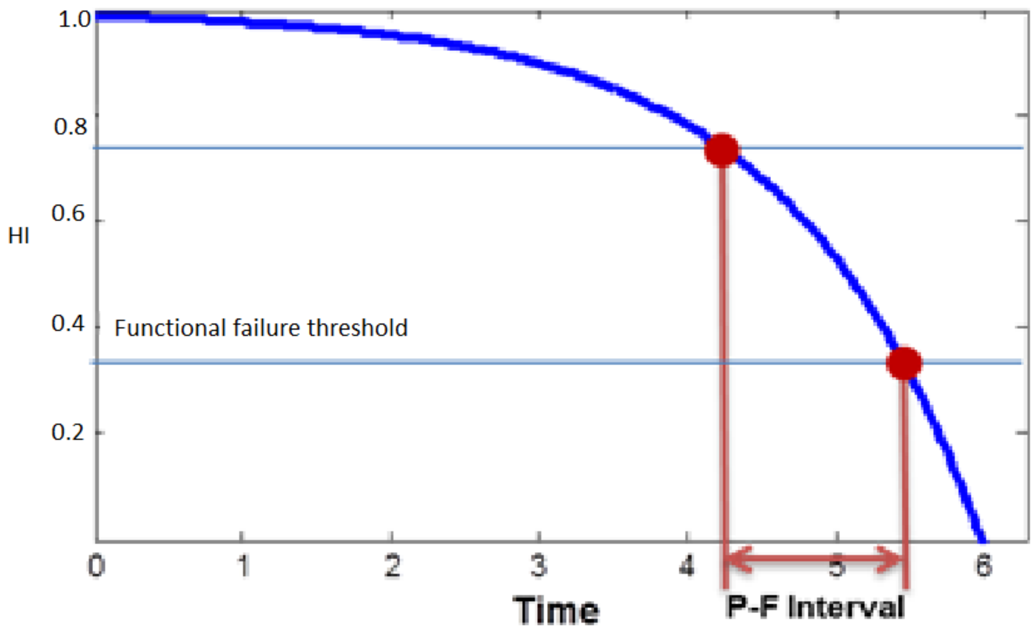


Figure 5.2. Real HI (Health Index) Degradation for Component

The Figure 5.2 shows the general degradation of the health index in real life environments. The physics of the system usually degrades far faster towards the EoL. However, in this chapter the health index has been simulated as a linear degradation of health to make the calculation less computational and simpler to simulate as shown in the Figure 3.4. Initial HI (Health Index) for Component. The initial state of the system as the HI is calculated from the equation 5.2. However when the sensor readings are received the HI calculation will be dynamic according to the sensor data received.

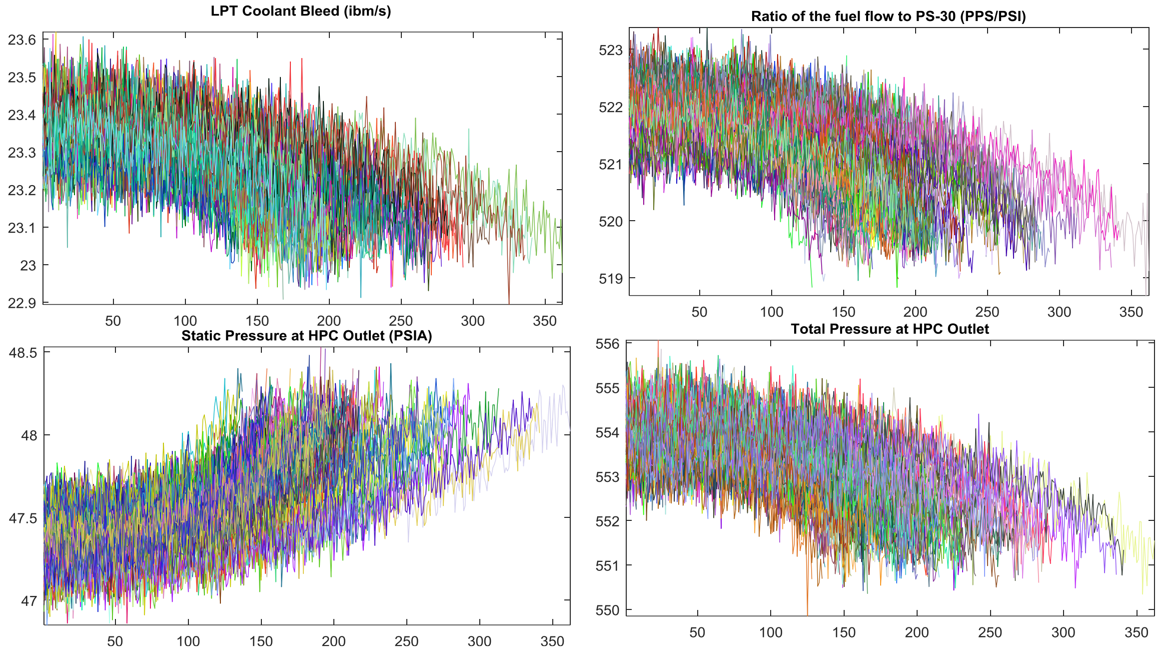


Figure 5.3a Four Sensors data over the time

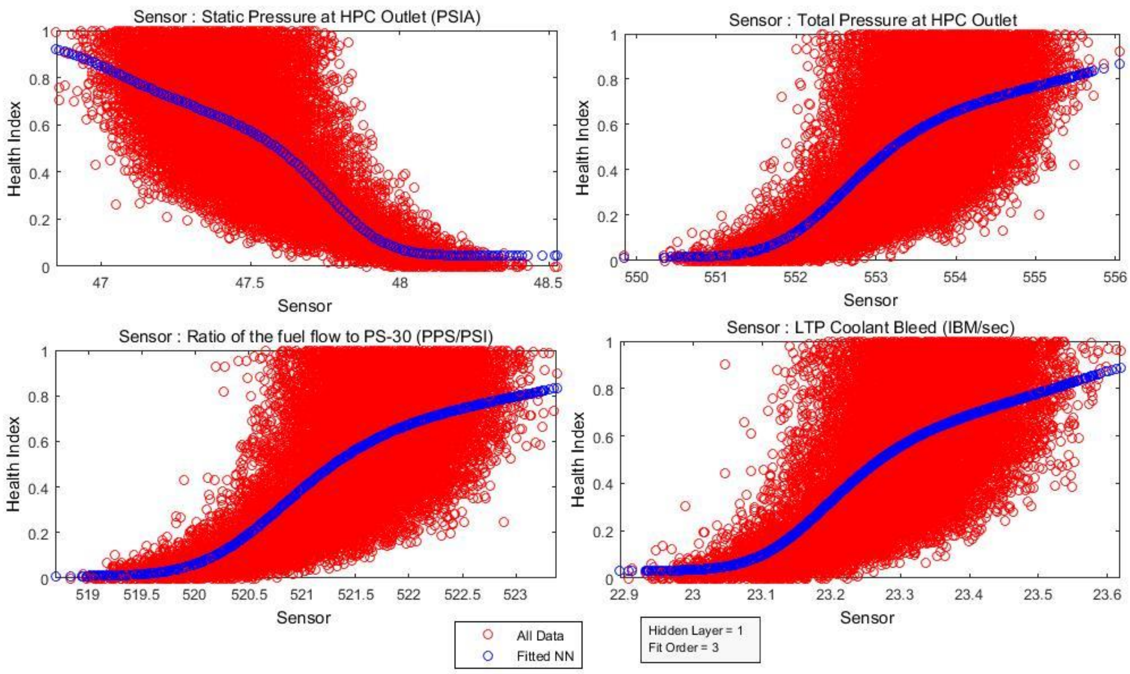


Figure 5.3 b Degradation of the of health index of four sensor over data.

The Figure 5.3a shows the sensor data on the Y-axis and X-axis is the time. Whereas in Figure 5.3b shows the HI on the Y-axis and the sensor reading on the X-axis. Figure 5.3b demonstrates the difference as to how the HI is degrading over time. The way to calculate the HI from sensor values is discussed in the next section.

5.5.1 Health index modelling using Neural Network

Radial basis functions (RBF) are kind of neural networks used for function approximations. RBFs are usually used in time series prediction, forecasting problems, regression, classification, non-linear system control etc. RBFs are three-layer feed-forward neural networks, where the hidden nodes implement a set of radial basis functions (e.g. Gaussian functions) as shown in Figure 5. On the other hand, the output nodes implement linear summation functions as in Multi-Layer Perceptron (MLP) [81]. The approximation function can be described in the following form in Eq. 5.3:

$$y(x) = \sum_{i=1}^{N} w_i \emptyset(\|x - x_i\|) \quad (5.3)$$

Where, $y(x) \in \mathbb{R}^n$ represents the actual approximation function with N numbers of radial basis. Each basis function is associated with different centres. Centre mean and standard deviation is symbolised as 'x_i' along with the associated weights 'w_i'. 'x' in the equation stands for the measurement inputted in the system. The norm and the basis function (i.e. '$\emptyset$') are typically taken to be the Euclidean distance respectively. In this research the Gaussian basis function is selected. Gaussian basis function can be formulised as given in Eq. 5.4:

$$\emptyset(x) = e^{-\frac{\|x - x_i\|^2}{2\sigma_i^2}} \quad (5.4)$$

The Neural RBF networks are universal approximators. A RBF network with enough hidden neurons can approximate continuous function with arbitrary precision. The weights 'w_i' and 'x_i' are determined in a manner that optimises the fit between the output and input.

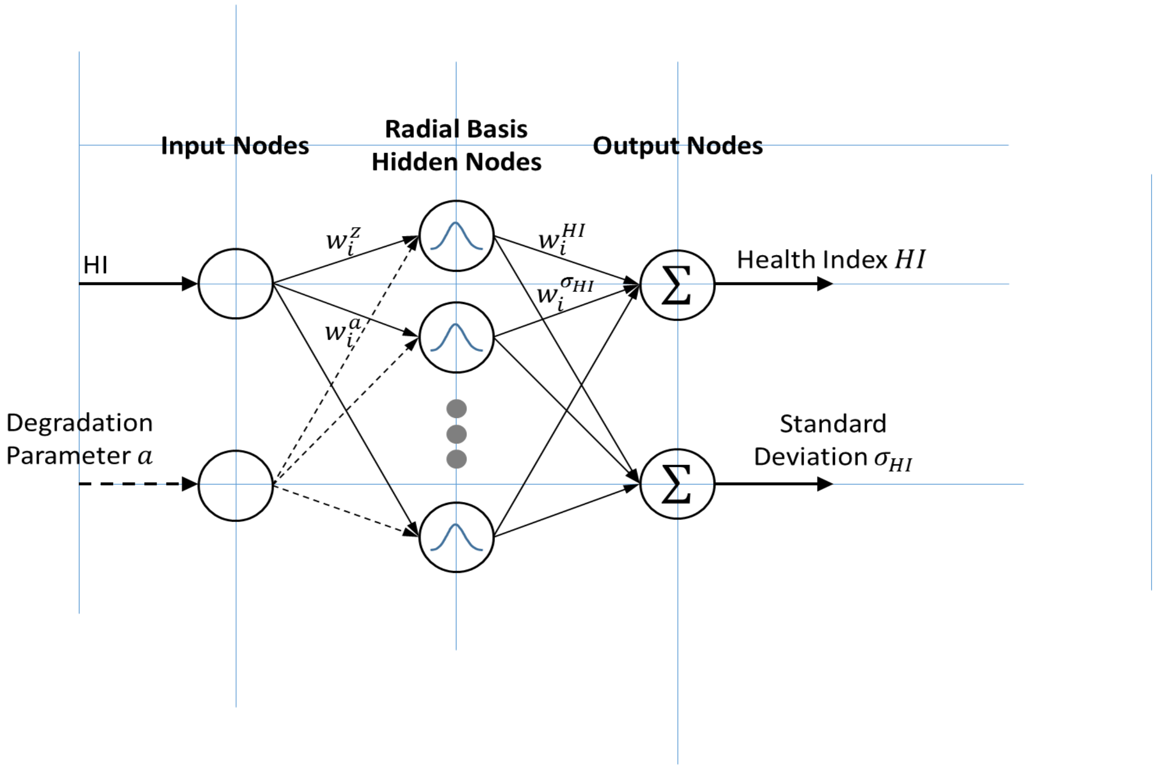


Figure 5.4. Radial Network Measurement Function.

The Neural Network has been fed the sensor data and has been mapped with the HI index along with the bound confidence (standard deviation) of 30%. The RBF also calculates and maps the data to the probability levels, because normal NN do not provide the probability/ confidence levels. The outputs of the RBF of the entire sensors are shown in the Figure 5.6. The Figure 5.4 shows the basic architecture of the Neural Network with the RBF function used. There are two inputs and two outputs to the NN, one output is the nominal model and the other is a standard deviation of the nominal model.

5.5.2 Cost function

In the previous section, Neural Network modelling is explained. However, for the model to closely approximate the degradation process, the model parameter values need to be tuned (or system identified) from the collected data samples. The aim is to minimise/optimise the prediction errors between the experimental degradation data and simulated data from the NN model.

The centres and associated weights are determined using a constrained nonlinear optimisation algorithm (i.e. 'fmincon') provided within MatLab. To do that, an object function script is written to calculate the error regarding the optimisation output.

In this chapter, Maximum likelihood Estimation (MLE) is used as a performance measure for system identification of the Neural Network degradation model defined in (eq 5.3). For a given data sample, MLE is obtained by maximising the log-likelihood metric, which is given in Eq. 5.5:

$$L = -\frac{1}{2}\ln(2\pi) - \sum \ln(\sigma_i) - \sum \frac{(\mu_i - x)^2}{2\sigma_i^2} \quad (5.5)$$

Where 'μ_i' and 'σ_i' are the basis function centre mean and standard deviation, whereas 'x' is the input measurement from the training data.

5.6 Particle Filter

5.6.1 Probabilistic Model

In order to predict EOL of a system, the system's current state of degradation has to be continuously estimated from the measurement updates. The basic system state transition equation is given in Eq. 5.6:

$$HI_k = HI_{k-1} - a \quad (5.6)$$

In this case, the fault precursor 'HI' and the parameter 'a' are required to be estimated. Kalman and Particle Filters are commonly used for estimating the degradation state. In this chapter, PF is used as it is more applicable to general non-linear systems. The estimation process of PF is based on the discrete stochastic state transition and measurement equations (Eq. 5.7) & (Eq. 5.8):

$$HI_k = f(HI_{k-1}, a_{k-1}) + \mathcal{N}(0, \sigma_{HI}^2) \qquad (5.7)$$

$$\widetilde{HI}_k = HI_k + \mathcal{N}(0, \sigma_{\widetilde{HI}}^2) \qquad (5.8)$$

where $\widetilde{HI}$, $\mathcal{N}(\cdot,\cdot)$, σ_{HI}^2 and $\sigma_{\widetilde{HI}}^2$ are noisy health index obtained from the NN using the sensory data, Gaussian random function, process uncertainty variance and measurement noise variance, respectively.

ALGORITHM 1: SIR Particle Filter [73]

Inputs: $\{(HI_{k-1}^i, a_{k-1}^i)\}_{i=1}^{N_P}$ and $\widetilde{HI}_k$

Outputs: $\{(HI_k^i, a_k^i), w_k^i\}_{i=1}^{N_P}$

Step 1 (Update)

for $i = 1$ to N_P **do**

$a_k^i \sim \mathcal{N}(\theta_k^i, \sigma_\theta)$ or $\mathcal{N}(m(a_k^i), h^2 V_k)$

$HI_k^i \sim p(HI_k | HI_{k-1}^i, a_k^i)$

end for

Step 2 (Resampling)

for $i = 1$ to N_P **do**

$w_k^i \leftarrow L(\widetilde{HI}_k | HI_k^i, a_k^i)$

end for

$W \leftarrow \sum_{i=1}^{N_P} w_k^i$

for $i = 1$ to N_P **do**

$w_k^i \leftarrow w_k^i / W$

end for

$\{(HI_k^i, a_k^i)\}_{i=1}^{N_P} \leftarrow$
$\text{Resample}\left(\{(HI_k^i, a_k^i), w_k^i\}_{i=1}^{N_P}\right)$

5.6.2 State and Parameter Estimation

PF uses a statistical method called Bayesian inference, in which measurements are used to estimate and update the HI (state) variable and model parameter a in a form of probability density function (pdf). In this chapter the measurement function is the RBF network methodology itself. Hence, to simplify we use a simplest form of the particle filter, named Sequential Importance Resampling (SIR) [73, 71], to demonstrate the concept in this chapter. The pseudo code of a generic SIR particle filter algorithm is listed in algorithm 1 and graphically illustrated in Figure 10. In PF, pdf is not explicitly defined, but instead N_P number of samples$\{(HI^i, a^i)\}_{i=1}^{N_P}$, so called particles, are used as an approximation of the pdf. A prior probability information of (HI_{k-1}, a_{k-1}) and a measurement update $\widetilde{HI}_k$ are the algorithm inputs. To initialise the algorithm, the particles (HI_0^i, a_0^i) are often sampled uniformly from the possible (or arbitrary) intervals of HI and a. If N_P is sufficiently large, then $\{(HI^i, a^i)\}_{i=1}^{N_P}$can be regarded as the representative draws of (HI, a), which effectively ensures consistent results between runs.

PF consists of two main steps: update and resampling. In the update step, the prior pdf $\{(HI_{k-1}^i, a_{k-1}^i)\}_{i=1}^{N_P}$ is propagated forward to time k by some random processes. HI_k is evolved using (5.8), which is obtained using the RBF network. However, some type of evolution needs to be defined for the parameter a_k. The typical solutions are to use either a random walk [74, 18] or kernel smoothing function [75, 76, 73]. The random walk can be defined by Eq. 5.9:

$$a_k = a_{k-1} + \mathcal{N}(0, \sigma_a^2), \quad (5.9)$$

Where σ_a defines a random walk step size. σ_a determines the rate and estimation performance of the parameter a_k. A large σ_a will give fast convergence but high fluctuations, whereas a small value of $\sigma_a$will produce a smoother (but slower) convergence of a_k.

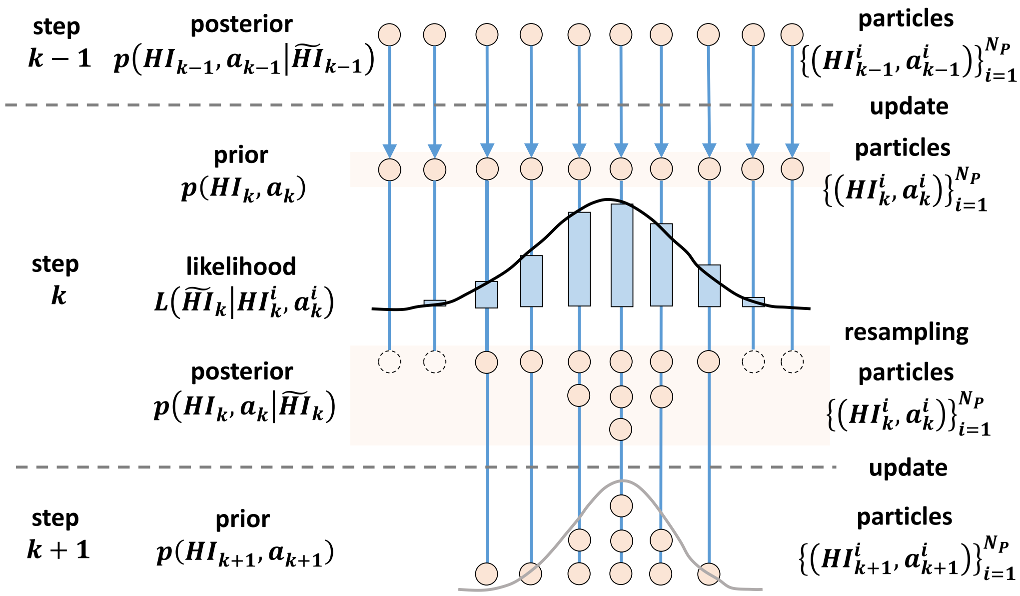


Figure 5.5. Illustration of Particle Filtering Process [71, 39].

In the resampling step, the likelihood of the particles $\{(HI_{k-1}^i, a_{k-1}^i)\}_{i=1}^{N_P}$ are evaluated using

$$L(\widetilde{HI}_k|HI_k^i, a_k^i) = \frac{1}{\sqrt{2\pi}\sigma_{\widetilde{HI}}} \exp\left(-\frac{\left(\widetilde{HI}_k - HI_k^i\right)^2}{2\sigma_{\widetilde{HI}}^2}\right) \tag{5.10}$$

.

In equation (5.10) quantitatively determines how likely a measurement $\widetilde{HI}_k$ is produced by a particle (HI_k^i, a_k^i). The particles are weighted in which their weights w_k^i are proportionally (or equal) to the computed likelihood values. This is in a way proposing a pdf for (HI_k, a_k). The weights are then normalised and used to systematically resample (or commonly known as roulette wheel) the particles, i.e. $\{(HI_k^i, a_k^i)\}_{i=1}^{N_P} \leftarrow \text{Resample}\left(\{(HI_k^i, a_k^i), w_k^i\}_{i=1}^{N_P}\right)$ in algorithm 1. Intuitively, a particle that has a higher weight will have a higher probability of being duplicated, and vice versa (see Figure 5.5). The resampled particles (posterior) are used to estimate the state $\overline{HI}_k$ and model parameter $\bar{a}_k$, which can be by either taking the mean or median of the particles, and then form the prior pdf for the next filtering iteration.

5.7 End of Life Prediction

At a given time k, the future state of HI of a particle (HI^i, a^i) can be predicted by propagating forward using (5.8) where HI_k^i and a_k^i are an initial condition and fixed model parameter, respectively. To compute RUL, we then propagate each particle until HI reaches the HI_{EOL} threshold to obtain the probability distribution of predicted EOL'. The distribution of predicted RUL' can then be obtained by simply subtracting the pdf of EOL' with the time index k. The estimated $\overline{RUL}$ can be calculated by either taking the mean or median of the distribution.

5.8 Results

To demonstrate the approach, we tested the developed algorithm based on C-MAPSS dataset. In this chapter, we have implemented our algorithms in MATLAB. The measurement data is rendered point by point to simulate online estimation and prediction.

The resulting state estimation, parameter estimation and remaining useful life prediction are shown in Figures 5.7, 5.8, 5.9 and 5.10 respectively.

The results were generated using two different algorithms: 1) nominal RBF NN model + PF parameter estimation, 2) similarity based prognostics for benchmarking the results.

The results highlight the differences in prediction performance of the non-scalable/adaptable data-driven models.

One of the commonly used performance metrics for prognostics is Root Mean Square Error ($RMSE$) [77]. For a given RUL prediction profile, $RMSE$ can be offline evaluated using

$$RMSE = \sqrt{\frac{1}{N_P}\sum_{k=1}^{N_P}(\overline{RUL}'_k - RUL_k)^2} \qquad (5.11)$$

Where $\overline{RUL}'_k$ and RUL_k are predicted mean (or median) RUL (either mean or median) and true RUL at time k, respectively.

The MAPE (Mean Absolute Percent Error) measures the prediction precision, the size of the error in percentage terms. It calculates as the average of the unsigned percentage error.

$$MAPE = \sqrt{\frac{1}{N_P}\sum_{k=1}^{N_P}\left(\frac{RUL_k - \overline{RUL}'_k}{RUL_k}\right) * 100} \qquad (5.12)$$

Table 5.2 to 5.5 shows the RUL prediction errors measured in $RMSE$ and $MAPE$ for the 4 test samples. It can be seen that the RUL prediction results are significantly better in the case of adaptable NN+PF technique comparatively to the normal data driven technique similarity based prognostics (SBP).

The Figure 5.6 demonstrators the nominal models which have been created from the data by using the RBF NN which models have been fed to the Particle Filter for RUL predictions. The solid blue line shows the mean of the nominal model and the dashed lines are the confidence bounds of the nominal models. The red dots are the data from the sensor which has been used to create the nominal model.

The measurement model has been created by using RBF NN and used as a measurement equation in the particle filter. Once the observed sensor value is provided to the particle filter it's fed to the measure equation which takes sensor observed values and provides HI mean with standard deviation which is needed to form a Gaussian distribution. The particle filter evaluates the likelihood of each particle and randomly resample according to its likelihood by using a roulette wheel genetic algorithm and reiterate the process. This way the particle with more likelihood will stay and the one with the lowest likelihood will be eliminated.

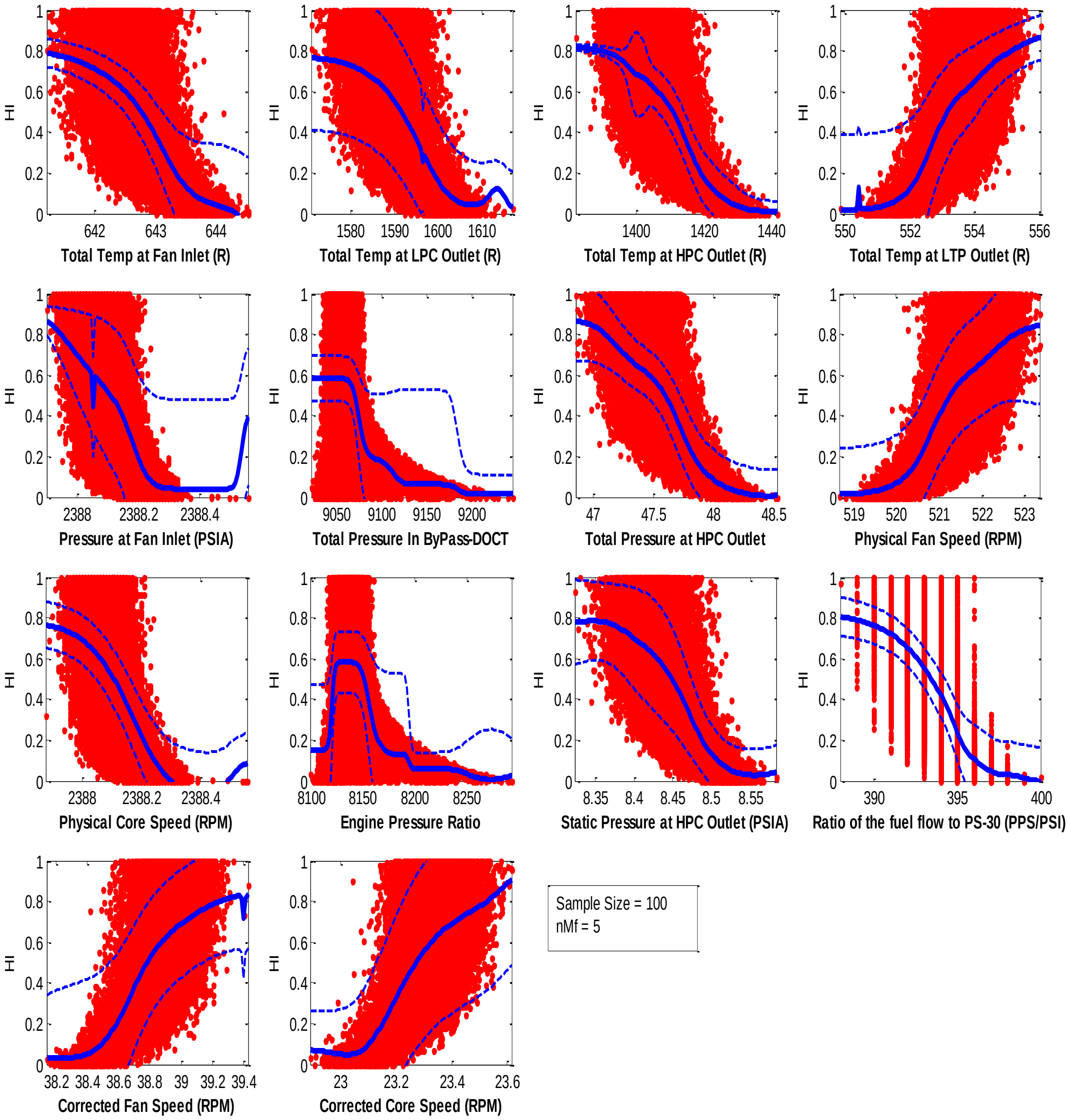


Figure 5.6. The NN results of 14 sensors of Turbofan Jet Engine

The presented technique has been tested rigorously to test the adaptability and robustness of the technique. The four specimens that have been selected to test the technique, are the longest flight lasted engine, small flight lasted engine and randomly selected middle two flights lasted engines specimens. These specimens have been selected to check the adaptability of the presented technique, as the longest lasted engine will more likely to be under less stresses comparatively to the shortest flight

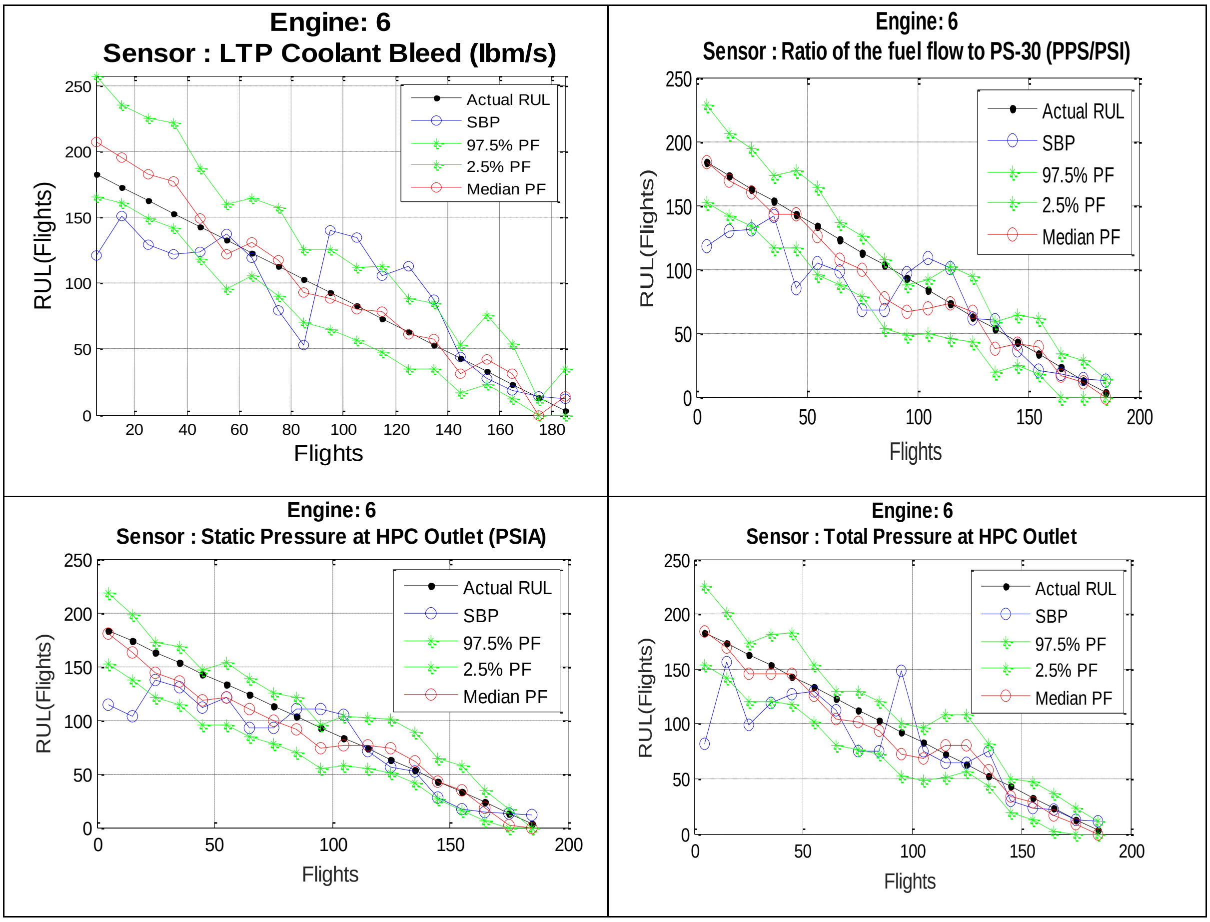


Figure 5.7. RUL Calculation of 4 Selection Sensors for Engine 6

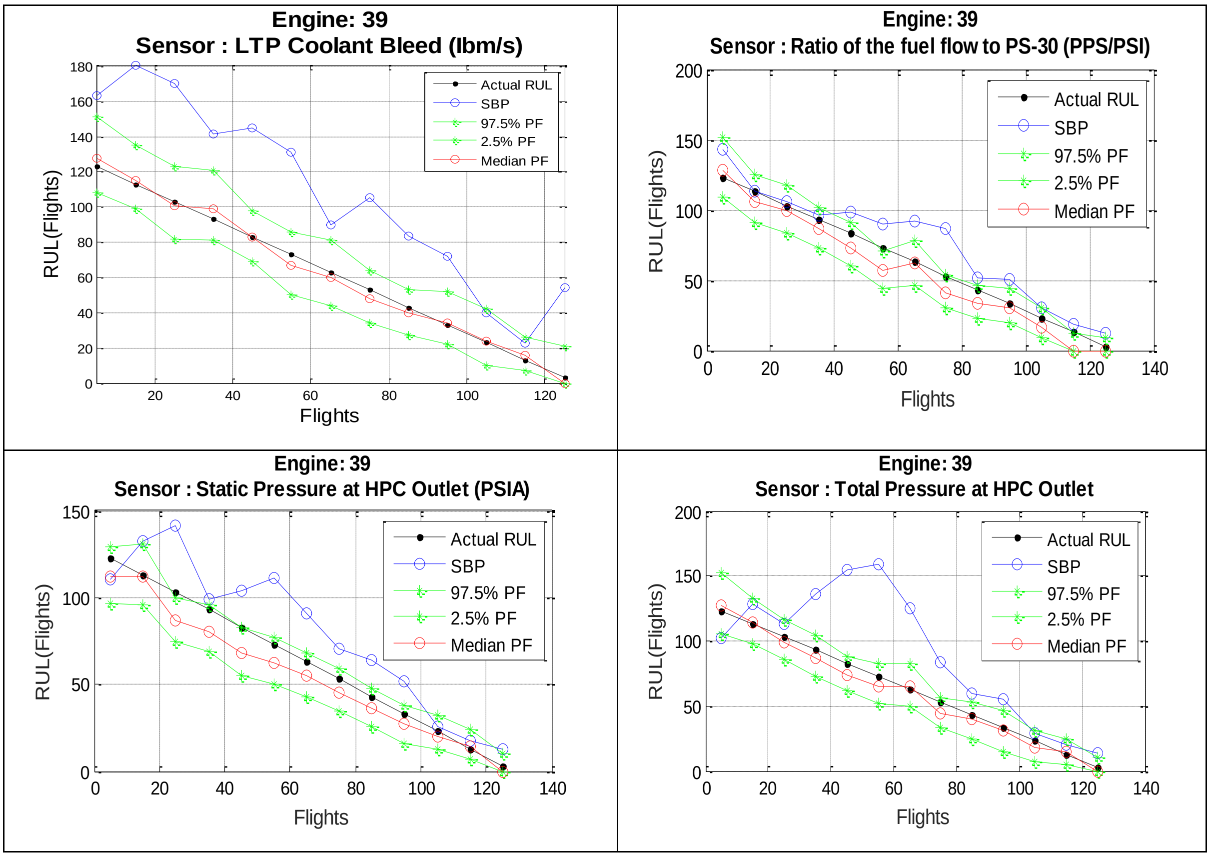


Figure 5.8. RUL Calculation of 4 Selection Sensors for Engine 39

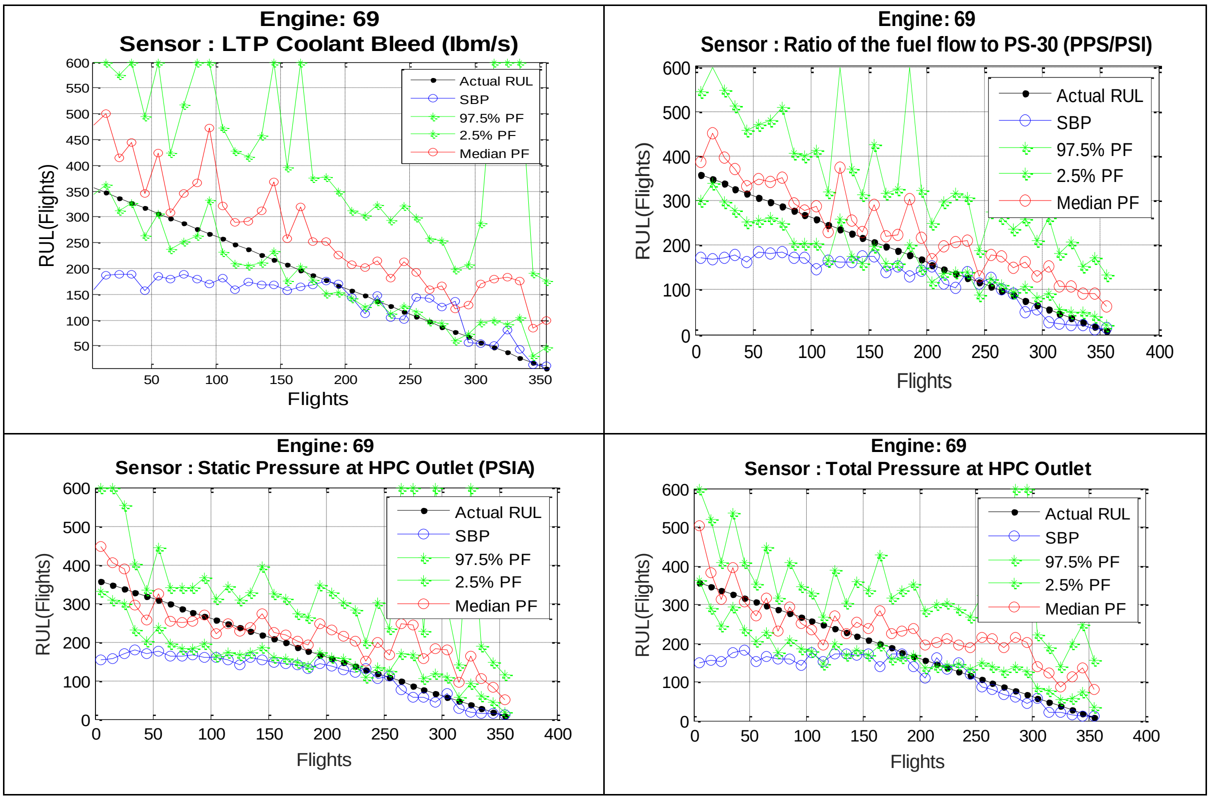


Figure 5.9. RUL Calculation of 4 Selection Sensors for Engine 69

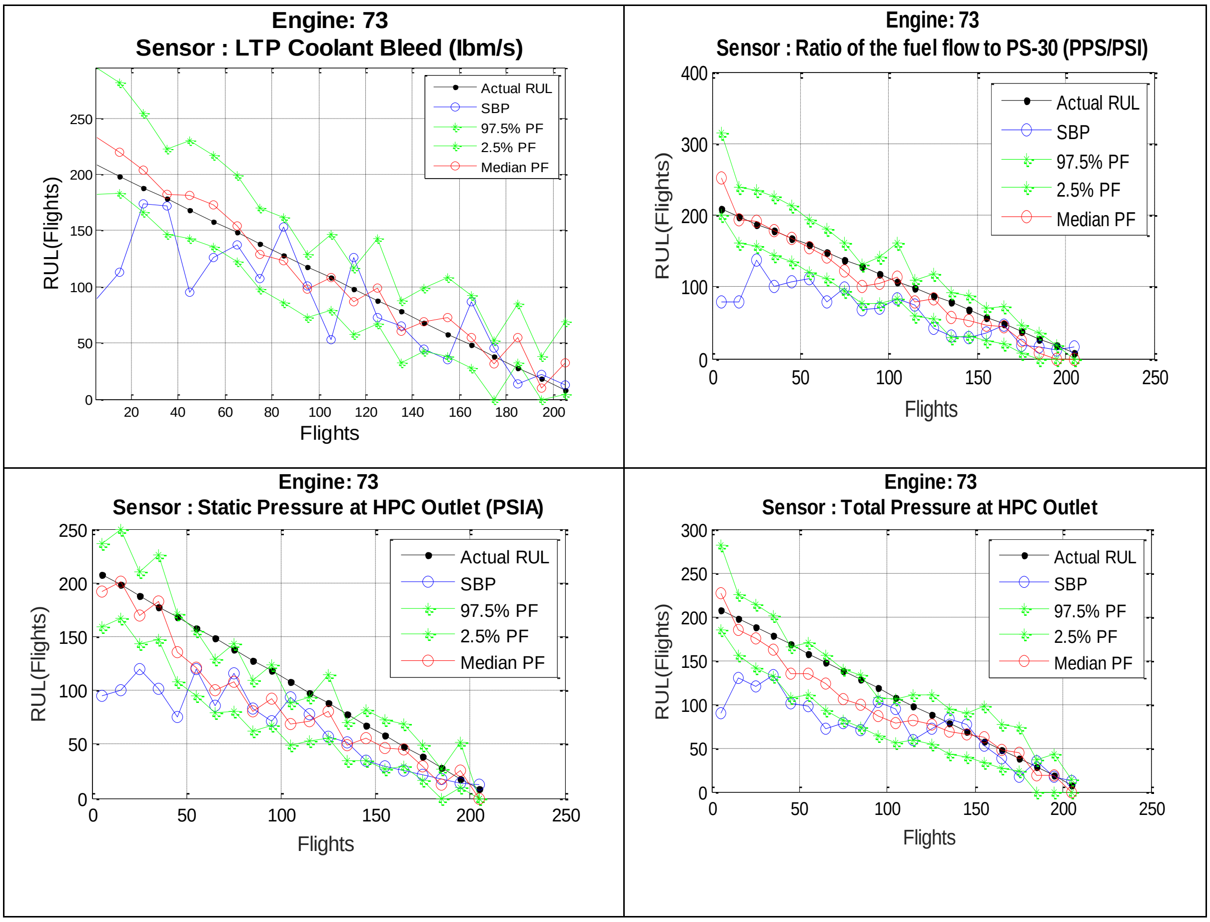


Figure 5.10. RUL Calculation of 4 Selection Sensors for Engine 73

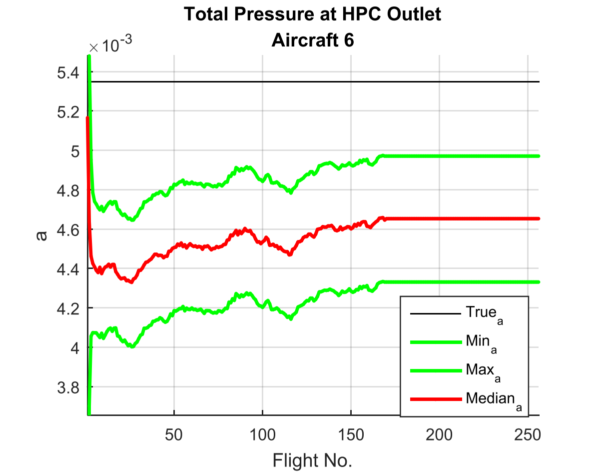


Figure 5.11. Parameter ‘a’ learning for Engine 6 sensor Total Pressure at HPC Outlet

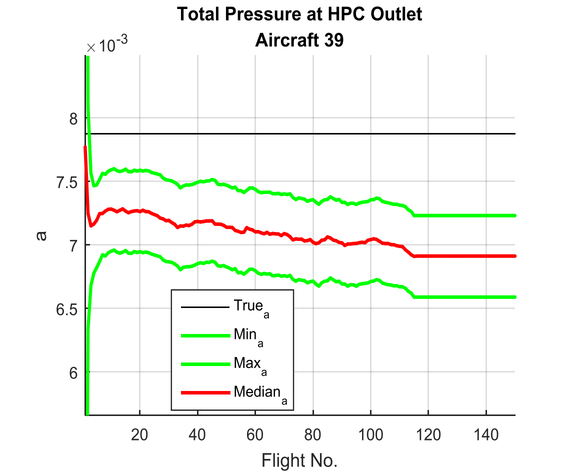


Figure 5.12. Parameter ‘a’ learning for Engine 39 sensor Total Pressure at HPC Outlet

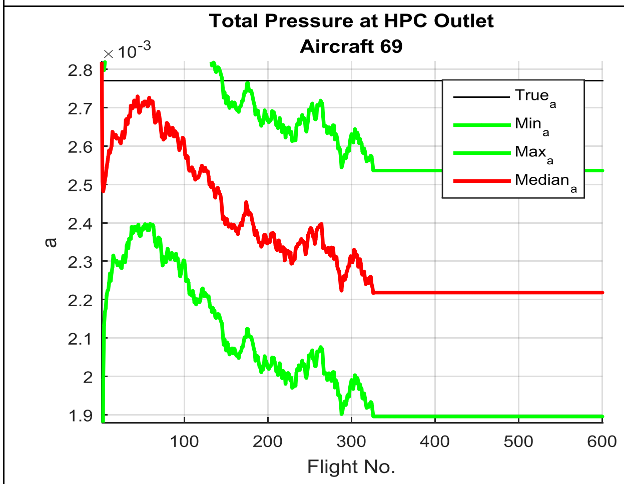


Figure 5.13. Parameter ‘a’ learning for Engine 69 sensor Total Pressure at HPC Outlet

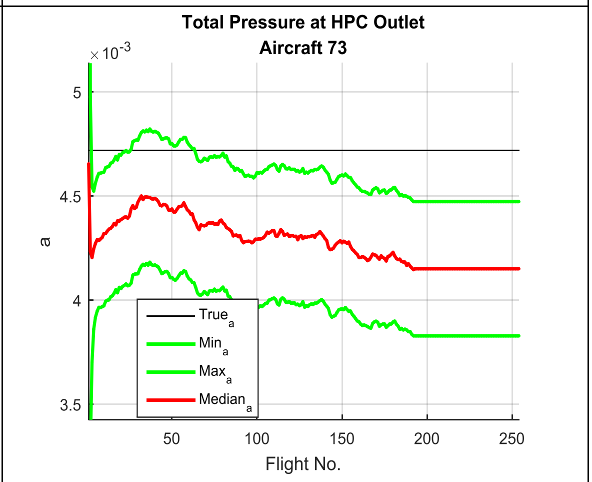


Figure 5.14. Parameter ‘a’ learning for Engine 73 sensor Total Pressure at HPC Outlet

The Figures 5.11 to 5.14 show the way parameter ‘a’ learned in different engines, how the particle filter is attempting to learn and how it’s evaluating over time. Once the value of ‘a’ provides the state estimations which are coinciding with what has been observed, it will not have enough errors to be updated dramatically to adapt.

5.9 Benchmarking

In benchmarking the techniques that has been used is Similarity-Based Prognostics (SBP). The result shows that the presented techniques have achieved better results in almost all tested cases. The results comparisons are demonstrated in the RUL Figures 5.7 to 5.10. The presented technique is proven to be an adaptable data driven technique in comparison the other data driven methods. The details on the SBP technique are provided in the next section.

5.9.1 Similarity-Based Prognostics

Similarity-based Prognostics (SBP) is a generic type of prognostic approach where the test specimen signal segments, consisting of sequential raw measurements or processed data are correlated to the previously collected data (i.e. historical data) segments by using a similarity concept. Unlike traditional data-driven models, in SBP, RUL is calculated by aggregating the weighted average of the training sample RUL values rather than extrapolating the test sample's current health level to a predefined threshold.

Similarity-based Prognostic approach is a powerful algorithm for RUL estimations, notably when the historical training sample size is relatively abundant. In addition, they are suitable for the cases where the degradation path is not necessarily exhibiting a monotonic propagation pattern which is difficult to model using parametric approaches [81]. Wang et al [82] won the Prognostic and Health Management Society's data challenge competition in 2008 where they employed a similarity-based prognostic approach to predict the RUL of turbofan engines created by C-MAPSS simulation.

In similarity based prognostics the estimations of RUL involve calculating the similarity between the test sample (i.e. 'q') and the training samples (i.e. '$r = 1{:}R$') as shown in Eq. (5.13). The similarity index is based on the calculated point wise Euclidean distances in between '$n - long$' sequences of observations. Distance score calculation in between training samples and the test sample at the 'i^{th}' time point formulated in Eq. (5.12). Final RUL estimation of a test sample at a time instance (i.e. 'I') is achieved by aggregating the weighted average of training samples' corresponding remaining useful life values as formulated in Eq. (5.14). To be more precise, 'rul_i^r' symbolises the remaining useful life of the 'r^{th}' training sample at 'i^{th}' time point which is obtained by calculating the difference between the training sample's end-of-life time and the 'i^{th}' time point. The most similar segment to the test segment is specified for each training sample whereas the RUL of the test sample is obtained by taking the weighted average of these training RUL values. In fact, the weights are obtained using the bell-shaped similarity functions.

$$d_i^r = \sqrt{\sum_{j=1}^{n} \left\| z_{I-n+j}^q - z_{i-n+j}^r \right\|^2} \tag{5.12}$$

$$s_i^r = e^{-\frac{(d_i^r)^2}{\lambda}} \tag{5.13}$$

$$RUL_I^q = \frac{\sum_{r=1}^{R} s_i^r rul_i^r}{\sum_{r=1}^{R} s_i^r} \tag{5.14}$$

'λ' is an arbitrary parameter which can be set to shape the desired interpretation of similarity, whereas 'n' defines the number of latest consecutive observations involved in similarity calculations.

The smaller the 'λ' is, the stronger the definition of similarity. For instance, a small value for 'λ' signifies that the training segment should be very similar to the test segment so that it will be appointed with a similarity value reasonably higher than zero. However, when working with higher decimal point precision systems, this concept becomes trivial as the similarity ratio in between training samples remain the same.

TABLE 5-2. Remaining Useful Life (RUL) Prediction Error for Sensor Static Pressure at HPC Outlet

Data Driven Techniques and Models	Root-Mean-Square-Error (RMSE) and Mean Absolute Percent Error (MAPE)				
	Engine 6	ENGINE 39	ENGINE 69	ENGINE 73	Error Calculation Method
RBF NN + PF	**12.12**	**9.30**	**68.91**	**24.80**	**RMSE**
	19.53	19.65	86.32	25.87	MAPE (%)
SBP	**31.04**	**24.21**	**88.04**	**49.95**	**RMSE**
	36.15	50.81	38.78	45.18	MAPE (%)

TABLE 5-3. Remaining Useful Life (RUL) Prediction Error for Sensor Total Pressure HPC Outlet.

Data Driven Techniques and Models	Root-Mean-Square-Error (RMSE) and Mean Absolute Percent Error (MAPE)				
	Engine 6	ENGINE 39	ENGINE 69	ENGINE 73	Error Calculation Method
RBF NN + PF	**10.78**	**5.17**	**72.78**	**19.04**	**RMSE**
	18.43	15.78	104.89	18.85	MAPE (%)
SBP	**36.86**	**31.86**	**86.25**	**53.07**	**RMSE**
	36.62	69.46	37.22	46.22	MAPE (%)

TABLE 5-4. Remaining Useful Life (RUL) Prediction Error for Sensor Ratio of fuel flow to Ps30.

Data Driven Techniques and Models	Root-Mean-Square-Error (RMSE) and Mean Absolute Percent Error (MAPE)				
	Engine 6	ENGINE 39	ENGINE 69	ENGINE 73	Error Calculation Method
RBF NN + PF	**11.57**	**8.43**	**63.35**	**19.04**	**RMSE**
	16.32	25.69	81.42	23.59	MAPE (%)
SBP	**31.86**	**22.49**	**84.08**	**55.76**	**RMSE**
	36.48	49.71	40.67	49.73	MAPE (%)

TABLE 5-5. Remaining Useful Life (RUL) Prediction Error for Sensor LPT Coolant Bleed.

Data Driven Techniques and Models	Root-Mean-Square-Error (RMSE) and Mean Absolute Percent Error (MAPE)				
	Engine 6	ENGINE 39	ENGINE 69	ENGINE 73	Error Calculation Method
RBF NN + PF	**12.91**	**3.59**	**95.75**	**14.76**	**RMSE**
	35.44	13.39	124.24	30.03	MAPE (%)
SBP	**30.81**	**44.17**	**81.10**	**40.9**	**RMSE**
	40.82	104.76	39.30	41.00	MAPE (%)

The compassion between the presented adaptable data driven technique to SBP has been shown in the Figures 5.7 to 5.10 and the error also has been presented in tables 5.2 to 5.5.

The $RMSE$ ratio between nominal and adaptable models is more significant if the operating condition "a" of a test sample is more different from the nominal model. Hence this explains the better $RMSE$ results as observed in this experiment.

5.10 Summary

The future health prediction enables to minimise the risk of system failure, which could lead to catastrophic accidents. Prognostics will be most useful when the component's operating condition varies from its nominal value and cannot be pre-determined in advance. In most cases, the actual end-of-life of a component can be expected to be significantly different from its manufacturer-defined mean time before failure.

There are two most common approaches reported in the literature, data-driven and model-based. The data driven approach allows complex degradation patterns to be easily captured in the model, but has a severe limitation in the operating conditions in which the model is valid. In the model-based approach, a degradation model is explicitly described using mathematical equations with physical meaning. Explicit mathematical representation (often not possible to derive economically) allows the relevant model parameters to be scaled online to match a current operating condition. Therefore, a combination of data-driven model and online parameter estimation can be used to address a prognostics problem which has a complex degradation model and a large variation in operation conditions, and this is what is proposed in this chapter.

How fast a system degrades is proportional to its degradation rate which will essentially depend on the operating conditions. If the rate is multiplied by some factors, then this also means we are scaling the time by a proportion of that factor. In order for a data-driven model to be scalable, a scaling parameter will have to be explicitly part of the model. Instead of overall fitting the data, the model (the model presented in this chapter) has to learn a nominal degradation pattern where different degradation profiles can be fitted by using some appropriate model parameters. This way the data-driven model is able to be updated to match the current operating conditions, and hence a better RUL prediction because of a more accurate model.

In this case study, it can be seen that the proposed prognostic approach has potential to be applicable in a real world environment where operating conditions can vary significantly and cannot be accurately defined in advance. The approach can potentially be used in applications where only accelerated testing data are available for the algorithm development. However, its claim as a generic adaptable approach has to be further researched in comparison with other engineering examples and with other data-driven techniques, like Time Delayed Neural Network (TDNN) or similarity based prognostics (SBP).

Chapter 6

Conclusions

Vehicle Level Reasoning System (VLRS) is a relatively new concept as compared to a system level diagnostics system especially in aircraft research and industry. In aircraft field, system level diagnostics and detection systems have the limitation as they only monitor single sub-systems of the aircraft. Whereas VLRS doesn't suffer this limitation, however VLRS does encounter different problems compared to sub-system diagnostic systems. The next section will discuss the VLRS philosophy, effectiveness and limitation.

6.1 The VLRS Effectiveness and their Philosophy

There are some issues when designing the VLRS, such as the integration of the VLRS with the subjected systems, which will provide the vital sensor data in order to perform the reasoning to produce vehicle level results. Secondly, most of the algorithm have convergence issue on vehicle level because of the vast amount of sensors data which needed to be processed. Thirdly, to monitor the enormous amount of data, it requires a huge computational processing power, which could be a problem if the reasoning is done on-board.

In this thesis the VLRS presented addresses all the above issues and presents a novel VLRS design, which overcomes above listed issues. However there are different limitations of the presented VLRS (presented in chapter 3). The presented VLRS utilises the HI and vehicle behaviour along with the sub-system diagnostics results, which are more computational efficient and have less problems in order to deploy the present aircraft without having huge integration issues.

In VLRS the prognostics engine's output and health index information are very important, as the components that have a lower health index are more likely to be faulty or working at the lowest threshold which could cause problems or future problems. In this thesis there are three techniques presented, one conceptual design of VLRS (chapter 3), operational scalable prognostic framework (chapter4) and adaptive data driven techniques (chapter 5).

6.2 Objectives and Evaluation

The outlined aims and objectives in the chapter 1 have been fulfilled entirely. The objective 1 (**Review and identify key modelling and prediction aspects of state-of-the-art model-based and data-driven prognostic approaches**) and objective 2 (**Identify key technical requirements to a predictive power of a prognostic reasoner if to be deployed in a real-world operation uncertain environment**) has been met and demonstrated in chapter 2 and 4 as well as in the published one journal and two conference article. Full review of VLRS and data driven prognostics techniques has been discussed. The real life problems have been heighted for prognostics framework.

The objective 3 (**Design a generic prognostic framework (or problem formulation) that can be operationally scaled to usage and imperfect manufacturing uncertainties and yet can be simply constructed using standard data-driven and parameter estimation building blocks**.) has been fulfilled and showed in the chapter 4 and published journal article. The design of the framework which can scaled to the real environment and with different usage profiles has been presented in the chapter 4 and in the journal article. The usage of the framework and the problems has bean also evaluated. The framework works very well and it scales to the usage profile as the results shows, however huge amount of data is required in order the train the framework.

The prognostics allow anticipation of system failures well ahead of time so that maintenance can be planned, scheduled and performed in the most optimal way. Prognostics will be most useful when the component's operating condition varies from its nominal value and cannot be pre-determined in advance. In most cases, the actual end-of-life of a component can be expected to be significantly different from its manufacturer-defined mean time before failure (MTBF).

Data-driven and model-based are the most common prognostic approaches reported in the literature. The former allows complex degradation patterns to be easily captured in the model, but has a severe limitation in the operating conditions in which the model is valid. In the model-based approach, a degradation model is explicitly described using mathematical equations with physical meaning. Explicit mathematical representation (often not possible to be derived) allows the relevant model parameters to be scaled online to match a current operating condition. To solve this problem two different data driven techniques have been presented in this thesis. In one of the techniques a combination of data-driven model and online parameter estimation can be used to address a prognostic problem which has a complex degradation model and large variation in operation conditions. How fast a system degrades is proportional to its degradation rate which will essentially depend on the operating conditions. If the rate is multiplied by some factors, then this also means we are scaling the time by a proportion of that factor. In order for a data-driven model to be scalable, a scaling parameter will have to be explicitly part of the model. Instead of overall fitting the data, the model (which is a TS fuzzy model in chapter 4) has to learn a nominal degradation pattern where different degradation profiles can be fitted by using some appropriate model parameters. This way the data-driven model is able to be updated to match the current operating conditions, and hence a better RUL prediction because of a more accurate model.

In the first data driven technique (chapter 4), a TS fuzzy (where any data driven technique can be used to create the model) model is used to approximate the nominal degradation pattern from the data. In overall, the model parameter will be able to converge close to its true and thus a good RUL prediction. However, at a given time, how well the prediction tracks the true RUL will depend on how much the modelling error (deviation of the scaled degradation pattern) is from the degradation profile at an interval around that particular time step. The keys to this approach therefore are: the data for different operating conditions have a

similar degradation pattern and different data can be closely approximated by rate-scaling the nominal model.
For this technique the filter clogging case study has been used. It can be seen that the proposed prognostic approach has potential to be applicable in a real world environment where operating conditions can vary significantly and cannot be accurately defined in advance. The approach can potentially be used in applications where only accelerated testing data are available for the algorithm development. However, its claim as a generic operational-scalable approach has to be further researched in comparison with other engineering examples and with other data-driven techniques, for example Time Delayed Neural Network (TDNN) Bayesian belief etc. can be used.

The objective 4 (**The conceptual design of VLRS by providing prognostics engine output as a Health Index to VLRS**) has been fulfilled and showed in the chapter 3 and published journal article and a conference article. The novel design of the VLRS has been presented and the usage of the behaviour and health index are used in the design. The design also has some limitations, such as the behaviour engine could give not every accurate fault list as well as it also need to process lot of sensory data to measure health index.
The objective 5 (**Proof-of-concept demonstration and performance evaluation of the developed framework using adaptive/scalable machine learning and online estimation models based on aerospace related prognostic case studies**) has been fulfilled and showed in the chapter 5 and published journal article. In the chapter 5 the adaptiveness of the second technique has been demonstrated by using several different dataset of different jet engines. The results are also presented and benchmarked.
This data driven technique (chapter 5) shows that adaptation to the condition. In this technique Neural Network (NN RBF) has been used to create the nominal model from the data and the nominal model with uncertainty bounds has been given to the particle filter (PF). In this technique the NASA Jet turbofan data has been used. This dataset has the data of 21 sensors of jet engines. The Jet engines are classified as complex systems as their degradation is very difficult to model,

therefore the sensor data has been used to convert into health index information, which means what sensor information really means to operator rather than just number. To calculate the HI from the sensor readings data-driven technique have been employed. Data driven techniques are very good in modelling the complex systems, therefore this task has been achieved by the Neural Network using radial basis function which are typical data driven techniques. Once the health degradation model has been created it has been given to the particle filter (PF). The particle filter has adapted the model to real sensor data. Many different samples have been tested in order to confirm whether the adaptation technique is working as intended (Please see results of chapter 5).

The technique which is presented in this chapter has been benchmarked by similarity based prognostics approach and the results are compared please see results are provided in the tables 5-2 to 5-5.

The scalable and adaptive, both data driven techniques shows that it has potential to be used at the industrial level, as in a real environment the conditions are very dynamic which is hardly the case with the laboratory generated data.

Generally, the presented prognostics techniques use laboratory generated data which are created in very controlled environment and do not contains much variations in the data. Both of these techniques focus on how to solve these issues as well which have been noted previously in this thesis.

6.3 Key Novelty

The key novelty of this are listed below:

1. The novel design has been presented of VLRS which utilise the usage of prognostics framework provided health index and vehicle behaviour.

2. The scalable data driven techniques has also presented by using the fuzzy logic and particle filter. This technique provides accurate results even if the usage of the system is very much different from the nominal model usage.

3. The adaptive data driven technique is also novel the way it has been implemented it also calculate the health index as well the RUL of the component. The also adapts to the mission context while calculates the RUL which makes this technique adaptive and novel.

6.4 Future Work

The Future works could be pursued into a few different directions, and one could be in VLRS, secondly the scalable prognostics framework presented in chapter 4 and thirdly could be the adaptable prognostics technique for complex system.

In VLRS, further design configurations and testing and results would be the 1st step and more sub-system information could be shared with VLRS and further development of the VLRS, where sub-system results could be provided along with the behaviour and health index. Automatically detection of the vehicle behaviour would be very beneficial for VLRS, therefore the vehicle behavioural engine would be another further work. The implementation of the VLRS on small sized drone would be real beneficial and economical to proof of concept. The behaviour engine and health index calculation could be improved after the results on the implementation in real life environment. Usually, each vehicle would have different fault allocation with different behaviour therefore real system testing will provide intensive amount of data which could be used in order to update the VLRS.

Secondly, In terms of prognostics framework, future development could be implemented with several different datasets in order to evaluate further performance which might be able to enhance by using of different deep learning algorithms along with more than just 3 number of degradation stages as presented in chapter 4.

Thirdly, the RBF Neural Network has been used to create the nominal model to map the sensor values to health index. This process also can be further develop by using cascading Neural Network or even Bayesian regression algorithm.

The fuel rig data (presented in chapter 4) also could be used on this technique, this task would be very simple to achieve and it would provide plausible results and it will add information to raw sensor data.

Furthermore, in complex system such as jet engine many different components are involved in the degradation of the whole system. The dataset has been used in the chapter 5 it has 21 different sensor data which are providing indirect condition monitoring of the system, one sensor information/data would be not enough to monitor the whole health of the system. Therefore the sensor fusion technique would be application which will provide the system level prognostics results. There are several different technique available for sensor fusion, however the sensor fusion technique must follow the overall health and RUL of system to validate if the sensor fusion technique is valid for prognostics proposes

References

[1] O. Benedetti, T. S. Baines, H. W. Lightfoot and R. M. Greennough, “State-of-the-art in integrated vehicle health management,” *G: J. Aerospace Engineering,* vol. 223, pp. 157-171, 2009.

[2] S. Ofsthun, “Integrated Vehicle Health Management for Aerospace Platforms,” *Instrumentation & Measurement Magazine, IEEE,* vol. 5, no. 3, pp. 21-24, 2002.

[3] J. H. MacConnell, “ISHM & Design: A review of the benefits of the ideal ISHM system,” in *Aerospace Conference IEEE*, Seattle, 2007.

[4] P. Brown and M. Sondalini, “The Evolution of Maintenance Practices,” [Online]. Available: http://www.lifetime-reliability.com/free-articles/maintenance-management/Evolution_of_Maintenance_Practices.pdf.

[5] L. Pintelon and A. Parodi-Herz, “Maintenance: An Evolutionary Perspective,” in *Complex System Maintenance Handbook*, Springer Series in Reliability Engineering, 2008, pp. 21-48.

[6] R. Kothamasu, S. H. Huang and W. H. Verduin, “System Health Monitoring and Prognostics- A review of current paradigms and practices,” *International Journal of Advanced Manufacturing Technology ,* vol. 28, no. 9, pp. 1012-1024, 2006.

[7] D. N. P. Murthy, A. Attrens and J. A. Essleston, “Strategic Maintenance Management,” *Journal of Quality in Maintenance Engineering,* vol. 8, no. 4, pp. 287-305, 2002.

[8] J. Z. Sikorska, M. Hodkiewicz and L. Ma, “Prognostic Modelling Options for Remaining Useful Life Estimation by Industry,” *Mechanical Systems and Signal Processing,* vol. 25, no. 5, pp. 1803-1836, 2011.

[9] J. Lee, J. Ni, D. Djurdjanovic, H. Qui and H. Liao, “Intelligent Prognostics Tools and E-maintenance,” *Computers in Industry ,* vol. 57, no. 6, pp. 476-489, 2006.

[10] F. Camci and R. B. Chinnam, “Health-state estimation and prognostics in machining processes,” *IEEE Transactions on Automation Science and Engineering ,* vol. 7, no. 3, pp. 581-597, 2010.

[11] O. F. Eker, F. Camci, A. Guclu, H. Yilboga, M. Sevkli and S. Baskan, “A simple sate-based prognostic model for railway turnout systems,” *IEEE Transactions on Industrial Electronics,* vol. 58, no. 5, pp. 1718-1726, 2011.

[12] I. Jennions, Integrate Venhicle Health Management, Perspectives on an Emerging Field, Warrendale, Pennsylvania, USA: SAE International, 2011.

[13] H. Hecht, “Why prognostics for avionics?,” in *Aerospace Conference 2006 IEEE*, Big Sky, MT, 2006.

[14] C. Kwan, X. Zhang, R. Xu and L. Haynes, “A novel approach to fault diagnostics and prognostics,” in *IEEE International Conference on Robotics and Automation*, 2003.

[15] Y. Xiong, X. Cheng, Z. J. Shen, C. Mi, H. Wu and V. K. Garg, “Prognostic and warning system for power-electronic modules in electric, hybrid electric, and fuel-cell vehicles,” *IEEE Transactions on Industrial Electronics,* vol. 55, no. 6, pp. 2268-2276, 2008.

[16] L. Zhang, X. Li and J. Yu, “A review of fault prognostics in condition based maintenance,” in *Proceeding of SPIE- The International Society for Optical Engineering*, 2006.

[17] Y. Peng, M. Dong and M. J. Zuo, “Current status of machine prognostics in condiction-based maintenance: A review,” *International Journal of Advanced Manufacturing Technology*, vol. 50, no. 1-4, pp. 297-313, 2010.

[18] M. Daigle and K. Goebel, “Model-based prognostics under limited sensing,” in *IEEE Aerospace Conference*, Big Sky, MT, 2010.

[19] O. F. Eker, F. Camci and I. Jennions, “Major challenges in prognostics: Study on benchmarking prognostics datasets,” in *PHM*, Europe, 6th- 8th July 2012.

[20] L. Liao and F. Kottig, “Review of Hybrid Prognostics Approaches for Remaining Useful Life Prediction of Engineered Systems, and an Application to Battery Life Prediction,” *IEEE Transactions on Reliabilty* , vol. 63, no. 1, pp. 191-207, 2014.

[21] G. Vachtsevanos, L. Lewis, M. Roemer, A. Hess and B. Wu, “Intelligent Fault Diagnosis and Prognosis for Engineering System,” Wiley, 2006.

[22] A. Heng, S. Zhang, A. C. C. Tan and J. Mathew, “Rotating machinery prognostics: State of the art, challenges and opportunities,” *Mechanical Systems and Signal Processing* , vol. 23, no. 3, pp. 724-739, 2009.

[23] X. Si, W. Wang, C. Hu and D. Zhou, “Remaining useful life estimation - A review on the statistical data driven approaches,” *European Journal of Operational Research,* vol. 213, no. 1, pp. 1-14, 2011.

[24] J. Luo, A. Bixby, K. Pattipati, L. Qiao, M. Kawamoto and S. Chigusa, “An interacting multiple model apprach to model-based prognostics,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2003.

[25] A. K. S. Jardine, D. Lin and D. Banjevic , “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” *Mechanical Systems and Signal Processing* , vol. 20, no. 7, pp. 1483-1510, 2006.

[26] H. Zhang, R. Kang and M. Pecht, “A hybrid prognostics and health management approach for condition-based maintennance,” in *IEEE International Conference on Industrial Engineering and Engineering Management* , 2009.

[27] M. Daigle, *Model based prognostics,* PHM Society , 2014.

[28] G. J. Kacprzynski, M. J. Roemer, G. Modgil, A. Palladino and K. Maynard, “Enhancement of physics-of-failure prognostic models with system level features,” in *Aerospace Conference Proceedings*, 2002.

[29] C. S. Byington, M. Watson, D. Edwards and P. Stoelting, “A model-based approach to prognostics and health management for flight control actuators,” in *IEEE Aerospace Conference Proceedings*, 2004.

[30] C. Cempel, H. G. Natke and M. Tabaszewski, “A passive diagnostic experiment with ergodic properties,” *Mechanical Systems and Signal Processing,* vol. 11, no. 1, pp. 107-117, 1997.

[31] J. Qiu, B. B. Seth, S. Y. Liang and C. Zhang, “Damage mechanics approach for bearing lifetime prognostics,” *Mechanical Systems and Signal Processing* , vol. 16, no. 5, pp. 817-829, 2002.

[32] D. Chelidze and J. P. Cusumano, “A dynamical systems approach to failure prognosis,” *Journal of Vibration and Acoustics, Transactions of the ASME,* vol. 126, no. 1, pp. 2-8, 2004.

[33] G. J. Kacprzynski, A. Sarlashkar, M. J. Roemer, A. Hess and W. Hardman, “Predicting remaining life by fusing the physics of failure modeling with diagnostics,” *JOM,* vol. 56, no. 3, pp. 29-35, 2004.

[34] G. A. Lesieutre, L. Fang and U. Lee, “Hierarchical Failure Simulation for Machinery Prognostics,” in *Biennial Conference; 12th, Reliability, stress analysis and failure prevention: A critical link*, Virginia Beach, VA , 1997.

[35] D. C. Swanson, “A general prognostic tracking algorithm for predictive maintenance,” in *IEEE Aerospace Conference Proceedings* , 2001.

[36] E. Phelps, P. Willett and T. Kirubarajan, “A statistical approach to prognostics,” in *Proceedings of SPIE-The International Society for Optical Engineering* , 2001.

[37] C. Hu, B. D. Youn and J. Chung, “A multiscale framework with extended Kalman filter for lithium-ion battery SOC and capacity estimation,” *Applied Energy* , vol. 92, pp. 694-704, 2012.

[38] E. Zio and G. Peloni, “Particle filtering prognostic estimation of the remaining useful life of nonlinear components,” *Reliability Engineering and System Safety,* vol. 96, pp. 403-409, 2011.

[39] D. An, J. Choi and N. H. Kim, “Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab,” *Reliabilty Engineering & Systems Safety,* vol. 115, pp. 161-169, 2013.

[40] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess and B. Wu, “Intelligent Fault Diagnosis and Prognosis for Engineering Systems,” in *John Wiley & Sons, Inc.*, Hoboken, New Jersey, USA, 2006.

[41] E. Lembessis, G. Antonopoulos and R. E. King, “CASSANDRA: an on-line expert system for fault prognosis,” 1989.

[42] T. Biagetti and E. Sciubba, “Automatic diagnostics and prognostics of energy conversation processes via knowledge-based systems,” *Energy,* vol. 29, no. 12-15, pp. 2553-2572, 2004.

[43] K. L. Butler, “Expert system based framework for an incipient failure detection and predictive maintenance system,” in *International Conference on Intelligent Systems Applications to Power Systems* , 1996.

[44] S. S. Choi, K. S. Kang, H. G. Kim and S. H. Chang, “Devopment of an on-line fuzzy expert system for integrated alarm processing in nuclear power plants,” *IEEE Transactions on Nuclear Science,* vol. 42, no. 4 pt 2, pp. 1406-1418, 1995.

[45] K. Dmitry and V. Dmitry, “An algorithm for rule generation in fuzzy expert systems,” in *Proceedings of the 17th international Conference on Pattern Recognition*, 2004.

[46] C. C. Lee, “Fuzzy logic in control systems: fuzzy logic controller,” *IEEE Transactions on Systems, Man and Cybernetics,* vol. 20, no. 2, pp. 404-418, 1990.

[47] E. Feng, H. Yang and M. Rao, “Fuzzy expert system for real-time process condition monitoring and incident prevention,” *Expert systems with Applications* , vol. 15, no. 3-4, pp. 383-390, 1998.

[48] A. Majidian and M. H. Saidi, “Comparison of Fuzzy logic and Neural Network in life prediction of boiler tubes,” *International Journal of Fatigue,* vol. 29, no. 3, pp. 489-498, 2007.

[49] A. Saxena , B. Wu and G. Vachtsevanos, “A hybrid reasoning architecture for fleet vehicle maintenance,” *Instrumentation & Measurement Magazine, IEEE,* vol. 9, no. 4, pp. 29-36, 2006.

[50] F. Khan , I. Jennions and T. Screenuch, “Integration Issues for Vehicle Level Distributed Diagnostic Reasoners,” SAE Technical Paper, 2013.

[51] F. Khan, O. F. Eker, T. Sreenuch and A. Tsourdos, “Multi-Domain Modeling and Simulation of an Aircraft System for Advanced Vehicle-Level Reasoning Research and Development,” *International Journal of Advanced Computer Science and Applications (IJACSA),* vol. 5, no. 4, pp. 86-96, 2014.

[52] A. Del Amo, K. Keller and K. Swearingen, “General reasoning system for health management,” in *Fuzzy Information Processing Society, IEEE*, 2005.

[53] G. Ramohalli, “The Honeywell on-board diagnostic and maintenance system for the Boeing 777,” in *Digital Avionics Systems Conference IEEE/AIAA*, 1992.

[54] A. N. Srivastava, D. Mylaraswamy, M. W. R and E. G. Cooper, ““Vehicle-level Reasoning Systems” in IVHM Perspectives on an Emerging Field,” in *SAE International.*, 2011..

[55] A. Srivastava , R. Mah and C. Meyer, *Automated Detection Diagnosis, and Prognosis to enable mitigation of adverse Events during flight,* NASA, 2009.

[56] A. Hess, G. Calvello and T. Dabney, “PHM a Key Enabler for the JSF Autonomic Logistic Support Concept,” in *IEEE Aerospace Conference*, 2004.

[57] S. Liao, “Expert system methodologies and application- a decade review from 1995 to 2004,” *Expert Systems with Applications,* pp. 93-103, 2005.

[58] Jianzhong Sun, H. Zou, W. Wenbion and P. G. Michael, “Application of a state space modeling techbique to system prognostic based on a health index for condition-based maintenance,” *ELSEVIER , Mechnical System and Sinal Processing,* vol. 28, pp. 585-596, 2012.

[59] C. E. Balin and J. Stankunus, “Investigation of Fault Detection and Analysis Methods for Central Maintenance Systems,” in *Aviation Technology*, 2013.

[60] S. Gillet, A. Nuric and V. Mouillet, “ENHANCEMENT IN REALISM OF ATC SIMULATIONS BY IMPROVING,” in *European Organisation for the Safety of Air Navigation -*, Brétigny/Orge, France.

[61] N. Gorjian, L. Ma, M. Mittinty, P. Yarlagadda and Y. San, “A review on degradation models in reliability analysis,” in *The 4th World Congress on Engineering Asset Management*, Athens, 2009.

[62] M. Jones, *Engine Fuel Filter Contamination,* Aero Magazine , 2008.

[63] X. Wang, “Aircraft fuel system prognostics and health management,” PhD and Masters by research theses (School of Engineering) , Cranfield University, 2012.

[64] O. F. Eker, F. Camci and I. K. Jennions, “Filter clogging data collection for prognostics,” in *Annual Conference of the Prognostics and Health Management Society*, New Orleans, 2013.

[65] *Signal Processing Toolbox: User's Guide,* Natick, MA: Mathworks, 2014.

[66] T. A. Johansen, R. Shorten and R. Murray-Smith, “On the interpretation and identification of dynamic Takagi-Sugeo fuzzy models,” *IEEE Transactions on Fuzzy Systems,* vol. 8, no. 3, pp. 297-313, 2000.

[67] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man and Cybernetics,* vol. 15, no. 1, pp. 116-132, 1985.

[68] R. Marti, M. Laguna and F. Glover, “Principles of scatter search,” *European Journal of Operational Research,* vol. 169, no. 2, p. 359–372, 2006.

[69] D. G. T. Denison, C. Holmes, B. K. Mallick and A. F. M. Smith, Bayesian Methods for Nonlinear Classification and Regression, New York: Wiley, 2002.

[70] A. Blake and H. Mumtaz, Technical Handbook - No. 4 Applied Bayesian econometrics for central bankers, London: Centre for Central Banking Studies, Bank of England, 2012.

[71] Z. Chen, “Bayesian filtering: from Kalman filters to particle filters and beyond,” *Statistics,* vol. 182, no. 1, pp. 1-69, 2003.

[72] I. Yildirim, “Bayesian inference: Metropolis-Hastings sampling,” Department of Brain and Cognitive Sciences, University of Rochester, Rochester, 2012.

[73] M. P. Rios and H. F. Lopes, “The extended Liu and West filter: parameter learning in Markov switching stochastic volatility models,” in *State-Space Models: Applications in Economics and Finance*, New York, Springer, 2013, pp. 23-61.

[74] M. J. Daigle and K. Goebel, “Model-based prognostics with concurrent damage progression processes,” *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 43, no. 3, pp. 535-546, 2013.

[75] C. Nemeth, P. Fearnhead, L. Mihaylova and D. Vorley, “Particle learning methods for state and parameter estimation,” in *IET Data Fusion & Target Tracking Conference*, London, 2012.

[76] J. Lui and M. West, “Combined parameter and state estimation in simulation-based filtering,” in *Sequential Monte Carlo Methods in Practice*, New York, Springer, 2001, pp. 197-223.

[77] A. Saxena, J. Celaya, B. Saha, S. Saha and K. Goebel, “Metrics for offline evaluation of prognostic performance,” in *Annual Conference of the PHM Society*, Portland, OR, 2010.

[78] A. Saxena and K. Goebel, “Turbofan Engine degradation simulation data set, NASA Ames Prognostics Data Repository,” [Online]. Available: http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/. [Accessed 15 August 2010].

[79] D. Frederick , J. DeCastro and J. Litt, “User's Guide for the Commercial Modular Acro-Propulsion System Simulation,” in *(C-MAPSS) NASA*, Cleveland, Ohio, 2007.

[80] A. Saxena, K. Goebel, D. Simon and N. Eklund, “Damage propagation modeling for aircraft engine run-to-failure simulation,” *International Conference on Prognostics and Health Management,* pp. 6-9, 2008.

[81] T. Wang, “ Trajectory Similarity Based Prediction for Remaining Useful Life Estimation (PhD thesis),,” UNIVERSITY OF CINCINNATI, UNIVERSITY OF CINCINNATI, CINCINNATI, 2010.

[82] T. Y. J. S. D. L. J. Wang, “A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems,” in *International Conference on Prognostics and Health Management, PHM 2008,*, 2008.

[83] F. O. Heimes, “Recurrent neural networks for remaining useful life estimation,” in *International Conference on Prognostics and Health Management*, Denver, CO, 2008.

[84] R. Huang, L. Xi, X. Li, C. R. Liu, H. Qiu and J. Lee, “Residual life predictions for ball bearing based on self-organizing map and back propagation neural network methods,” *Mechanical Systems and Signal Processing,* vol. 21, pp. 193-207, 2007.

[85] C. S. Byington and M. Watson, “Data-driven neural network methodology to remaining life predictions for aircraft actuator components,” in *IEEE Aerospace Conference*, Big Sky, MT, 2004.

[86] A. M. Riad, H. K. Elminir and H. M. Elattar, “Evaluation of neural networks in the subject of prognostics as compared to linear regression model,” *International Journal of Engineering & Technology,* vol. 10, no. 6, pp. 50-56, 2010.

[87] T. Benkedjouh, K. Medjaher, N. Zerhouni and S. Rechak, “Remaining useful life estimation based on nonlinear feature reduction and support vector regression,” *Engineering Applications of Artificial Intelligence,* vol. 26, pp. 1751-1760, 2013.

[88] K. Medjaher, D. A. Tobon-Mejia and N. Zerhouni, “Remaining useful life estimation of critical components with application to bearing,” *IEEE Transactions on Reliability,* vol. 61, no. 2, pp. 292-302, 2012.

[89] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni and G. Tripot, “A data-driven failure prognostics method based on mixture of Gaussions Hidden Markov models,” *IEEE Transactions of Reliability,* vol. 61, no. 2, pp. 491-503, 2012.

[90] D. Liu, J. Pang, J. Zhou, Y. Peng and M. Pecht, “Prognostics for state of health estimation of lithium-io batteries based on combination Gaussian process functional regression,” *Microelectronics Reliability,* vol. 53, pp. 832-839, 2013.

[91] Y. Liu, S. Mohanty and A. Chattopadhyay, “A Gaussian Process based prognostics framework for composite structures,” in *Proceeding of SPIE, Smart Structures and Materials & Nondestructive*, San Diego, CA, 2009.

[92] X. S. Si, W. Wang, C. H. Hu and D. H. Zhou, “Remaining useful life estimation - A review on the statistical data driven approaches,” *European Journal of Operational Research,* vol. 213, pp. 1-14, 2011.

[93] Y. Li, S. Billington, C. Zhang, T. Kurfess, S. Danyluk and S. Liang, "Adaptive prognostics for rolling element bearing condition," *Mechanical Systems and Signal Processing,* vol. 13, no. 1, pp. 103-113, 1999.

[94] W. He, N. Williard, M. Osterman and M. Pecht, "Prognostics of lithium-ion batteries using extended Kalman filtering," in *IMAPS Advanced Technology Workshop on High Reliability Microelectronics for Military Applications*, Linthicum Heights, MD, 2011.

[95] M. Chookah, M. Nuhi and M. Modarres, "A probabilistic physics-of-failure model for prognostic health management of structures subject to pitting and corrosion-fatige," *Reliability Engineering and System Safety,* vol. 96, pp. 1601-1610, 2011.

[96] T. Tinga, "Application of physical failure models to enable usage and load based maintenance," *Reliability Engineering and System Safety,* vol. 95, pp. 1061-1075, 2010.

[97] Y. G. Li and P. Nilkitsaranont, "Gas turbine performance prognostic for condition-based maintenance," *Applied Energy,* vol. 86, no. 10, pp. 2152-2161, 2009.

[98] O. Sasahara, "JT9D engine/module performance deterioration results from back to back testing," in *Proceedings of the 7th International Symposium on Air Breathing Engines,* Beijing, China, 1986.

[99] J. Crosby, "Factors relating to deterioration based on Rolls-Royce RB211 in service performance," 1986.

[100] C. Wang, L. Xu and W. Peng, "Conceptual design of remote monitoring and fault diagnosis systems," *Information Systems,* vol. 32, no. 7, pp. 996-1004, 2007.

[101] W. Donat, W. Choi, S. Singh and K. Pattipati, "Data visualization, data reduction and classifier fusion for intelligent fault diagnosis in gas turbine engines," *Journal of Engineering for Gas Turbines and Power,* vol. 130, no. 4, p. Article ID 041602, 2008.

[102] H. Bassily, R. Lund and J. Wagner, "Fault detection in multivariate signals with applications to gas turbines," *IEEE Transactions on Signal Processing,* vol. 57, no. 3, pp. 835-842, 2009.

[103] K. Young, D. Lee, V. Vitali and Y. Ming, "A fault diagnosis method for industrial gas turbines using Bayesian data analysis," *Journal of Engineering for Gas Turbines and Power,* vol. 132, p. Article ID 041602, 2010.

[104] B. Yu, D. Liu and T. Zhang, "Fault diagnosis for micro-gas turbine engine sensors via wavelet entropy," *Sensors,* vol. 11, no. 10, pp. 9928-9941, 2011.

[105] S. Wu, P. Wu, C. Wu, J. Ding and C. Wang, "Bearing fault diagnosis based on multiscale permutation entropy and support vector machine," *Entropy,* vol. 14, no. 8, pp. 1343-1356, 2012.

[106] S. Wu, C. Wu, T. Wu and C. Wang, "Multi-scale analysis based ball bearing defect diagnostics using Mahalanobis distance and support vector machine," *Entropy,* vol. 15, no. 2, pp. 416-433, 2013.

[107] W. T. Thomson and R. J. Gilmore, "Motor current signature analysis to detect faults in induction motor drives-Fundamentals, data interpretation, and Industrial case histories," in *Proceedings of the Thiry-second Turbo machinery Symposium*, 2003.

[108] D. Gorinevsky, S. Boyd and S. Poll, "Estimation of faults in dc electrical power system," in *American Control Conference,* St. Louis, 2009.

[109] W. R. Finley, M. M. Hodowanec and W. G. Holter, "An Analytical Approach to Solving Motor Vibration Problems," in *Petroleum and Chemical Industry Conference, Industry Applications Society 46th Annual* , San Diego, 1999.

[110] N. Gorjian, L. Ma, M. Mittinty, P. Yarlagadda and Y. Sun, "A review on degradation models in reliability analysis," in *Engineering Asset Management- Springer* , 2010.

APPENDIX