Health Monitoring of Gas Turbine Engine Framework Design and Strategies

Rishi Relan

A Dissertation Submitted In Partial Fulfilment of the Requirements for the Degree of Master of Philosophy in the Department of Automatic Control and Systems Engineering

June 2015

DECLARATION

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is majorly my own work. The work done in collaboration with others is duly acknowledged as well as specified in the text at appropriate palces.

Rishi Relan

ACKNOWLEDGEMENTS

This I would like to express my deep gratitude to my supervisors, Professor Visakan Kadirkamanathan, who is also the present Director of Rolls-Royce UTC, and Professor Robert. F. Harrison for giving me the opportunity to purse my MPhil research under their guidance. I especially thank Professor Visakan Kadirkamanathan for his excellent guidance, helpful insights and providing great encouragement as well as time during my research. I feel honoured to have worked under such an experienced supervisor. I sincerely thank Professor Robert. F. Harrison for passing on his scientific knowledge and expertise on prognostics. I would like to thank the former Director of Rolls-Royce UTC, Professor Peter Fleming for his constant encouragement and advice.

I would like to thank, Dr. Martha Arbayani Bin Zaidan, my dear friend and colleague in the Rolls-Royce UTC for the technical discussions as well as the collaboration, which produced our joint work on “Integrated Prognostic Approach” and contributed towards the design of better health monitoring approach for gas turbine engines.

I am grateful to Rolls-Royce plc and Engineering and Physical Sciences Research Council (EPSRC) for providing me the necessary financial support through Dorothy Hodgkin Postgraduate Award (DHPA).

I would like to thank Andy Mills for all the helpful discussions, and many thanks to Debbie Proctor, Renata Ashton for being able to solve just about anything administrative. My profound thanks to my dear friends Vignesh, Rui and members of the UTC for their unconditional support and with whom I had the great luxury of working especially, Hasanin, Maszatul, Ioannis, Arun and others.

I would like to express my greatest appreciation and respect to my mother, thousands of miles away in New Delhi, India. Special thanks goes to my wife, being a PhD student at the University of Edinburgh in Scotland and also taking care of our recently born beloved daughter alone, takes a lot of effort. This work would not have been possible without constant love, prayers and sacrifice from them. I am truly indebted to all their patience, support and understanding. Thank you!

I would also like to express my immense gratitude to the staff of the Department of Automatic Control and Systems Engineering. Last but not the least; I thank the University of Sheffield for giving me this great opportunity. I feel proud to have been part of this excellent institution. To all those cited by name here and many others I did not mention by name who supported me during this wonderful journey, I will forever owe you this achievement.

Thank you all very much!

LIST OF FIGURES

List of TABLES

LIST OF ACRONYMS

EHM	Equipment Health Management
SatCom	Satellite communication
FMMEA	Failure Mode Mechanism Effects Analysis
GTE	Gas Turbine Engine
SATAA	Sense, Acquire, Transfer, Analyse, And Act
BIT	Built-in-test
PIMENTO	Prognostic Inference Method Embedded in Novel Toolset
FMU	Fuel Metering Unit
LRU	Line Replaceable Unit
TGT	Turbine gas temperature
M.I	Mutual information
KDE	Kernel density estimation
FCU	Fuel control unit
EEC	Electronic engine control
FADEC	Full authority digital engine control
AFD	Active fault diagnosis
FDI	Fault detection and isolation
CS	Compressed sensing
ApEn	Approximate entropy
Div	Divergence

LPC	Low pressure combustion
HPT	High pressure turbine
AFD	Active fault diagnosis
FMECA	Failure mode, effects, and criticality analysis
RMS	Root mean square
MAP	Maximum a posteriori
CTG	Cardiotocography
MMF	Magneto-motive force
PSD	Power spectral density
EEG	Electro-Encephalogram
FMV	Fuel metering valve
SpEn	Spectral entropy
RUL	Remaining useful life
CPD	Change point detection
CBM	Condition Based Maintenance
KDE	Kernel density estimation
TM	Torque motor
TMD	Torque motor demand
SV	Servo valve
RuLSIF	Relative unconstrained Least-Squares Importance Fitting
i.i.d	Independently Identically distributed
BR-3	Bayesian Regression 3

LIST OF SYMBOLS

μ	Mean
σ, std	Standard deviation
σ^2, var	Variance
RMS	Root mean square
X	Vector of sampled signal or random variable
x_n	Sampled value of a signal or i.i.d samples of a random variables, where $n = 1,2,3 \dots \mathbb{R}$
$p(x)$	Probability density function
$K(\blacksquare)$	Kernel function
$\widehat{p(x)}$	Estimated probability density function
S	Covariance Matrix
det	Determinant
δS	Elementary change of entropy
δQ	Reversibly received elementary heat
$Temp$	Temperature
$H(\blacksquare)$	Entropy
h, γ	kernel bandwidth
χ	Finite support set discrete random variable
$I(X;Y)$	Mutual Information
$H(X;Y)$	Joint Entropy
$p(x;y)$	Joint probability distribution function

$P(x\|y)$	Condition probability distribution function of x given y
$p(x\|y)$	Condition probability density function of x given y
$H(x\|y)$	Condition entropy of x given y
m	Embedding dimension
r	Predetermined tolerance value for approximate entropy
$d(X(i), X(j))$	Distance between two vectors $X(i), X(j)$
k	Constant
$S_{xx}(w)$	Power spectral density
w	Angular frequency
$\mathbb{E}$	Expectation operator
$x(t)$	Signal from length up to time t
P_f	Power level
H_{SpEn}	Spectral Entropy
$\sum \blacksquare$	Summation
$\mathbb{R}^d$	d-dimensional vector in Euclidian space of real numbers
$'$	Transpose operator
$\mathbb{Z}_t$	Set of retrospective subsequence samples starting at time t
α	Relative Pearson divergence(PE)
$P(\blacksquare)$	Probability distribution
$f(\blacksquare)$	Convex function

θ	Set of parameters
$J(\blacksquare)$	Cost function
λ	Regularization parameter
$g(\blacksquare)$	Density ratio
$\hat{g}(\blacksquare)$	Density ratio estimator
d	Dimension of the Euclidian space / state-space/ dimensionality of vectored time series
SS	Support Set

ABSTRACT

This thesis develops the research focus for the System's health monitoring DHPA project being undertaken at Sheffield Control and Systems Engineering UTC. The research aims to develop a framework for extracting and maximizing the information in the measured data for system health monitoring/analysis.

This research project focuses on the development of a methodology to extract as much as information (or features) from measured data including transient signals obtained during test rig testing, normal operation (e.g. typical flight phases), transient manoeuvres or a specific functional test protocol to diagnose (and if feasible use for prognosis) the condition of a system accessories/sub-systems or components.

The following thesis briefly describes the problem formulation, the motivation behind this research. A broad overview of the possible areas for research as well as significance of proposed research is presented and the fundamental issues related to health monitoring of any complex system are discussed. This work has been divided in to three different parts namely framework design for system health monitoring, feature/information extraction for system health monitoring and design of a change detection algorithm.

There exist lots of technological gaps in existing state of the art architecture/framework of Equipment Health Monitoring (EHM) system e.g. present day EHM systems do not utilize the available transient data generated at various stages of the gas turbine engine flight cycle. Hence to fulfil those gaps a concise and generic framework for the system level health monitoring of gas turbine engine has been proposed.

Based on the proposed framework some feature extraction methods have been developed based on information theory and complexity theory. These methods have been applied to extract features from a real data obtained from a test rig of a fuel metering valve.

The performance of a system degrades over time due to deterioration mechanisms and single fault events. While deterioration mechanisms occur gradually, single fault events are characterized by occurring accidentally. Sometime during the normal operation of the system a change/event may occur at system / subsystem or component level sensor signal, which can be due to initiation of an incipient fault which can take a long time to appear in the representative sensor signatures hence it may not be easily detectable using naked eye. Identifying these changes as soon as possible is referred to as change detection.

A trend monitoring algorithm based on spectral entropy has been developed and also applied to the oil debris monitoring problem. This trend monitoring status can act a continuous input to the intelligent equipment health monitoring system. Furthermore, a change detection algorithm based on direct density ratio estimation also has been developed and applied to low pressure turbine data in order to detect change point and include this information in to the prognosis process. This detected change point can also acts as triggering point for data acquisition a higher rate as well as for longer time period. The data acquired higher sampling rate and or longer time can be used further better fault diagnosis and prognosis.

TABLE OF CONTENTS

INTRODUCTION

1.1 Background

With the advanced technological developments, a large amount of complex and expensive machinery is operated in today's world and it is also expected to meet very high demands for productivity, efficiency and quality. Due to these high expectations the traditional preventive and corrective maintenance approaches have been found insufficient to meet these goals, therefore a more efficient maintenance strategy is needed to handle today's ever growing and demanding situation.

As the complexity of industrial systems increases, fault diagnosis and failure prognosis become more and more important, since they are crucial means to maintain system safety and high reliability. Hence, the desire for reliable systems is a major objective for any production houses including operators and manufacturers. Consequently, the maintenance of critical machinery is always a major expense for the organizations and an essential activity of administration and operation.

In general, a fault refers to an abnormal condition that may lead to reduction or loss of the capability of a system or its component to perform a required function at a pre-specified level of efficiency and reliability. On the other hand, a failure means the inability of a system or its component to perform its required functions within specified performance requirements. A loosen belt is an example of fault in a mechanical system where the belt is still working but the transmission efficiency is decreased. However, a broken belt is a failure since the belt is not working anymore and must be replaced.

System health monitoring is a key feature for failure prevention and Condition Based Maintenance (CBM). A health monitoring system needs to detect a fault or failure in a timely manner so that and the faulty components can be replaced effectively to ensure system's normal operations. In the last few decades, various maintenance strategies have evolved such as from reactive maintenance, to age-based preventive maintenance, then evolving further to the modern condition-based maintenance strategies. Reactive maintenance is usually performed after the system breakdown and is not able to operate anymore. In order to deal with system shut down and also to prevent catastrophic failures, which can cause emergency shutdowns, an age-based preventive maintenance strategies was introduced. In this policy, the health check of the system was carried out based on system's operating time regardless of the health condition of the system. Actually, an Age-based preventive maintenance may sometimes reduce unexpected failures, but it is not cost effective and one cannot completely eliminate or rule out the possibility or the risk of the major failures. All these different (very) conventional maintenance strategies do not necessarily satisfy the demands of high reliability as well as efficiency of the modern complex engineering systems. Fortunately, CBM is considered as an effective alternative. In this strategy unnecessary maintenance is avoided by only taking maintenance actions, when there is a clear and plausible evidence of abnormality in a monitored system is observed**(Lee, Ni, Djurdjanovic, Qiu, & Liao, 2006; Vachtsevanos, Lewis, Roemer, Hess, & Wu, 2006)**. The monitoring is usually based on combining the information from different sensor measurements and it does not really interrupt the normal operation of the monitored plant. The main idea behind this strategy is to avoid excessive or insufficient maintenance in order to significantly increase the actual system availability or run time for high production activities.

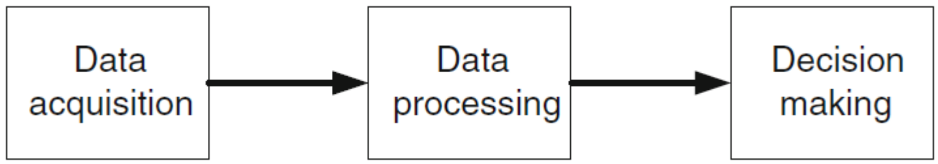


Figure 1: Three steps of CBM **(M. Y. M. Yu, Wang, & Huang, 2010)**

In general, CBM strategy includes three key steps namely: data collection, data processing and decision-making **(M. Yu, Wang, Luo, & Huang, 2011; M. Y. M. Yu, Wang, & Huang, 2010)**.These steps are shown in the Figure 1. Data collection step involves obtaining the data related to system condition/health from the sensors installed at various levels of system's hierarchy. Data processing step is about pre-processing, handling and analysing the data or signals collected for better understanding and interpretation of system health or condition. Finally, the purpose of the decision-making step is to recommend or develop the efficient maintenance strategies based on the analysis performed in the previous step. As mentioned before, Fault diagnosis and prognosis are two critical steps/factors in any modern CBM, and they are complementary tasks. Diagnosis is actually a "static" indicator whereas failure prognosis is more a "dynamic" indicator. The main objective of the diagnosis step is to indicate whether or not a fault/anomaly has occurred and at the same time provide some useful information about the severity or extent of the fault/anomaly **(Samantaray & Bouamama, 2010)**. Prognosis step tries to track fault degradation process (by modelling using data driven or physics based approach) and predict as accurately as possible the Remaining Useful Life (RUL) of a faulty component or subsystem. Prognosis in CBM has received more attention in the recent past for different applications and has become a hot topic of research for various industrial segments. Actually, prognosis is often much more efficient than the diagnosis to achieve zero-downtime performance, which is very important when a delayed information about the possible failure can be catastrophic in some applications such as (e.g., helicopter gearbox and nuclear power plant etc.). But in order to do an accurate prognosis, the status of system's health/condition must be known as accurately as possible. Therefore, the monitoring of system's health is very crucial for any CBM.

1.2 Motivation

Many engineering systems, including gas turbine engines, rely on operational data to monitor the systems health. The operational data is derived predominantly from steady-state conditions at a predetermined time or operational condition. Such a strategy, in the gas turbine engine case, is successful in detecting a significant number of faults. Difficulties detecting the remaining faults relate to the steady-state data not containing characteristics of these faults. The location of the fault is also an important factor to correct monitoring of the system's health, hence continuous monitoring of subsystem's / component's health is also required.

1.2.1 Aerospace gas turbine engine maintenance

For many of the world's largest manufacturers, aftermarket service and parts operations essentially define a significant part of their business revenue. According to the statistics, after-sales services and parts contribute only 25% of revenues across all manufacturing companies but are responsible for 40-50% of profits. Table 1.1 shows that the aerospace and defense business accounts for about 47% percent of revenue, the largest in comparison to other global industries (**Koudal, 2006**). Engine manufacturers, e.g. General Electric, Pratt & Whitney and Rolls-Royce, all have performance-based contracts with commercial airlines in which their compensation is tied to product availability (hours flown) **(S.-H. Kim, Cohen, &**

Netessine, 2007); **(Marinai, Probert, & Singh, 2004**). Services, such as TotalCarer and power by the hour arrangements, are now regarded as an essential element of delivering asset operation **(King, Bannister, Clifton, & Tarassenko, 2009**).

Global industry	Share of service and parts business in overall sales
Aerospace and defence	47%
Automotive and commercial vehicle	37%
Diversified manufacturing and industrial products	20%
High technology and telecommunications equipment	19%
Life science/medical devices	21 %
All companies	26%

Table 1: Revenue impact of service and parts business by global industry (**Koudal, 2006**)

The economic impact of such service contracts is significant. For example, Rolls-Royce, one of the world's largest jet engine and gas turbine makers, has more than 14,000 aerospace engines in service, operated by more than 500 airlines and powering more than 5.5 million commercial flights per years (**OSyS, 2014**). The considerable number of the engines to be maintained, in terms of service and providing proper spare-parts, enable this company to generate revenue about 55% of the more than US$11 billion in total revenues (**Rolls-Royce, 2014**).This evidence emphasizes the significant benefits of applying as well as developing health monitoring techniques for civil aerospace gas turbine engines, which is one of the effective ways to reduce life cycle costs, improve engine reliability as well as availability **(Y. G. Li & Nilkitsaranont, 2009; Marinai et al., 2004**). Figure 2 illustrates the benefits of health monitoring in civil aerospace industry, based on (**Leao, Fitzgibbon, Puttini, & de Melo, 2008**). Aircrafts are highly valuable assets and large budgets are spent in aircraft support, maintenance and logistics. The application of health monitoring technologies in civil aerospace can potentially yield profits to commercial aircraft operators.

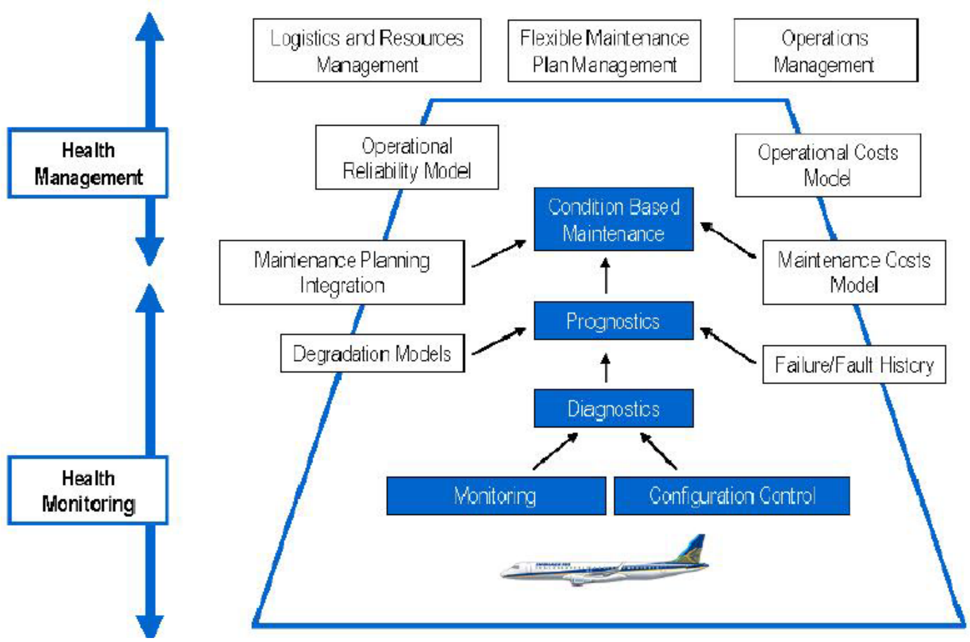


Figure 2: Health monitoring benefits in civil aerospace industry **(Leao et al., 2008**)

1.2.2 Gas Turbine Engine Principle

The concept of the gas turbine has been acknowledged to an English coalmaster and inventor, named John Barber (1734-1801), who patented his idea about gas turbine in 1791 **(Davey, 2003**). He established the basic principle of a gas turbine engine, despite the lack of technology at that time **(P. G. Hill & Peterson, 1965**). Material, design and manufacturing techniques needed to put this principle into a working machine were not fully available until the early parts of the 20th century. The first patent was granted to *Frank Whittle (1907-1996)* in 1930 for using a gas turbine to produce a propulsive jet (**Rolls-Royce plc, 1996**).

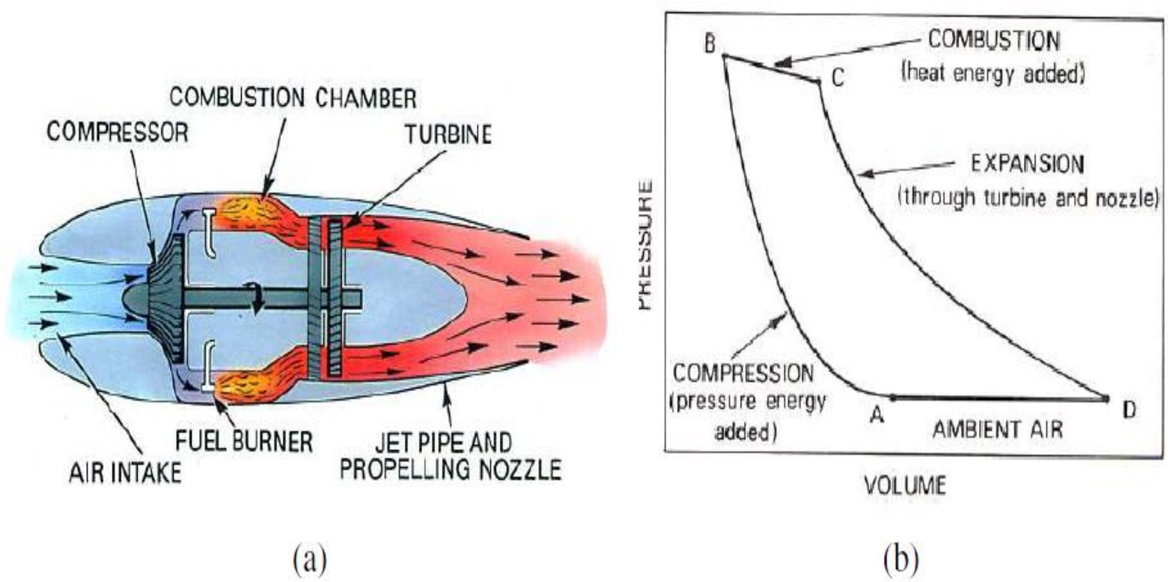


Figure 3: A Whittle-type turbo-jet engine schematic and its working cycle (**Rolls-Royce plc, 1996**)

Gas turbine engines are widely used in different fields to generate energy. They are also very commonly used in aircraft. A gas turbine engine is essentially a power plant which utilises air as a working fluid to produce power in the form of thrust, shaft-power or compressed air. Figure 3(a) shows a Whittle-type turbo-jet engine (**Rolls-Royce plc, 1996**).

It can be seen that there are several internal sections inside, including inlet section, compressor section, combustor section and turbine section. In addition, Figure 3(b) illustrates the working cycle on a pressure-volume diagram. The engine cycles show that in each instance there is induction, compression, combustion and exhaust. In the inlet section, the air intake and fan directs the air into the engine (point A), then the compressor compresses the air to a high pressure (point A to B). After the compressor section, the high-pressure air is directed into the combustion chamber, where fuel (kerosene) is spread as small particles and burned at high temperature and constant pressure, thereby considerably increasing the volume and velocity of air (point B to C). The high velocity air is then directed towards the turbine and driving it using the kinetic energy from the high-speed gas. A portion of the high-velocity air is expanded through the exit nozzle, producing thrust (point C to D) **(Rolls-Royce plc, 1996**).

When gas turbine engines are run, they become fouled with airborne contaminants such as oil, pollen, soot, unburned fuel, soils and salt which encrust compressor components **(Rainer Kurz & Brun, 2007**). Therefore, gas turbine engines show the effects of damage and deterioration in its lifetime of service. The degradation of an engine has an adverse effect on the engine's overall performance (**Khani et al., 2012**). Various factors affect degradation in gas turbine engine performance (**R. Kurz & Brun, 2001**): including Dust/dirt ingestion and further accumulation on fan blades/compressor air foils. Increased air seal, compressor and turbine blade-tip clearances because of rub other mechanisms such as erosion of air foils and seals, hot section oxidation, foreign object damage.

1.2.3 Gas Turbine Sensor and Monitoring Systems

To reduce maintenance cost and avoid service disruption, equipment health monitoring (EHM) has been employed in modern gas turbine engines. EHM is a pro-active technique for predicting when something might go wrong (prognostics) and preventing a potential threat before it has a chance to develop into a real problem, e.g. fault. EHM can be used to estimate the health of thousands of engines operating worldwide, using on-board sensors and live satellite feeds (**Nick Waters, 2009**).

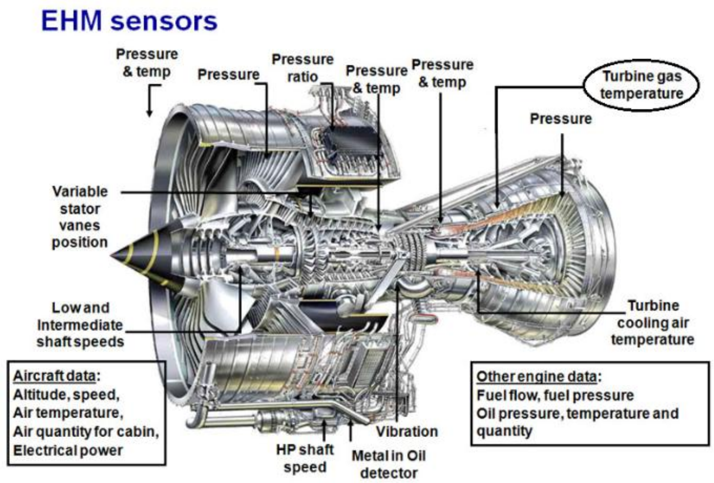


Figure 4: The location of the EHM sensors on the Rolls-Royce Trent 900 engine(**Rolls-Royce, 2014**)

There are several sensors fitted to monitor critical engine characteristics, such as temperatures, pressures, speeds, flows and vibration levels, to ensure they are within acceptable tolerances and to highlight when they are not. Figure 4 shows the typical parameters measured on the Rolls-Royce Trent 900 engine (**Rolls-Royce, 2014**).

1.3 Outline & contribution of the thesis

The key premise for this thesis is that additional fault characteristics can be detected when more informative data can be gathered by utilizing various methodologies such as exciting the systems using suitable transient manoeuvres to produce dynamic signals and/or by injecting an additional signal to produce additional dynamic signals and acquiring the data intelligently to capture maximal health information. An important element of this thesis is to design a generic framework dealing with various issues such as data generation, data collection, communication and decision making, for the health monitoring of complex engineering systems such as e.g. civil gas turbine engine. Other major focus is on the extraction of indicative information/features from system dynamic characteristics without interfering with the functionality of the system. Another generic but very important aspects of trend monitoring and change point detection (change in system's state or fault) for system health monitoring are also tackled in this thesis and later an integrated prognostic methodology combining the change point detection algorithm and Bayesian prognostics is proposed for better remaining useful life calculation of the system.

Chapter 2

Description: This chapter provides an introduction to state-of-art monitoring system and its capabilities of the gas turbine engine. It gives a first-hand description of the hierarchical & modular structure of the gas turbine engine. It points out the challenges and limitations in the current generation health monitoring systems with respect to the technologies involved in data collection, communication and decision making.

Contribution: A comprehensive literature review, which covers all the as aspects such as data generation, data collection, data communication as well as decision making, which are critical towards designing an integrated health monitoring system is done. Use of the transient information (in the various sensor signals as well as system's operation) for the purpose of system's health monitoring is discussed in detail. A new architecture for the next generation equipment health monitoring systems for the gas turbine engine is proposed, which also includes the information from various ancillary/accessory system. A detailed discussion about technological impact and the suitability of the available methodologies, for the use in next generation gas turbine engine of the available methodologies is provided.

Chapter 3

Description: This chapter describes a generic framework for health monitoring of gas turbine engines which takes in to consideration aspects of data collection, communication, compression and decision making etc. A proposal is also made to modify and extend the Rolls-Royce "Sense-Acquire-Transfer-Analyse-Act Paradigm"

Contribution: A systematic framework for the design of health monitoring for inclusion of transient information in accessing the overall health of the system is proposed. Furthermore various suggestions to include new technologies/methodologies to extend the framework are made, so that existing bottlenecks, technological gaps in the existing equipment health monitoring system with respect to data collection, communication, compression and decision making etc. can be addressed in the next generation monitoring systems for the civil gas turbine engines. The Rolls-Royce "Sense-Acquire-Transfer-Analyze-Act Paradigm" is also extended to include the suggested changes proposed the chapter.

Chapter 4

Description: This chapter gives a brief introduction of the challenges associated with the identification of the faulty system in a complex system. A complex system can fail in a multiple ways. In a physical system, most of the faults that are manifested as system-level failures are initiated at the component-level, and a gas turbine engine is no exception. In the engine, there are a large number of components, each of which can have multiple failure modes. Furthermore, each failure mode is a product of many failure mechanisms that are simultaneously active. In short, there can be a large number of failure scenarios in the engine. Ideally, a system-level health assessment methodology should take all these possibilities into consideration. However, in most of the practical cases, it is not possible to cover all of these cases. Hence it is important to choose a representative set of suitable system/subsystems or components which are most relevant in terms of time, cost and effort.

Contribution: In this chapter, a systematic way to select a candidate subsystem based on the criticality of the problem, time, cost and effort is described. In the case of gas turbine engine it is emphasized, how the various factors such as operational mode of the flight, expert or stakeholder's knowledge of high value faults (coming from Failure Mode Mechanism Effects Analysis, FMMEA study), which can contribute to the selection of suitable sub-system for fault investigation can be utilized and combined with already existing knowledge about the working of engine subsystem to select a candidate sub-system. Based on the described approach, Fuel metering valve (FMV) is selected as the candidate subsystem for further investigations.

Chapter 5

Description: In order to design a robust equipment health monitoring system, data selection plays a critical role. For the better decision making of the system's state, suitable sensor signals should be selected and information hidden in those sensor signals must be properly extracted in order to make intelligent decisions. Feature extraction is always a crucial step for information gathering as well as health monitoring of a system. Whenever any change or faults occur, most of the systems always manifest abnormal and sometimes nonlinear dynamic behaviour. Hence it is necessary to extract the features hidden in the sensory signals for more accurate health monitoring and diagnosis. Various methods based on information theory and complexity theory are described in the chapter. These methods were further tested on the real test rig signal of fuel metering valve.

Contribution: In this chapter, feature extraction methods based on Shannon entropy, mutual information, and approximate entropy are proposed and tested on the real-life test-rig designed to imitate the oil debris building up problem in a fuel metering valve of a gas turbine engine. The proposed techniques are found accurate as well as robust enough to distinguish between the healthy and unhealthy system.

Chapter 6
Description: This chapter presents an information theoretic trend monitoring as well as change point detection (CPD) algorithm to solve the generic problems mentioned below encountered in the system lifecycle with hierarchical and modular structures.

- Continuous trend monitoring of the system/sub-system or any other asset health.
- Dealing with irregular events such as information arising from irregular events occurring during the life cycle of an asset and rapid degradation in the state of health parameter.

A simple demonstration case study is performed to test the effectiveness of the spectral entropy based trend monitoring approach. Later on direct-density ratio based concept is applied to two case studies. In the first case study, the CPD algorithm is used to detect the change in a vibration signal (covariate), which affects the slope of degradation. Furthermore, in the second case study, CPD is applied directly to the degradation data to discover when an unknown maintenance event took place. This information directs the main prognostic algorithm to reset the prediction. Based on this information, an integration approach is discussed to track and predict the health of the system.

Contribution: The proposed trend monitoring as well as integrated prognostic approach combining change detection and previously proposed Bayesian technique proves to be promising. These methods demonstrate several advantages e.g. continuous tracking of a health index, utilisation of the available multiple engine data as well as data available at various levels of a system's hierarchy. Trend monitoring methods is able to continuous track, whereas change detection method is able to detect changes or faults in multiple covariates (e.g. vibration, ambient temperature) at any level of the system to then informing main prognostic algorithm to update its belief about degradation. The recovery in degradation data due to maintenance action can also be handled automatically based on integrated prognostic concept.

Chapter 7
Description: Main conclusions drawn from the research work are presented in this chapter.

Chapter 8
Description: In this chapter, concrete suggestions are made about the future work which may be carried out in order to design a fully integrated and robust equipment health monitoring system.

2 Gas Turbine Monitoring Systems

2.1 State-of-art monitoring system

Monitoring systems technologies log the actions, performance and status of the components in the electrical and control systems. They collect data from some sensor signals deemed indicative of performance and mechanical elements of the engine, which are then used to draw certain conclusions, based on algorithms programmed into the monitoring system. The monitoring system can be ground-based, assessing systems flying in the air, floating on the sea, or generating electricity on another continent. The aim of a monitoring system is to maximize availability and minimize operational disruption. Some state of the art capabilities of the present Rolls-Royce Equipment Health Management (EHM) system are mentioned below in Figure 5. The main bottlenecks lie on on-board data compression for data transfer via SatCom as well as capturing data from different parts of engine systems at different bandwidth which eventually result in different action times because data (information) arrives at external service provider or Operations at different times. Some state of the art capabilities of the present Rolls-Royce Equipment Health Management (EHM) systems are mentioned below:

- Performance monitoring – e.g. COMPASS, DAC
- Oil debris
- Engine Accessory Built-In-Test – detection of failure state only (i.e. testability)
- Vibration monitoring

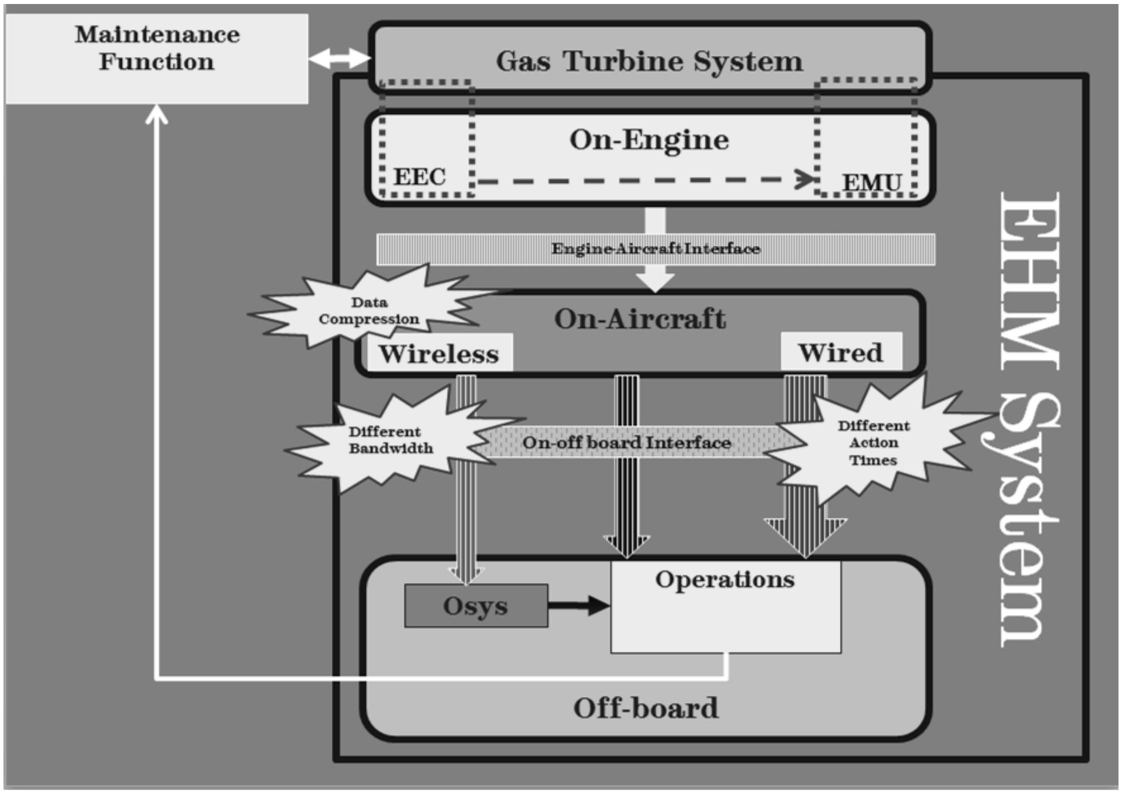


Figure 5: Overview of EHM Architecture

- For the purposes of this report a number of attributes of these EHM systems are apparent.
- Fixed schedule to the tests (i.e. at specific point within the flight profile or engine modes)
- Incipient fault detection is limited to a sub-set of core engine components
- Full system bandwidth captured only for vibration
- Monitoring is ambivalent to system condition / risk, the same fidelity of monitoring is always completed.

The design modern of GTE follows a modular concept. The basic engine by itself is not operable and cannot serve all the functions the airframe depends on. Additionally to its main components the basic engine needs various accessory systems to become an operable engine. Accessory systems are seen as data rich, potentially providing information on both their own condition and the wider system. Hence it is also important to monitor the health of critical auxiliary sub-systems/components attached with the engine in order to improve the overall health monitoring of aerospace engines.

This work seeks to maximize the health state information in the measured data (transient as well as steady state) that is used in analysis. One of the main emphases of this research project is to test hypothesis that utilization the all data available including the use of transient information provides increased prognostic horizon (for GTE accessories / LRUs). The first few questions which come to mind when considering the inclusion of transient information in the design of health monitoring system are explored below as well as the motivation behind the present hypothesis of including transient or dynamic information for the design of health monitoring is also discussed below:

2.2 Masking of fault by controller at steady state

Many complex engineering systems like Gas turbine engines use feedback control systems as a means to achieve desired dynamic and steady-state performance. The main advantage of a simple feedback control system lies in the fact that the feedback control system can perform efficiently even in the presence of un-measurable disturbances such as system faults and degradations. However this ability/utility is in total contrast to the functionality of a system's condition monitoring function, which seeks to observe the current state of the system. There can be cases where intermittent/incipient faults in sensors result in oscillations of the selected value without the steady state deviation being large enough to be classified as a fault.

2.3 Transient information: How good it is?

With advances in technology, machines/systems have become increasingly complicated and there is a need to look at not only stationary machine conditions but also transient machine conditions. Some machine/systems failures, if not most of them, happen during transition periods (such as during machine start-up, machine shutdown, acceleration or deceleration phases etc.) & can sometimes provide information about machine conditions which cannot be revealed from stationary signals. The points below provide a very brief overview of the condition machine/systems can go through during transient conditions.

- Many types of equipment have natural resonant vibration characteristics through which the equipment has to pass on start-up, shutdown, and sometimes even some operating ranges.

- Loads, pressures, temperatures and vibration are changing rapidly, especially during start-ups, shutdowns and rapid speed/acceleration changes e.g. Jerk and jounce.

  It's during these transitions or "transient events" where much can be learned about the health and performance of the machine/turbo-machinery. The following case studies, carried out by different research groups, also included transient information for the purpose of health assessment of a Gas turbine engine.

- In some situations good quality steady state measurements are difficult to obtain, such as from military aircraft engines that operate up to 70% of the total mission time at unsteady conditions,(**G. L. Merrington, 1989**).

- (**White, 1988**) has summarized four main (diagnostic performances) differences between steady state and transient conditions as mentioned below:

  1. "*During a transient condition the shaft inertia will either demand or produce power depending on whether it is accelerated or retarded*".

  2. "*Pressure and temperature gradients during transients cause different mass flows into and out of components depending on the rate of change of the transient*".

  3. "*The heat balance is not satisfied during transient operation. Heat is either transferred to or given out by engine components adjacent to the gas stream. This means that expansion and compression are no longer adiabatic*".

  4. "*During transient operations the properties and dimensions of various turbine components can change. The main reasons for this are the physical properties of the materials which are subject to expansion and strain due to temperature and centrifugal forces. This can detrimentally affect tip clearances and leakage of bleed air flows*".

- An overview of the use of both performance and mechanical transient analysis as a means to detect gas turbine problems and the need for transient analysis and transient analysis techniques has been discussed in **(Meher-Homji, Cyrus B | Bhargava, 1994**).

- More specifically, it has been shown in **(Borguet, Dewallef, & Léonard, 2005**) that the use of measurements representative of transient behaviour significantly improves the diagnosis accuracy. The study compares the diagnostic efficiency of identification of 14 different fault cases by the use steady state data and transient data. Comparison results clearly show the improved diagnostic efficiency by the inclusion of transient data. The study was carried out in the framework of Obidicote project-work package 4: steady state test cases.

- (**G. Merrington, Kwon, Goodwin, & Carlsson, 1991**) showed a marked improvement in fault identification by applying analytical redundancy methods to gas turbine engine transient data. (**G. L. Merrington, 1989**) developed a method for estimating the effects of unmeasured fault parameters from input/output transient measurements and discussed the effects of sampling rate and the measurement noise on resultant sensitivity of the technique. (**G. L. Merrington, 1994**) applied model based technique to the problem of detecting degraded performance in a military turbofan engine from take-off acceleration type transients and established that good fault coverage can be gleaned from the resultant pseudo-steady state gain estimate. **(G. L. Merrington, 1994**).

- **(Eustace, Woodyatt, Merrington, & Runacres, 1994**) compared the fault diagnostic technique based on use of steady-state engine data with technique based on transient data and concluded that for range of faults examined, not only is there similar fault information contained within transient data, but faults can be detected with increasing sensitivity using these data.

- In addition, some gas turbine component faults, such as bearing faults (based on bearing temperature readings) **(Meher-Homji, Cyrus B | Bhargava, 1994**) and mis-scheduled nozzle control during transients (**Merrington, 1988**), contribute little to steady state performance deviation but significant to transient performance change.

Furthermore, performance shift due to engine faults is very likely to be magnified during transients compared to that at corresponding steady state conditions. Even though most of the case studies mentioned above do consider the advantage of including the transient information in to the prognostic/diagnostic framework but unfortunately, some difficulties arise when processing signals acquired during transient events, non-stationary phenomena or when the systems are working in an unsteady operating conditions. Moreover, a loss of effectiveness can be observed for many signal processing techniques **(Chatterton, Pennacchi, Ricci, Borghesani, & Vania, 2013)** in the case of complex systems such as gas turbine engine, in which the measured signals can be a mixture of different sources, and signals are affected by high level of noise. The main difficulty lies in fact that the main transient signal features are often submerged in the background of noise, especially in the early stage of failure development. Enhancement and extraction of these features or components are the main tasks in detecting the defect.

2.4 Available transients in real GTE operation

It is additionally worth noting that a normal operation exposes a Gas Turbine engine to levels of transients above those suggested by idealised flight profiles. Sources of transients may include compensation for wind gusts during manual and auto-throttle control, climbs during flight to achieve different altitudes, taxi behaviour, go-arounds, descent phase reacceleration due to stacking aircraft near to an airport. Figure 6 shows the variation of turbine pressure ratio (TPR) along with the its altitude (ALT) profile in a typical civil aircraft flight cycle.

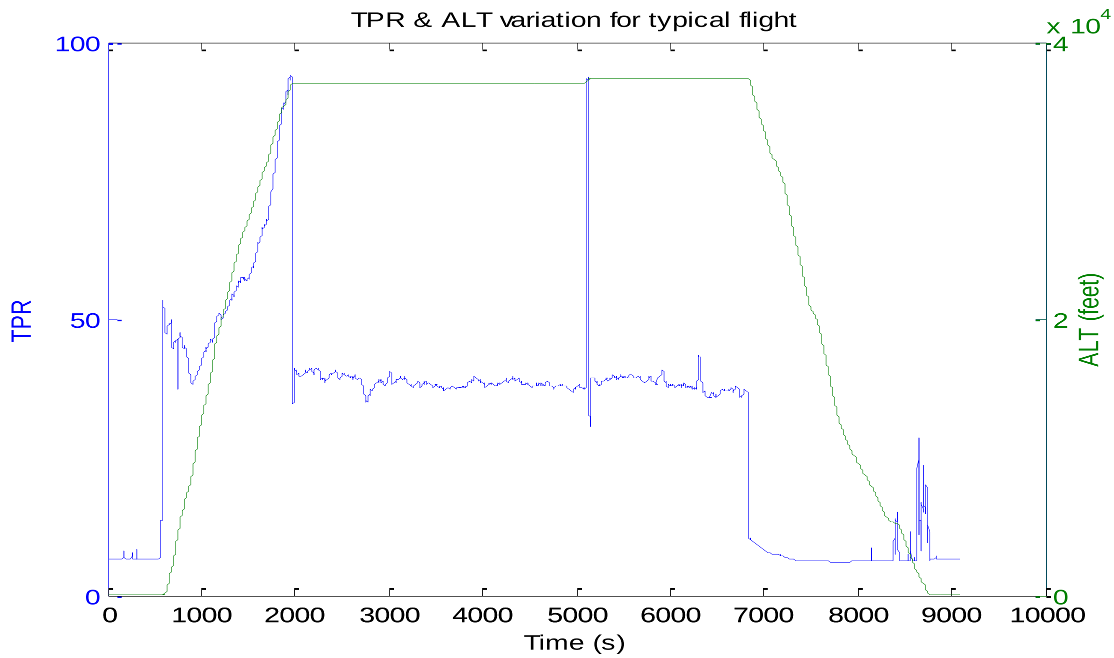


Figure 6: Examples of transients in normal flight profile

2.5 Fundamental Questions One needs to Answer?

There are several influencing factors to maximise the health state information in the measured data that is used in analysis e.g. Where, what, when & how do we measure? The challenges regarding the selection of sufficiently informative measurements or samples of data is a key issue in the design of any health monitoring system. A few questions which one might have to answer while considering the design of any health monitoring system are mentioned below:

- Where do we need to measure? (e.g. engine environment)
- What do we need to measure?

- When do we need to measure? (e.g. specific tests or during operation)
- What is the current EHM capability?
- What are the practical, as well as operational, constraints?
- How to configure measurements? (Sampling rate, data compression etc.)
- What makes it difficult/expensive?
- Is it feasible to get the data to an analyst?

This thesis introduces methodologies sourced from the literature which aims to address the above questions. In the following sections first a framework will be proposed for health monitoring of the gas turbine engine based on transient and steady-state data as well as the techniques and few approaches available to answer/tackle the issues mentioned above will be discussed. It is worth noting that there are many ways to extract maximal information/features from data (signal processing, machine learning concepts etc.) but the upstream operation of obtaining the right data/information/features is the focus here.

2.6 Hierarchical and Modular Structure

Complex engineering system is defined as a group of interrelated, interacting or interdependent constituents (components) forming a complex whole (**Rebovich, 2008**). Many engineering systems are comprised of hundreds or thousands of components. Intermediate groupings, or various levels of subsystems, are necessary to describe or depict these systems manageably. Such an engineering system that requires one or more levels of definition intermediate to system and component is characterized as a complex system in this thesis. Thus, a complex system is a system composed of a number of subsystems, each of which is embodied by a particular set of components, or sub-subsystems.

A turbine engine consists of its main components, which change the state of the gas flow in the sequence of the thermodynamic working cycle. The design of modern turbofan engines follows a modular concept (**Linke-Diesinger, 2008**). In Figure 7 an aircraft engine is presented as an example of a complex system, which comprises of subsystems (e.g., LPC, HPT). Further down the hierarchy, the subsystems are composed of components (e.g., HPT blades).

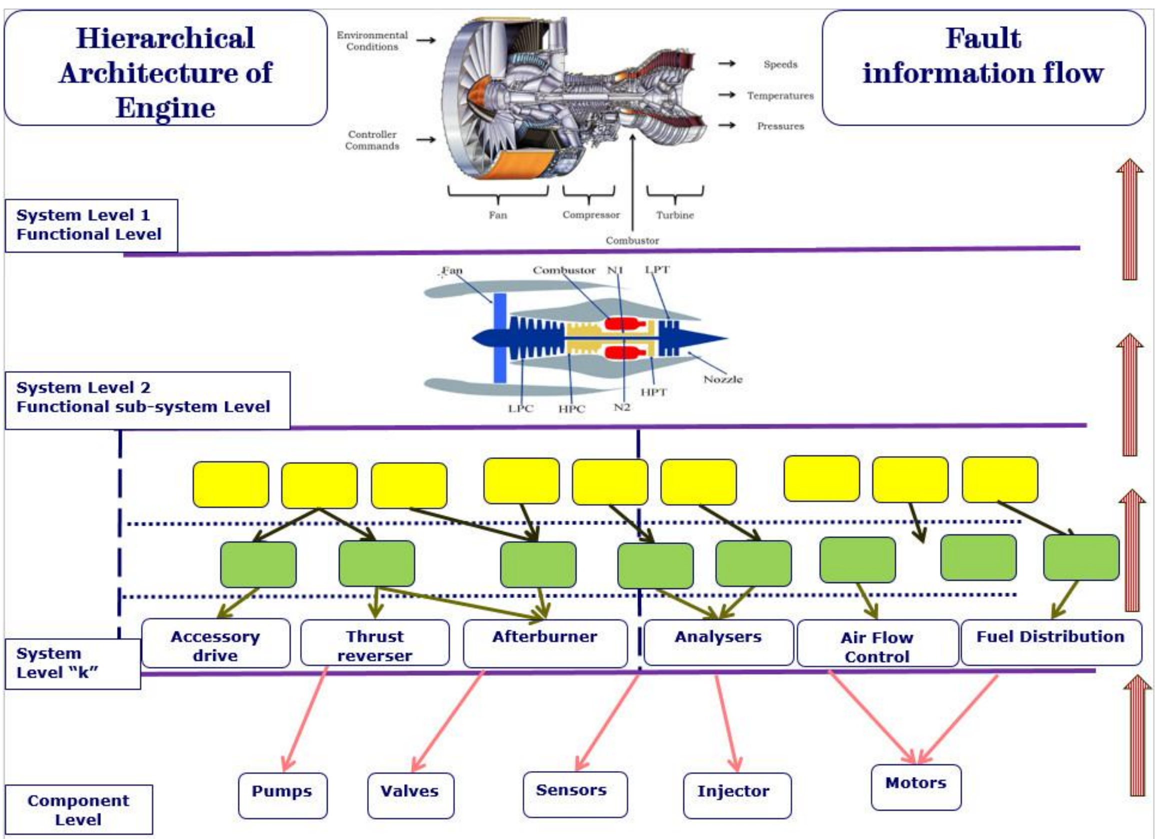


Figure 7: Hierarchical Architecture of an Engine

The basic engine by itself is not operable and cannot serve all the functions the airframe depends on. Additionally to its main components the basic engine needs various systems to become an operable engine.

Furthermore the modular structure (**Manzar abbas, 2009**) of the gas turbine engine can further be extended as shown in Figure 8.

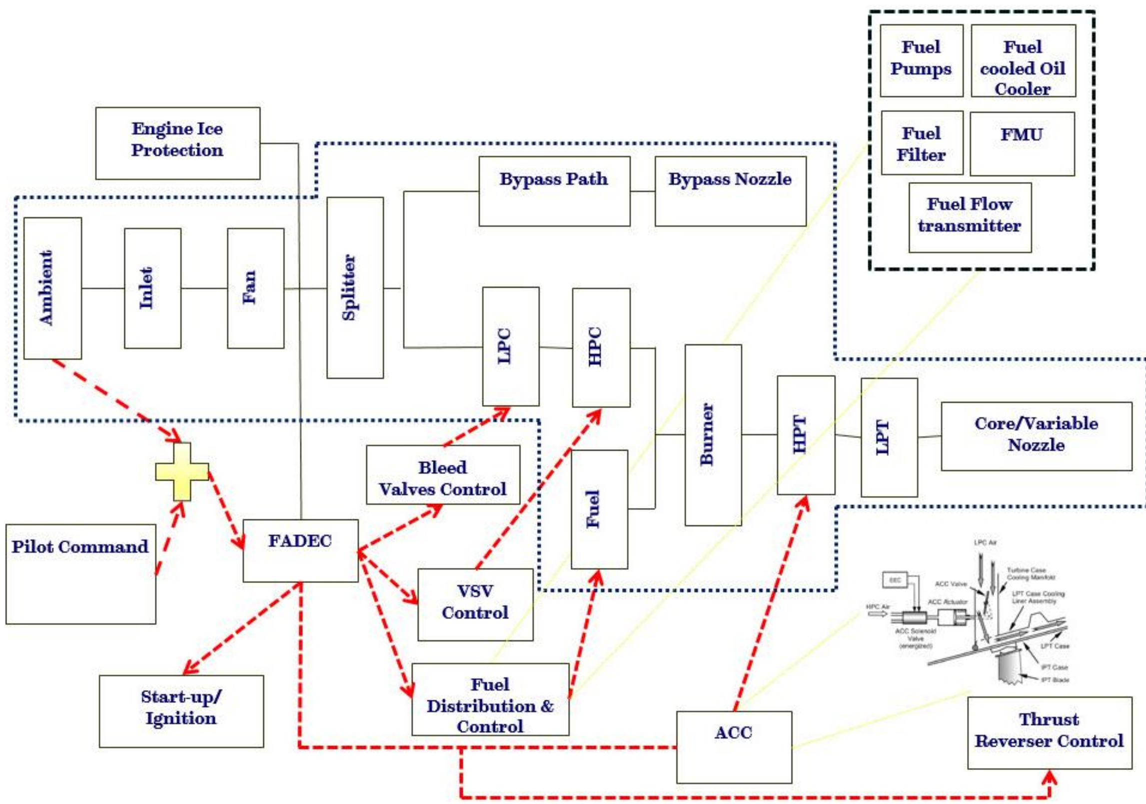


Figure 8: Engine Modules & Accessories

While faults arise at component level, sensing capabilities are limited to subsystem level, and system operations and maintenance practices are scheduled based on system level parameters. As mentioned in the section above, most state of the art EHM available in the market do not consider the inclusion of state of the any subsystem or component individually. This thesis proposes the inclusion of accessories or sub-system health in to the overall health assessment of the engine as shown in the Figure 9.

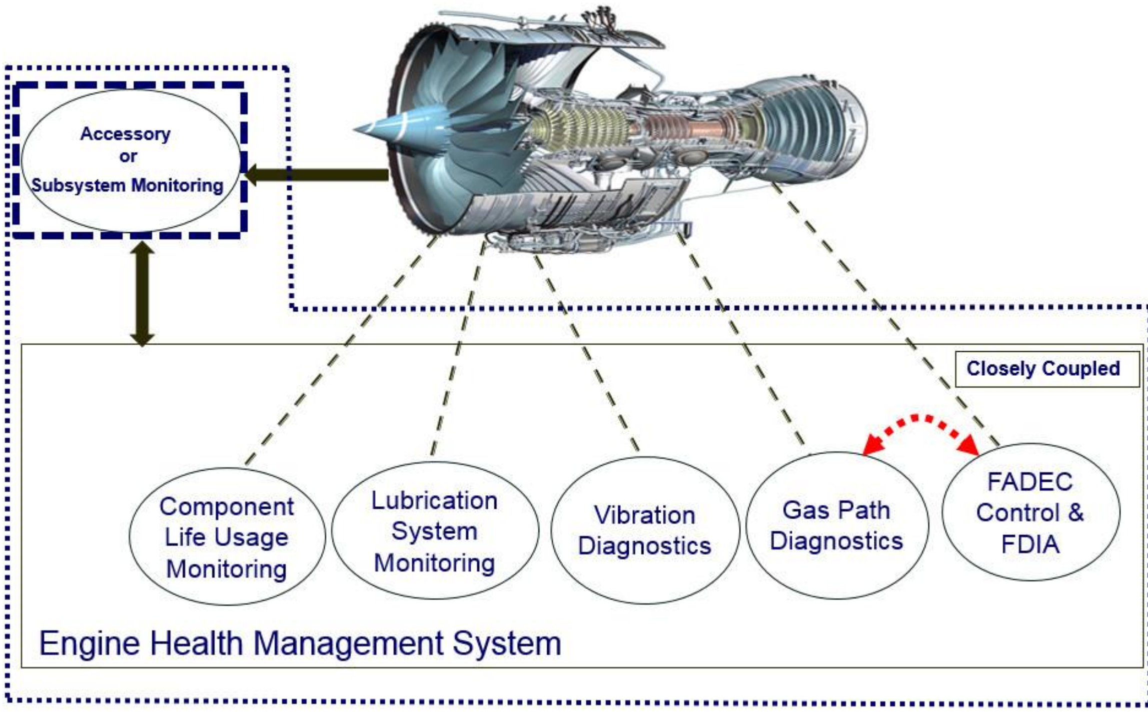


Figure 9: Proposed EHM

2.7 Factors affecting design of integrated health monitoring system

In order to design an efficient integrated health monitoring system, one needs to address the issues of data collection, data selection as well as data generation. There are many ways in which these issues can be addressed. The following section presents a concise literature survey about the techniques which may help to find answers to some of the fundamental questions associated with design of such an integrated system and may also help to bridge the technological gap.

2.7.1 Data collection and selection

Acquiring a parsimonious data set for health assessment

Often the most time-consuming and costly task in any scientific investigation is the gathering of data. In many cases we have limited resources for collecting such data. Hence, it is particularly valuable to determine ways in which we can make use of these resources as much as possible as. Sufficiently informative data is a key to the success of any analysis:

"The more informative the data, the simpler is the analysis"

Using the literature as a starting point, a number of data collection approaches have been considered. Broadly these are classified as:

Active information fusion, optimal sensor selection & sensor management – Active information fusion is about selectively choosing the sensors so that the information gain can compensate the cost spent in information gathering.

Active data selection - Learning efficiently by actively selecting particularly salient data points.

Active Sensing – can be stated as a problem of controlling strategies applied to the data acquisition.

All of the above mentioned approaches are closely related and deal with the issue of most informative way to select, collect or classify data. Other important factors in data collection are data transfer, data compression and where to collect the data from as well as combining the all the information gathered from number of sources in an efficient as well most informative way. The approaches mentioned below deal with such issues:

Compressive sensing - compressed sensing is the process of acquiring and reconstructing a signal that is supposed to be sparse or compressible. The first two approaches are active research areas in field of machine learning whereas active sensing, compressive sensing and active information fusion, optimal sensor selection, sensor management are active areas of research in the field of signal processing. In the next few sections, the approaches mentioned above will be described very briefly and their relevance this project research will be discussed.

a. Active information fusion, optimal sensor selection & sensor management

For the purpose of information gathering many information fusion applications especially in aerospace and military domains are often characterized as a high degree of complexity due to three challenges:

- Data are often acquired from sensors of different modalities e.g. pressure, fluid flow, temperature and with different degrees of uncertainty
- Decision must be timely - made quickly enough to be useful
- The world (flight as well as load conditions in this present case) situations as well as sensory observations evolve over time.

Multi-sensor management system/active controller or process seeks to manage or coordinate the usage of a suite of sensors or measurement devices in a dynamic, uncertain environment, to improve the performance of data fusion and ultimately that of perception. It is also beneficial to avoid overwhelming storage and computational requirements in a sensor and data rich environment by controlling the data gathering process such that only the truly necessary data are collected and stored. A generic sensor management is responsible for answering questions like:

- Which observation tasks are to be performed and what are their priorities?
- How many sensors are required to meet an information request?
- When are extra sensors to be deployed and in which locations?
- Which sensor sets are to be applied to which tasks?
- What is the action or mode sequence for a particular sensor?
- What parameter values should be selected for the operation of sensors?

A dynamic active information fusion framework tries to simultaneously address the challenges mentioned above. In a simpler way, an active information system is to selectively choose those information sources that are most informative to the problem while minimizing the associated costs in terms of computational complexity, time, and required resources in acquiring the information. Overall efficiency can be achieved by aggregating only a subset of the most relevant sensory data to address current problem. Architecture of the active fusion framework is problem specific. One of the generic representations of the active information fusion architecture is shown in [**Figure 10**].

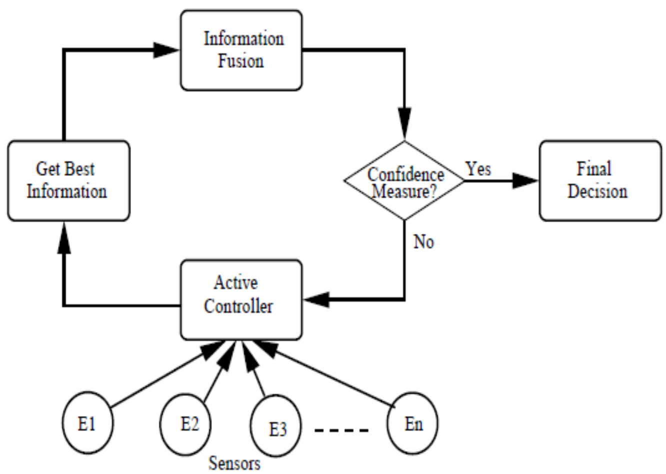


Figure 10: Active Information Fusion (**Looney, 2002**)

A good introduction to data fusion and mathematical techniques can be found in (**D.L. Hall & Llinas, 1997; David L. Hall & McMullen, 2004; Liggins, Hall, & Llinas, 2008**). **(Xiong & Svensson, 2002**) give a very good survey of the issues and approaches related to multi-sensor management for information fusion. A key to success for any active information fusion framework is proper sensor selection or management of different sensor so that the information gain can compensate the cost spent in information gathering. There are numerous applications of sensor selection including computer vision, control systems and sensor networks etc. which can be found in **(Denzler & Brown, 2000, 2002a, 2002b**); (**Miller & Runggaldier, 1997**). **(Y. Zhang & Ji, 2005, 2010**) discuss the issue of sensor selection in active information fusion (**Looney, 2002**) framework.

(**Rowaihy et al., 2007**) presents a survey of sensor selection schemes in wireless sensor networks. **(Yongmian Zhang and Qiang Ji, 2006**) applied the concept of Dynamic Bayesian Network for active and dynamic information fusion. **(Roemer, Kacprzynski, & Orsagh, 2001**) provides an assessment of data and knowledge fusion strategies for prognostics and health management.

b. Active data selection

For any scientific analysis, informative data is the most important factor for generating sensing results as well as validating hypothesis. The main question we are trying to answer is

How do we select the data?

Most theories for data modelling often assume that the data are provided by a source that we do not control. However, there are two scenarios in which active selection of data can be useful.

- Data measurements are relatively expensive or time-consuming (e.g. bore scope),
- We can determine where to look 'next' [Active Sensing] so as to learn [Active Learning] as much as possible.

c. Active sensing

Active sensing is a large field aimed at providing systems with tools and methods for decision making under uncertainty, e.g. in a changing environment and lack of sufficient information. This has been extensively researched in the field of robotics. Active sensing can be thought both in terms of hardware sensor capabilities where sensors send signals (e.g. ultrasonic waves or electrical signals) to environment and receive response by itself as well as for monitoring strategies for gathering most informative sensory data from environment. Here we are concerned with the latter case. Some of the aspects of active sensing incorporate the following:

- How to make decisions for next actions in order to extract maximum information from the sensor data.
- Minimize costs such as operational effects and energy.

Hence, Active Sensing can be stated as a problem of controlling strategies/policies/actions applied to the data acquisition process which will depend on the current state of the data interpretation and the goal or the task of the process. An action is a particular kind of event leading to a change in system state or in the state of the world. The states capture all the information relevant to the system decision-making process. The field of active sensing is closely related to the work in experiment design (**Lindley.D.V, 1956**); (**Fedorov, 1972**), active learning (David J. C. MacKay, 1992); (**Seung, Opper, & Sompolinsky, 1992**) and sensor placement (**A Krause, Guestrin, Gupta, & Kleinberg, 2006)**, (**Andreas Krause, Singh, & Guestrin, 2008**). The active sensing, active perception paradigm is introduced by **(Bajcsy, 1985, 1988**), (**Aloimonos, Weiss, & Bandyopadhyay, 1988**). **(Mihaylova, Lefebvre, Bruyninckx, Gadeyne, & Schutter, 2002**) provide a survey of active sensing in robotics. **(Kreucher, Kastella, & Hero III, 2005**) presented a sensor management approach using the active sensing principle. **(Denzler & Brown, 2000, 2002b**) discussed an information theoretic approach to data and parameter selection. (**S. Liu & Holloway, 2002**) explained the general principle of active sensing for stochastic system.

d. Compressive/Distilled sensing

Conventional approaches to sampling signals or images follow Shannon's celebrated theorem: the sampling rate must be at least twice the maximum frequency present in the signal (the Nyquist rate). Compressed sensing (CS) is a novel sensing/sampling paradigm that goes against the common wisdom in data acquisition. CS theory suggests that one can recover certain signals and images from far fewer samples or measurements than traditional methods use.

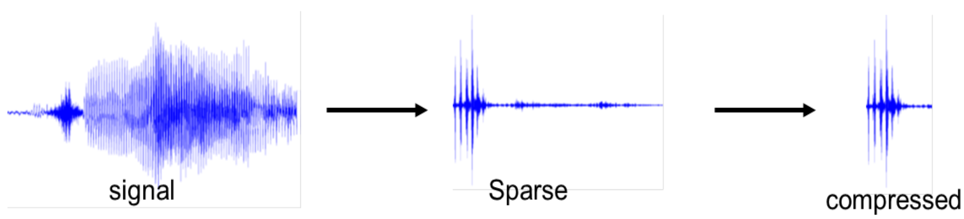


Figure 11: Basic Idea of Compressed Sensing

The key idea of compressed sensing(CS) is given by **(Baraniuk, 2007; Candes & Wakin, 2008; Donoho, 2006**). CS relies on the empirical observation that many types of signals or images can be well-approximated by a sparse expansion in terms of a suitable basis, that is, by only a small number of non-zero coefficients. This is the key to the efficiency of many lossy compression techniques such as JPEG, MP3 etc.

A compression is obtained by simply storing only the largest basis coefficients. When reconstructing the signal the non-stored coefficients are simply set to zero. This is certainly a reasonable strategy when full information of the signal is available. However, when the signal first has to be acquired by a somewhat costly, lengthy or otherwise difficult measurement (sensing) procedure, this seems to be a waste of resources.

First, large efforts are spent in order to obtain full information on the signal, and afterwards most of the information is thrown away at the compression stage. One might ask whether there is a clever way of obtaining the compressed version of the signal more directly, by taking only a small number of measurements of the signal. It is not obvious at all whether this is possible since measuring directly the large coefficients requires knowing a priori their location. Compressive sensing provides nevertheless a way of reconstructing a compressed version of the original signal by taking only a small amount of linear and non-adaptive measurements.

The theory of compressive sensing uses a lot of tools from probability theory. Another important feature of compressive sensing that practical reconstruction can be performed by using efficient algorithms. Since the interest is in the vastly under-sampled case, the linear system describing the measurements is underdetermined and therefore has infinitely many solutions. The key idea is that the sparsity helps in isolating the original signal. The theoretical foundation of compressed sensing has links to and also explores methodologies from various other fields such as, for example, applied harmonic analysis, frame theory, geometric functional analysis, numerical linear algebra, optimization theory, and random matrix theory.

2.7.2 Data generation

How do we generate data in order to get maximum information about the system?

Data generation deals with the fields of design of experiments, dynamic system identification and statistics to name the most relevant. Design of experiment is a fundamental process in every research field and equivalently has practical applicability in various areas specifically in engineering science, medical sciences and social sciences. For this research, a significant amount of literature has been explored and a few relevant approaches are described below:

Experimental design – deals with issues such as choice/design input and sampling rate selection for system identification. A specific area under experimental design most relevant to this research is the design of an optimal input signal for system identification and model discrimination.

The next few sections describe very briefly each of the approaches mentioned above and their relevance to this research.

a. Experimental design

System identification deals with constructing mathematical models of dynamical systems from observed input/output data through an experiment. An experiment here is a process or study that results in the collection of data. The results of experiments are not known in advance. Usually, statistical experiments are conducted in situations in which researchers can manipulate the conditions of the experiment and can control the factors that are irrelevant to the research objectives.

Experimental design is the process of planning a study to meet specified objectives. Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient sample size and power are available to answer the research questions of interest as clearly and efficiently as possible.

In order to obtain the maximal information from the observation data, the idea of optimal experimental design can originally be traced back in early statistics literature (**Wald, 1943**) ;**(Cox, 1958**); (**Fedorov, 1972**); (**Whittle, 1973**); (**P.Wynn, 1972**) as well as in the engineering literature (**Levadi, 1966**);(**Gagliardi, 1967**); (**Goodwin, G., Payne, 1973**); **(Goodwin, G., Payne, R., Murdoch, 1973**);(**Arimoto & Kimura, 1971**), (**Mehra, 1974a, 1974b**); (**Goodwin & Payne, 1977**); **(M. B. Zarrop, 1974, 1979; M. Zarrop, 1979**); (**Hildebrand & Gevers, 2003**). A recent survey is contained in (**Gevers, 2005**) where many additional references can be found.

b. Optimal Input design

The purpose of the design of identification experiments is to make the collected data maximally informative with respect to the intended use of the model, subject to constraints that might be at hand. Optimal input design for system identification was an active area of research in the 1970's, with different quality measures of the identified model being used for this optimal design Over several decades, a large body of literature has developed on the topic of optimal input design (see, e.g., **(Mehra, 1974a, 1974b**), (**M. B. Zarrop, 1979**);(**Ljung, 1999, 2010**), (**Jansson & Hjalmarsson, 2005**), (**Bombois & Gilson, 2006**), (**Rivera, Lee, Mittelmann, & Braun, 2009; Steenis & Rivera, 2009**) and references therein.

2.7.3 Fault Diagnosis and learning

There are various approaches to fault detection and diagnosis. Each approach has its own strengths and weaknesses. In most of the practical applications, multiple approaches are combined to design a suitable fault detection methodology. There are two main categories: Passive and Active. The passive approach where the detector monitors the inputs and the outputs of the system and decides whether (and if possible what kind of) a fault has occurred. This is done by comparing the measured input - output behaviour with normal behaviour of the system. The passive approach is used to continuously monitor the system in particular when the detector has no way of acting upon the system, for material or security reasons, whereas active approach relies on the fact that an external signal can be modulated or send to system to extract important information. In this section, we highlight some of the major differentiating factors between the passive and active technique. In this section first a short introduction to passive approaches is given and later the active approaches are discussed.

a. Passive Fault Detection Approach

Traditional passive fault detection is well established and present in our in everyday life, like e.g. warning lights and sounds, fire and smoke alarms etc. As described above, the main idea is to observe (measure or estimate) certain important parameters or states (observable or estimated) of the system, and if those parameters or states deviate away from the predefined expected values (thresholds), the event would be declared as a fault. Thresholds are typically chosen using a simple probabilistic approach to the system, i.e. if a parameter is out of predefined specific bounds, there is a small probability that the observed part of the system is operating in healthy conditions.

There are numerous references on research and implementation of this approach **(M Blanke, B, & Lunau, 1997; M. Blanke, Staroswiecki, & Wu, 2001; Mogens Blanke, Lunze, Kinnaert, Staroswiecki, & Schröder, 2006; Ron J. Patton, Frank, & Clark, 2000). (Gerfler, 1997; P. Zhang & Ding, 2008)** described a fault diagnosis approach in which the diagnosis is only based on already existing signals in the system such as e.g., the control input, the disturbance noise and the measurement output. As a consequence of this, the faults in the system can only be detected / isolated, when either it is excited by either the control input or the disturbance noise. In other words, the control inputs or noise should be persistently exciting in order to excite the system in same frequency band in which the fault is present. **(Isermann & Ballé, 1997; Isermann, 1994, 1995; R.J. Patton & Chen, 1997).** Most of the work in the area of fault detection is geared towards this type of approach **(Basseville, 1988; Benveniste & Basseville, 1984; Chen & Patton, 1998; R. Patton, Clark, & Frank, 1989; Ron J. Patton et al., 2000; A S Willsky, 1976; Alan S Willsky, 1986)**. This kind of approach will not in general give an optimal fault diagnosis of the system. The issues and drawbacks of this approach are discussed below.

b. Issues with passive approaches

- Complexity of the system can make observation of all parameters impossible such as e.g. in the gas turbine engine, observation of all system parameter at all, level of system hierarchy is difficult.
- The requirement of continuous observation of parameters itself increases the complexity of the system as more sensors and alarm mechanisms to perform observations are required and a fault in these additional equipment can destroy passive fault detection scheme.
- One of the most desired requirements of today's complex systems is robustness. Robustness can be achieved by the use of parallel or redundant modules, systems of observers and controllers, specially designed feedback loops, etc. Inclusion of such additional requirements and modules can

have an undesired effect of straining one part of the system, if something is wrong with another part of the system, while at the same time, observed parameters might remain within the pre-defined range. This kind of additional strain can lead to the more disastrous collapse of the system in the future; therefore it presents potential danger, and can lead to reduced operational time of the equipment.

In fact, what is common for all passive methods is that they are not acting on the system. Instead, if the system is excited by let us say some auxiliary signals, then the detection and isolation of faults can be done in a much more systematic and efficient way. As discussed by many researchers in many cases, the fault diagnosis will also be much faster**(H Niemann, 2006; Henrik Niemann, 2006; Poulsen & Niemann, 2008)**. The section below, gives an idea about the active diagnosis approach.

c. Active diagnosis

Active diagnosis – The active approach to failure detection consists in acting upon the system on a periodic basis or at critical times using a test signal in order to exhibit abnormal behaviours which would otherwise remain undetected during normal operation. All of the above mentioned approaches are closely related and deal with the issue of most informative way to select or collect or classify data. Other important factors in data collection are data transfer, data compression and where to collect the data from as well as combining the all the information gathered from number of sources in an efficient and most informative way. The approaches mentioned below deal with such issues:

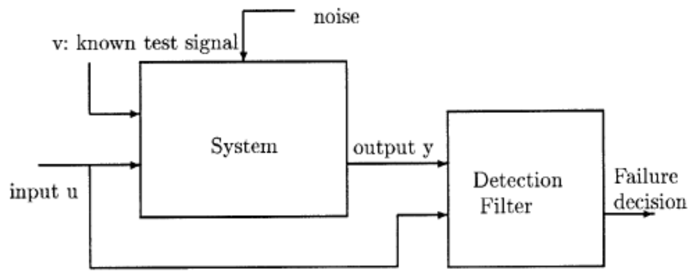


Figure 12: Active Diagnosis(**R Nikoukhah & Campbell, 2002**)

The idea of injecting a signal into the system for identification purposes, that is, to determine the values of various physical parameters, has been widely used and is a fundamental part of engineering design. The design of test signals has been a major issue in system identification, but their use in failure detection has been introduced in (**X. J. Zhang, 1989**); **(Keresteciogˇlu & Zarrop, 1994; Kerestecioglu, 1993**) and **(Uosaki, Tanaka, & Sugiyama, 1984**). The basic idea behind active [**Figure 12**] fault diagnosis (**Ramine Nikoukhah, 1998**) is that, at some point of time (scheduled intervals) during normal operation of the system, a specific test signal is injected into the system for a finite period of time. This signal is supposed to expose the failure modes of the system which are then detected by the detection filter. The idea of active fault diagnosis has further been explored by different groups in different contexts such as in **(R Nikoukhah, Campbell, & Delebecque, 2000**) perturbations are modelled as bounded energy signals, where the aim is to find minimum energy auxiliary input signal which enables to decide surely in which mode the process operates.

Robust multimodal identification problem has been explored by (**R. Nikoukhah, Campbell, Horton, & Delebecque, 2002**); with priori information by **(Ramine Nikoukhah & Campbell, 2006**); for nonlinear systems by **(Andjelkovic, Sweetingham, & Campbell, 2008**); for closed loop systems by **(Esna Ashari, Nikoukhah, &**

Campbell, 2009); for detection of incipient faults in (**Fair, Campbell, & Carolina, 2009**) and **(R. Nikoukhah, Campbell, & Drake, 2010**). **(Esna Ashari, Nikoukhah, & Campbell, 2011**) studied the design of auxiliary signal for active fault detection for linear discrete-time systems whereas **(Esna Ashari, Nikoukhah, & Campbell, 2012**) studied the effect of feedback system on active fault detection.(**H Niemann, 2005**) used the new approach of active fault diagnosis in connection with Fault Tolerant Control. (**Henrik Niemann & Kjølstad, 2008**) applied active fault diagnosis (AFD) approach for fault detection and isolation (FDI) of parametric faults in dynamic systems. The fault diagnosis performance was investigated with respect to different information levels from the external inputs to the systems.

A controller reconfiguration based approach rather than an exogenous excitation signal for active fault diagnosis for additive or parametric faults has been proposed in **(Stoustrup & Niemann, 2010**). (**Šimandl & Herejt, 2003**) deals with the information processing strategies and considers their influence on fault detection performance index. (**Simandl, 2005**) extends the idea presented in (**Šimandl & Herejt, 2003**) where information about the auxiliary signal generator was given a priori and depended on decision in known manner, to more general case for which the generator of the auxiliary input is also searched.

Since the inputs of a system are usually utilized for control, it is necessary to tackle the problem of simultaneous control and active detection. Early works **(Stoustrup, Grimble, & Niemann, 1997**), (**Jamouli, 2003**) studied the relationships and limitations of the integrated fault detection and control, and more detailed treatment was given in **(Khosrowjerdi, Nikoukhah, & Safari-Shad, 2004**), (**H Niemann, 2006**) and **(Blackmore & Williams, 2006**). (**Siroky, Simandl, Axehill, & Puncochar, 2011**) presents an optimization approach to resolve the competing aims of active fault detection and control.

d. Optimal Input design for model discrimination/ fault diagnosis

In order to estimate (identify) the parameters of a given model, one may have to choose the proper model structure among a set of available candidate models, which may correspond, for instance, to competing scientific hypotheses about the description of some phenomenon e.g. nominal and fault mode operation of any system. Choosing between model structures is called model discrimination.

In practice, of course, the ability to discriminate distinguishable model structures depends on the informational content of the data collected. This is why optimal experiment design for model discrimination has received some attention in the statistical literature (see, e.g., (**G. E. P. B. and W. J. Hill, 1967**);**(A. C. Atkinson and D. R. Cox, 1974; Atkinson & Fedorov, 1975**); **(Dette & Titoff, 2009**).Applications of experiment design for discrimination can be found in domains as diverse as chemistry **(Alberton, Schwaab, Lobão, & Pinto, 2011; Schwaab et al., 2006**), machine learning (**Rajamoney, 1993**), systems biology **(Kreutz & Timmer, 2009**); (**Skanda & Lebiedz, n.d., 2010**); **(Mélykúti, August, Papachristodoulou, & El-Samad, 2010**) and psychology (**Cavagnaro, Myung, Pitt, & Kujala, 2010; Myung & Pitt, 2009**).

Active approach also has some drawbacks. In such approaches the system is excited by auxiliary signals along with the signals present in the loop. These signals might disturb the performance of the system in the fault free case. Hence, the design of an auxiliary signal is the most important part of modern active fault diagnosis system.

e. Active learning

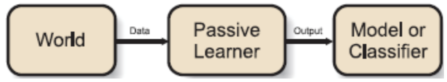


Figure 13: Passive Learner

Figure 14: Active Learner

What is Active Learning?

Active learning – It can be described as a closed-loop phenomenon of a learner selecting actions or making queries that influence what data are added to its information content from which associations and rules are learned.

The primary goal of machine learning is to derive general patterns or models from a limited amount of data. An active learner Figure 14 gathers information about the world by asking queries and receiving responses. It then outputs a classifier or model depending upon the task that it is being used for. An active learner differs from a passive learner which simply receives a random data set from the world and then outputs a classifier or model. One analogy is that a standard passive learner Figure 13 is a student that gathers information by sitting and listening to a teacher while an active learner is a student that asks the teacher questions, listens to the answers and asks further questions based upon the teacher's response. It is plausible that this extra ability to adaptively query the world based upon past responses would allow an active learner to perform better than a passive learner. Furthermore we need not set out our desired queries in advance. Instead, we can choose our next query based upon the answers to our previous queries. This process of guiding the sampling process by querying for certain types of instances based upon the data that we have seen so far is called active learning.

In practice, active learning offers its greatest rewards in situations where data are expensive or difficult to obtain, or when the environment is complex or dangerous. In industrial settings each training point may take days to gather and cost thousands of dollars; a method for optimally selecting these points could offer enormous savings in time and money. There is also a significant body of work on the design of experiments in the field of optimal experimental design (**Atkinson, 2001**). The general theory of active learning has been studied in the area of optimal experimentation design (**Chaloner & Verdinelli, 1995**); (**Sebastiani & Wynn, 2000**) and (**Fedorov, 1972**). A system proposed to drive robots that will perform queries whose results would be fed back into the active learning system was investigated in the context of refining theories found with Inductive Logic Programming by (**Muggleton, S.H.; Bryant, C.H.; Page, C.D.; Sternberg, 1999**).

One other major area of machine learning is reinforcement learning (**Kaelbling, Littman, & Moore, 1996**). For tackling the reinforcement problem there are some model based algorithms ,that explicitly have two modes of operation: an explore mode that tries to estimate and refine the parameters of the whole model and an exploit mode that tries to maximize the reward given the current model **(Kearns & Koller, 1999; Kearns & Singh, 2002**). The explore mode can be regarded as being an active learner; it tries to learn as much about the domain as possible, in the shortest possible time.

2.7.4 Metrics for decision making

In mathematical statistics and information theory there are various techniques available for decision making based on entropy (Information) or distance based metrics. Few of the general techniques for distinguishing between different probability distributions are mentioned below for reference:

Statistical distances

- Kullback–Leibler divergence(Ali & Silvey, 1966a; Csiszár, 1967; Hershey & Olsen, 2007; B. Liu, Yan, Li, & Wang, 2010; David J C MacKay, 2002)
- Rényi's divergence**(Van Erven & Harrëmos, 2014)**
- Jensen–Shannon divergence**(Grosse et al., 2002; P. W. Lamberti, Majtey, Borras, Casas, & Plastino, 2008; Pedro W. Lamberti & Majtey, 2003)**
- Bhattacharyya distance**(Choi & Lee, 2003; Guorong, Peiqi, & Minhui, 1996; Kailath, 1967)**

Criteria for information based model, observation, feature, sensor selection

- Akaike information criterion(**Bozdogan, 2000)**
- Bayesian information criterion**(Y. Zhang & Ji, 2010)**
- Deviance information criterion**(Spiegelhalter, Best, Carlin, & Van der Linde, 2014)**
- Focused information criterion(**Claeskens & Hjort, 2003)**
- Entropy, relative entropy, cross-entropy, Mutual Information**(Cover & Thomas, 2006)**
- Fisher-information matrix (sometimes simply called information) can be defined as the variance of the score, or as the expected value of the observed information**(Borguet & Léonard, 2008; Lu, Ye, & Neuman, 2011)**.
- Approximate entropy, Sample Entropy and permutation entropy**(Bandt & Pompe, 2002; He, Huang, & Zhang, 2012a; Kosmidou & Hadjileontiadis, 2009)**

The section below provides a reference to few of the criteria's used for data or model selection in the literature for different purposes:

Data Selection

The criterion for how informative a new datum will depend on what we are interested in. Experimental design within a Bayesian framework using the Shannon information as an objective function has been studied by (**Lindley.D.V, 1956**) and by (**Luttrell, 1985**). (David J. C. MacKay, 1992) discusses three different alternatives objective functions based on Information measure for data selection. (**Chater & Oaksford, 1999**) explains the information gain and decision theoretic approaches to data selection. (**Osborne, Garnett, & Roberts, 2010**) recently discusses a Bayesian formalism for the intelligent selection of observations from sensor networks that may intermittently undergo faults or change points.

Active Leaning

In the regression setting, active learning has been investigated by **(Cohn, Ghahramani, & Jordan, 1996**). (David J. C. MacKay, 1992) also explores the effects of different information-based loss functions for active learning in a regression setting, including the use of KL-divergence. Another related area to active learning is the notion of value of information in decision theory. The value of information of a variable is the expected increase in utility that we would gain if we were to know its value. For example, in a printer troubleshooting task (**Heckerman et al., 1994**), where the goal is to successful diagnose the problem, we may have the option of observing certain domain variables (such as “ink warning light on”) by asking the user questions. We can use a value of information computation to determine which questions are most useful to ask.

2.8 Critical Evaluation of methodologies

In this section, a critical evaluation of the particular research methodology from an industrial perspective is provided. The main conclusions which can be drawn from above discussion in terms of technological impact of the particular chosen research methodology are summarized below.

Active sensing is a very recent field offers a new research direction for data acquisition. The main benefit lies in that less storage space is required during data acquisition and its technological impact can be high but following limitations such as the requirement of a real time optimal online optimization algorithm, optimal sensor selection and the order of operations for calculation of criteria for optimality grows exponentially as the number of sensors increase hinder the use of active sensing in prognostic framework.

Passive data fusion or optimal sensor field has been applied to various engineering system but Active information fusion field offers a good research opportunity for the use in a prognostic framework. It offers potential benefit of using existing sources of information for informative data collection in an optimal way which can later be used of fault diagnosis with high technological impact but it requires an efficient online optimization of available data acquisition/ monitoring strategies.

Although the use of Compressed sensing in data acquisition process is very promising and once mature its technological impact will be very high but currently it is not a not mature technology (even though highly researched field) and the restrictive hardware capabilities makes it difficult to use in near future.

Active data selection heavily researched area in the machine learning community for training data selection for function approximations, model selection etc. In future, its technological impact can be medium to low depending upon the particular problem. It can potential be used for selecting interesting sample of data (fault data) for further analysis. The main limitation attached to the use of this methodology is that it can mostly be applied to offline data selection for function approximation.

Experiment design/ Optimal Input design/ Optimal input for model discrimination is a heavily researched field in system theory, system modelling, and control system. It offers a unique mix of scientific and application driven research. It can have a very high technology Impact and offer benefits as many industrial systems rely on experiment design based system identification for modelling. The limitations of these methodologies lies in the fact that it requires efficient experiment design processes, requires the exact knowledge of model or sometimes very high amount of data.

Active diagnosis is a very recent fault diagnosis methodology. It offers a very good research opportunity and a potentially a very high Technology impact with the benefits by providing an accurate identification of faults during run-time but practical applications are limited due to system constraints.

Active learning like Active data selection is heavily researched topic in the field of machine learning. Its technology Impact can be considered from medium to low. It may offer few benefits when applied to some industrial problem. Main restriction to the use of this methodology is due to the fact that it is very problem specific and has not been really tested on industrial problem.

2.9 Summary

In this chapter, an introduction to state-of-art monitoring system and its capabilities of the gas turbine engine is given. The hierarchical & modular structure of the gas turbine engine along with the challenges and limitations in the current generation health monitoring systems, with respect to the technologies involved in data collection, communication and decision making of the gas turbine engine is discussed. Based on comprehensive literature survey, use of the transient information (in the various sensor signals as well as system's operation) for the purpose of system's health monitoring is discussed in detail. A new architecture for the next generation equipment health monitoring systems for the gas turbine engine is proposed, which also includes the information from various ancillary/accessory system. A detailed discussion about technological impact and the suitability of the available methodologies, for the use in next generation gas turbine engine of the available methodologies is provided.

3 Framework Design

3.1 Generic framework for system health monitoring

In this section, a generic framework for the health monitoring of complex engineering systems such as gas turbine engines is proposed. First of all, a systematic methodology for the design of health monitoring for inclusion of transient information in accessing the overall health of the system is proposed as shown in Figure 15. The proposed framework clearly distinguishes between the data collection, data generation and decision making boundary for designing a robust equipment health monitoring system for civil aerospace gas turbine engines.

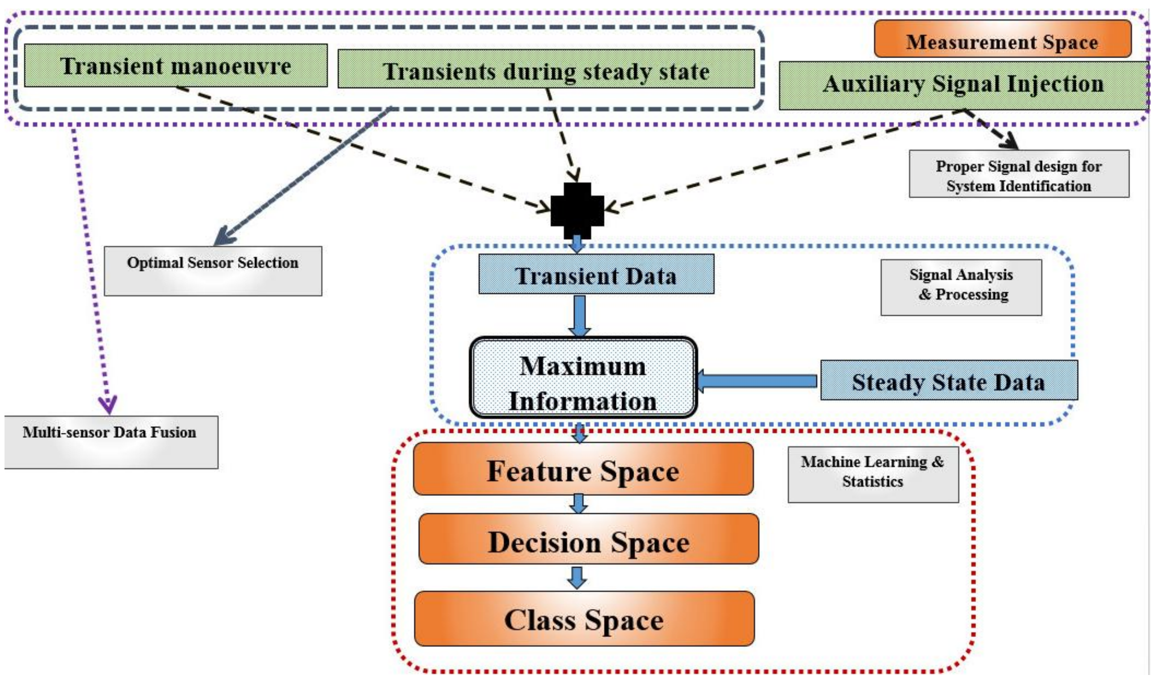


Figure 15: Proposed Framework

In the modern day civil gas turbine engines, the data is normally collected as snap-shots (taken at various pre-specified intervals during the whole flight) or as the steady-state data i.e. at the ground stations/flight parking bays (when the aircraft is at rest).

Hence, present day EHM systems do not utilize the available transient data generated at various stages of the gas turbine engine flight cycle. Therefore, in order to maximize the information or minimize the uncertainty about the system/state of the chosen subsystem/component and later this information can be used for accurate fault diagnosis, a comprehensive data generation, data collection, communication as well as intelligent decision making capabilities are required in the next generation EHM. Figure 15 depicts one of the possible ways, how the transient information can be included in to data collection process and which technological questions need to be answered in order to design a fully robust informative data collection methodology in order to include the transient information for maximizing the information about the state of the system.

As shown in the Figure 15, transient data can be collected, either during the different phases of the flight, during different transient manoeuvres e.g. acceleration, deceleration etc., transients occurring during the steady-state operation of the flight e.g. turbulences, sudden loss in the altitude or change of direction or by designing a dedicated auxiliary signal, which pokes the systems at pre-specified interval in order to gather/dynamic signals without interfering with the normal operations of the system. There are many technological questions such as optimal sensor selection for data collection, sampling rate for data acquisition, bandwidth available for communication at the ground station and redundancy of the data available from various sources/sensors, which one needs to address before finally deciding to pursue a particular methodology for data acquisition and processing. The above shown framework can further be refined to Figure 16 depict various approaches available for capturing most informative data:

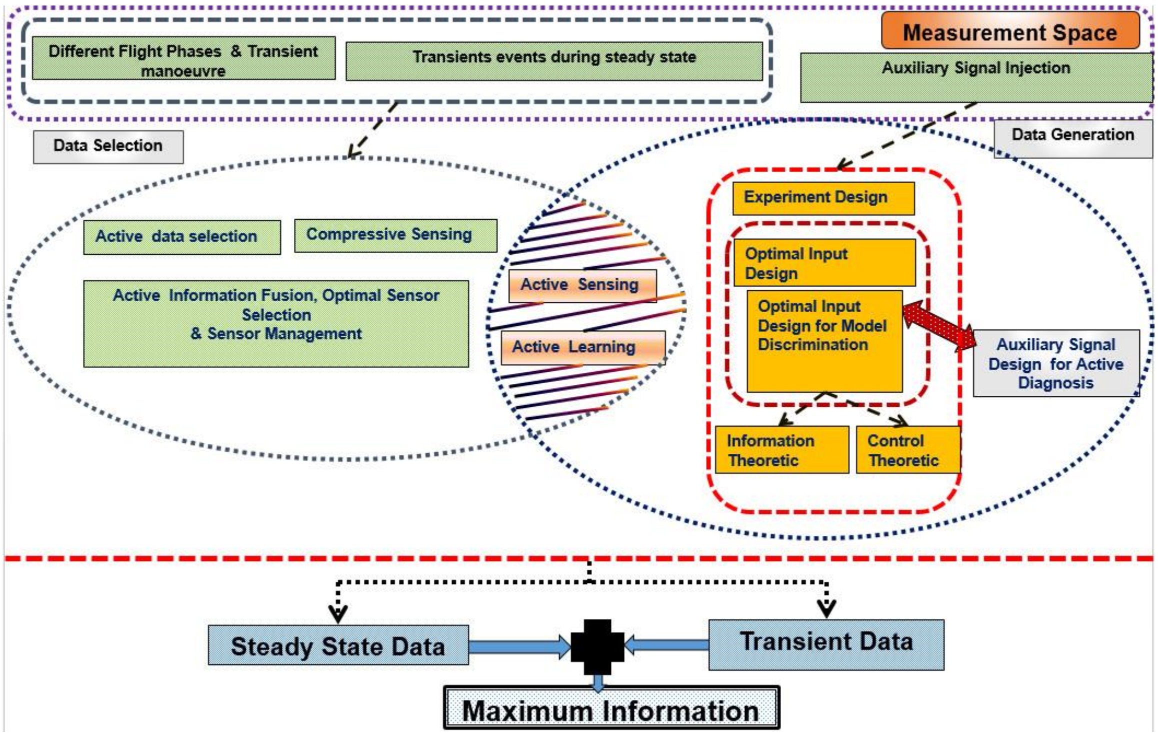


Figure 16: Framework for health monitoring

In this proposed extended framework following key aspects of system health monitoring have been considered as shown in Figure 16:

- Data Selection: Acquiring a parsimonious data set for health assessment
- Sensor selection and placement
- Active[1] sensing – data acquisition configuration
- Compression
- Data generation: Generating health-informative data
- Signal/experiment design

[1] The word **Active** has been extensively used in different contexts in the literature from different scientific domains. Other options could be e.g. Responsive, receptive etc.

- Optimal Input signal design for model discrimination (E.g. optimal signal design for discriminating between nominal and faulty mode of the same physical system or between two different fault modes of system with multiple fault modes.

This is underpinned by multi-sensor data fusion, information fusion and other signal analysis methodologies. These categories lead to the following approaches for fault diagnosis and learning.

- Using the informative data for fault diagnosis (getting information to analytics)
- Auxiliary signal design for active fault diagnosis (Combines the field of optimal experimental design and fault diagnosis)
- Active learning (When the system is diagnosed to be in faulty mode, the selecting different queries/signal to learn more about the faulty mode of the system)

Figure 16 shows various possibilities in which the proposed framework can be expanded and how different techniques can be included to accommodate the factors mentioned above. It also shows clearly, where different techniques lie on different side of paradigm as well as which techniques share the boundary between data collection and data generation.

The following section would how the above mentioned aspects fit as well as extend/contribute to overall Rolls-Royce equipment health management paradigm of "**Sense-Acquire-Transfer-Analyse-Act (SATAA)**".

3.2 Sense-Acquire-Transfer-Analyse-Act Paradigm

With an electrical system connecting all the equipment with power, the control system controlling all the actions of the system, a monitoring system is needed to log the actions, performance and status of the components in these systems. As mentioned before, the monitoring collection system can be ground-based, while monitoring systems fly in the air, float on the sea, or generate electricity on another continent. All monitoring systems work on the simple basis of (**Rolls-Royce, 2014**):

Sense - Acquire - Transfer - Analyse - Act.

The aim of a monitoring system is to maximise reliability and availability. A monitoring system will not stop a system from malfunctioning, but will log system data from which system characteristics can be deduced. The monitoring system collects data from all over the system and provides feedback at a specific location depending on the product the equipment is part of. The monitoring systems provides us with the freedom of knowing what our system is doing, how well it is doing it and helps in predicting how it will react next time we run it.

Fundamental issues in any closed loop system are its sensing capabilities, data acquisition, data selection, data transmission and analysis. Figure 17 below presents a possible mapping of the above discussed fields to the Rolls-Royce Engine Health Monitoring "Sense-Acquire-Transfer-Analyse-Act" Paradigm **(Rolls-Royce,**

2014). This mapping can be modified or improved in future according to the needs/limitations as well as chosen/developed approaches. This section will briefly describe the scope as well as limitations of few of the above mentioned research fields and their relevance / applicability to the Rolls-Royce engine health monitoring paradigm of "**Sense-Acquire-Transfer-Analyse-Act (SATAA)**" **(Rolls-Royce, 2014**). The presented mapping shows, how the current day on-board, off-board as well as communication capabilities can be enhanced by the possible adaptation of the methodologies discussed above at the different stages of the paradigm.

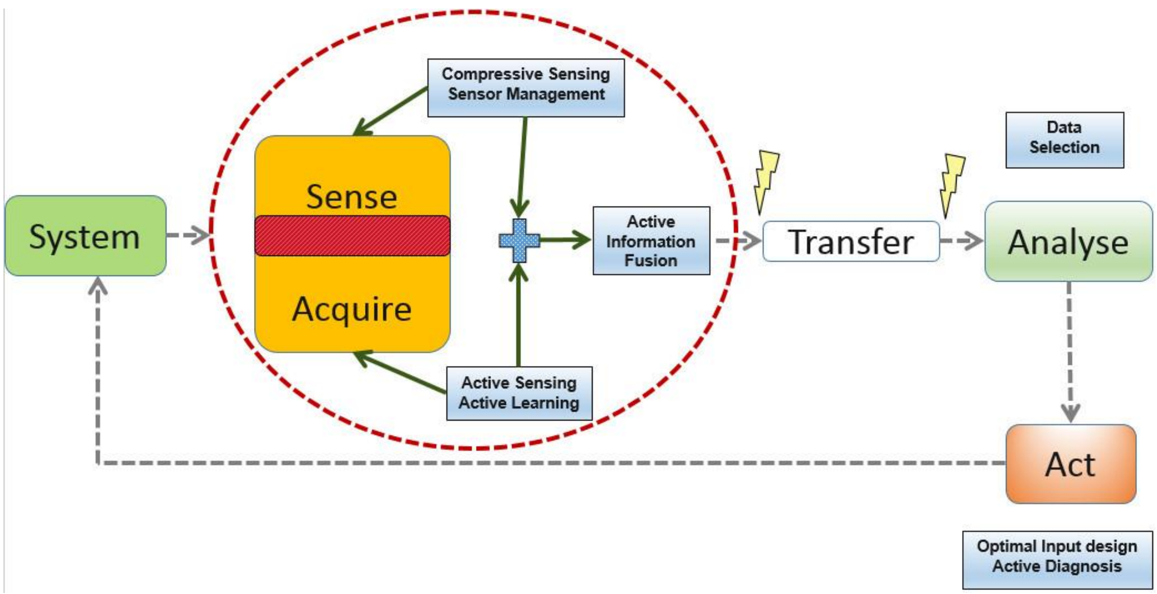


Figure 17: Sense-Acquire-Transfer-Analyse-Act paradigm

3.3 Summary

A systematic framework for the design of health monitoring for inclusion of transient information in accessing the overall health of the system is proposed. Furthermore various suggestions to include new technologies/methodologies to extend the framework are made, so that existing bottlenecks, technological gaps in the existing equipment health monitoring system with respect to data collection, communication, compression and decision making etc. can be addressed in the next generation monitoring systems for the civil gas turbine engines. The Rolls-Royce "Sense-Acquire-Transfer-Analyse-Act Paradigm" is also extended

4 Case Study Identification

4.1 Pimento Tool & Expert knowledge

A complex system can fail in a multiple ways. In a physical system, most of the faults that are manifested as system-level failures are initiated at the component-level, and a gas turbine engine is no exception. In the engine, there are a large number of components, each of which can have multiple failure modes. Furthermore, each failure mode is a product of many failure mechanisms that are simultaneously active. In short, there can be a large number of failure scenarios in the engine. Ideally, a system-level health assessment methodology should take all these possibilities into consideration. However, in most of the practical cases, it is not possible to cover all of these cases. Hence it is important to choose a representative set of suitable system/subsystems or components which are most relevant in terms of criticality of the fault, time, cost and effort.

In order to choose a candidate system or subsystem as shown in Figure 18, one must consider following factors:

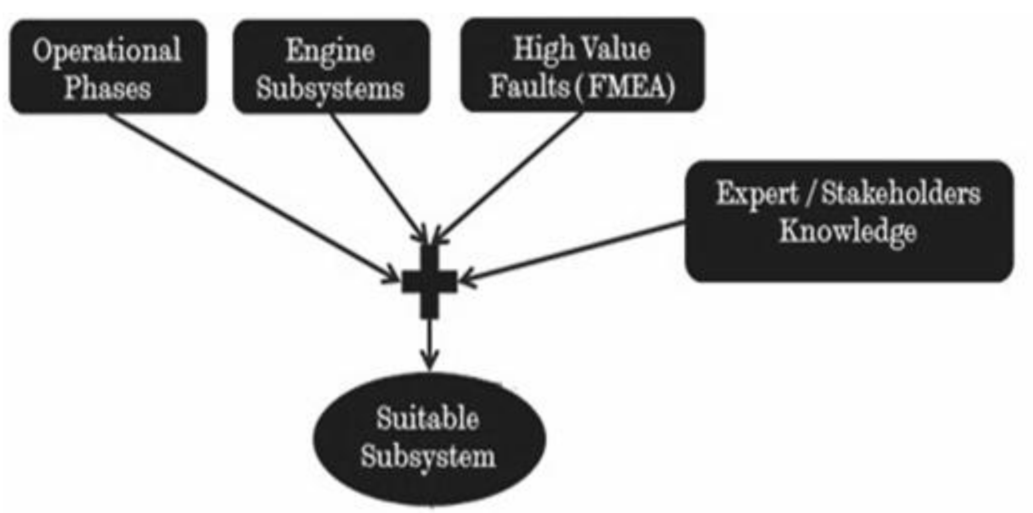


Figure 18: Selection of subsystem selection

- Mission profile or operating phase of the aircraft
- Failure mechanisms of each of these components are known a priori. This information can be obtained through the failure modes, mechanisms, and effects analysis (FMMEA) of the system.
- Expert knowledge

a) Aircraft Mission Profile

A mission profile is a detailed description of an aircraft's fight path and it's in-flight activities **(Jayaram & Rivera, 1994**). These profiles are broken into more specialized segments known as phases, which focus on specific flight operations. Figure 19 shows various phases of a typical flight cycle of a commercial aircraft. Operational flight phases play an important role in the selection of a suitable system as loads and stresses under which some these system/components operate depend on the phase of flight they are operating which in turn effect the probability of them developing a fault or at later stage a failure.

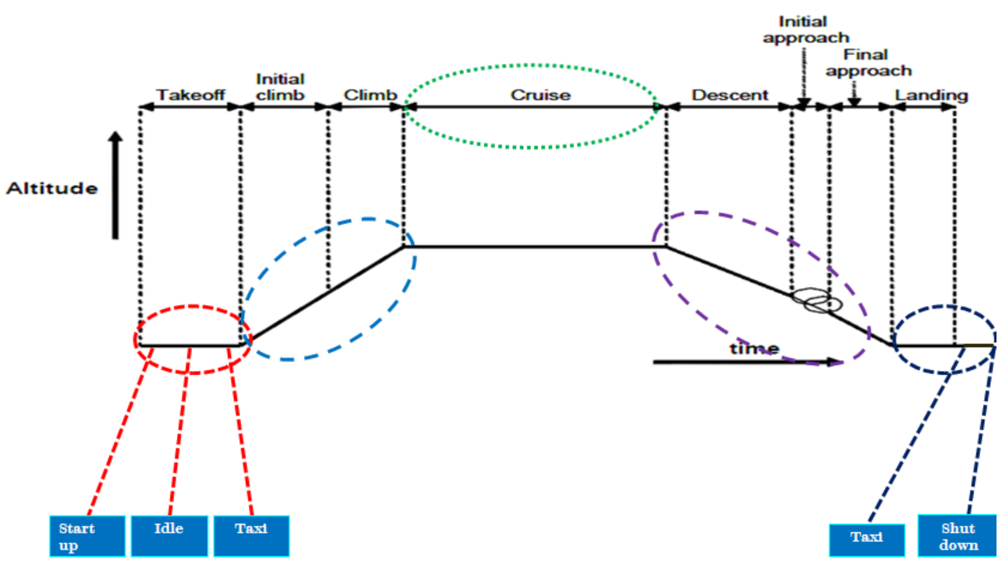
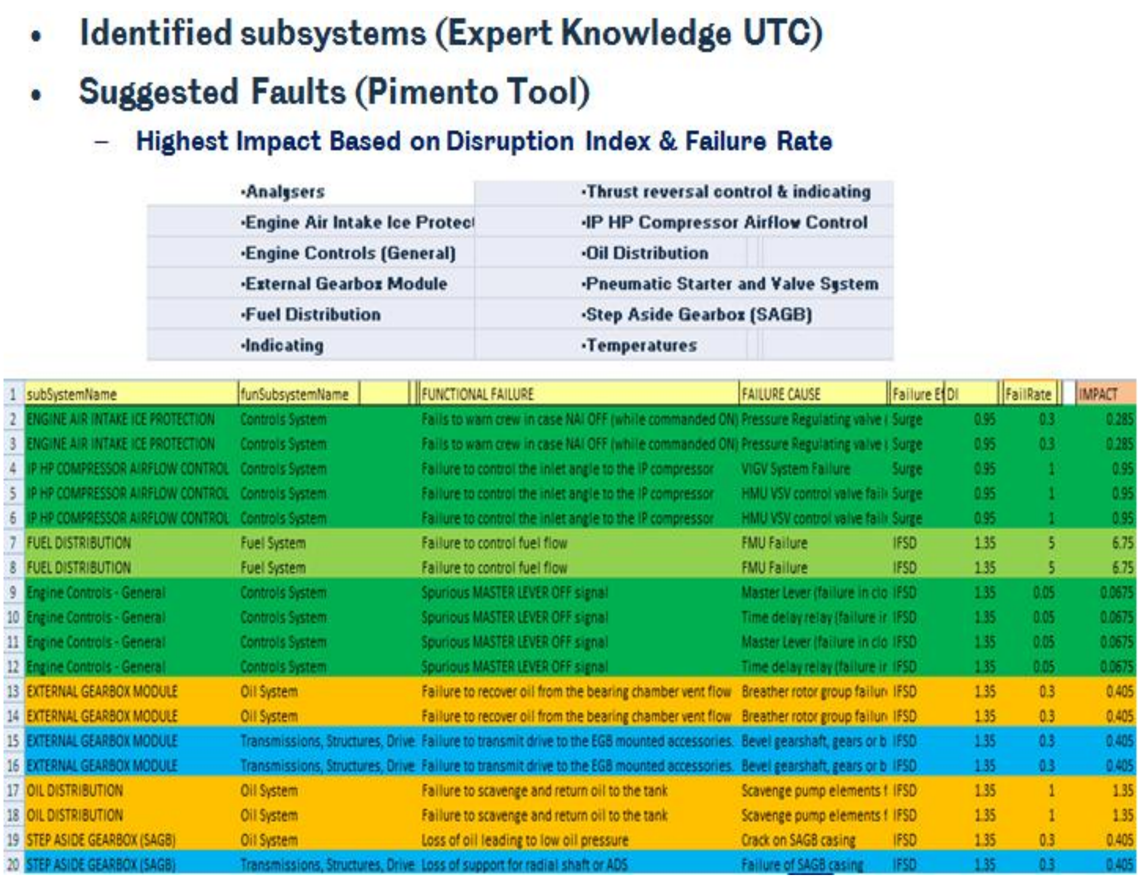


Figure 19: Civil Aircraft Flight Cycle

b) FMECA Study (PIMENTO TOOL)

In this work the components that are most likely to fail are identified through FMECA study using PIMENTO TOOL and expert knowledge. FMECA studies identify potential failure of a component/subsystem, determine the effects of this failure, and identify actions that can eliminate or reduce the likelihood of potential failures to occur (**Bowles & Ph.D, 2012**). The failure mechanisms of each of these components are known a priori. This information can be obtained through the failure modes, mechanisms, and effects analysis (FMMEA) (**Pecht, 1995**) of the system. The purpose of FMMEA is to identify potential failure mechanisms for all potential failures modes, and to prioritize failure mechanisms.



- Identified subsystems (Expert Knowledge UTC)
- Suggested Faults (Pimento Tool)
  - Highest Impact Based on Disruption Index & Failure Rate

·Analysers	·Thrust reversal control & indicating
·Engine Air Intake Ice Protec	·IP HP Compressor Airflow Control
·Engine Controls (General)	·Oil Distribution
·External Gearbox Module	·Pneumatic Starter and Valve System
·Fuel Distribution	·Step Aside Gearbox (SAGB)
·Indicating	·Temperatures

1	subSystemName	funSubsystemName	FUNCTIONAL FAILURE	FAILURE CAUSE	Failure E	DI	FailRate	IMPACT
2	ENGINE AIR INTAKE ICE PROTECTION	Controls System	Fails to warn crew in case NAI OFF (while commanded ON)	Pressure Regulating valve i	Surge	0.95	0.3	0.285
3	ENGINE AIR INTAKE ICE PROTECTION	Controls System	Fails to warn crew in case NAI OFF (while commanded ON)	Pressure Regulating valve i	Surge	0.95	0.3	0.285
4	IP HP COMPRESSOR AIRFLOW CONTROL	Controls System	Failure to control the inlet angle to the IP compressor	VIGV System Failure	Surge	0.95	1	0.95
5	IP HP COMPRESSOR AIRFLOW CONTROL	Controls System	Failure to control the inlet angle to the IP compressor	HMU VSV control valve fail	Surge	0.95	1	0.95
6	IP HP COMPRESSOR AIRFLOW CONTROL	Controls System	Failure to control the inlet angle to the IP compressor	HMU VSV control valve fail	Surge	0.95	1	0.95
7	FUEL DISTRIBUTION	Fuel System	Failure to control fuel flow	FMU Failure	IFSD	1.35	5	6.75
8	FUEL DISTRIBUTION	Fuel System	Failure to control fuel flow	FMU Failure	IFSD	1.35	5	6.75
9	Engine Controls - General	Controls System	Spurious MASTER LEVER OFF signal	Master Lever (failure in clo	IFSD	1.35	0.05	0.0675
10	Engine Controls - General	Controls System	Spurious MASTER LEVER OFF signal	Time delay relay (failure ir	IFSD	1.35	0.05	0.0675
11	Engine Controls - General	Controls System	Spurious MASTER LEVER OFF signal	Master Lever (failure in clo	IFSD	1.35	0.05	0.0675
12	Engine Controls - General	Controls System	Spurious MASTER LEVER OFF signal	Time delay relay (failure ir	IFSD	1.35	0.05	0.0675
13	EXTERNAL GEARBOX MODULE	Oil System	Failure to recover oil from the bearing chamber vent flow	Breather rotor group failur	IFSD	1.35	0.3	0.405
14	EXTERNAL GEARBOX MODULE	Oil System	Failure to recover oil from the bearing chamber vent flow	Breather rotor group failur	IFSD	1.35	0.3	0.405
15	EXTERNAL GEARBOX MODULE	Transmissions, Structures, Drive	Failure to transmit drive to the EGB mounted accessories.	Bevel gearshaft, gears or b	IFSD	1.35	0.3	0.405
16	EXTERNAL GEARBOX MODULE	Transmissions, Structures, Drive	Failure to transmit drive to the EGB mounted accessories.	Bevel gearshaft, gears or b	IFSD	1.35	0.3	0.405
17	OIL DISTRIBUTION	Oil System	Failure to scavenge and return oil to the tank	Scavenge pump elements f	IFSD	1.35	1	1.35
18	OIL DISTRIBUTION	Oil System	Failure to scavenge and return oil to the tank	Scavenge pump elements f	IFSD	1.35	1	1.35
19	STEP ASIDE GEARBOX (SAGB)	Oil System	Loss of oil leading to low oil pressure	Crack on SAGB casing	IFSD	1.35	0.3	0.405
20	STEP ASIDE GEARBOX (SAGB)	Transmissions, Structures, Drive	Loss of support for radial shaft or ADS	Failure of SAGB casing	IFSD	1.35	0.3	0.405

Figure 20: Identified subsystem using Pimento Tool

The purpose of FMMEA is to identify potential failure mechanisms for all potential failures modes, and to prioritize failure mechanisms. Figure **20** shows the suitable subsystem suggested by performing FMECA as well as FMMEA study using PIMENTO Tool. After considering the factors important for the selection of suitable subsystems the following subsystems have been found suitable for analysis.

- Fuel distribution system
- Bleed valves
- Oil supply system

For this study fuel distribution system has been considered. The section below explains briefly the functioning of fuel distribution unit as well deals with a specific problem of oil debris monitoring in a component namely fuel metering valve of fuel distributions system.

4.2 Fuel system

In the fuel supply of an engine two systems are involved. These are the aircraft fuel supply and storage system and the engine fuel distribution system. The aircraft fuel system is also called the primary fuel system and has the functions of fuel storage as well as the supply of all engines with low pressure fuel. The engine mounted system is called the secondary fuel system. It is responsible for the fuel supply of the individual engine including the metering of the fuel for combustion. The engine-mounted fuel distribution system delivers clean pressurized fuel for combustion and hydraulic purposes. The system ensures that the fuel has the proper temperature and pressure for its use. The fuel for combustion is metered by a metering device. This is controlled by the computer of the FADEC system, the electronic engine control (EEC). In older hydro-mechanical systems the metering device is a component of the hydro-mechanical fuel control unit (FCU).

4.3 Fuel distribution system description

The basic fuel distribution system of a typical engine consists of the following components (listed in the sequence in which they are passed by the fuel): Fuel pump, Fuel-cooled oil cooler, Fuel filter, Fuel metering device (FMU), Fuel flow transmitter, Fuel manifold components and Fuel nozzles.

Fuel systems provide the engine with the necessary fuel to support the combustion process and the flow control such that the required quantity of fuel to enable an easy start, acceleration and deceleration in all different flight conditions. To achieve this, the fuel pump is used for sending the fuel into the fuel spray nozzles, which further inject the fuel in the form of an atomised spray into the combustion chamber for combustion. At this stage, the fuel is further mixed with an appropriate quantity of air and burnt together to produce hot gas to drive the turbine **(Soares, 2008)**.

The main controlling devices in the fuel systems are fully automatic except the selection of engine power output, which is normally achieved by a manual throttle or a power lever, because the flow rate must vary accordingly in order to balance/properly mix with the amount of air passing through the engine, while maintaining a constant selected engine speed or pressure ratio **(Soares, 2008)**., it is necessary to have An automatic safety control is usually needed to prevent crossing any maximum limit on the engine gas temperature, compressor delivery pressure, and the assembly speed. A governor provided in the fuel system also prevents the over-speeding for the safety reason.

4.4 Case Study: Oil debris monitoring in Fuel metering device (FMU)

For the purpose of this research a case study has been undertaken to investigate the applicability of the information theoretic techniques in identification of the faults and to extract the information about developed fault by analysing information from sensors signals. The main objective of this research was to investigate the inclusion of various available signals (directly measured, derived from measured signals or estimated signals) in identification of faults. Fuel metering system has been selected as a suitable subsystem/component for a test case study.

4.5 Fuel metering unit

The main function of the FMU and Fuel Pump is to deliver a regulated fuel flow to the combustor, where the fuel is mixed with the air. The fuel pump consists of a low pressure centrifugal pump and a high-pressure gear pump, which are driven by the high-speed gearbox. The fuel input from the aircraft fuel tanks will flow through an aircraft fuel pump to the Low-Pressure (LP) pump.

Regulation of the quantity of fuel injected into the combustion system is important because it has the most significant effect on the engine behaviour. One of these effects is on the gas turbine engine control of power or thrust. The relationship between the airflow through the engine and the fuel supplied is complicated by changes in altitude, air temperature and aircraft speed. Varying these factors change the density of the air within the engine intake and consequently the mass of the air through the engine is also varied. There is an electronic system control, which measures and translates changing engine conditions to automatically adjust the fuel pump output electronically. Electronic Engine Control (EEC) system continuously monitors shaft speeds, temperatures and pressures along the engine to ensure its safe operation. The EEC commands the FMU to increase or decrease the flow of fuel to the engine to activate the desired level of thrust. Figure 21 shows the block diagram of a typical fuel metering unit.

This case study deals with identification of fault developed in fuel metering valve units due to accumulation of debris in fuel filters over a period of time. The data is collected from a test rig of a fuel metering unit. Over a period of time the quantity of the debris was increased to simulate the similar operating conditions as in real operation. As shown in Figure 21 the control signals are the signals which are available on a real system during real flight conditions & extra rig sensors were installed to measure the additional signals available at various stages of fuel metering valve. Figure 22 shows the torque motor current signal which is used to drive the FMV spool valve, Figure 23 shows the spool valve position and Figure 24 shows the effect magneto-motive force acting on the spool. These signals are used as the available signals for further analysis.

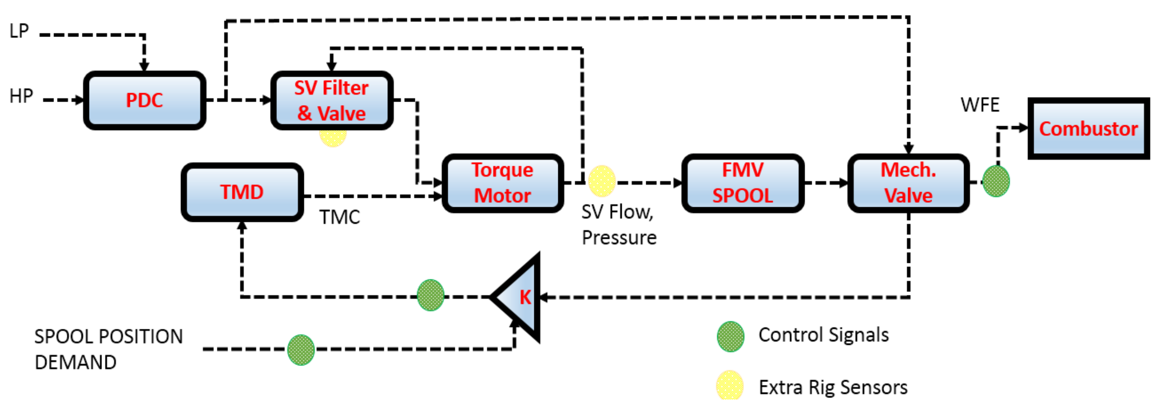


Figure 21: Fuel Metering Device

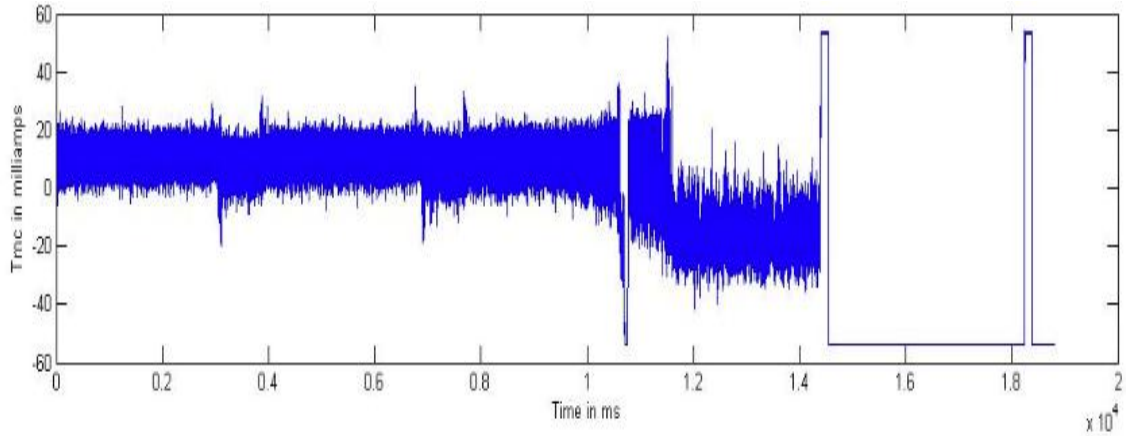


Figure 22: TMC Current

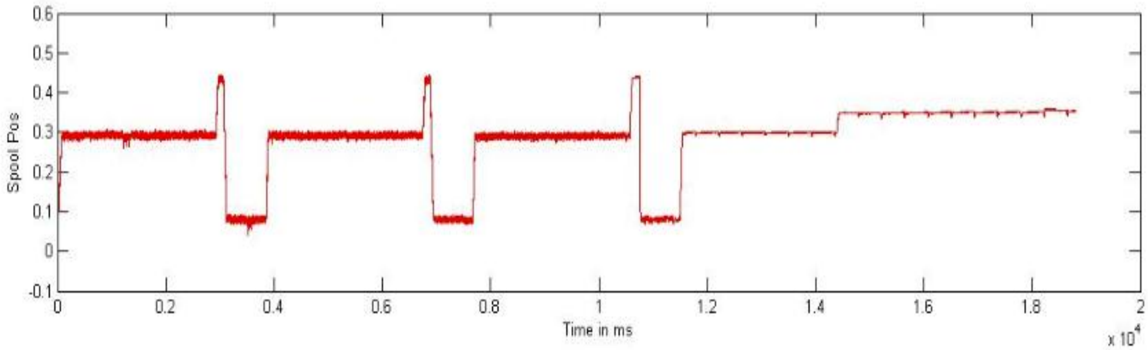


Figure 23: Spool Valve Position

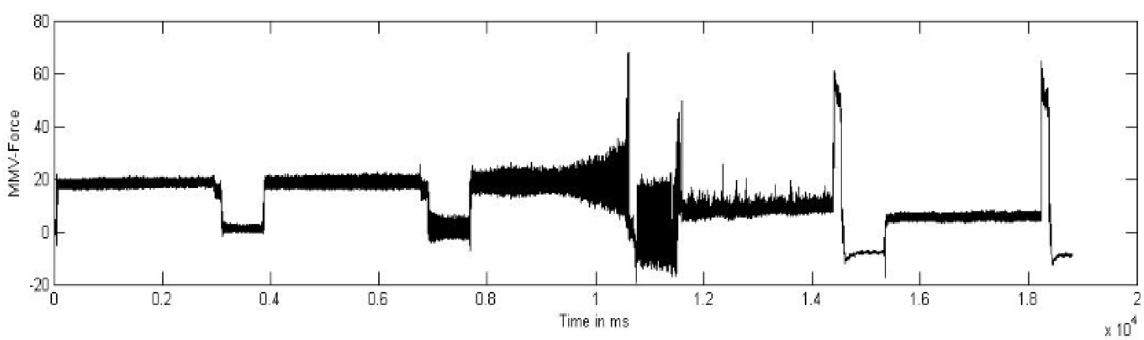


Figure 24: Resultant Force on Spool

Disclaimer: This block diagram shown in Figure 21 is based on an existing real platform based at Rolls-Royce plc, Derby and the data for this application was acquired at 40Hz sampling frequency.

4.6 Summary

This chapter gives a brief introduction of the challenges associated with the identification of the faulty system in a complex system. A complex system can fail in a multiple ways. In a physical system, most of the faults that are manifested as system-level failures are initiated at the component-level, and a gas turbine engine is no exception. In the engine, there are a large number of components, each of which can have multiple failure modes. Furthermore, each failure mode is a product of many failure mechanisms that are simultaneously active. Hence it is important to choose a representative set of suitable system/subsystems or components which are most relevant in terms of time, cost and effort. A systematic way to select a candidate subsystem based on the criticality of the problem, time, cost and effort is described. In the case of gas turbine engine it is emphasized, how the various factors such as operational mode of the flight, expert or stakeholder's knowledge, knowledge of high value faults (coming from Failure Mode Mechanism Effects Analysis, FMMEA study), which can contribute to the selection of suitable sub-system for fault investigation can be utilized and combined with already existing knowledge about the working of engine subsystem to select a candidate sub-system. Based on the described approach, Fuel metering valve (FMV) is selected as the candidate subsystem for further investigations.

5 Information collection

In order to design a robust equipment health monitoring system, data selection plays a critical role. For the better decision making of the system's state, suitable sensor signals should be selected and information hidden in those sensor signals must be properly extracted in order to make intelligent decisions. Feature extraction is always a crucial step for information gathering as well as health monitoring of a system. When-ever any change or faults occur, most of the systems always manifest abnormal and sometimes nonlinear dynamic behaviour. Hence it is necessary to extract the features hidden in the sensory signals for more ac-curate health monitoring and diagnosis. Feature extraction is always a crucial step for information gathering as well as health monitoring of a system. Whenever any change or faults occur, most of the systems always manifest abnormal and sometimes nonlinear dynamic behaviour. Hence it is necessary to extract the features hidden in the sensory signals for more accurate health monitoring and diagnosis. In this study apart from common statistical time domain features, other features are extracted based on the principles as well as tools/techniques available from the field of information theory and complexity theory. In this chapter all computations are performed on discrete-time data but that continuous-time definitions may be used for explanatory ease on occasion.

5.1 Statistical features extraction.

When the original discretized time domain signal is considered, some basic discriminative information which can be extracted in the form of statistical parameters from the *n* time domain samples, which can later be used for health monitoring are, root mean square (RMS), mean, peak value, crest factor, Skewness, kurtosis, Variance, Standard Deviation, Impulse Factor, Shape Factor etc. In the section below first a formal mathematical definition of these terms is given and later on 3 different datasets have been used to emphasize the usefulness of including some of these parameters in the wider set of feature set for health monitoring. Below the statistical definition are given for a signal $X = \{x_{1,} x_{2,} x_{3,} \dots \dots x_n\}$:

a. Statistical definition of time domain quantities

- Mean: $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$
- Standard deviation $\sigma = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^2}$
- Variance $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^2$
- Root-Mean Square $RMS = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n} x_i}$
- Kurtosis $= \dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^2\right)^2}$
- Skewness $= \dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^3}{\left(\sqrt[2]{\left(\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\mu)^2\right)}\right)^3}$
- Peak Value $= \frac{1}{2}(max(x) - min(x))$
- Crest Value $= \mathcal{P}eak\ value/RMS$
- Impulse factor $= \mathcal{P}eak\ value/\mu$
- Shape factor $= RMS/\mu$

	Mean Value	Standard Deviation	Variance	RMS Value	Skew-ness	Kurtosis	Crest Factor	Shape Factor	Impulse Factor
Seg1	**18.482**	0.6751	0.4557	18.4944	-0.0779	2.7763	0.1428	1.0006	0.1429
Seg2	**18.7717**	0.9992	0.9985	18.7983	0.0723	2.3768	0.1850	1.0014	0.1852
Seg3	**16.8362**	5.5139	30.4038	17.7162	-2.1240	6.9395	0.9423	1.0522	0.9916
Seg4	**12.9630**	10.5172	110.6124	16.6928	-0.3454	3.7555	2.6080	1.2877	3.3584

Table 2: Data Set 1

	Mean Value	Standard Deviation	Variance	RMS Value	Skew-ness	Kurtosis	Crest Factor	Shape Factor	Impulse Factor
Seg1	**18.3749**	0.7134	0.509	18.3887	-0.0921	2.7175	0.1475	1.0007	0.1476
Seg2	**18.7317**	0.9704	0.9417	18.7568	0.0808	2.4184	0.1877	1.0013	0.188
Seg3	**18.6229**	2.1225	4.505	18.7434	0.1301	2.0314	0.3685	1.0064	0.3709
Seg4	**18.4434**	5.5447	30.7442	19.2588	1.9292	14.61	1.8364	1.0442	1.9176

Table 3: Data Set 2

	Mean Value	Standard Deviation	Variance	RMS Value	Skew-ness	Kurtosis	Crest Factor	Shape Factor	Impulse Factor
Seg1	**18.5723**	0.6279	0.3943	18.5829	-0.0464	2.8974	0.1325	1.0005	0.1325
Seg2	**18.7755**	1.0043	1.0086	18.8023	0.077	2.3786	0.1849	1.0014	0.1852
Seg3	**18.5629**	2.1066	4.438	18.682	0.1293	2.0623	0.3697	1.0064	0.3721
Seg4	**14.6766**	9.9532	99.067	17.7333	-0.1948	4.8845	2.455	1.2082	2.9663

Table 4: Data Set 3

b. Results & Discussions

In this investigation various statistical and indicative time domain quantities have been calculated, which can be potentially be included in the wider set of features. Three different segmentations have been used for performing these investigations. One of the possible segmentation options can be seen in Figure 25. The data in each segment is then used to calculate the above mentioned statistical quantities. It can be easily observed from the Table 2, Table 3 and Table 4 that the second, third and fourth order statistical moments of data segments (segment 3 and 4 represents the onset of the fault conditions) can be included as few of the possible features for the detection of the faults a change (marginal sometimes) is visible, whereas there robustness depends largely on the number of sample points used to calculate the statistics. Features such as Impulse factor, crest factor and shape factor can be bit more reliable as a significant change is evident and they largely depend on the peak value, mean value and RMS value of the signal. Even though these time domain statistical parameters can be included in the feature set but their robustness as well as reliability needs to be tested for different scenarios. Therefore an extended feature set must be constructed containing more robust and discriminating features/information sources which may help in distinguishing between a faulty and non-faulty condition.

5.2 Features extraction based information theory

An entropy as well as mutual information calculator based on kernel density estimation is developed for this purpose, which can distinguish between the faulty and non-faulty states when we can measure/estimate a signal which has a direct relationship with the faulty signal/physical phenomenon. The basic idea is that the entropy of health parameter and mutual information between the health parameter and other available sensor signal changes when a fault is present.

One of the fundamentals problems in application of information-theoretic framework is in the calculation of univariate or multivariate probability density function (pdf). The estimation of pdf of a random variable can be done either in a parametric or a non-parametric way. The section below presents one of the many different non-parametric methods available for calculating the multivariate pdf of the random variable.

c. Estimation of Probability Density Function using Kernel density estimation

In statistics, kernel density estimation (KDE) is a non-parametric way to estimate the probability density function of a random variable. Let X = $[x_1, x_2, \ldots\ldots x_n]$ be an i.i.d sample drawn from some distribution with an unknown density $p(x)$. We are interested in estimating the shape of this function. Its kernel density estimator **(Moon, Rajagopalan, & Lall, 1995)**is

$$\hat{p}(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad (1)$$

Where $K(\bullet)$ is the kernel — a symmetric but not necessarily positive function that integrates to one — and $h > 0$ is a smoothing parameter called the bandwidth. A kernel with subscript h is called the scaled kernel **(Moon et al., 1995)** and defined as

$$K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right) \qquad (2)$$

A range of kernel functions are commonly used: uniform, triangular, Quartic, Epanechnikov, normal etc. In case of multivariate kernel density estimation, the kernel density estimate **(Moon et al., 1995)** is defined to be

$$\hat{p}(x) = \frac{1}{n}\sum_{i=1}^{n} K(u) \qquad (3)$$

Where

$$u = \frac{(x - x_i)^T S^{-1} (x - x_i)}{h^2} \qquad (4)$$

$K(u)$ is a multivariate kernel function (Moon et al., 1995), Let $(x_1, x_2, ... x_d)$ be a sample of d-dimensional random vectors whose density is being estimated, $x_i = [x_{1i}, x_{2i}, ... x_{di}]'$, $i = 1 - n$ are the n sample vectors, h is kernel bandwidth, and S is the covariance matrix on x_i. The kernel function $K(u)$ must be a valid probability density function. In this present case the multivariate Gaussian probability density function is used for $K(u)$, which is defined as **(Moon et al., 1995)**

$$K(u) = \frac{1}{(2\pi)^{d/2} h^d \det(S)^{1/2}} \exp(-u/2) \qquad (5)$$

Here "*d*" is the dimensionality of the vectored time series. There are many methods for choosing the bandwidth h. The "optimal" Gaussian bandwidth corresponding to kernel defined in equation (5) is given by **(Moon et al., 1995)**

$$h = \left\{ \frac{4}{(d+2)} \right\}^{1/(d+4)} n^{-1/(d+4)} \qquad (6)$$

d. Entropy of a signal

There are many notions of entropy which have been proposed in the literature. "*The etymology of the word entropy dates back to the famous German physicist **Rudolf Clausius** in second half of 18th century who defined entropy as a thermodynamic state variable*". ***Rudolf Clausius*** originally defined it as

$$\delta S = \frac{\delta Q}{Temp} \qquad (7)$$

Where δS an elementary change of entropy, δQ is a reversibly received elementary heat, and $Temp$ is an absolute temperature. Obviously this definition has no meaning in the field of signal processing. However, it started a diffusion of entropy as a term into the other areas."(**Ekštein & Pavelka, 2004**).

The entropy as a measure of system disorganization/disorder appeared for the first time in the field of thermodynamics. It is scientifically known that the systems usually tend to go from a state of order (low entropy) to a state of maximum disorder (high entropy). If we try to put this in other terms, it means that the entropy of a system can also be related to the amount of information it actually contains. A highly ordered system can be described using fewer bits of information (See later for more information) than a disordered system.

"*As proposed by scientists, the relation between entropy and signal processing is based on the hypothesis that the noise (white noise) is a projection of a system in the thermodynamic equilibrium into a signal. Hence, the noisy signal is supposed to have the highest entropy value, while the speech signal (which mainly contains periodic sounds like e.g. vowels) has significantly lower entropy value as it is supposed to be more organized, and it usually requires an extra energy to be produced in such an organized form*[2]".

The idea of entropy of random variables and processes by **Claude E. Shannon** is a cornerstone of the modern information theory and of the modern age of ergodic theory. Shannon introduced in his paper *"A Mathematical Theory of Communication"* (**Shannon, 1948**). According to his definition/concept, Entropy is a measure of the average information content one is missing when one does not know

[2] This principle reflects the Second thermodynamic postulate saying that entropy can be lowered if an energy is exerted into the task of organizing the examined system

the value of the random variable. For a discrete random variable X with values in a finite set χ (support set in case of continuous random variable), Shannon entropy (**Shannon, 1948**) is defined as below

$$H(X) = -\sum_{x \in \chi} p(x) \log_b p(x) \geq 0 \tag{8}$$

Where "b" is the base of the logarithm. Common values of b are 2 (for discrete random variables), Euler's number e, and 10, When b = 2, the units of entropy are also commonly referred to as bits.

Differential entropy also referred to as continuous entropy (**Cover & Thomas, 2006**) is a concept in information theory that extends the idea of Shannon entropy to continuous probability distribution. Let X be a random variable with cumulative distribution function $P(x) = Pr(X \leq x)$. If $P(x)$ is continuous, the random variable is said to be continuous.

Let $p(x) = P^{'}(x)$ when the derivative is defined. If $\int_{-\infty}^{\infty} p(x) > 0$, then $p(x)$ is called the probability density function for X. The set where $p(x) > 0$ is called the support set of X. The differential entropy $h(x)$ of a continuous random variable X with a density $f(x)$ is defined as (**Cover & Thomas, 2006**)

$$H(X) = -\int_{SS} p(x) log\ p(x) dx, \tag{9}$$

Where, ss is the support set of the random variable. As in the discrete case, the differential entropy depends only on the probability density of the random variable, and hence the differential entropy is sometimes written as $H(f)$ rather than $H(X)$.

Entropy has been used in multiple ways in the field science and engineering. The section below will give a brief overview of the different ways the entropy has been used in various scientific domains.

The principle of maximum entropy was introduced 1957 by Edwin Jaynes in his articles on Information Theory and Statistical Mechanics **(Jaynes, 1957a, 1957b, 1968**). In this paper Jaynes explains the approach in statistical mechanics by a principle which he called "maximum-entropy principle". The principle of maximum entropy is invoked or can be used, when we have some / partial piece(s) of information about a probability distribution of the data samples, but not enough to characterize it completely—it is quite likely, because we do not have the means or resources to know it completely. The principle of maximum entropy, in the very simplistic terms postulates that we should always choose the probability distribution that maximizes the amount of unpredictability/randomness contained in the distribution, under the constraint that the distribution matches the average that we measured (assumed).

(**Kullback, 1959**) also introduced a measure of relative information (relative entropy) between two probability distributions with respect to another. This also considered as a measure of the distance (in statistical sense) between these two probability distributions.

The identification of the model order in signal processing and system identification is an important factor and still an open area of research. When the order of the model (size of the parameter vector θ) is fixed, the estimation of an optimal value (in the sense of maximum likelihood, maximum a posteriori (MAP) or other Bayesian estimators) is well established, but the determination of the order of the model is still an active area of investigation and big challenge. Among the various tools used for determination of model order, entropy is used as one of the tools.

The famous Akaike criterion (**Hirotugu Akaike, 1969; H Akaike, 1974; Farrier, 1984; M. Wax & Kailath, 1985; Mati Wax, 1991**) uses this amount to determine the optimal order of the model in the specific context of linear models (in the parameters).

Entropy is used in multiple ways in spectral analysis. The classic example can be found in (**Burg P.J, 1967**). Entropy has been used in multiple ways for spectral estimation such as minimum cross-entropy spectral analysis (**Shore, 1981**), multidimensional power spectral estimation **(McClellan, 1982**).

e. Entropy calculation of the force signal

The data is obtained from a test rig developed to simulate the scenario of fault caused by debris build up in the fuel metering servo-valve has been used for analysis and validation purposes. Figure 22 shows the Torque motor current and [Figure 23] valve spool position signals respectively which are directly available on-board as well as on test rig.

It can be easily seen from the Figure 22 and Figure 23 respectively that not much information can be obtained from the torque motor current signal and valve spool position signals during the build of debris in the fuel system (simulation of faulty conditions) until the actual fault has occurred where as it can be observed from the Figure 24 that resultant force on the spool also changes (increases) as more and more debris is introduced in to the system (fuel filter). Hence In this present case, resultant force (calculated from the pressure signals time series data obtained from test rig) on the spool of the valve is found to be a suitable health index for further analysis. For the analysis the magneto-motive net force signal has been divided in to segments as can be seen below in the Figure 25.

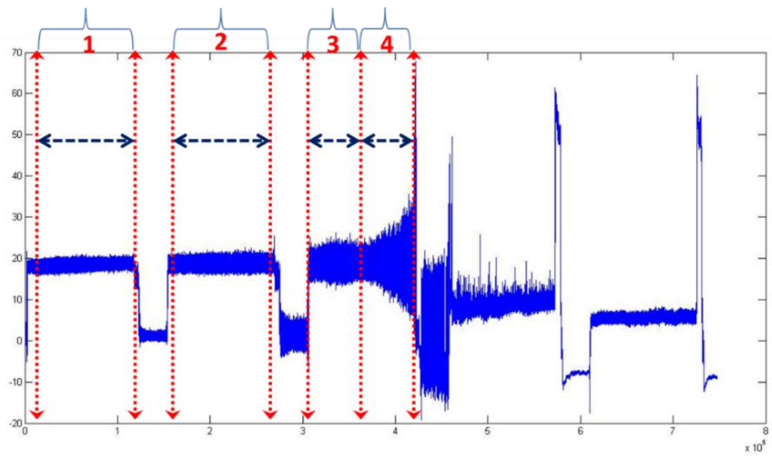


Figure 25: Segments of magneto-motive force

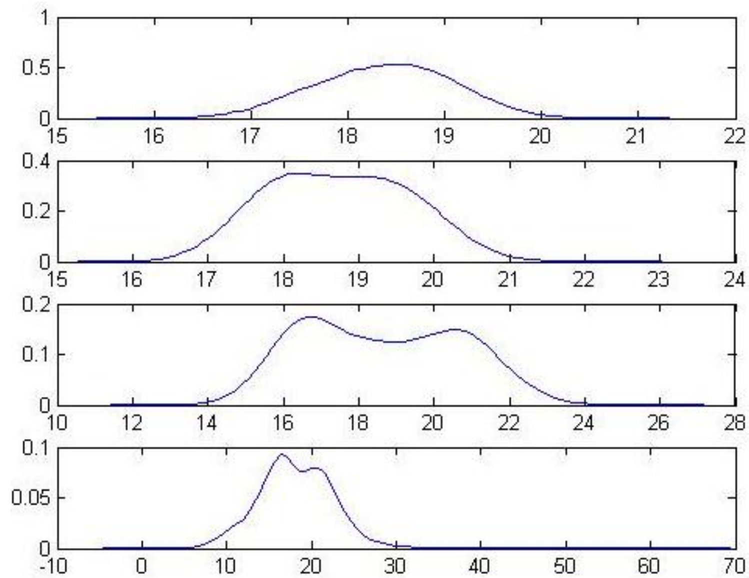


Figure 26: Probability density functions of different segments

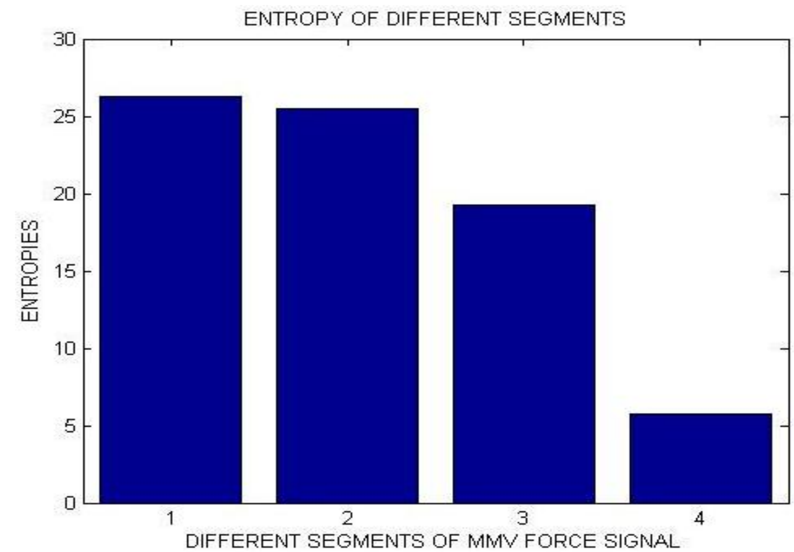


Figure 27: Entropies of Segments

	Data Set 1	Data Set 2	Data Set 3	Data Set 4	Data Set 5
Seg1	**26,2443**	**31,1825**	**29,6588**	**30,7628**	**29,1215**
Seg2	**25,5035**	**30,1130**	**28,6692**	**24,4049**	**28,6692**
Seg3	**19,2674**	**23,3082**	**21,5811**	**20,8741**	**20,0679**
Seg4	**5,7473**	**20,0679**	**9,4750**	**12,2861**	**12,2861**

Table 5: Entropy Calculation for Different Data Sets

f. Results & Discussions

In this investigation differential entropy of the different segments of the net magneto- motive force signal Figure 24 has been calculated. For this investigation the time series data is considered to be stationary and all other operating conditions are assumed to be same. The data can further be divided into the segments as shown in the Figure 25. This choice of data is only specific to this application because the data shown in the Figure 25 from a real platform which tries to emulate the situation in the air. Different troughs in the data are not present in the real data and they are actually present due to repetition of testing cycle for the oil debris build up problem. The main idea about this segmentation choice to actually combine the data without these troughs in order to emulate the situation in the air e.g. continuous build-up of oil debris in the fuel metering valve filter. The data in each segment is then used to calculate the probability density function (pdf) as shown in Figure 26. For the calculation of the probability density functions data is assumed to be stationary in each segment. From each probability density function then differential entropy is calculated. It can be easily seen from the Figure 27 that entropy of signal decreases as more and more debris is introduced in to the system. Further, in order to consider differential entropy as a possible discriminating feature, more similar experiments were carried out with different choice of data lengths and segmentation. Table 5 shows the result of such experiments. It can be easily observed that similar to first experiment the entropy of signal decreases as more and more debris is introduced in to the system. Hence differential entropy can easily be considered as one of discriminating features/information sources which may help in distinguishing between a faulty and non-faulty condition.

Disclaimer: These calculation were performed on MATLAB release 2013b, on a Windows 7 Professional PC with the following specifications.

Processor:	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz 3.40 GHz
Installed memory (RAM):	8.00 GB
System type:	64-bit Operating System

g. Mutual Information

Mutual information (M.I) is one of many quantities that measures how much one random variable tells us about another. It is a dimensionless quantity with (generally) units of bits, and can be thought of as the reduction in uncertainty about one random variable given knowledge of another. High mutual information indicates a large reduction in uncertainty; low mutual information indicates a small reduction; and zero mutual information between two random variables means the variables are independent.

Formally, the mutual information of two discrete random variables X and Y can be defined as (**David J C MacKay, 2002**):

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)} \right) \qquad (10)$$

In the case of continuous random variables, the summation is replaced by a definite double integral (**David J C MacKay, 2002**):

$$I(X;Y) = \int\int p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)} \right) dxdy \qquad (11)$$

Where $p(x,y)$ is now the joint probability density function of x and y, and $p(x) = \int_y p(x,y)$, $p(y) = \int_x p(x,y)$ are the marginal probability density functions of X and Y respectively. Qualitatively, entropy is a measure of uncertainty – the higher the entropy, the more uncertain one is about a random variable. To understand what $I(X;Y)$ actually means, we first need to define entropy, conditional entropy & joint entropy. Following quantity are defined as in (**David J C MacKay, 2002**)

The entropy of a random variable X is given by

$$H(X) = -\int p(x) \log p(x)\; dx \qquad (12)$$

The joint entropy of a set of random variables X & Y is given by

$$H(X;Y) = -\int\int P(x,y) \log P(x,y)\; dxdy \qquad (13)$$

The conditional entropy is the average uncertainty about X after observing a second random variable Y, and is given by (**David J C MacKay, 2002**)

$$H(X|Y) = -\iint p(x,y)\log\frac{p(x,y)}{p(y)}\,dxdy \qquad (14)$$

$$H(X|Y) = -\int p(y)\,H(X|Y=y)\,dy \qquad (15)$$

With the definitions of $H(X)$ and $H(X|Y)$, equation (11) can be written as (**David J C MacKay, 2002**)

$$I(X;Y) = H(X) - H(X|Y) \qquad (16)$$

Hence Mutual information is therefore the reduction in uncertainty about variable X, after observing Y and satisfies $I(X;Y) = I(Y;X)$. [Figure 28] shows how the total entropy $H(X;Y)$ of a joint ensemble can be broken down.

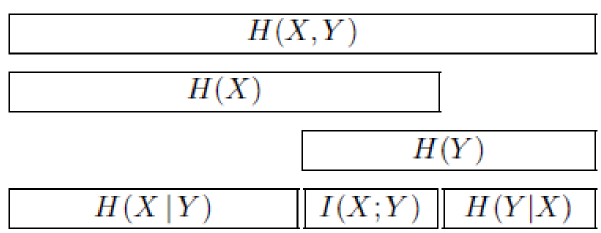


Figure 28: The relationship between joint information, Marginal entropy, conditional entropy and mutual entropy **(David J C MacKay, 2002**).

h. Estimation of Mutual information

As discussed in the previous section, estimating the mutual information (MI) between X and Y requires the estimation of the joint probability density function of (X,Y). This estimation of the joint probability function has to be carried on the known data set. This technique was first developed for discrete random variables later it has been extended to continuous random variables. For this reason, the MI has to be estimated.

There are two basic approaches to estimate the Mutual Information namely parametric and nonparametric methods. In the nonparametric estimation methods, usually no meaningful parameters are associated, whereas the parametric estimation methods on the other hand usually make assumptions about the functional form of the regression (mostly in least squares sense) and then the estimate is of actually only of those parameters that are available as free parameters. Some of the techniques available for estimating the MI include histogram based, adaptive partitioning, splines, kernel density and nearest neighbour etc.**(Walters-Williams & Li, 2009)**

Histogram-based (**Moddemeijer, 1989a, 1989b**) and kernel-based **(Mars & Arragon, 1982; Moon et al., 1995)** Pdf estimations are among the most commonly used methods. However, their use is usually restricted to one-dimensional or two-dimensional probability density functions. K-nearest neighbour method is also analyzed by **(Kraskov, Stögbauer, & Grassberger, 2004**) and is able to handle multi-dimensional probability density functions. A survey about various methods available to estimate the mutual information can be found in **(Walters-Williams & Li, 2009**).

i. Use of Mutual information

Mutual information is used in several areas of science and engineering. Mutual information is used in the field of language processing by **(Brown, Desouza, & Mercer, 1992; Lankhorst, Moddemeijer, Box, & Approach, 1993**), speech processing & speech recognition by **(Bahl, Brown, de Souza, & Mercer, 1986; Normandin, Cardin, & De Mori, 1994; Okawa, Kobayashi, & Shirai, 1994; Povey et al., 2008**). **(Mars & Arragon, 1982; Moddemeijer,**

1989a, 1989b) used it for the estimation of time delay. Mutual information was used for lag identification in nonlinear time series (**Fraser, 1989; Granger & Lin, 1994; Harvill & Ray, 2000; Kantz & Schreiber, 2003; Mars & Arragon, 1982**). Mutual information was used to analyze non-linear system in **(H. S. Kim, Eykholt, & Salas, 1999; Papana & Kugiumtzis, 2008, 2009**).

Use of mutual information in image processing can be found in (**Hastreiter, Freund, Greiner, & Ertl, 1997; Maes, Collignon, Vandermeulen, Marchal, & Suetens, 1997**). A very good survey about the use of mutual information in field of image processing can be found in (**Pluim, Maintz, & Viergever, 2003**). Other scientific field in which mutual information is extensively used is machine learning. In the specific area of machine learning various methods have been developed for feature selection using mutual information which can be found in (**Estévez, Tesmer, Perez, & Zurada, 2009; Kwak, 2002; H. Liu, Sun, Liu, & Zhang, 2009; Peng, Long, & Ding, 2005; Zaffalon & Hutter, 2002**).

(**Joshi, Deignan, Meckl, & Jennings, 2005**) proposed a fault detection algorithm for multi-input multi-output (MIMO) systems based on a clustering algorithm developed using mutual information. A fault diagnosis procedure based on discriminant analysis and mutual information was developed by (**Verron, Tiplica, & Kobi, 2008**). **(Munawar, Reidemeister, & Ward, 2011**) used in their work normalized mutual information (NMI) as to automatically monitor the health of complex software systems and localize faulty components when faults occur.

A nonparametric signal detection and classification technique was proposed for condition-based maintenance of Helicopter Drivetrains using mutual information measures in the time–frequency domain by (**Coats et al., 2011**). In (**Zugasti & Arrillaga,**

2012) mutual information was used to identify sensor fault for the structural health monitoring. (J. Yu, Chen, & Rashid, 2013) applied a multi-way independent component analysis (MICA) mixture model and mutual information based fault detection and diagnosis approach for batch process monitoring.

j. Results & Discussions

In this investigation a method for fault identification has been implemented. This method is based on the information theory, and more precisely, on Mutual Information. Mutual Information analyses the dependence between sensors information. If that information dependence changes between 2 different states, we can say that there is probably faulty condition in the system provided all other operating conditions remain same? In order to have that information, the dependence between two healthy sensors must be known under various operating conditions, that is why a learning phase is needed.

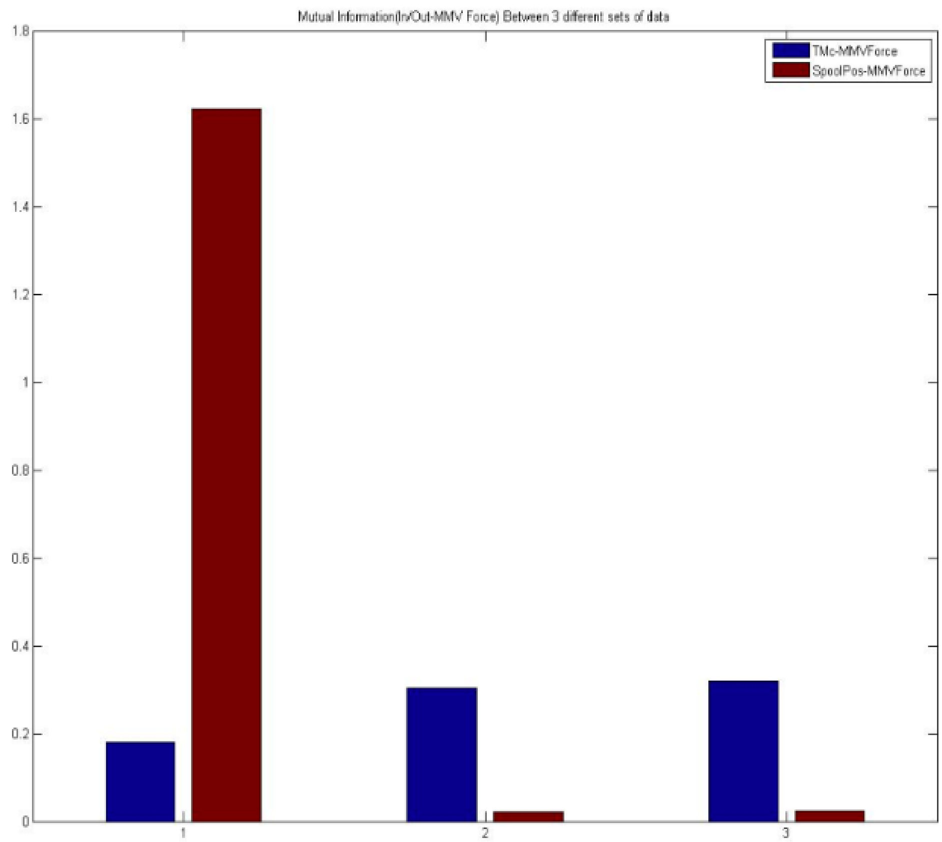


Figure 29: Mutual Information under different Conditions

As it can be seen in the Figure 29 the results obtained from testing of the developed algorithm. It can be seen that mutual information between the Torque motor current as shown in Figure 22 and resultant force do not change much but mutual information between valve spool position signal as shown in Figure 23 and resultant force changes drastically during faulty and non-faulty conditions. Hence the possible outcome of this investigation can easily be further utilized to develop a

a) Technique which can be used a selection procedure for optimal sensors/features.

b) Technique for fault identification If we can estimate/measure/calculate force correctly.

Note: Similarly like the calculation of entropy, here it is also assumed that for the calculation of the probability density functions data is assumed to be stationary throughout each segment.

Disclaimer: These calculation were performed on MATLAB release 2013b, on a Windows 7 Professional PC with the following specifications.

Processor:	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz 3.40 GHz
Installed memory (RAM):	8.00 GB
System type:	64-bit Operating System

5.3 Feature based on Complexity theory

Due to instantaneous variations in any system parameters such as friction, damping, or any external loading conditions over the time, dynamical systems tend to show / exhibit different dynamical properties (time-variation, nonlinearities etc.) which can often be characterized as non-linear behaviours **(Pintelon, Schoukens, & Vandersteen, 1997)**. Therefore, techniques for non-linear parameter estimation or non-linear dynamics provide a very good alternative to extracting defect-related features hidden in the measured sensor signals that may not be effectively identified using other techniques **(Yan & Gao, 2007)** In the past, various researchers put a lot of effort and investigated a number of non-linear parameter identification techniques such as e.g. Correlation Dimension **(Ruelle, 1990**) and Complexity Measure (Lempel & Ziv, 1976; Steven M Pincus, 1991). Correlation dimension has been successfully applied to gearbox tooth defect diagnosis by (**Jiang, Chen, & Qu, 1999**) and for rolling bearing defect detection by **(D. B. Logan & Mathew, 1996; D. Logan & Mathew, 1996**). As stated in (L. A. Smith, 1988; Theiler, 1990), In order to reliably estimate of the Correlation Dimension of a time series, it is often required that a large quantity of data points should be calculated, which requires high computational time and unsuitability for on-line, real-time applications.

The Complexity Measure, in comparison, is computationally more efficient as claimed by various authors in the past. The complexity of a signal can be described by two different measures: the Lempel–Ziv Complexity as defined in **(Kaspar & Schuster, 1987; Lempel & Ziv, 1976**) and Approximate Entropy (ApEn) which is proposed by **(Steven M Pincus, 1991)**. Approximate Entropy expresses the regularity of a time series in multiple dimensions, and contains more time-related information **(Steven M Pincus, 1991)**. Therefore, this makes ApEn an attractive tool for monitoring system dynamics, as information on the temporal progression (time evolution) of a defect is valuable not only for properly diagnosing the current system/machine health condition, but also in accurately predicting the future behaviour of the machine health.

a) Approximate Entropy

Approximate Entropy (ApEn) is a regularity statistic that quantifies the unpredictability of fluctuations in a time series and can be used to classify complex systems. As the working condition/health of a system/machine deteriorates due to the various initial conditions and/or progression of defects due to various operating conditions, it would eventually result in a decrease in regularity (periodic behaviour in case of rotating machine, phenomenon like whirl appear due to unbalance) of health parameter e.g. Such a change in the specific shape of a waveform or pattern being observed, would also tend to initiate a change in its corresponding Approximate Entropy (ApEn) value.

Studies done by (**Diambra, de Figueiredo, & Malta, 1999**) on the Electro-Encephalogram (EEG) signals revealed that an increase of the ApEn values can be used to detect and characterize eplieptic activities. In the field of biomedical engineering, studies on fetal heart rates using cardiotocography (CTG) done by (**Signorini, Magenes, Cerutti, & Arduini, 2003)** have shown correlation between the increase of ApEn values and pathological conditions. (**Abásolo et al., 2005**) found that the EEG analysis with ApEn could be a useful tool to increase the insight into brain dysfunction in Alzheimer's disease.(**Yan & Gao, 2007**) developed a machine health monitoring system based on the Approximate Entropy (ApEn).

(**Awwad, Hasan, Dyson, Balli, & Gan, 2008**) studied whether approximate entropy (ApEn) analysis provides a suitable method of detecting differences induced by a motor preparation task in time-ordered inter-spike intervals (ISIs) recorded in monotonically firing motoneurons. ApEn was used for analysing the regularity and complexity of the acoustic emission ssignals (AE) signals for crack monitoring by (**Taylor, Lin, & Chu, 2011**). ApEn was used to extract the nonlinear information and features of the vibration signal of the four typical faults of rotating machinery by (He, Huang, & Zhang, 2012b). **(Fang, Chen, Zheng, & Harrison, 2012**) have shown that ApEn can be used to extract different features in left-hand and right-hand motor imagery EEG efficiently. An EEG analysis system of seizure detection based on a cascade of wavelet-approximate entropy for feature selection was proposed by (**Shen et al., 2013**).

Theoretical background

For a time series X containing N data points $\{x(1), x(2),....,x(n)\}$, its ''regularity'' can be measured by ApEn in a multiple dimensional space, in which a series of vectors are constructed and expressed as (**Steven M Pincus, 1991**)

$$
\begin{aligned}
&X(1) = \{x(1), x(2), \ldots\ldots, x(m)\} \\
&X(2) = \{x(2), x(3), \ldots\ldots, x(m+1)\} \\
&\ldots\ldots \\
&X(n-m+1) = \{x(n-m+1), x(n-m+2), \ldots\ldots, x(n)\}
\end{aligned} \tag{17}
$$

In equation (17), each of the vectors is composed of m consecutive and discrete data points of the time series S (Steven M Pincus, 1991). The distance $d(X(i), X(j))$ between two vectors $X(i)$ and $X(j)$ can be defined as the maximum difference in their respective corresponding elements (**Steven M Pincus, 1991**):

$$
d(X(i), X(j)) = \max_{k=1,2,\ldots,m}(|x(i+k-1) - x(j+k-1)|), \tag{18}
$$

Where $i = 1, 2, \ldots., N-m+1,\ j = 1, 2, \ldots., n-m+1,$ and n is the number of data points contained in the time series. For each vector $X(i)$, a measure that describes the similarity between the vector $X(i)$ and all other vectors $X(j),\ j = 1, 2, \ldots., n-m+1,\ j \neq i$ can be constructed as(**Steven M Pincus, 1991**)

$$
C_i^m(r) = \frac{1}{n-(m-1)} \sum_{j \neq i} \Theta\{r - d[X(i), X(j)]\} \tag{19}
$$

Where

$$
\Theta\{x\} = \begin{Bmatrix} 1, & x \geq 0, \\ 0, & x < 0. \end{Bmatrix} \tag{20}
$$

The symbol r in equation (19) represents a predetermined tolerance value, defined as r (**Steven M Pincus, 1991**)

$$r = k * std(X), \tag{21}$$

Where, $(k > 0)$ is a constant and $std(\bullet)$ is the standard deviation of the time series. By defining (**Steven M Pincus, 1991**)

$$\phi^m(r) = \frac{1}{n-m+1}\sum_i \ln[C_i^m(r)],\ i = 1,2,....,n-m+1 \tag{22}$$

the ApEn value of the time series can be calculated as (**Steven M Pincus, 1991**)

$$ApEn(m,r) = \lim_{N\to\infty} [\phi^m(r) - \phi^{m+1}(r)] \tag{23}$$

For practical applications, a finite time series consisting N data points is used to estimate the ApEn value of the time series, which is defined as (**Steven M Pincus, 1991**)

$$ApEn(m,r,n) = \phi^m(r) - \phi^{m+1}(r) \tag{24}$$

As explained in **(B. Liu et al., 2010)** , the value of the estimate depends on the choice of m and r. These two parameters must be fixed apriori before ApEn can be calculated. As suggested by **(Steven M Pincus, 1991)**, m can be taken as 2 and r can be taken as $(0.1 \sim 0.25) * std$, where std is the standard deviation of the original data.

Equation (24) indicates the similarity among the reconstructed vectors within the time series, when the dimension of the vectors has increases from m to $m+1$**(Steven M Pincus, 1991)**. In this way the regularity of the time series being analysed is affected and consequently, the associated ApEn value changes. The greater the regularity is, the lower the ApEn value as suggested by the various researchers mentioned above. Take for an example, a periodic time series containing only a single frequency component will have a low ApEn value (which eventually approaches zero), due to the high regularity and periodic nature of the signal. In contrast, a relatively complex time series containing multiple frequency components e.g. a random noise sequence, random phase multisines etc. will have a high ApEn values, due to a low level of regularity **(Yan & Gao, 2007)**

To calculate the ApEn value of a given time series X, data points contained within the time series are first rearranged into a series of m and $m+1$ dimensional vectors, respectively, as illustrated in Figure 30. Then the distances between two corresponding data points from each corresponding vector is calculated. As mentioned in the equation (24) above, the similarity measure score for the reconstructed vectors for each dimension m and $m+1$ for a given value of r is obtained.

Subsequently, using equation (24), the ApEn value for the time series X with n data points is calculated. To ensure consistency of the ApEn calculation, a minimum data length n, as well as appropriate dimension m and tolerance r needs to be (pre-) determined by performing various training runs or by lots of inputs from domain experts. In this case study, ApEn is used to extract features from force signal obtained from the fuel metering valve spool.

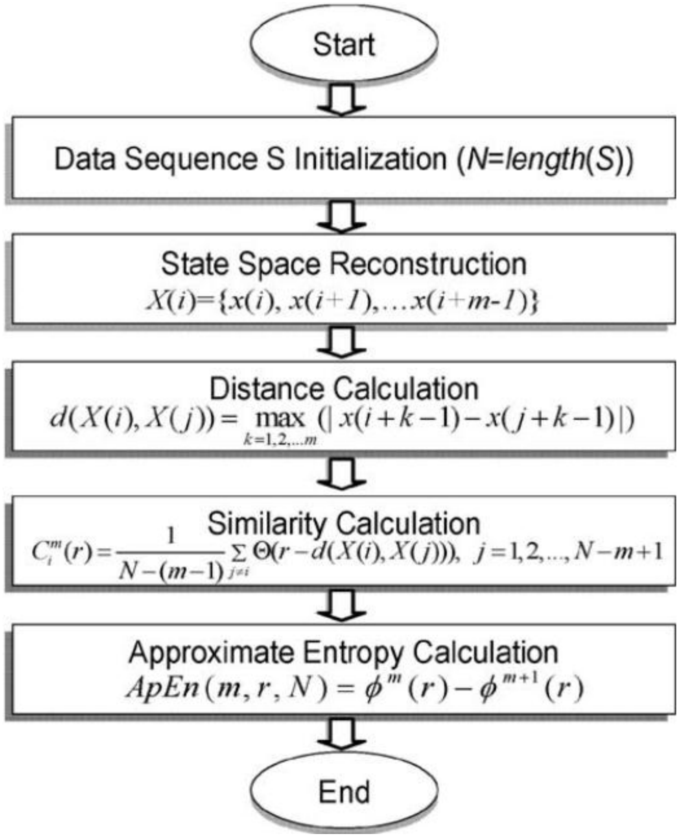


Figure 30: Steps in calculation of Approximate Entropy Here $\boldsymbol{X} = \boldsymbol{S}\ \boldsymbol{and}\ \boldsymbol{N} = \mathbf{n}$ (**Yan & Gao, 2007**)

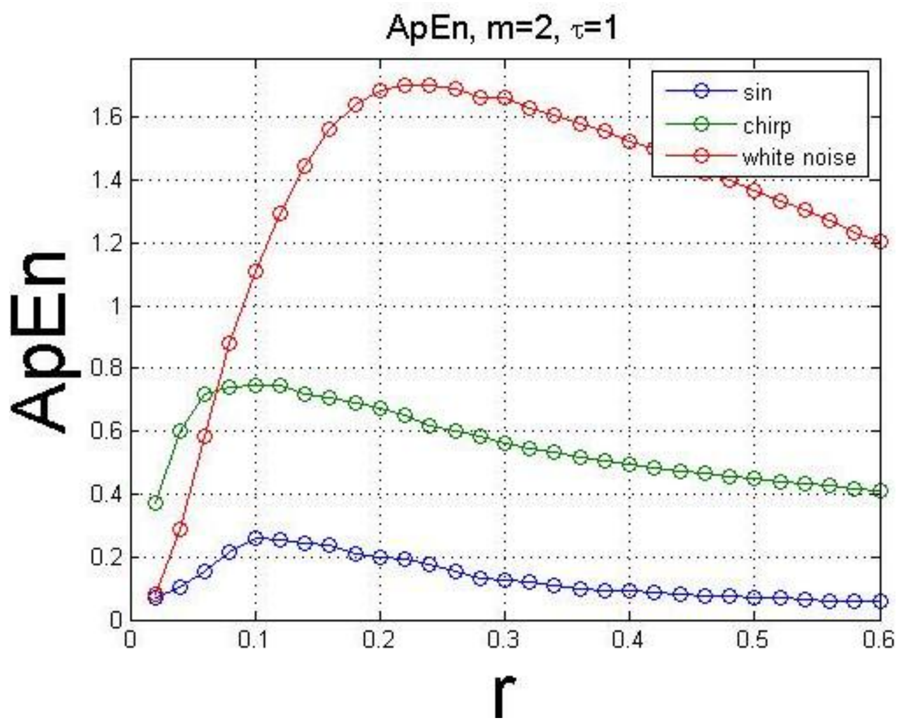


Figure 31: Approximate Entropy of simulated signals

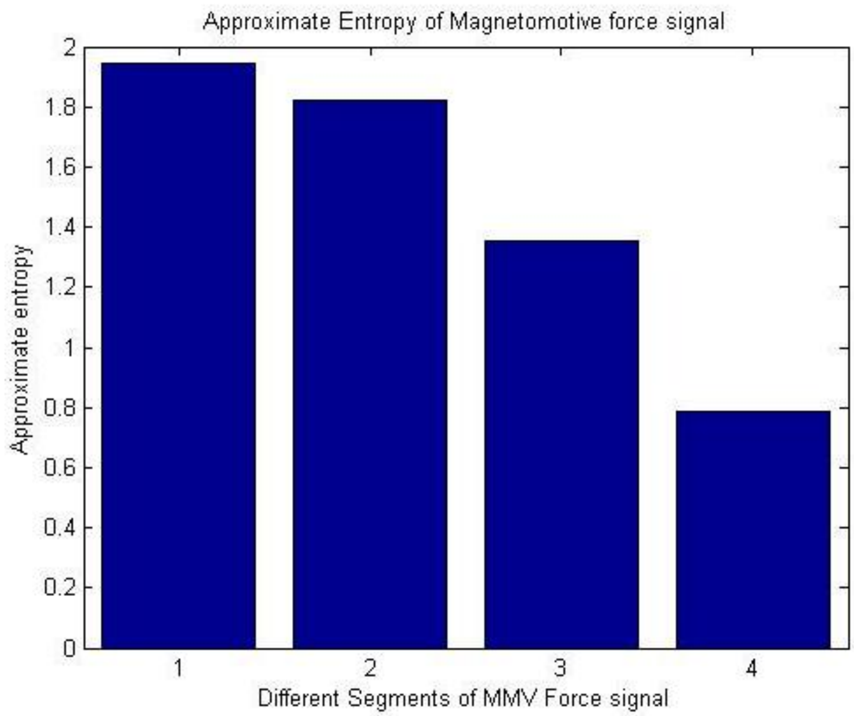


Figure 32: Approximate Entropy of Different Segments

Number of data sample in each segment = 102300

Down sampling	Embedded Dimension (m) = 1	Embedded Dimension (m) = 2	Embedded Dimension (m) = 3
1	Seg1 = 1.9653	Seg1 =1.9433	Seg1 =1.8835
	Seg2 =1.8581	Seg2 =1.8222	Seg2 =1.7826
	Seg3 =1.4435	Seg3 =1.3522	Seg3 =1.3253
	Seg4 =0.8342	Seg4 =0.7870	Seg4 =0.7698
3	Seg1 =2.2758	Seg1 =2.2343	Seg1 =1.9904
	Seg2 =2.2507	Seg2 =2.2177	Seg2 =1.9975
	Seg3 =2.0505	Seg3 =2.0034	Seg3 =1.7784
	Seg4 =1.5351	Seg4 =1.4739	Seg4 =1.3728
5	Seg1 =2.9773	Seg1 =2.6407	Seg1 =1.1461
	Seg2 =2.9717	Seg2 =2.6645	Seg2 =1.1529
	Seg3 =2.8626	Seg3 =2.4425	Seg3 =0.8684
	Seg4 =2.4023	Seg4 =2.1248	Seg4 =1.3155

Table 6: Case Study 1

Number of data sample in each segment = 60000

Down sampling	Embedded Dimension (m) = 1	Embedded Dimension (m) = 2	Embedded Dimension (m) =3
1	Seg1 = 2.0134	Seg1 =1.9812	Seg1 = 1.8862
	Seg2 = 1.8479	Seg2 =1.8072	Seg2 = 1.7525
	Seg3 = 0.5919	Seg3 =0.5884	Seg3 = 0.5796
	Seg4 = 0.4205	Seg4 =0.4471	Seg4 = 0.4397
3	Seg1 = 2.2970	Seg1 =2.2374	Seg1 = 1.8723
	Seg2 = 2.2415	Seg2 =2.1903	Seg2 = 1.8904
	Seg3 = 1.1146	Seg3 =1.0971	Seg3 = 1.0722
	Seg4 = 0.9535	Seg4 =0.9318	Seg4 = 0.8951
5	Seg1 = 2.9736	Seg1 =2.4731	Seg1 = 0.8412
	Seg2 = 2.9582	Seg2 =2.5164	Seg2 = 0.8321
	Seg3 = 1.9309	Seg3 =1.8521	Seg3 = 1.6137
	Seg4 = 1.8168	Seg4 =1.6800	Seg4 = 1.3733

Table 7: Case Study 2

500 randomly selected samples in each segment

Data Sets	Embedded Dimension (m) = 1
1	Seg1 = 1.4033
	Seg2 = 1.3092
	Seg3 = 1.1634
	Seg4 = 1.0160
2	Seg1 = 1.3788
	Seg2 = 1.3298
	Seg3 = 1.1634
	Seg4 = 1.1111
3	Seg1 = 1.3724
	Seg2 = 1.3256
	Seg3 = 1.1634
	Seg4 = 0.5696

Table 8: Test for minimum number of samples

b) Results & Discussions

In this investigation, first a toy test case simulation has been performed on three different kinds of signals to test the effectiveness of approximate entropy in differentiating patterns of regularity/complexity. Figure 31 shows the result of the approximate entropy calculated on the signals with different complexity namely, sine, chirp and white noise signal at different values of r with embedding dimension $m = 2$. It clearly shows that, predictability of white noise is less or complexity is more as compared to more periodic signals like sine or chirp. Approximate entropy of the different segments shown in Figure 25 of the net magneto-motive force (MMV) signal are calculated. For this investigation data is considered to be stationary and all other operating conditions are assumed to be same. It can be easily seen from the Figure 32 that approximate entropy of signal decreases as more and more debris is introduced in to the system. In order to test the robustness of the approximate entropy based feature to distinguish between healthy and unhealthy state, two different case studies were performed, results of which are summarised in the Table 6 & Table 7. Two different data sets were randomly selected from the segments shown in the Figure 25 and a scan of different embedding dimensions as well as different sampling rate was performed. It can be concluded that, a proper choice of embedding dimension as well as sampling rate is crucial for the success of this approach in different application. Hence along with information-theoretic features calculated in Section 3.1 approximate entropy can also be considered as one of discriminating features/information sources which may help in distinguishing between a faulty and non-faulty condition. The advantage of using approximate entropy as feature for system's health lies in the fact that it requires relatively less data samples for discriminating between healthy and unhealthy state, which is very useful in the real-time scenario. This is evident from the investigation performed on three different data sets as shown in the **Table 8**. The samples for each segments were chosen randomly from 3 different data sets. It can be observed that for a particular set of embedding dimensions and sample rate, the ApEn decreases continuously as the oil debris starts building up. The elapsed time for the calculation of the ApEn was 0.092527 seconds. 0.130959 seconds. 0.107242 seconds for different datasets respectively. These calculation were performed on MATLAB release 2013b, on a Windows 7 Professional PC with the following specifications.

Processor:	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz 3.40 GHz
Installed memory (RAM):	8.00 GB
System type:	64-bit Operating System

5.4 Summary

For the better decision making of the system's state, suitable sensor signals should be selected and information hidden in those sensor signals must be properly extracted in order to make intelligent decisions. Feature extraction is always a crucial step for information gathering as well as health monitoring of a system. Whenever any change or faults occur, most of the systems always manifest abnormal and sometimes nonlinear dynamic behaviour. Hence it is necessary to extract the features hidden in the sensory signals for more accurate health monitoring and diagnosis. In this chapter, feature extraction methods based on, Shannon entropy, mutual information, spectral entropy and approximate entropy are proposed and tested on the real-life test-rig designed to imitate the oil debris building up problem in a fuel metering valve of a gas turbine engine. The proposed techniques are found accurate as well as robust enough to distinguish between the healthy and unhealthy system provided some assumptions on the properties like stationarity of time series data are met The advantages of ApEn are its lower computational demand as it does not really needs the explicit calculation of the probability density functions and it can be applied in real time after properly understanding the effect from noise **(Rhea et al., 2011)** on the calculation of entropy because for a poor signal to noise ratio calculation of the pdf's is relatively more difficult as well as more computationally expensive step. In ApEn proper selection of the r parameter helps to improve the performance of the algorithm on data contaminated by experimental noise **(S M Pincus & Goldberger, 1994; Steven M Pincus, 1991)**, , since this parameter act as a filter parameter. Conceptually, if the r parameter is larger than the experimental noise, then the effect of the experimental noise on the analysis should be reduced **(Deffeyes, Harbourne, Stuberg, & Stergiou, 2011)**.

Disclaimer: These finding are only applicable to FMV calculation and may not be directly extended to another subsystems, components or system.

6 Trend Monitoring & Change Detection

6.1 Continuous trend/change monitoring

Monitoring the status of various systems and sub-systems in the gas turbine engine provides a unique and challenging environment for the design, function and use of sensor-based monitoring equipment. Trend or change monitoring is an integral part of the process together with the maintenance in an intelligent equipment health monitoring system. Trend and change monitoring is therefore a tool that provides early indication of changing state of health of the various important as well as critical system components, and allows for early intervention, but is also a means by which the effect of interventions and control actions may be recorded, evaluated and controlled. In the section below, first few very specific cases are discussed, which are often encountered in aerospace industry, especially in civil gas turbine engine logistics and maintenance cycle and later on various techniques are introduced, which can be used both in the case of batch or continuous trend monitoring of critical quantities. In this chapter all computations are performed on discrete-time data but that continuous-time definitions may be used for explanatory ease on occasion.

6.2 Dealing with system hierarchy and irregular events

As discussed in Section 1.5 a gas turbine engine can be described as a hierarchical system, most the decision about overall health of a gas turbine engine are taken based on the parameters/sensors measurements obtained from the top level system sensor e.g. Turbine gas temperature (TGT) margin but in order to design an efficient integrated health monitoring system one needs to include information about any change (online or offline) such as any maintenance event e.g. compressor on-line washing, information about any faults/change occurring at various levels of hierarchy e.g. information about low pressure turbine vibration data, fault occurring in fuel metering valve etc. before further analysis can be carried out and a comprehensive as well as concrete information about state of health of the system can be provided or predicted. Therefore, continuous trend monitoring as well as accurate detection of a change point in the states of system, system parameters or system health index and timely inclusion of this information plays an important as well as crucial role in designing any advanced health monitoring system.

In the following section, first a brief introduction of two different techniques, first for trend monitoring and second for change point detection is given as well as its suitability for application in two typical scenarios observed in the civil gas turbine engine health monitoring is tested. Later, an integrated prognosis approach **(Martha A Zaidan, R.Relan, Harrison, & Mills, 2014**) is proposed and is applied in combination with an existing prognosis method proposed by (**Skaf, Zaidan, Harrison, & Mills, 2013; MA Zaidan, Harrison, Mills, & Fleming, 2013; Martha Arbayani Zaidan, Mills, & Harrison, 2013**) to address the problem of predicting gas turbine engine global health index (TGT margin).

6.3 Spectral Entropy based trend monitoring

In this section, we consider the use of a running measure of power spectrum disorder to track changes continuously in the magnetomotive force signal acting on the fuel metering valve (FMV). Along with the usual tracking of the time domain statistical parameters such mean, standard deviation, variance and higher moments of the signal property, here we propose spectral entropy (SpEn) as another simple but robust way for the purpose of trend monitoring.

Spectral entropy (SpEn), describes the irregularity of the signal spectrum and is a normalized form of the famous Shannon's entropy defined in the previous section. It quantifies the spectral complexity of the time series. It makes use of the amplitude components of the power spectrum of the given signal as probabilities for entropy calculations. A variety of spectral transformations exist. Of these available transformations, the Fourier transformation (FT) is probably the most well-known transformation method from which the power spectral density (PSD) can be obtained/calculated. The PSD is a function that represents the distribution of power as a function of frequency. For analysing the frequency content of the signal $x(t)$, one might like to compute the ordinary Fourier transform ; however, for many signals of interest this Fourier transform does not exist. Because of this, it is advantageous to work with a truncated (continuous) Fourier transform $\widehat{x_T}(w)$, where the signal is integrated only over a finite interval $[0, Time]$: (**Brigham & Morrow, 1967; S. W. Smith, 1997)**

$$\widehat{x_{Time}}(w) = \frac{1}{\sqrt{Time}} \int_0^{Time} x(t)\, e^{-i\omega t} dt \qquad (25)$$

For the Discrete Fourier Transform the integration in the above equation will be replaced by the summation sign Then the power spectral density can be defined as below **(S. W. Smith, 1997)**:

$$S_{xx}(w) = \lim_{Time \to \infty} \mathbb{E}\left[\left|\widehat{x_{Time}}(w)\right|^2\right] \qquad (26)$$

For each frequency in the frequency band of interest, the power level P_f obtained from Fourier Transform is summed and then the total power, P_f is calculated. Normalization of P_f with respect to the total spectral power $S_{xx}(w)$, yields a probability density function. The power level at each frequency is divided by the total power $P_f = P_f / P_{Total}; P_{Total} =$ total power, which in the end yields the total, $P_f = 1$. After this, the Spectral Entropy is computed by multiplying the power at each frequency by the logarithm of the same power, $P_f * \log(P_f)$ and then multiplying the result by -1.**(Kannathal, Choo, Acharya, & Sadasivan, 2005).** Total entropy is the sum of entropy computed over entire frequency band of interest. Thus, the spectral entropy **(Kannathal et al., 2005)** is given by:

$$H_{SpEn} = \sum_f P_f * log_2 * (1/P_f) \qquad (27)$$

The ability of spectral entropy (SpEn) to show the irregularity is independent of amplitude or frequency of the signal. Spectral entropy has been applied in diverse disciplines for various applications. (**Misra & Ikbal, 2004**) applied spectral entropy based method to design automatic speech recognition system. (**Chechetkin & Lobzin, 2004**) applied Spectral entropy criteria for structural segmentation in genomic DNA sequences. Various entropy based criteria including the spectral entropy has been used for detection of epilepsy in EEG in (**Kannathal et al., 2005**). **(Martorano, Facco, Falzetti, & Pelaia, 2007**) used Spectral entropy assessment with auditory evoked potential in neuro-anesthesia. Time-varying spectral entropy is used by **(Papo, Caverni, Douiri, Podlipsky, & Baudonnière, 2007**) to differentiates between positive and negative feedback-related EEG activity in a hypothesis testing paradigm. Analysis of depth of anesthesia with Hilbert-Huang spectral entropy was performed by **(X. Li, Li, Liang, Voss, & Sleigh, 2008**). A spectral entropy based study to identify cardiac arrhythmias is carried out by (**Staniczenko, Lee, & Jones, 2009**). Entropy and complexity measures were used by **(Sabeti, Katebi, & Boostani, 2009**) for EEG signal classification of schizophrenic and control participants. Entropies based detection method for epileptic seizures was proposed by **(Pravin Kumar, Sriraam, Benakop, & Jinaga, 2010**). A hybrid spectral-entropy approach was used in **(Han, Muniandy, &**

Dayou, 2011) for acoustic classification of Australian anurans. Evaluation of spectral entropy was recorded by **(Morgaz et al., 2011**) to measure anaesthetic depth and antinociception in sevoflurane-anaesthetised Beagle dogs. A normalized spectral entropy related index is able to measure a part of the structural complexity of an ecological time series in **(Zaccarelli, Li, Petrosillo, & Zurlini, 2013**).

Figure 33 shows the evolution of the spectral entropy of the net magneto-motive force acting on the spool of the fuel metering valve (FMV). For the purpose of analysis, the segment of the signals measurement from the test rig are concatenated for emulating a scenario of the continuous operation of the FMV and oil debris build-up phenomenon. It can be easily observed that spectral entropy (SpEn) is clearly able to track the change in the nature of the magneto-motive force due build-up of the oil debris in the filter of fuel metering valve, hence it can be used as one of the system health feature to continuously monitor the trend/evolution of the signal under consideration.

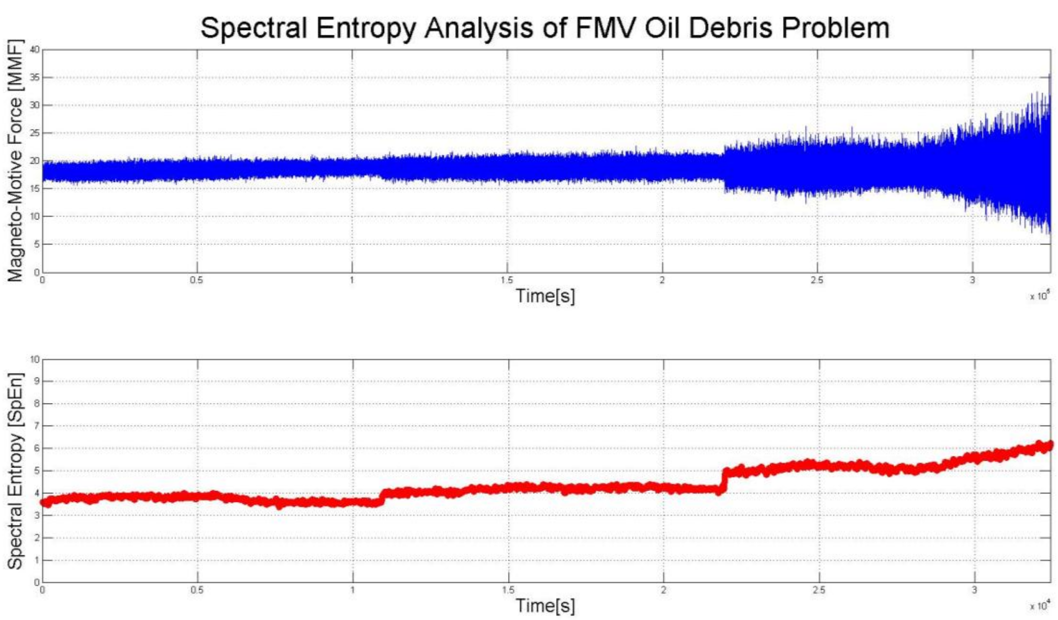


Figure 33: Spectral Entropy Analysis

6.4 Change point detection

The aim of change point detection (CPD) algorithm is to discover points at which sudden changes occur in a time-series data (**Kawahara & Sugiyama, 2012; Song Liu, Yamada, Collier, & Sugiyama, 2013**).These method can be classified based on the delay in detection: real-time detection or retrospective detection. Real-time detection is used for applications which require immediate response. On the other hand, retrospective detection can be used for applications, which can usually tolerate a bit longer reaction periods/time. As proposed by (**Song Liu et al., 2013**), latter algorithm tends to give more robust and accurate detection normally.

Here we propose a CPD method, named relative unconstrained least-squares importance fitting (RuLSIF) (**Song Liu et al., 2013; Yamada, Suzuki, Kanamori, Hachiya, & Sugiyama, 2013**), for the detection of anomalies (maintenance and step change/fault events), as mentioned earlier. In this approach, we do not require to estimate probability densities, such as kernel density estimation (**Brodsky & Darkhovsky, 1993; Csörg\Ho & Horvath, 1988**) as required in the last chapter, but instead estimates the ratio of probability densities directly **(Vapnik, 1998**). The following section summarise the main idea behind RuLSIF (**Yamada et al., 2013**) CPD algorithm,.

6.4.1 Change point detection by relative density-ratio estimation

Let us assume, z_t a time-series of data to be monitored for detecting changes in the statistical properties of the data. In the case of a gas turbine engine, the observed data can be either a global health index such as Thrust gas temperature(TGT Margin) or any sensor signal originating from various levels of hierarchy in the system architecture.

z_t is d-dimensional time-series sample at time t (**Yamada et al., 2013**), where

$$z_t = [z_1, z_2, z_3, z_4, \dots \dots, z_d\ ] \in \mathbb{R}^d. \tag{28}$$

Let us assume Z_t as a sample of time series at time t with length $\mathcal{K}$, given by (**Yamada et al., 2013**):

$$Z_t = [z_t', z_{t+1}', z_{t+2}', z_{t+3}', \dots \dots, z_{t+k-1}'] \in \mathbb{R}^{d\mathcal{K}} \tag{29}$$

Here $'$ is the transpose operator. Next, let $\mathbb{Z}_t$ be a set of n retrospective subsequence samples starting at time t (**Yamada et al., 2013**):

$$\mathbb{Z}_t = [z_t', z_{t+1}', z_{t+2}', z_{t+3}', \dots \dots, z_{t+n-1}'] \in \mathbb{R}^{d\mathcal{K}*n} \tag{30}$$

As illustrated in Figure 34, this forms a $d\mathcal{K} * n$ Hankel matrix, which is an important part of this algorithm, is used for change detection utilizing the concept of the subspace learning **(Kawahara & Sugiyama, 2012; Moskvina & Zhigljavsky, 2003**).

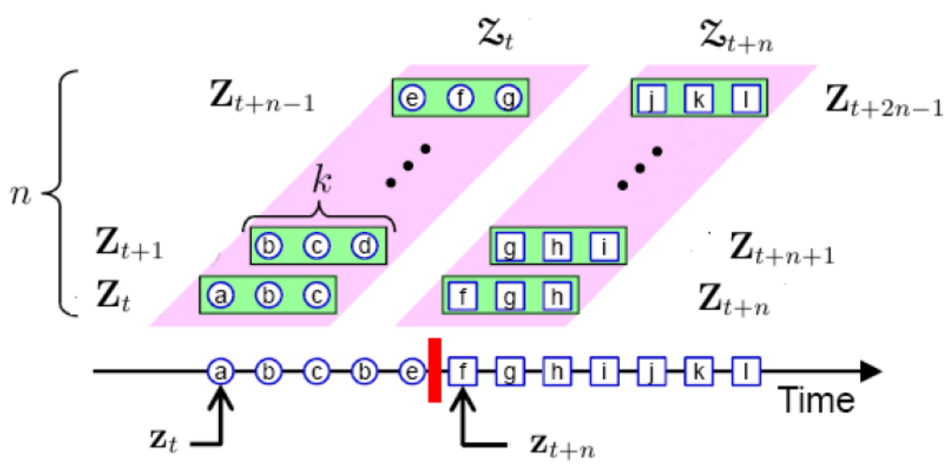


Figure 34: One-dimensional time-series data

The main idea behind this algorithm is to calculate a dissimilarity measure between two consecutive segments of the time series data, $\mathbb{Z}_t$ and $\mathbb{Z}_{t+n}$. The higher the dissimilarity measure, the more likely it is that, a significant change has occurred in the statistical property of the time series. In RuLSIF, the dissimilarity measure is defined by (**Yamada et al., 2013**):

$$Div(P_t||P_{t+n}) + Div(P_{t+n}||P_t), \tag{31}$$

Where P_t and P_{t+n} are denoted by $P(Z)$ and $P_\alpha^*(Z)$, are the probability distributions of the samples in sequence $\mathbb{Z}_t$ and $\mathbb{Z}_{t+n}$ respectively. Here "Z" represent argument of the probability density function. $Div(P_t||P_{t+n})$ is the α-relative Pearson divergence(PE) defined by (**Song Liu et al., 2013**) (**Yamada et al., 2013**),

$$PE(P||P^*) = \frac{1}{2}\int P^*(Z)\left(\frac{P(Z)}{P_\alpha^*(Z)} - 1\right)^2 dZ \tag{32}$$

which is a special case of f-divergence **(Ali & Silvey, 1966b; Csiszár, 1967)**, where f is a convex function.

$$Div(P_t||P_{t+n}) = \int P^*(z) f\left(\frac{P(z)}{P_\alpha^*(z)}\right) dZ, \tag{33}$$

$$\begin{aligned} Div(P_t||P_{t+n}) &= PE_\alpha(P||P^*) \\ &= \frac{1}{2} PE_\alpha(P||\alpha P + (1-\alpha)P^*) \\ &= \frac{1}{2}\int P_\alpha^*(Z)\left(\frac{P(Z)}{P_\alpha^*(Z)} - 1\right)^2 dZ \end{aligned} \tag{34}$$

Where $\alpha-$ relative Pearson divergence (PE) measures the difference between two probability distributions $P(Z)$ and $P_\alpha^*(Z)$, for $0 \leq \alpha < 1$. Where $P_\alpha^*(Z) = \alpha P(Z) + (1-\alpha)P^*(Z)$ is the $\alpha-$mixture density and the $\alpha-$relative density ratio is bounded above by $1/\alpha$ for $\alpha > 0$. The $\alpha-$ relative density ratio,r_α is defined by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$r_\alpha(Z) = \left(\frac{P(Z)}{\alpha P(Z) + (1-\alpha)P^*(Z)} = \frac{P(Z)}{P_\alpha^*(Z)}\right) \qquad (35)$$

The expectation of $f(Z)$ under $P(Z)$ is denoted by $\mathbb{E}_{P(Z)}[f(Z)]$, given by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\begin{aligned} Div(P_t||P_{t+n}) &= PE_\alpha(P||P^*) \\ &= \frac{1}{2}\int P_\alpha^*(Z)(r_\alpha(Z)-1)^2 dZ \\ &= \frac{1}{2}\int \mathbb{E}_{P_{\alpha^*(Z)}}(Z)[(r_\alpha(Z)-1)^2] \end{aligned} \qquad (36)$$

From equation (36), it can be concluded that, only the density ratio needs to be calculated. Furthermore, the density ratio can be expressed as a kernel, described below (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$r_\alpha(Z) = g(Z) = \sum_{l=1}^{n} \theta_l K(Z, Z_l) \qquad (37)$$

Where, where $\theta = (\theta_1, \theta_2, \theta_{3 \ldots\ldots} \theta_n)'$are learning parameters and $K(Z, Z_l)$ is a Gaussian kernel, defined by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$K(Z, Z_l) = e^{\left(-\frac{\|Z-Z^*\|^2}{2\gamma^2}\right)} \qquad (38)$$

Here $\gamma > 0$ is the kernel width, which is determined based on cross validation. The parameters θ can be learned by minimising the squared loss between true relative ratio, $r_\alpha(Z)$, and estimated relative ratio, $\hat{g}(Z)$, given by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$J(Z) = \frac{1}{2}\int P_{\alpha}^{*}(Z)(r_{\alpha}(Z) - \hat{g}(Z))^{2}dZ$$
$$= -\,\mathbb{E}_{P(Z)}[\hat{g}(Z)] + \frac{\alpha}{2}\mathbb{E}_{P(Z)}[\hat{g}(Z)^{2}] + \frac{1-\alpha}{2}\mathbb{E}_{P(Z)}[\hat{g}(Z)^{2}] + constant. \quad (39)$$

By replacing $\hat{g}(Z)$, by a kernel, in equation (39), and approximating the expectations by their empirical averages, RuLSIF optimisation problem can be reformulated as (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\frac{min}{\theta \in R^{n}} \left[\frac{1}{2}\theta^{*}\hat{H}\theta - \hat{h}'\theta + \frac{\lambda}{2}\theta^{*}\theta\right] \quad (40)$$

Where, $\frac{\lambda}{2}\theta^{*}\theta$ is the penalty term for the regularization, ($\lambda > 0$) is a regularization parameter. Parameters $\hat{v}$ is the n −dimensional vector with the l^{th} element given by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\hat{v}_{l} = \frac{1}{n}\sum_{i=1}^{n} K(Z_{i}, Z_{l}) \quad (41)$$

And $\hat{V}$ is the $n * n$ −matrix with the $(l, l^{*})^{th}$ element given by (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\hat{V}_{l,l^{*}} = \frac{\alpha}{n}\sum_{i=1}^{n} K(Z_{i}, Z_{l})K(Z_{i}, Z_{l^{*}}) + \frac{1-\alpha}{n}\sum_{j=1}^{n} K(Z_{j}{}^{*}, Z_{l})K(Z_{j}{}^{*}, Z_{l}{}^{*}) \quad (42)$$

Hence, the analytic solution of the equation (40) can be found by solving the following problem (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\hat{\theta} = \left(\hat{V} + \lambda I_n\right)^{-1} \hat{v} \tag{43}$$

Where, I_n is the n-dimensional identity matrix. Finally, a density ratio estimator can be written as (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$r_\alpha(Z) = \hat{g}(Z) = \sum_{l=1}^{n} \hat{\theta}_l K(Z, Z_l) \tag{44}$$

when $\alpha = 0$, this method is reduced to unconstrained least-squares importance fitting (uLSIF) described in **(Kanamori, Hido, & Sugiyama, 2009**). Now further this density ratio estimator can be for change point detection of a time series. By substituting the equation (44) in the equation (36) the α −relative Pearson divergence can be expressed as (**Song Liu et al., 2013**) (**Yamada et al., 2013**):

$$\begin{aligned} \widehat{PE}_\alpha = &-\frac{\alpha}{2n}\sum_{l=1}^{n} \hat{g}(Z_i)^2 \\ &-\frac{(1-\alpha)}{2n}\sum_{j=1}^{n^*} \hat{g}(Z_i^*)^2 \\ &+\frac{1}{n}\sum_{i=1}^{n} \hat{g}(Z_i) - \frac{1}{2} \end{aligned} \tag{45}$$

The indicator score/value obtained from the (symmetric version, equation (31) is used as basis of change point in time series (e.g. TGT margin, vibration signals etc.). The main advantages of using this technique for change/anomaly/fault detection (for gas turbine engine problems here) over previously proposed techniques can be summarised as below:

- This approach is quite simple **(Vapnik, 1998)**, as it does not require the estimation of probability density function explicitly (e.g. kernel density estimation for each segment of the sequence, as described in the previous chapter), but this approach estimates the ratio of probability densities directly without going through density estimation step. In this regard, this technique can easily be adapted to be used in quasi-real time scenarios.

- The acquired data (global health index of a gas turbine engine like TGT margin as well as other sensor signals (covariates) from different level of hierarchy) normally have large amount of noise attached to it. This method is based on a non-parametric method, where it does not need to assume a specific functional form for the distribution of the data sequences. Therefore, such models are better equipped to handle the model's complexity as well as uncertainty(**Bishop, 2006; Rasmussen & Williams, 2006**).

- The solution of this problem can be computed analytically **(Kanamori et al., 2009; Song Liu et al., 2013; Yamada et al., 2013**), hence can be used for real time applications.

- The basic idea of RuLSIF is to consider relative density ratios, which are smoother and always bounded from above (**Yamada et al., 2013**). [3]

[3] Thresholds and value of α needs to be defined a-priori in order to determine maintenance event as well as covariate changes, therefore density ratios must be bounded for detecting change point.

6.5 Application of Change point detection

In the section below two specific case studies have been performed which uses the methodology shown in the Figure 39 . An event/change detection algorithm based on direct density ratio estimation has been applied.

- To detect maintenance events in recoverable systems.
- To low pressure turbine vibration signal. This information about change in vibration signal has later been combined with TGT margin (Global health index) in order to do better prognosis and calculation of remaining useful life of the gas turbine engine.

a. Change detection for recoverable system

An aerospace gas turbine engine is a recoverable system which means its performance can be improved by a properly scheduled maintenance action **(Schneider, Demircioglu Bussjaeger, Franco, & Therkorn, 2010**). Elements of gas turbine degradation, such as compressor fouling, are recoverable through compressor washing. These actions increase the useful life of a gas turbine engine and help in optimizing the performance of the gas turbine over a longer period. However, these maintenance actions are performed by a separate organization to those performing fleet management, leading to uncertainty in the maintenance state of the asset (**Skaf et al., 2013**). One of the ways to include information about such maintenance events is by detecting accurately these maintenance events directly from the measured service data of a global health index like e.g. Turbine gas temperature (TGT) margin. A change point detection algorithm can be used to detect a significant increase in TGT margin. The event detection information can be subsequently passed on to a prognosis algorithm. Figure 35 shows a general concept how the change point detection algorithm provides information to reset the prognostic algorithm. The change point detection is applied directly to TGT margin to detect the significant change in the degradation.

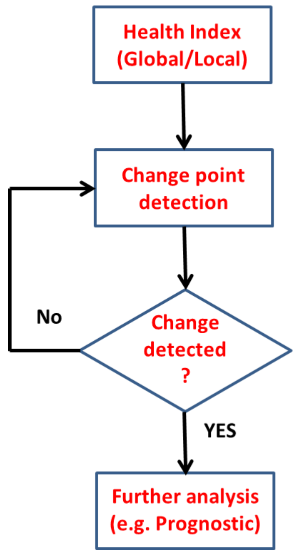


Figure 35: Concept of a change point detection for recoverable system

Figure 36 shows an example of TGT margin data and the score of change-point detection algorithm. The top figure shows the real TGT margin (Blue +) and its ground truth (Red solid), respectively. The bottom figure shows the result of the change point detection algorithm based on direct density ratio estimation **(Kawahara & Sugiyama, 2012; Song Liu et al., 2013**).

A significant increase in change point score e.g. > 0.6 (chosen heuristically) in this present case can be observed clearly at time index around 30, which indicates the occurrence of a maintenance event.

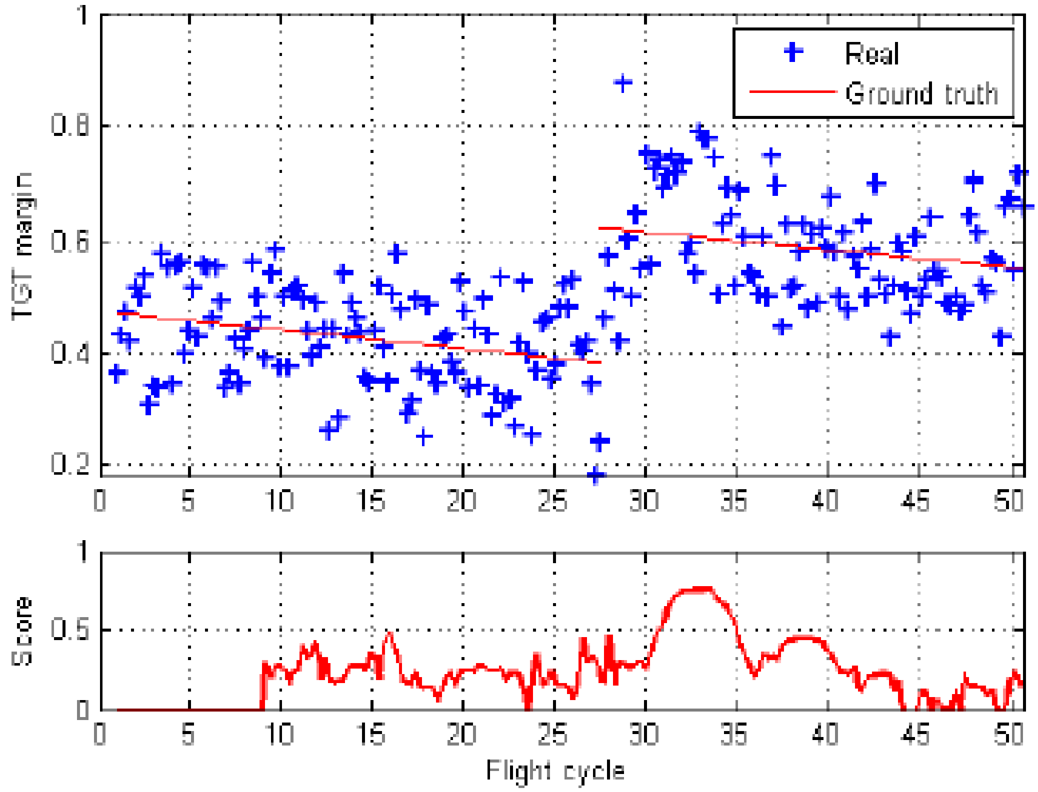


Figure 36: Change point detection in TGT Margin data after maintenance event(**Martha A Zaidan, R.Relan, Harrison, & Mills, 2014**)

b. Change detection for capturing variation in slope of global health index

During the condition monitoring of a gas turbine engine a considerable change in slope of degradation of the global health index like TGT margin can be observed. This change in slope of degradation can be due to various factors (covariates) such as operating conditions, due to occurrence of an event (abnormal behaviour, a shift change in any local health index) and faults occurring at local or global level of gas turbine engine hierarchy. This section discusses two unique concepts using two different covariate effects.

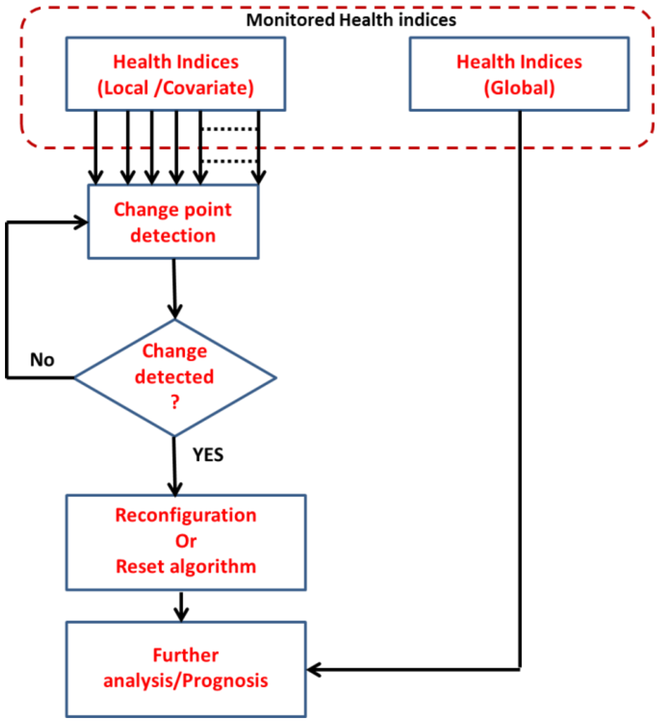


Figure 37: Incorporating information from covariates

There are several key operational factors/parameters (covariates) which are used in order to monitor the performance of an engine, including engine's operating speed, temperature, pressure, fuel flow and vibration levels (**Ackert, 2010**) at different stages in a gas turbine engine. In this case study, after deep discussions with subject matter experts the vibration of lower pressure turbine shaft is used as a main factor/covariate. As Turbine Gas Temperature (TGT) is measured near the exit station of low pressure turbine, hence it also makes vibration of low pressure turbine shaft a good candidate for further analysis. Figure 38 shows how detecting the change in vibration level correlate with the rapid change in TGT margin. The top subfigure illustrates the real TGT margin and its ground truth, represented by cross and solid line, respectively. The crossing of vertical dash line with the TGT margin data represent the point where the slope of the degradation starts to change. From the Figure 38 it can be seen that the first phase of normal deterioration is represented by time span between flight cycle 0 and 45, while the second phase of rapid deterioration is represented by time span after flight cycle 45. The middle figure shows the measured low pressure turbine vibration signal, whereas the bottom figure is the scores (which represent the level of change occurred) obtained from change point detection algorithm.

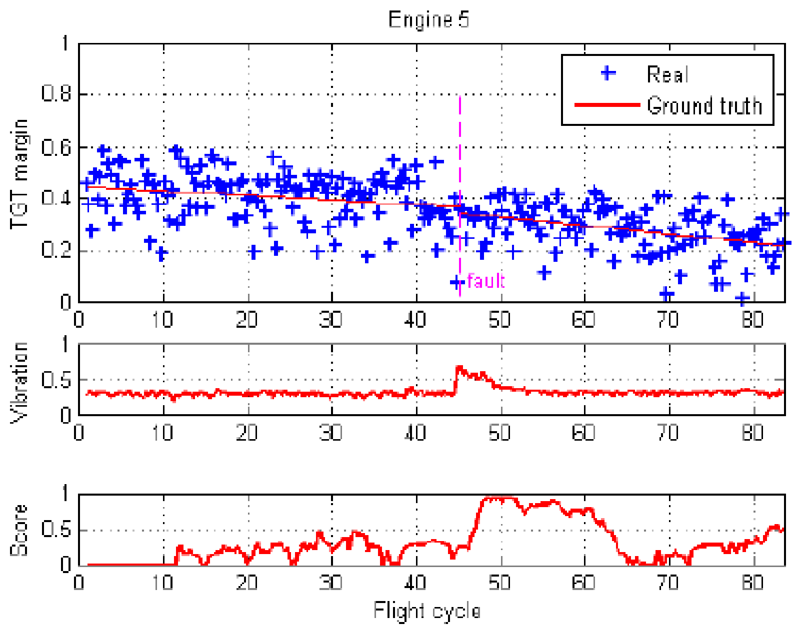


Figure 38: Change point detection using information from covariates(**Martha A Zaidan, R.Relan, et al., 2014**).

A threshold value is calculated heuristically based on offline implementation of change point detection on multiple sets of vibration (covariate) data From several experiments, this threshold can be specified. The value of this threshold will depend on fleet to fleet data, chosen subsystem/component as well as the chosen covariate. A lower threshold may be chosen carefully and only if when the signal to noise ratio is good. A properly chosen threshold would eventually decide the impact of whole prognostic algorithm on the RUL as one of the task of CPD is use the information about this threshold to calculate the indicator score and eventually inform the main prognostic algorithm. Figure 37 illustrates a block diagram of general concept for including obtained by applying the change detection algorithm on different factors/covariate affecting the performance of the engine. When a change point detection algorithm detects the change in one of the factors, that information is passed on to other algorithm for further analysis.

A specific case study of prognosis of TGT margin utilizing the concept developed above and Bayesian hierarchical modelling prognosis concept developed by (**MA Zaidan, Harrison, et al., 2013; MA Zaidan, Mills, & Harrison, 2013**) is discussed in the section below and also can be found in **(Martha A Zaidan, R.Relan, et al., 2014**).

c. Integrated prognostics: Combining change point detection with remaining useful life calculation

In the proposed integrated prognostics approach **(Martha A Zaidan, R.Relan, et al., 2014**) several monitored signals, e.g. global health index and other covariates, are fed into a prognostic algorithm proposed by **(Martha a. Zaidan, Harrison, Mills, & Fleming, 2014**) to estimate RUL of gas turbine engine as shown in the Figure 39. At the same time, these parameters are also monitored continuously by CPD algorithm. If CPD algorithm detects significant increase in TGT margin (the indicator score is bigger than the defined threshold), it considers that maintenance action is just performed and prognostic algorithm should be restarted. Furthermore, whenever the CPD algorithm detects abnormality in one of monitored systems parameters/covariates (the indicator score is bigger than the defined threshold), it considers that there is a fault or other changes in slope of health index (degradation parameters) due to various operating conditions, which would affect the degradation process. The prognostic model is then reset or reconfigured to improve the RUL estimation

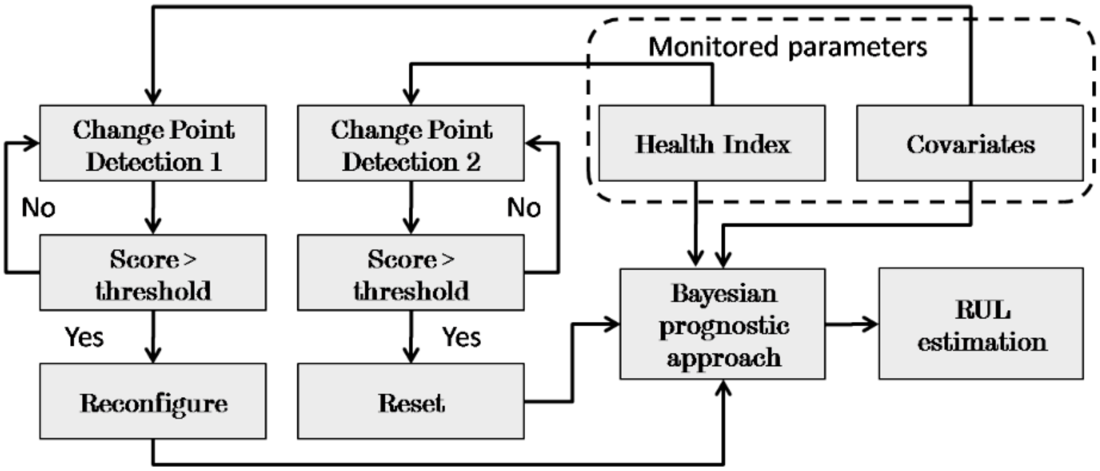


Figure 39: Integrated prognostics: Combining Bayesian approach and CPD(**Martha A Zaidan, R.Relan, et al., 2014**)

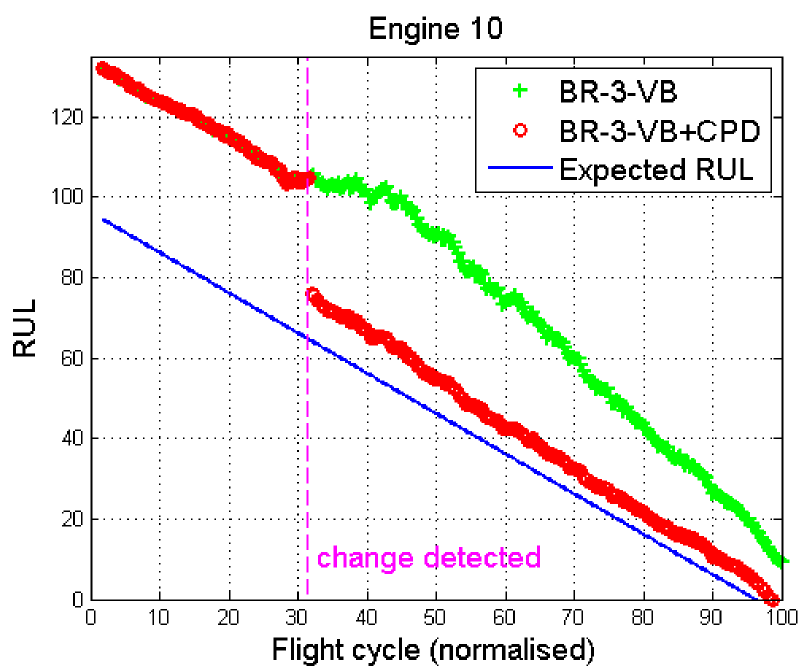


Figure 40: Result of CPD + Bayesian Algorithm(**Martha A Zaidan, R.Relan, et al., 2014**)

Figure 40 illustrates the results of a remaining useful life prediction problem where the slope of degradation parameters/ health index under consideration is decays rapidly. It clearly shows that the integrated prognostics approach performs much better as compared to the previously proposed approach BR-3 (Variational Bayesian) and it converges very fast to the “expected” RUL. The flexibility of resetting or reconfiguring the prognostics algorithm at the onset of a significant change in the degradation parameters/health index or a change in the covariate, results in better RUL estimation.

6.6 Summary

This chapter presents two different techniques for generic problems like continuous trend monitoring and change point detection. The results described in this chapter are partially a result of joint work done in collaboration to propose an integrated prognostic approach. The main contribution of the author in are:

- Development of **spectral entropy** based trend monitoring scheme.
- Development and proposal of **direct-density ration based change point detection (CPD) technique** for the detection of change in global health index or any of the other systems parameters.
- Development of **methodology for integrated prognostic approach**. [4]

This chapter discusses, how the system health monitoring systems can make use information available from continuously tracking the

[4] *The result of this approach are produced as a scientific paper (jointly co-authored) and are under review in the journal of “**Experts systems with applications**”, Elsevier, 2014.*

change in the status of a health parameter of an asset as well as how the prognostic performance can be improved by utilising the information about irregular events, such as maintenance event, slope change in degradation and by including that information into the prognostic framework.

The proposed integrated prognostic concept is promising for use in a complex hierarchical system. This method is able to detect any changes or faults in multiple covariates (e.g. vibration, ambient temperature) at any level of the system's hierarchy, including sub-system as well as component level. The challenge is to select optimally the information from multiple covariates which may affect the degradation of health parameter indirectly as well as performance of prognostics algorithm.

Disclaimer: The simulation and calculation were performed on MATLAB release 2013b, on a Windows 7 Professional PC with the following specifications.

Processor:	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz 3.40 GHz
Installed memory (RAM):	8.00 GB
System type:	64-bit Operating System

7 Conclusions

This chapter summarises several major conclusions which can be drawn from this thesis. The main objective of this research is to identify the technological gaps in the existing state-of-the-art Equipment Health Management (EHM) system as well as present day's "Sense-Acquire-Transfer-Analyse-Act Paradigm" of Rolls-Royce Engine Health Monitoring system.

The methodologies, framework as well as algorithms proposed in this thesis aim to deal with various challenges, which arise in developing a robust and intelligent health monitoring system for complex systems such as civil gas turbine engines. Selection of an appropriate framework/methodology/algorithm for a particular application is crucial to the ultimate success of an Equipment Health Management (EHM) system. Therefore, this requires a good understanding of challenges associated with a specific application in such a complex system in order to make intelligent decisions on-board or off-board regarding the health of a system.

As described in before, Monitoring systems technologies log the actions, performance and status of the components in the electrical and control systems. They collect data from some sensor signals deemed indicative of performance and mechanical elements of the engine, which are then used to draw certain conclusions, based on algorithms programmed into the monitoring system. The aim of a monitoring system is to maximize availability and minimize operational disruption.

In **chapter 1**, a brief background on the existing health monitoring as well as maintenance paradigms is provided. A clear motivation for developing a new framework is stated. Along with the commercial benefits of the, various advantages are also discussed, which support the idea of developing an intelligent health monitoring system. A brief introduction of the working of gas turbine engine as well as state-of-the-art gas turbine engine equipment health monitoring system is given. In the end, outline and contribution of the thesis are stated.

In **chapter 2**, a thorough analysis of existing state-of-art Equipment Health Management (EHM) system has been performed and various existing technological gaps and bottlenecks in architecture/framework of EHM have been pointed out. A thorough analysis based on exhaustive literature review of the factors affecting the next generation integrated Equipment Health Management (EHM) system has been performed and technological as well as scientific impact of incorporating various suggested techniques and technologies in to the new Equipment Health Management (EHM) system architecture has also been discussed.

To overcome these bottlenecks and fill those technological gaps discussed in previous chapter, in **chapter 3**, a concise and generic framework for health monitoring of complex systems such as gas turbine engine has been developed. A clear emphasis is put on the use the transient information (along with the steady-state information) in designing a framework for equipment health monitoring system. It is also clearly stated, how and when this transient information can be collected for gathering better information about the state of health of the asset under consideration. To overcome the fundamental issues associated with any closed loop system such as its sensing capabilities, data acquisition, data selection, data transmission and analysis a clear picture has been presented, how and where the above proposed approaches would fit in or contribute/ extended the “Sense-Acquire-Transfer-Analyse-Act Paradigm” of Rolls-Royce Engine Health Monitoring system.

In order to implement and test some of proposed methods, a case study identification study has been performed for choosing a suitable system/subsystem or component in **chapter 4**. A clear pathway is proposed to design a methodology which combines information from various sources such as knowledge about high value fault based on Failure mode effect cause analysis (FMECA), operational phases of flight, expert knowledge and knowledge about the working of a particular sub-systems. Based on this approach the three sub-system were selected as candidate sub-system, out of which fuel metering value (FMV) was finally selected for further investigation.

In order to gain as much information as possible, about the state of system/asset under consideration and later, to classify between the faulty and non-faulty state, in **chapter 5** some feature extraction methods (keeping in mind their applicability in real-time scenarios) based on information theory e.g. entropy, mutual information and complexity theory e.g. approximate entropy have developed and applied to the data obtained (in batch mode scenario) from a test rig developed to simulate the scenario of fault caused by debris build up in the fuel metering servo-valve.

To address the fundamental problem of extracting a much information as possible about the state of the system under consideration (be it global or local system) and later use this information in order to better prognosis about the state/remaining useful life of the system, in **chapter 6**, an information-theoretic spectral entropy based method is proposed for trend monitoring of a signal/health index. Later on, a relative direct-density ration based change point detection algorithm has been developed and is also applied to problem of prognosis of recoverable system and to detect change in the state of global health index by incorporating the information about the state of subsystem at a local level.

8 Future Work

Despite the proposed framework for intelligent data collection/generation as well as the extended (closed loop) version of the "Sense-Acquire-Transfer-Analyse-Act" paradigm and the proposed algorithms dealing with information collection/feature selection and anomaly detection problem prove to be promising, there are still a number of research areas, which are required to be improved and technological gaps which need to be filled, in order to design a next generation equipment health monitoring system for the future gas turbine engines. This section provides several concrete suggestions for the future work. To develop a robust and effective health monitoring system, following aspects related to the method use must be investigated:

1. Effect of data length

As the proposed algorithm based on the information theory require calculation of the probability density function, the length of available data samples therefore play a really important part in the overall success of the algorithm. Hence, it must be ensured that there are sufficient number of data samples available before quantities like entropy, mutual information can be accurately enough calculated/estimated. Effect of data length on robustness of feature extracted to distinguish between healthy and unhealthy state of system under consideration must be thoroughly investigated.

2. Feature Extraction and Selection of dominant features

Although this thesis proposes a few different methods based on information theory and complexity theory, to distinguish between the healthy and unhealthy state of system under consideration. There are multiple other ways to extract different features from the signal. Once a set of different features have been extracted, the selection of dominant features for good classification performance is very critical to design a robust and intelligent equipment health monitoring system.

3. Further application of the Change point detection algorithm

Application of the **Change point detection algorithm** to develop a **recommender system** for better data acquisition and sensor configuration.

a. Reconfiguration of sensor or data acquisition system

As already discussed in Section 1.6 that often the most time-consuming and costly task in any scientific investigation is the gathering of data. In order to design an efficient integrated health monitoring system, one needs to address, one of the fundamental the issues of data collection i.e.
When and how do we need to measure?

According to the framework proposed in Section 2 one can acquire the data in following ways:

a. During operation
b. During an event e.g. change detection problem
c. specific tests (Active Fault diagnosis or optimal input design for system identification)

In the section 4.4., data was collected from a test rig imitating the conditions of the civil aircraft during a normal operating flight cycle. The change point information gathered at different level of system's hierarchy can act as a triggering point for the data acquisition at higher sampling rates and/or for longer acquisition times in order to do better data analysis at a later stage or can be combined with information already available at different levels of hierarchy to do better prognosis about the remaining useful life of a system/sub-system or a component.

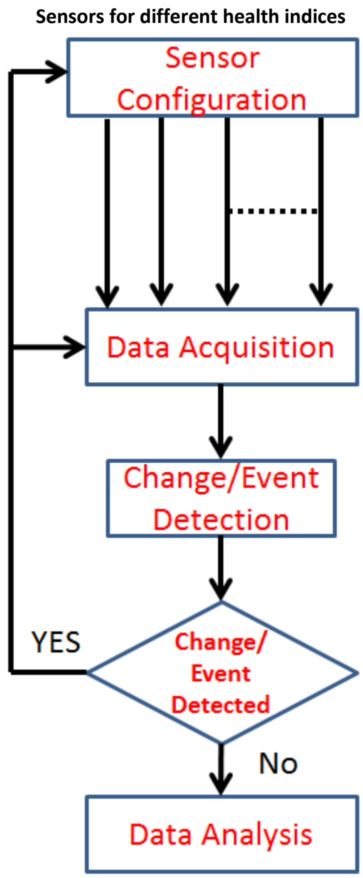


Figure 41: Change detection for data collection

In the Figure 41, on the similar principle a methodology is proposed based on the change point detection algorithm to tackle with other fundamental problem of data collection in any health monitoring system. As shown above, a change detection algorithm can be used to trigger and/or reconfigure the data acquisition system and/or sensory system

4. Development of Information transfer algorithm

As discussed above, many complex engineering like gas turbine engine have different level of hierarchy as well as modular structure, a change at one of level of hierarchy can affect the dynamics at other level of hierarchy. There might or might not be direct interaction between different levels, but one can try to track or detect the changes as well as information transfer between various sensors (on-board and/or virtual sensors) in the different level of hierarchy during a faulty and normal operating condition. There are various techniques employed in neuroscience to track the changes in brain by observing the response at different part of brain by stimulating in the different and tracking the change in information of the brain activity or by quantifying the information flow. Hence, field of causality research is a strong candidate for application in to fault diagnosis etc.

5. Development of an Intelligent Fault Diagnosis and Prognosis

A critical part of developing and implementing an effective as well as reliable health monitoring system is actually based on the ability to detect faults/anomalies in early enough stages/phases of system's operation and later on to do something useful with the acquired information. Fault isolation and diagnosis process uses these detection events as the start of the process for classifying the fault for the system being monitored **(Vachtsevanos et al., 2006)** Condition monitoring and/or failure prognosis then forecasts the remaining useful life (the approximate operating time between the detection of the any anomaly or fault and an unacceptable level of degradation of any of the system's health parameters). If the identified anomaly or fault affects the life of any critical component, then the failure prognosis models also must take in to consideration and should reflect this diagnosis. As a minimum, the following probabilities should be used to specify fault detection and diagnostic accuracy according to (**Vachtsevanos et al., 2006**):

- "*The probability of anomaly detection, including false-alarm rate and real fault probability statistics*".
- "*The probability of specific fault diagnosis classifications using specific confidence bounds and severity predictions*".

To specify the accuracy of any prognostic algorithm, the developer/end user must first define following points:

- The level of condition/health parameter degradation beyond which operation of the asset/system is considered unsatisfactory or undesirable.
- A minimum amount of warning time which should provide the operator and maintainer, the required information so that he/she can be act before the onset of actual failure.
- A minimum probability level that remaining useful life will be equal to or greater than the minimum warning level.

Such an approach will involve synergistic deployments of component health monitoring technologies, as well as integrated reasoning capabilities for the interpretation of fault-detect outputs as described in (**Khawaja, Vachtsevanos, & Wu, 2005; Tang et al., 2008; Vachtsevanos et al., 2006**). Furthermore, it will also involve the introduction of various learning (machine) technologies to support the continuous improvement of the reasoning capabilities. Finally, a plausible maintenance and logistics architecture is required that can govern integration and interoperation within the system, for example between various on-board elements and their counterpart ground-based support functions, as well also between various health management system functionalities, external maintenance and operation functions. These kind of Decision-Making Support Systems (DMSS) are mostly computer-based systems, which can support any individual or an organisational decision-making processes. Recent advances in information technology and artificial intelligence especially machine learning are continuously enhancing the capabilities of these systems and giving rise to so called, intelligent-DMSS. A typical intelligent-DMSS will usually contain various decision layers, generally making decision about sensor placements as well as sensing strategies, signal processing and database management systems, fault diagnosis, fault prognosis and logistics etc.

9 References

A. C. Atkinson and D. R. Cox. (1974). Planning Experiments for Discriminating Between Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, *36*(3), 321–348.

Abásolo, D., Hornero, R., Espino, P., Poza, J., Sánchez, C. I., & de la Rosa, R. (2005). Analysis of regularity in the EEG background activity of Alzheimer's disease patients with Approximate Entropy. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *116*(8), 1826–34. doi:10.1016/j.clinph.2005.04.001

Ackert, S. (2010). Engine Maintenance Concepts for Financiers. *Macquarie AirFinance, Mar*.

Akaike, H. (1969). Power spectrum estimation through autoregressive model fitting. *Annals of the Institute of Statistical Mathematics*, *21*(1), 407–419. doi:10.1007/BF02532269

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1100705

Alberton, A. L., Schwaab, M., Lobão, M. W. N., & Pinto, J. C. (2011). Experimental design for the joint model discrimination and precise parameter estimation through information measures. *Chemical Engineering Science*, *66*(9), 1940–1952. doi:10.1016/j.ces.2011.01.036

Ali, S. M., & Silvey, S. D. (1966a). A General Class of Coefficients of Divergence of One Distribution from Another. *Journal of the Royal Statistical Society. Series B (Methodological)*, *28*(1), 131–142. Retrieved from http://www.jstor.org/stable/2984279

Ali, S. M., & Silvey, S. D. (1966b). A General Class of Coefficients of Divergence of One Distribution from Another. *Journal of the Royal Statistical Society. Series B (Methodological)*, *28*(1), 131–142.

Aloimonos, J., Weiss, I., & Bandyopadhyay, A. (1988). Active vision. *International Journal of Computer Vision*, *1*(4), 333–356. doi:10.1007/BF00133571

Andjelkovic, I., Sweetingham, K., & Campbell, S. L. (2008). Active fault detection in nonlinear systems using auxiliary signals. *2008 American Control Conference*, (v), 2142–2147. doi:10.1109/ACC.2008.4586809

ARIMOTO, S., & KIMURA, H. (1971). Optimum input test signals for system identification—an information-theoretical approach. *International Journal of Systems Science*, *1*(3), 279–290. doi:10.1080/00207727108910011

Atkinson, A. C. (2001). One hundred years of the design of experiments on and off the pages of Biometrika. *Biometrika*, *88*(1), 53–97. doi:10.1093/biomet/88.1.53

ATKINSON, A. C., & FEDOROV, V. V. (1975). The design of experiments for discriminating between two rival models. *Biometrika*, *62*(1), 57–70. doi:10.1093/biomet/62.1.57

Awwad, B., Hasan, S., Dyson, M., Balli, T., & Gan, J. Q. (2008). A Study via Feature Selection on the Separability of Approximate Entropy for Brain-Computer Interfaces, (m), 189–194.

Bahl, L., Brown, P., de Souza, P., & Mercer, R. (1986). Maximum mutual information estimation of hidden Markov model parameters for speech recognition. In *ICASSP '86. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 11, pp. 49–52). Institute of Electrical and Electronics Engineers. doi:10.1109/ICASSP.1986.1169179

Bajcsy, R. (1985). Real-time obstacle avoidance algorithm for visual navigation. In *Proc. of the 3rd Workshop on Computer Vision: Representation and Control* (pp. 55–59).

Bajcsy, R. (1988). Active perception. *Proceedings of the IEEE*, *76*(8), 966–1005. doi:10.1109/5.5968

Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical Review Letters*, *88*(17), 174102. doi:10.1103/PhysRevLett.88.174102

Baraniuk, R. (2007). Compressive Sensing [Lecture Notes]. *IEEE Signal Processing Magazine*, *24*(4), 118–121. doi:10.1109/MSP.2007.4286571

Basseville, M. (1988). Detecting changes in signals and systems—A survey. *Automatica*, *24*(3), 309–326. doi:10.1016/0005-1098(88)90073-8

Benveniste, A., & Basseville, M. (1984). Detection of abrupt changes in signals and dynamical systems : Some statistical aspects. In A. Bensoussan & J. Lions (Eds.), *Analysis and Optimization of Systems* (Vol. 62, pp. 143–155). Springer Berlin / Heidelberg. Retrieved from http://dx.doi.org/10.1007/BFb0004951

Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

Blackmore, L., & Williams, B. (2006). Finite Horizon Control Design for Optimal Discrimination between Several Models. In *Proceedings of the 45th IEEE Conference on Decision and Control* (pp. 1147–1152). IEEE. doi:10.1109/CDC.2006.377045

Blanke, M., B, S. A., & Lunau, C. P. (1997). Fault-tolerant control systems-- A holistic view. *Control Engineering Practice*, *5*(5), 693–702. doi:http://dx.doi.org/10.1016/S0967-0661(97)00051-8

Blanke, M., Lunze, J., Kinnaert, M., Staroswiecki, M., & Schröder, J. (2006). *Diagnosis and fault-tolerant control*. *Diagnosis and Fault-Tolerant Control* (pp. 1–672). doi:10.1007/978-3-540-35653-0

Blanke, M., Staroswiecki, M., & Wu, N. E. (2001). Concepts and methods in fault-tolerant control. *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, *4*. doi:10.1109/ACC.2001.946264

Bombois, X., & Gilson, M. (2006). 14 th IFAC Symposium on System Identification , Newcastle , Australia , 2006 WITH GUARANTEED ACCURACY IN THE. *System*, 505–510.

Borguet, S., Dewallef, P., & Léonard, O. (2005). On-line transient engine diagnostics in a Kalman filtering framework. *ASME Turbo Expo*.

Borguet, S., & Léonard, O. (2008). The Fisher Information Matrix as a Relevant Tool for Sensor Selection in Engine Health Monitoring. *International Journal of Rotating Machinery*, *2008*, 1–10. doi:10.1155/2008/784749

Bowles, J. B., & Ph.D. (2012). Fundamentals of Failure Modes and Effects Analysis Table of Contents. In *Annual RELIABILITY and MA INTAINABILITY Symposium* (p. Tutorial Notes).

Bozdogan, H. (2000). Akaike's Information Criterion and Recent Developments in Information Complexity. *Journal of Mathematical Psychology*, *44*(1), 62–91. doi:10.1006/jmps.1999.1277

Brigham, E. O., & Morrow, R. E. (1967). The fast Fourier transform. *Spectrum, IEEE*, *4*, 63 –70. doi:10.1109/MSPEC.1967.5217220

Brodsky, B. E., & Darkhovsky, B. S. (1993). *Nonparametric methods in change point problems*. The Netherlands: Kluwer Academic Publishers Group.

Brown, P., Desouza, P., & Mercer, R. (1992). Class-based n-gram models of natural language. *Computational Linguistics*, *18 (4)*(1950), 467–477. Retrieved from http://dl.acm.org/citation.cfm?id=176316

Burg P.J. (1967). Maximum Entropy Spectral Analysis. In *Proc. of the 37th Meeting of the Society of Exploration Geophysicists* (pp. 34–41). Oklahoma City, Oklahoma.

Candes, E. J., & Wakin, M. B. (2008). An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, *25*(2), 21–30. doi:10.1109/MSP.2007.914731

Cavagnaro, D. R., Myung, J. I., Pitt, M. a, & Kujala, J. V. (2010). Adaptive design optimization: a mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, *22*(4), 887–905. doi:10.1162/neco.2009.02-09-959

Chaloner, K., & Verdinelli, I. (1995). Bayesian Experimental Design: A Review. *Statistical Science*, *10*(3), 273–304. Retrieved from http://www.jstor.org/stable/2246015

Chater, N., & Oaksford, M. (1999). Information gain and decision-theoretic approaches to data selection: Response to Klauer (1999). *Psychological Review*, *106*(1), 223–227. doi:10.1037/0033-295X.106.1.223

Chatterton, S., Pennacchi, P., Ricci, R., Borghesani, P., & Vania, A. (2013). Development of a new signal processing diagnostic tool for vibration signals acquired in transient conditions. *Chemical Engineering Transactions*, *33*, 61–66. Retrieved from http://eprints.qut.edu.au/66491/

Chechetkin, V. R., & Lobzin, V. V. (2004). Spectral entropy criteria for structural segmentation in genomic DNA sequences. *Physics Letters, Section A: General, Atomic and Solid State Physics*, *328*, 79–86. doi:10.1016/j.physleta.2004.05.060

Chen, J., & Patton, R. J. (1998). *Robust Model-Based Fault Diagnosis for Dynamic Systems (The International Series on Asian Studies in Computer and Information Science)* (p. 380). Springer. Retrieved from http://www.amazon.com/Model-Based-Diagnosis-International-Computer-Information/dp/0792384113

Choi, E., & Lee, C. (2003). Feature extraction based on the Bhattacharyya distance. *Pattern Recognition*, *36*(8), 1703–1709. doi:10.1016/S0031-3203(03)00035-9

Claeskens, G., & Hjort, N. L. (2003). The Focused Information Criterion. *Journal of the American Statistical Association*. doi:10.1198/016214503000000819

Coats, D., Cho, K., Shin, Y.-J., Goodman, N., Blechertas, V., & Bayoumi, A.-M. E. (2011). Advanced Time–Frequency Mutual Information Measures for Condition-Based Maintenance of Helicopter Drivetrains. *IEEE Transactions on Instrumentation and Measurement*, *60*(8), 2984–2994. doi:10.1109/TIM.2011.2122370

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *J. Artif. Int. Res.*, *4*(1), 129–145. Retrieved from http://dl.acm.org/citation.cfm?id=1622737.1622744

Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience.

Cox, D. (1958). *Planning of Experiments*. New York, New York, USA: Wiley-Blackwell.

Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, *2*, 299–318.

Csörg\Ho, M., & Horvath, L. (1988). Nonparametric methods for change-point problems. *Handbook of Statistics*, *7*, 403–425.

Davey, N. (2003). *The Gas Turbine – Development and Engineering*. Watchmaker Publishing.

Deffeyes, J. E., Harbourne, R. T., Stuberg, W. A., & Stergiou, N. (2011). Approximate entropy used to assess sitting postural sway of infants with developmental delay. *Infant Behavior and Development*, *34*(1), 81–99. doi:10.1016/j.infbeh.2010.10.001

Denzler, J., & Brown, C. (2000). *Optimal selection of camera parameters for state estimation of static systems: An information theoretic approach*.

Denzler, J., & Brown, C. (2002a). An Information Theoretic Approach to Optimal Sensor Data Selection for State Estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, *24*.

Denzler, J., & Brown, C. M. (2002b). Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(2), 145–157. doi:10.1109/34.982896

Dette, H., & Titoff, S. (2009). Optimal discrimination designs. *The Annals of Statistics*, *37*(4), 2056–2082. doi:10.1214/08-AOS635

Diambra, L., de Figueiredo, J. C. B., & Malta, C. P. (1999). Epileptic activity recognition in EEG recording. *Physica A: Statistical Mechanics and Its Applications*, *273*(3-4), 495–505. doi:10.1016/S0378-4371(99)00368-4

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, *52*(4), 1289–1306. doi:10.1109/TIT.2006.871582

Ekštein, K., & Pavelka, T. (2004). Entropy and entropy-based features in signal processing. In *Proceedings of PhD workshop systems & control*.

Esna Ashari, A., Nikoukhah, R., & Campbell, S. L. (2009). Asymptotic behavior and solution approximation of Active Robust fault detection for closed-loop systems. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conference*, (2), 1026–1031. doi:10.1109/CDC.2009.5400754

Esna Ashari, A., Nikoukhah, R., & Campbell, S. L. (2011). Auxiliary signal design for robust active fault detection of linear discrete-time systems. *Automatica*, *47*(9), 1887–1895. doi:10.1016/j.automatica.2011.06.009

Esna Ashari, A., Nikoukhah, R., & Campbell, S. L. (2012). Effects of feedback on active fault detection. *Automatica*, *48*(5), 866–872. doi:10.1016/j.automatica.2012.02.020

Estévez, P. a, Tesmer, M., Perez, C. a, & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, *20*(2), 189–201. doi:10.1109/TNN.2008.2005601

Eustace, R. W., Woodyatt, B. A., Merrington, G. L., & Runacres, A. (1994). Fault Signatures Obtained From Fault Implant Tests on an F404 Engine. *Journal of Engineering for Gas Turbines and Power*, *116*(1), 178. doi:10.1115/1.2906789

Fair, M., Campbell, S. L., & Carolina, N. (2009). Active incipient fault detection with more than two simultaneous faults. In *… Man and Cybernetics, 2009. SMC 2009 …* (pp. 3322–3327). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5346202

FANG, Y., CHEN, M., ZHENG, X., & Harrison, R. (2012). Feature Extraction of Motor Imagery in BCI with Approximate Entropy⋆. *Journal of Computational …*, *6*(March), 2485–2491. Retrieved from http://www.jofcis.com/publishedpapers/2012_8_6_2485_2491.pdf

Farrier, D. (1984). Jaynes' principle and maximum entropy spectral estimation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *32*(6), 1176–1183. doi:10.1109/TASSP.1984.1164467

Fedorov, V. (1972). *Theory of Optimal Experiments*. Academic Press.

Fraser, A. (1989). Information and entropy in strange attractors. *IEEE Transactions on Information Theory*, *35*(2), 245–262. doi:10.1109/18.32121

Gagliardi, R. (1967). Input selection for parameter identification in discrete systems. *IEEE Transactions on Automatic Control*, *12*(5), 597–599. doi:10.1109/TAC.1967.1098682

Gerfler, J. (1997). Fault detection and isolation using parity relations, *5*(5), 653–661.

Gevers, M. (2005). Identification for Control: From the Early Achievements to the Revival of Experiment Design. In *Proceedings of the 44th IEEE Conference on Decision and Control* (pp. 12–12). IEEE. doi:10.1109/CDC.2005.1582109

Goodwin, G. C., & Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis (Google eBook)* (p. 291). Academic Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=oUY0bKa3WToC&pgis=1

Goodwin, G., Payne, R. (1973). Design and characterisation of optimal test signals for linear single input - single output parameter estimation. In *3rd IFAC Symposium*. The Hague/Delft.

Goodwin, G., Payne, R., Murdoch, J. (1973). Optimal test signal design for linear single input - single output closed loop identification. In *CACSD Conference*. Cambridge.

Granger, C., & Lin, J.-L. (1994). USING THE MUTUAL INFORMATION COEFFICIENT TO IDENTIFY LAGS IN NONLINEAR MODELS. *Journal of Time Series Analysis*, *15*(4), 371–384. doi:10.1111/j.1467-9892.1994.tb00200.x

Grosse, I., Bernaola-Galván, P., Carpena, P., Román-Roldán, R., Oliver, J., & Stanley, H. E. (2002). Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *65*(4). doi:10.1103/PhysRevE.65.041905

Guorong, X. G. X., Peiqi, C. P. C., & Minhui, W. M. W. (1996). Bhattacharyya distance feature selection. *Proceedings of 13th International Conference on Pattern Recognition*, *2*. doi:10.1109/ICPR.1996.546751

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE*, *85*(1), 6–23. doi:10.1109/5.554205

Hall, D. L., & McMullen, S. A. H. (2004). *Mathematical Techniques in Multisensor Data Fusion (Artech House Information Warfare Library)* (p. 466). Artech House. Retrieved from http://www.amazon.co.uk/Mathematical-Techniques-Multisensor-Information-Warfare/dp/1580533353

Han, N. C., Muniandy, S. V., & Dayou, J. (2011). Acoustic classification of Australian anurans based on hybrid spectral-entropy approach. *Applied Acoustics*, *72*, 639–645. doi:10.1016/j.apacoust.2011.02.002

Harvill, J. L., & Ray, B. K. (2000). Lag Identification for Vector Nonlinear Time Series. *Communications Statistics: Theory and Methods*, *29*.

Hastreiter, P., Freund, J., Greiner, G., & Ertl, T. (1997). Fast Mutual Information Based Registration And Fusion Of Registered Tomographic Image Data. In *Eds.), Workshop: Digitale Bildverarbeitung in der Medizin, Vol. 5, GI, Deutsche Gesellschaft fuer medizinische Informatik (GMDS* (pp. 146–151).

He, Y., Huang, J., & Zhang, B. (2012a). Approximate entropy as a nonlinear feature parameter for fault diagnosis in rotating machinery. *Measurement Science and Technology*, *23*(4), 045603. doi:10.1088/0957-0233/23/4/045603

He, Y., Huang, J., & Zhang, B. (2012b). Approximate entropy as a nonlinear feature parameter for fault diagnosis in rotating machinery. *Measurement Science and Technology*, *23*(4), 045603. doi:10.1088/0957-0233/23/4/045603

Hershey, J. R., & Olsen, P. A. (2007). Approximating the Kullback Leibler divergence between Gaussian mixture models. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* (Vol. 4). doi:10.1109/ICASSP.2007.366913

Hildebrand, R., & Gevers, M. (2003). Identification for control: optimal input design with respect to a worst-case v-gap cost function. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)* (Vol. 1, pp. 996–1001). IEEE. doi:10.1109/CDC.2003.1272697

Hill, G. E. P. B. and W. J. (1967). Discrimination among Mechanistic Models. *Technometrics*, *9*(1), 57–71.

Hill, P. G., & Peterson, C. R. (1965). *Mechanics and thermodynamics of propulsion*. Addison-Wesley Pub. Co. Retrieved from http://books.google.be/books?id=lINTAAAAMAAJ

Isermann, R. (1994). On the applicability of model-based fault detection for technical processes. *Control Engineering Practice*. doi:10.1016/0967-0661(94)90781-1

Isermann, R. (1995). Model BAsed Fault Detection and Diagnosis Methods. In *Proceedings of American Control Conference* (Vol. 3, pp. 1605–1609).

Isermann, R., & Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. In *Control Engineering Practice* (Vol. 5, pp. 709–719). doi:10.1016/S0967-0661(97)00053-1

Jamouli, P. J. F. H. C. A. H. (2003). Joint synthesis of control and fault detection algorithms: Study of P.I. controller influence. In *Proceedings of the European Control Conference ECC'03*. Cambridge.

Jansson, H., & Hjalmarsson, H. (2005). Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, *50*(10), 1534–1549. doi:10.1109/TAC.2005.856652

Jayaram, S., & Rivera, F. (1994). An object-oriented method for the denition of mis- sion proles for aircraft design. In *Proceedings of AIAA 32nd Aerospace Sciences Meeting* (pp. 94–0867).

Jaynes, E. (1957a). Information theory and statistical mechanics I. *Physical Review*. Retrieved from http://medcontent.metapress.com/index/A65RM03P4874243N.pdf

Jaynes, E. (1957b). Information theory and statistical mechanics. II. *Physical Review*. Retrieved from http://prola.aps.org/abstract/PR/v106/i4/p620_1

Jaynes, E. (1968). Prior probabilities. *… Science and Cybernetics, IEEE Transactions on*, (3). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4082152

Jiang, J. D., Chen, J., & Qu, L. S. (1999). THE APPLICATION OF CORRELATION DIMENSION IN GEARBOX CONDITION MONITORING. *Journal of Sound and Vibration*, *223*(4), 529–541. doi:10.1006/jsvi.1998.2161

Joshi, a., Deignan, P., Meckl, P., & Jennings, K. (2005). Information theoretic fault detection. *Proceedings of the 2005, American Control Conference, 2005.*, 1642–1647. doi:10.1109/ACC.2005.1470203

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: a survey. *Journal of Artificial Intelligence Research*, *4*, 237–285.

Kailath, T. (1967). The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transactions on Communication Technology*, *15*(1). doi:10.1109/TCOM.1967.1089532

Kanamori, T., Hido, S., & Sugiyama, M. (2009). A least-squares approach to direct importance estimation. *The Journal of Machine Learning Research*, *10*, 1391–1445.

Kannathal, N., Choo, M. L., Acharya, U. R., & Sadasivan, P. K. (2005). Entropies for detection of epilepsy in EEG. *Computer Methods and Programs in Biomedicine*, *80*, 187–194. doi:10.1016/j.cmpb.2005.06.012

Kantz, H., & Schreiber, T. (2003). *Nonlinear Time Series Analysis*. *Nonlinear Time Series Analysis*. Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9780511755798

Kaspar, F., & Schuster, H. (1987). Easily calculable measure for the complexity of spatiotemporal patterns. *Physical Review A*, *36*(2), 842–848. doi:10.1103/PhysRevA.36.842

Kawahara, Y., & Sugiyama, M. (2012). Sequential change-point detection based on direct density-ratio estimation. *Statistical Analysis and Data Mining*, *5*(2), 114–127. doi:10.1002/sam.10124

Kearns, M., & Koller, D. (1999). Efficient reinforcement learning in factored MDPs. In *Proceedings of the 16th international joint conference on Artificial intelligence - Volume 2* (pp. 740–747). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. Retrieved from http://dl.acm.org/citation.cfm?id=1624312.1624325

Kearns, M., & Singh, S. (2002). Near-Optimal Reinforcement Learning in Polynomial Time. *Mach. Learn.*, *49*(2-3), 209–232. doi:10.1023/A:1017984413808

KERESTECIOGˇLU, F., & ZARROP, M. B. (1994). Input design for detection of abrupt changes in dynamical systems. *International Journal of Control*, *59*(4), 1063–1084. doi:10.1080/00207179408923118

Kerestecioglu, F. (1993). *Change Detection and Input Design in Dynamical Systems (UMIST Control Systems Centre)* (p. 168). Research Studies Press. Retrieved from http://www.amazon.com/Change-Detection-Dynamical-Systems-Control/dp/0863801528

Khani, N., Segovia, C., Navaratne, R., Sethi, V., Singh, R., & Pilidis, P. (2012). Towards Development of a Diagnostic and Prognostic Tool for Civil Aero-Engine Component Degradation. In *ASME 2012 Gas Turbine India Conference* (p. 803). ASME. doi:10.1115/GTINDIA2012-9703

Khawaja, T., Vachtsevanos, G., & Wu, B. (2005). Reasoning about uncertainty in prognosis: a confidence prediction neural network approach. In *2005 Annual Meeting of the North American on Fuzzy Information Processing Society* (pp. 7–12).

Khosrowjerdi, M. J., Nikoukhah, R., & Safari-Shad, N. (2004). A mixed approach to simultaneous fault detection and control. *Automatica*, *40*(2), 261–267. doi:10.1016/j.automatica.2003.09.011

Kim, H. S., Eykholt, R., & Salas, J. D. (1999). Nonlinear dynamics, delay times, and embedding windows. *Physica D: Nonlinear Phenomena*, *127*(1-2), 48–60. doi:10.1016/S0167-2789(98)00240-1

Kim, S.-H., Cohen, M. A., & Netessine, S. (2007). Performance Contracting in After-Sales Service Supply Chains. *Management Science*, *53*(12), 1843–1858. doi:10.1287/mnsc.1070.0741

King, S., Bannister, P. R., Clifton, D. A., & Tarassenko, L. (2009). Probabilistic approach to the condition monitoring of aerospace engines. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, *223*(5), 533–541. doi:10.1243/09544100JAERO414

Kosmidou, V., & Hadjileontiadis, L. (2009). Sign language recognition using intrinsic-mode sample entropy on sEMG and accelerometer data. *… , IEEE Transactions on*, *56*(12), 2879–2890. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4760276

Koudal, P. (2006). *The service revolution in global manufacturing industries*.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, *69*(6), 066138. doi:10.1103/PhysRevE.69.066138

Krause, A., Guestrin, C., Gupta, A., & Kleinberg, J. (2006). Near-optimal sensor placements: maximizing information while minimizing communication cost. In *Information Processing in Sensor Networks, 2006. IPSN 2006. The Fifth International Conference on* (pp. 2–10). doi:10.1109/IPSN.2006.244031

Krause, A., Singh, A., & Guestrin, C. (2008). Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. *J. Mach. Learn. Res.*, *9*, 235–284. Retrieved from http://dl.acm.org/citation.cfm?id=1390681.1390689

Kreucher, C., Kastella, K., & Hero III, A. O. (2005). Sensor management using an active sensing approach. *Signal Processing*, *85*(3), 607–624. doi:10.1016/j.sigpro.2004.11.004

Kreutz, C., & Timmer, J. (2009). Systems biology: experimental design. *The FEBS Journal*, *276*(4), 923–42. doi:10.1111/j.1742-4658.2008.06843.x

Kullback, S. (1959). *Information Theory and Statistics*. New York: Wiley.

Kurz, R., & Brun, K. (2001). Degradation in Gas Turbine Systems. *Journal of Engineering for Gas Turbines and Power*, *123*(1), 70. doi:10.1115/1.1340629

Kurz, R., & Brun, K. (2007). Gas turbine tutorial—maintenance and operating practices effects on degradation and life. *36th Turbomachinery Symposium, Houston, TX, September*, 10–13.

Kwak, N. (2002). Input feature selection by mutual information based on Parzen window. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(12), 1667–1671. doi:10.1109/TPAMI.2002.1114861

Lamberti, P. W., & Majtey, A. P. (2003). Non-logarithmic Jensen-Shannon divergence. *Physica A: Statistical Mechanics and Its Applications*, *329*(1-2), 81–90. doi:10.1016/S0378-4371(03)00566-1

Lamberti, P. W., Majtey, A. P., Borras, A., Casas, M., & Plastino, A. (2008). Metric character of the quantum Jensen-Shannon divergence. *Physical Review A - Atomic, Molecular, and Optical Physics*, *77*(5). doi:10.1103/PhysRevA.77.052311

Lankhorst, M. M., Moddemeijer, R., Box, P. O., & Approach, A. I. (1993). Automatic Word Categorization : An Information-theoretic Approach Automatic Word Categorization. In P. G. M. B. (Eds. . K.A. Schouwhamer Immink (Ed.), *In Proc. 14th Symp. on Information Theory in the Benelux, Veldhoven* (pp. 62–69).

Leao, B. P., Fitzgibbon, K. T., Puttini, L. C., & de Melo, G. P. B. (2008). Cost-Benefit Analysis Methodology for PHM Applied to Legacy Commercial Aircraft. In *2008 IEEE Aerospace Conference* (pp. 1–13). IEEE. doi:10.1109/AERO.2008.4526599

Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., & Liao, H. (2006). Intelligent prognostics tools and e-maintenance. *Computers in Industry*, *57*, 476–489. doi:10.1016/j.compind.2006.02.014

Lempel, A., & Ziv, J. (1976). On the Complexity of Finite Sequences. *IEEE Transactions on Information Theory*, *22*(1), 75–81. doi:10.1109/TIT.1976.1055501

Levadi, V. (1966). Design of input signals for parameter estimation. *IEEE Transactions on Automatic Control*, *11*(2), 205–211. doi:10.1109/TAC.1966.1098296

Li, X., Li, D., Liang, Z., Voss, L. J., & Sleigh, J. W. (2008). Analysis of depth of anesthesia with Hilbert-Huang spectral entropy. *Clinical Neurophysiology*, *119*, 2465–2475. doi:10.1016/j.clinph.2008.08.006

Li, Y. G., & Nilkitsaranont, P. (2009). Gas turbine performance prognostic for condition-based maintenance. *Applied Energy*, *86*(10), 2152–2161. doi:10.1016/j.apenergy.2009.02.011

Liggins, M. E., Hall, D. L., & Llinas, J. (2008). *Handbook of Multisensor Data Fusion: Theory and Practice, Second Edition (Electrical Engineering & Applied Signal Processing Series)* (p. 872). CRC Press. Retrieved from http://www.amazon.co.uk/Handbook-Multisensor-Data-Fusion-Engineering/dp/1420053086

Lindley.D.V. (1956). On a Measure of the Information Provided by an Experiment. *The Annals of Mathematical Statistics*, *27*(No.4), 986–1005. Retrieved from http://www.jstor.org/discover/10.2307/2237191?uid=3738032&uid=2&uid=4&sid=21100964713253

Linke-Diesinger, A. (2008). *Systems of Commercial Turbofan Engines: An Introduction to Systems Functions (Google eBook)* (p. 236). Springer. Retrieved from http://books.google.com/books?hl=en&lr=&id=bgN2cAxF9OAC&pgis=1

Liu, B., Yan, L., Li, L., & Wang, W. (2010). Comparing study of nonlinear model for epileptic preictal prediction. In *2010 4th International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2010*. doi:10.1109/ICBBE.2010.5515760

Liu, H., Sun, J., Liu, L., & Zhang, H. (2009). Feature selection with dynamic mutual information. *Pattern Recognition*, *42*(7), 1330–1339. doi:10.1016/j.patcog.2008.10.028

Liu, S., & Holloway, L. E. (2002). Active sensing policies for stochastic systems. *IEEE Transactions on Automatic Control*, *47*(2), 373–377. doi:10.1109/9.983383

Liu, S., Yamada, M., Collier, N., & Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks : The Official Journal of the International Neural Network Society*, *43*, 72–83. doi:10.1016/j.neunet.2013.01.012

Ljung, L. (1999). *System Identification: Theory for the User (2nd Edition)* (p. 672). Prentice Hall. Retrieved from http://www.amazon.com/System-Identification-Theory-User-Edition/dp/0136566952

Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control*, *34*(1), 1–12. doi:10.1016/j.arcontrol.2009.12.001

Logan, D. B., & Mathew, J. (1996). USING THE CORRELATION DIMENSION FOR VIBRATION FAULT DIAGNOSIS OF ROLLING ELEMENT BEARINGS—II. SELECTION OF EXPERIMENTAL PARAMETERS. *Mechanical Systems and Signal Processing*, *10*(3), 251–264. doi:10.1006/mssp.1996.0019

Logan, D., & Mathew, J. (1996). USING THE CORRELATION DIMENSION FOR VIBRATION FAULT DIAGNOSIS OF ROLLING ELEMENT BEARINGS—I. BASIC CONCEPTS. *Mechanical Systems and Signal Processing*, *10*(3), 241–250. doi:10.1006/mssp.1996.0018

Looney, C. G. (2002). Active information fusion for decision making under uncertainty. In *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat.No.02EX5997)* (Vol. 1, pp. 643–650). Int. Soc. Inf. Fusion. doi:10.1109/ICIF.2002.1021215

Lu, D., Ye, M., & Neuman, S. P. (2011). Dependence of Bayesian Model Selection Criteria and Fisher Information Matrix on Sample Size. *Mathematical Geosciences*, *43*(8), 971–993. doi:10.1007/s11004-011-9359-0

Luttrell, S. P. (1985). The use of transinformation in the design of data sampling schemes for inverse problems. *Inverse Problems*, *1*(3), 199–218. doi:10.1088/0266-5611/1/3/006

MacKay, D. J. C. (1992). Information-Based Objective Functions for Active Data Selection. *Neural Computation*, *4*(4), 590–604. doi:10.1162/neco.1992.4.4.590

MacKay, D. J. C. (2002). *Information Theory, Inference & Learning Algorithms*. New York, NY, USA: Cambridge University Press.

Maes, F., Collignon, a, Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, *16*(2), 187–98. doi:10.1109/42.563664

Manzar abbas, G. vachtsevanos. (2009). A System-Level Approach to Fault Progression Analysis in Complex Engineering Systems. In *Annual Conference of the Prognostics and Health Management Society*. San Diego, CA: ISBN: 978-1-936263-00-4. Retrieved from http://icsl.gatech.edu/aa/images/6/68/Systemlevel.pdf

Marinai, L., Probert, D., & Singh, R. (2004). Prospects for aero gas-turbine diagnostics: a review. *Applied Energy*, *79*(1), 109–126. doi:10.1016/j.apenergy.2003.10.005

Mars, N., & Arragon, G. Van. (1982). Time delay estimation in non-linear systems using average amount of mutual information analysis. *Signal Processing*, *4*, 139–153. Retrieved from http://www.sciencedirect.com/science/article/pii/0165168482900172

Martorano, P., Facco, E., Falzetti, G., & Pelaia, P. (2007). Spectral entropy assessment with auditory evoked potential in neuroanesthesia. *Clinical Neurophysiology*, *118*, 505–512. doi:10.1016/j.clinph.2006.11.002

McClellan, J. (1982). Multidimensional spectral estimation. *Proceedings of the IEEE*, *70*(9), 1029–1039. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1456699

Meher-Homji, Cyrus B | Bhargava, R. (1994). Condition Monitoring and Diagnostic Aspects of Gas Turbine Transient Response. *International Journal of Turbo and Jet-Engines.*, *Vol. 11*(no. 1), 99–111. Retrieved from http://md1.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&recid=A9518123AH&q=Condition+Monitoring+and+Diagnostic+Aspects+of+Gas+Turbine+Transient+Response&uid=791962050&setcookie=yes

Mehra, R. (1974a). Optimal input signals for parameter estimation in dynamic systems--Survey and new results. *IEEE Transactions on Automatic Control*, *19*(6), 753–768. doi:10.1109/TAC.1974.1100701

Mehra, R. (1974b). Optimal inputs for linear system identification. *IEEE Transactions on Automatic Control*, *19*(3), 192–200. doi:10.1109/TAC.1974.1100554

Mélykúti, B., August, E., Papachristodoulou, A., & El-Samad, H. (2010). Discriminating between rival biochemical network models: three approaches to optimal experiment design. *BMC Systems Biology*, *4*, 38. doi:10.1186/1752-0509-4-38

Merrington, G., Kwon, O.-K., Goodwin, G., & Carlsson, B. (1991). Fault Detection and Diagnosis in Gas Turbines. *Journal of Engineering for Gas Turbines and Power*, *113*(2), 276. doi:10.1115/1.2906559

Merrington, G. L. (1988). Identification of dynamic characteristics for fault isolation purposes in a gas turbine using closed-loop measurements. In *AGARD, Engine Condition Monitoring: Technology and Experience* (Vol. SEE N89–16, pp. 36.1–36.13). AGARD. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=7122742

Merrington, G. L. (1989). Fault Diagnosis of Gas Turbine Engines From Transient Data. *Journal of Engineering for Gas Turbines and Power*, *111*(2), 237. doi:10.1115/1.3240242

Merrington, G. L. (1994). Fault diagnosis in gas turbines using a model-based technique. *Journal of Engineering for Gas Turbines and Power; (United States)*, *116:2*. Retrieved from http://www.osti.gov/energycitations/product.biblio.jsp?osti_id=7115419

Mihaylova, L., Lefebvre, T., Bruyninckx, H., Gadeyne, K., & Schutter, J. De. (2002). Active Sensing for Robotics – A Survey. In *5th International Conference On Numerical Methods and Applications* (pp. 316–324).

Miller, B. M., & Runggaldier, W. J. (1997). Optimization of Observations: a Stochastic Control Approach. *SIAM Journal on Control and Optimization*, *35*(3), 1030. doi:10.1137/S0363012995287878

Misra, H., & Ikbal, S. (2004). Spectral entropy based feature for robust ASR. *Acoustics, Speech, and …*, *1*, I–193–6. doi:10.1109/ICASSP.2004.1325955

Moddemeijer, R. (1989a). *Delay-Estimation with Application to Electroencephalograms in Epilepsy*. University of Twente, Enschede (NL). Retrieved from http://www.cs.rug.nl/~rudy/papers/abstracts/RM8901.html

Moddemeijer, R. (1989b). On estimation of entropy and mutual information of continuous distributions. *Signal Processing*, *16*, 233–248. Retrieved from http://www.sciencedirect.com/science/article/pii/0165168489901321

Moon, Y.-I., Rajagopalan, B., & Lall, U. (1995). Estimation of mutual information using kernel density estimators. *Physical Review E*, *52*(3), 2318–2321. doi:10.1103/PhysRevE.52.2318

Morgaz, J., Granados, M. del M., Domínguez, J. M., Navarrete, R., Fernández, A., Galán, A., … Gómez-Villamandos, R. J. (2011). Evaluation of spectral entropy to measure anaesthetic depth and antinociception in sevoflurane-anaesthetised Beagle dogs. *Veterinary Journal*, *188*, 352–355. doi:10.1016/j.tvjl.2010.06.001

Moskvina, V., & Zhigljavsky, A. (2003). An algorithm based on singular spectrum analysis for change-point detection. *Communications in Statistics-Simulation and Computation*, *32*(2), 319–352.

Muggleton, S.H.; Bryant, C.H.; Page, C.D.; Sternberg, M. J. . (1999). Combining active learning with inductive logic programming to close the loop in machine learning. In S. Colton (Ed.), *Proceedings of the AISB'99 Symposium on AI and Scientific Creativity*. The Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB).

Munawar, M. A., Reidemeister, T., & Ward, P. A. S. (2011). Efficient Fault Detection and Diagnosis in Complex Software Systems with Information-Theoretic Monitoring. *IEEE Transactions on Dependable and Secure Computing*, *8*(4), 510–522. doi:10.1109/TDSC.2011.16

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review*, *116*(3), 499–518. doi:10.1037/a0016104

Nick Waters. (2009). Engine Health Management. *Ingenia*.

Niemann, H. (2005). Fault tolerant control based on active fault diagnosis. In *American Control Conference, 2005. Proceedings of the 2005* (pp. 2224 – 2229 vol. 3). doi:10.1109/ACC.2005.1470299

Niemann, H. (2006). A Setup for Active Fault Diagnosis. *Automatic Control, IEEE Transactions on*, *51*(9), 1572–1578. doi:10.1109/TAC.2006.878724

Niemann, H. (2006). Active fault diagnosis in closed-loop uncertain systems. In *6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes* (pp. 587–592). doi:10.3182/20060829-4-CN-2909.00097

Niemann, H., & Kjølstad, N. (2008). Information Based Fault Diagnosis. In *Proceedings of the 17th World Congress The International Federation of Automatic Control* (pp. 8890–8895). Retrieved from http://www.nt.ntnu.no/users/skoge/prost/proceedings/ifac2008/data/papers/2163.pdf

Nikoukhah, R. (1998). Guaranteed Active Failure Detection and Isolation for Linear Dynamical Systems. *Automatica*, *34*(11), 1345–1358. doi:10.1016/S0005-1098(98)00079-X

Nikoukhah, R., & Campbell, S. L. (2002). Active failure detection: Auxiliary signal design and on-line detection. In *Proc. IEEE Med. Conf. Control & …*. Retrieved from http://med.ee.nd.edu/MED10/pdf/221.pdf

Nikoukhah, R., & Campbell, S. L. (2006). Auxiliary signal design for active failure detection in uncertain linear systems with a priori information. *Automatica*, *42*(2), 219–228. doi:10.1016/j.automatica.2005.09.011

Nikoukhah, R., Campbell, S. L., & Delebecque, F. (2000). Detection signal design for failure detection: a robust approach. *International Journal of Adaptive Control and Signal Processing*, *14*(7), 701–724. doi:10.1002/1099-1115(200011)14:7<701::AID-ACS617>3.0.CO;2-6

Nikoukhah, R., Campbell, S. L., & Drake, K. (2010). An active approach for detection of incipient faults. *International Journal of Systems Science*, *41*(2), 241–257. doi:10.1080/00207720903045817

Nikoukhah, R., Campbell, S. L., Horton, K. G., & Delebecque, F. (2002). Auxiliary signal design for robust multimodel identification. *IEEE Transactions on Automatic Control*, *47*(1), 158–164. doi:10.1109/9.981737

Normandin, Y., Cardin, R., & De Mori, R. (1994). High-performance connected digit recognition using maximum mutual information estimation. *IEEE Transactions on Speech and Audio Processing*, *2*(2), 299–311. doi:10.1109/89.279279

Okawa, S., Kobayashi, T., & Shirai, K. (1994). Automatic training of phoneme dictionary based on mutual information criterion. In *Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing* (Vol. i, p. I/241–I/244). IEEE. doi:10.1109/ICASSP.1994.389310

Osborne, M. A., Garnett, R., & Roberts, S. J. (2010). Active Data Selection for Sensor Networks with Faults and Changepoints. In *2010 24th IEEE International Conference on Advanced Information Networking and Applications* (pp. 533–540). IEEE. doi:10.1109/AINA.2010.36

OSyS. (2014). OSyS Company Information. Retrieved from http://www.o-sys.com/about-us/our-company/

P.Wynn, H. (1972). Results in the theory and construction of d-optimum experimental designs. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 133–147.

Papana, A., & Kugiumtzis, D. (2008). Evaluation of mutual information estimators on nonlinear dynamic systems. *Nonlinear Phenomena in Complex Systems*, *11*(2), 225–232. Retrieved from http://arxiv.org/abs/0809.2149

Papana, A., & Kugiumtzis, D. (2009). Evaluation of Mutual Information Estimators for Time Series. *International Journal Bifurcation and Chaos*, *19*(12), 4197–4215.

Papo, D., Caverni, J. P., Douiri, A., Podlipsky, I., & Baudonnière, P. M. (2007). Time-varying spectral entropy differentiates between positive and negative feed back-related EEG activity in a hypothesis testing paradigm. *International Journal of Psychophysiology*, *66*, 183–195. doi:10.1016/j.ijpsycho.2007.06.004

Patton, R., Clark, R., & Frank, P. M. (1989). *Fault Diagnosis in Dynamic Systems: Theory and Applications (Prentice Hall International Series in Systems and Control Engineering)* (p. 624). Prentice-Hall. Retrieved from http://www.amazon.co.uk/Fault-Diagnosis-Dynamic-Systems-International/dp/0133082636

Patton, R. J., & Chen, J. (1997). Observer-based fault detection and isolation: Robustness and applications. *Control Engineering Practice*, *5*(5), 671–682. doi:10.1016/S0967-0661(97)00049-X

Patton, R. J., Frank, P. M., & Clark, R. N. (2000). *Issues of Fault Diagnosis for Dynamic Systems* (p. 598). Springer. Retrieved from http://www.amazon.co.uk/Issues-Fault-Diagnosis-Dynamic-Systems/dp/3540199683

Pecht, M. (1995). *Product Reliability, Maintainability, and Supportability Handbook* (p. 448). CRC Press. Retrieved from http://www.crcpress.com/product/isbn/9780849394577

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1226–38. doi:10.1109/TPAMI.2005.159

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity, *88*(March), 2297–2301.

Pincus, S. M., & Goldberger, A. L. (1994). Physiological time-series analysis: what does regularity quantify? *The American Journal of Physiology*, *266*(4 Pt 2), H1643–H1656.

Pintelon, R., Schoukens, J., & Vandersteen, G. (1997). Frequency domain system identification using arbitrary signals. *IEEE Transactions on Automatic Control*, *42*(12), 1717–1720. doi:10.1109/9.650025

Pluim, J. P. W., Maintz, J. B. A., & Viergever, M. a. (2003). Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, *22*(8), 986–1004. doi:10.1109/TMI.2003.815867

Poulsen, N., & Niemann, H. (2008). Active Fault Diagnosis Based on Stochastic Tests. *Int. J. Appl. Math. Comput. Sci.*, *18*(4), 487–496. doi:http://dx.doi.org/10.2478/v10006-008-0043-6

Povey, D., Kanevsky, D., Kingsbury, B., Ramabhadran, B., Saon, G., & Visweswariah, K. (2008). Boosted MMI for model and feature-space discriminative training. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 4057–4060). IEEE. doi:10.1109/ICASSP.2008.4518545

Pravin Kumar, S., Sriraam, N., Benakop, P. G., & Jinaga, B. C. (2010). Entropies based detection of epileptic seizures with artificial neural network classifiers. *Expert Systems with Applications*. doi:10.1016/j.eswa.2009.09.051

Rajamoney, S. A. (1993). The design of discrimination experiments. *Machine Learning*, *12*(1-3), 185–203. doi:10.1007/BF00993066

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. Cambridge, Massachusetts: The MIT press.

Rebovich, G. (2008). *Systems of Systems Engineering: Principles and Applications*. (U. Jamshidi, Mo(University of Texas, San Antonio, Ed.) (p. 480). CRC Press. Retrieved from http://www.crcpress.com/product/isbn/9781420065886

Rhea, C. K., Silver, T. A., Hong, S. L., Ryu, J. H., Studenka, B. E., Hughes, C. M. L., & Haddad, J. M. (2011). Noise and Complexity in Human Postural Control : Interpreting the Different Estimations of Entropy, *6*(3), 1–9. doi:10.1371/journal.pone.0017696

Rivera, D. E., Lee, H., Mittelmann, H. D., & Braun, M. W. (2009). Constrained multisine input signals for plant-friendly identification of chemical process systems. *Journal of Process Control*, *19*(4), 623–635. doi:10.1016/j.jprocont.2008.08.006

Roemer, M. J., Kacprzynski, G. J., & Orsagh, R. F. (2001). Assessment of data and knowledge fusion strategies for prognostics and health management. In *2001 IEEE Aerospace Conference Proceedings (Cat. No.01TH8542)* (Vol. 6, pp. 2979–2988). IEEE. doi:10.1109/AERO.2001.931318

Rolls-Royce. (2014). Rolls-Royce: Our Corporate Story. Retrieved from http://www.rolls-royce.com/about/index.jsp

Rolls-Royce plc. (1996). *The jet engine*. (Rolls-Royce, Ed.) (5th ed.). Derby, UK.

Rowaihy, H., Eswaran, S., Johnson, M., Verma, D., Bar-noy, A., & Brown, T. (2007). A survey of sensor selection schemes in wireless sensor networks. In *In SPIE Defense and Security Symposium Conference on Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*.

Ruelle, D. (1990). The Claude Bernard Lecture, 1989. Deterministic Chaos: The Science and the Fiction. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *427*(1873), 241–248. doi:10.1098/rspa.1990.0010

Sabeti, M., Katebi, S., & Boostani, R. (2009). Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. *Artificial Intelligence in Medicine*, *47*, 263–274. doi:10.1016/j.artmed.2009.03.003

Samantaray, A. K., & Bouamama, B. O. (2010). Bond Graph Modeling in Process Engineering. In *Model-based Process Supervision - A Bond Graph Approacch* (pp. 13–80). doi:10.1007/978-1-84800-159-6

Schneider, E., Demircioglu Bussjaeger, S., Franco, S., & Therkorn, D. (2010). Analysis of Compressor On-Line Washing to Optimize Gas Turbine Power Plant Performance. *Journal of Engineering for Gas Turbines and Power*, *132*(6), 062001. doi:10.1115/1.4000133

Schwaab, M., Silva, F. M., Queipo, C. A., Barreto, A. G., Nele, M., & Pinto, J. C. (2006). A new approach for sequential experimental design for model discrimination. *Chemical Engineering Science*, *61*(17), 5791–5806. doi:10.1016/j.ces.2006.04.001

Sebastiani, P., & Wynn, H. P. (2000). Maximum Entropy Sampling and Optimal Bayesian Experimental Design. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *62*(1), pp. 145–157. Retrieved from http://www.jstor.org/stable/2680683

Seung, H. S., Opper, M., & Sompolinsky, H. (1992). Query by committee. In *Proceedings of the fifth annual workshop on Computational learning theory - COLT ’92* (pp. 287–294). New York, New York, USA: ACM Press. doi:10.1145/130385.130417

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(July 1948), 379–423;623–656. Retrieved from http://dl.acm.org/citation.cfm?id=584093

Shen, C. P., Chen, C. C., Hsieh, S. L., Chen, W. H., Chan, J. M., Chen, C. M., … Chiu, M. J. (2013). High-Performance Seizure Detection System Using a Wavelet-Approximate Entropy-fSVM Cascade With Clinical Validation. *Clinical EEG and Neuroscience : Official Journal of the EEG and Clinical Neuroscience Society (ENCS)*. doi:10.1177/1550059413483451

Shore, J. (1981). Minimum cross-entropy spectral analysis. *Acoustics, Speech and Signal Processing, IEEE …*, (2). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1163539

Signorini, M. G., Magenes, G., Cerutti, S., & Arduini, D. (2003). Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings. *IEEE Transactions on Bio-Medical Engineering*, *50*(3), 365–74. doi:10.1109/TBME.2003.808824

Simandl, M. (2005). Optimal input and decision in multiple model fault detection. In P. Zítek (Ed.), *Proceedings of the 16th IFAC World Congress* (pp. 1877–1877). doi:10.3182/20050703-6-CZ-1902.01878

Šimandl, M., & Herejt, P. (2003). Information processing strategies and multiple model fault detection. In *Proceedings of the 22nd IASTED International Conference Modelling, Identification, and Control* (pp. 347–352). Calgary: IASTED. Retrieved from http://www.kky.zcu.cz/en/publications/SimandlM_2003_Information

Siroky, J., Simandl, M., Axehill, D., & Puncochar, I. (2011). An optimization approach to resolve the competing aims of active fault detection and control. In *IEEE Conference on Decision and Control and European Control Conference* (pp. 3712–3717). IEEE. doi:10.1109/CDC.2011.6160869

Skaf, Z., Zaidan, M. A., Harrison, R. F., & Mills, A. R. (2013). Accommodating Repair Actions into Gas Turbine Prognostics. In *Annual Conference of the Prognostics and Health Management Society 2013*. New Orleans, Louisiana, USA.

Skanda, D., & Lebiedz, D. (n.d.). Bioinformatics - Supplementary Material biochemical systems Derivation of the optimization criterion I, (4), 1–9.

Skanda, D., & Lebiedz, D. (2010). An optimal experimental design approach to model discrimination in dynamic biochemical systems. *Bioinformatics (Oxford, England)*, *26*(7), 939–45. doi:10.1093/bioinformatics/btq074

Smith, L. A. (1988). Intrinsic limits on dimension calculations. *Physics Letters A*, *133*(6), 283–288. doi:10.1016/0375-9601(88)90445-8

Smith, S. W. (1997). The Complex Fourier Transform. In *The Scientist and Engineer’s Guide to Digital Signal Processing* (pp. 567–580). doi:10.1016/B978-0-7506-7444-7/50068-6

Soares, C. (2008). *Gas Turbines*. *Gas Turbines* (pp. 1–31). doi:10.1016/B978-075067969-5.50015-6

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *76*(3), 485–493. doi:10.1111/rssb.12062

Staniczenko, P. P. A., Lee, C. F., & Jones, N. S. (2009). Rapidly detecting disorder in rhythmic biological signals: A spectral entropy measure to identify cardiac arrhythmias. *Phys. Rev. E*, *79*(1), 11915. doi:10.1103/PhysRevE.79.011915

Steenis, R., & Rivera, D. E. (2009). Plant-friendly signal generation for system identification using a modified SPSA methodology. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) Held Jointly with 2009 28th Chinese Control Conference*, 470–475. doi:10.1109/CDC.2009.5400015

Stoustrup, J., Grimble, M. J., & Niemann, H. (1997). Design of integrated systems for the control and detection of actuator/sensor faults. *Sensor Review*, *17*(2), 138–149. doi:10.1108/02602289710170311

Stoustrup, J., & Niemann, H. (2010). Active fault diagnosis by controller modification. *International Journal of Systems Science*, *41*(8), 925–936. doi:10.1080/00207720903470197

Tang, L., Kacprzynski, G. J., Goebel, K., Saxena, A., Saha, B., & Vachtsevanos, G. (2008). Prognostics-enhanced automated contingency management for advanced autonomous systems. In *2008 IEEE Conference on Prognostics and Health Management* (pp. 1–9).

Taylor, P., Lin, L., & Chu, F. (2011). Nondestructive Testing and Evaluation Approximate entropy as acoustic emission feature parametric data for crack detection, (July 2013), 37–41. doi:10.1080/10589759.2010.521825

Theiler, J. (1990). Estimating fractal dimension. *Journal of the Optical Society of America A*, *7*(6), 1055. doi:10.1364/JOSAA.7.001055

Uosaki, K., Tanaka, I., & Sugiyama, H. (1984). Optimal input design for autoregressive model discrimination with constrained output variance. *IEEE Transactions on Automatic Control*, *29*(4), 348–350. doi:10.1109/TAC.1984.1103525

Vachtsevanos, G., Lewis, F. L., Roemer, F., Hess, A., & Wu, B. (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems* (pp. 1–434). doi:10.1002/9780470117842

Van Erven, T., & Harrëmos, P. (2014). Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, *60*(7), 3797–3820. doi:10.1109/TIT.2014.2320500

Vapnik, V. N. (1998). *Statistical learning theory*. New York: John Wiley & Sons.

Verron, S., Tiplica, T., & Kobi, A. (2008). Fault detection and identification with a new feature selection based on mutual information. *Journal of Process Control*, *18*(5), 479–490. doi:10.1016/j.jprocont.2007.08.003

Wald, A. (1943). On the Efficient Design of Statistical Investigations. *The Annals of Mathematical Statistics*, *14*(2), 134–140. Retrieved from http://www.jstor.org/discover/10.2307/2235815?uid=3738032&uid=2134&uid=2&uid=70&uid=4&sid=21100964983173

Walters-Williams, J., & Li, Y. (2009). Estimation of Mutual Information: A Survey. In P. Wen, Y. Li, L. Polkowski, Y. Yao, S. Tsumoto, & G. Wang (Eds.), *Rough Sets and Knowledge Technology* (Vol. 5589, pp. 389–396). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02962-2_49

Wax, M. (1991). Detection and localization of multiple sources via the stochastic signals model. *Signal Processing, IEEE Transactions on*, *39*(9102459), 2450–2456. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=98000

Wax, M., & Kailath, T. (1985). Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *33*(2), 387–392. doi:10.1109/TASSP.1985.1164557

WHITE, M. (1988). An Investigation of Component Deterioration in Gas Turbines Using Transient Performance Simulation. In *ASME, Gas Turbine and Aeroengine Congress and Exposition* (p. ASME Paper 88–GT–258). Amsterdam. Retrieved from http://md1.csa.com/partners/viewrecord.php?requester=gs&collection=TRD&recid=A8854334AH&q=An+Investigation+of+Component+Deterioration+in+Gas+Turbines+Using+Transient+Performance+Simulation&uid=791962050&setcookie=yes

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *Journal of the Royal Statistical Society. Series B (Methodological)*, *35*(1), 123–130.

Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *AUTOMATICA*, *12*, 601–611.

Willsky, A. S. (1986). Detection of abrupt changes in dynamic systems. In *Detection of Abrupt Changes in Signals and Dynamical Systems, number 77 in Lecture Notes in Control and Information Sciences* (pp. 27–49). Springer-Verlag.

Xiong, N., & Svensson, P. (2002). Multi-sensor management for information fusion: issues and approaches. *Information Fusion*, *3*(2), 163–186. doi:10.1016/S1566-2535(02)00055-6

Yamada, M., Suzuki, T., Kanamori, T., Hachiya, H., & Sugiyama, M. (2013). Relative density-ratio estimation for robust distribution comparison. *Neural Computation*, *25*(5), 1324–1370.

Yan, R., & Gao, R. X. (2007). Approximate Entropy as a diagnostic tool for machine health monitoring. *Mechanical Systems and Signal Processing*, *21*(2), 824–839. doi:10.1016/j.ymssp.2006.02.009

Yongmian Zhang and Qiang Ji. (2006). Active and dynamic information fusion for multisensor systems with dynamic bayesian networks. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *36*(2), 467–472. doi:10.1109/TSMCB.2005.859081

Yu, J., Chen, J., & Rashid, M. M. (2013). Multiway independent component analysis mixture model and mutual information based fault detection and diagnosis approach of multiphase batch processes. *AIChE Journal*, *59*(8), 2761–2779. doi:10.1002/aic.14051

Yu, M., Wang, D., Luo, M., & Huang, L. (2011). Prognosis of hybrid systems with multiple incipient faults: Augmented global analytical redundancy relations approach. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *41*(3), 540–551. doi:10.1109/TSMCA.2010.2076396

Yu, M. Y. M., Wang, D. W. D., & Huang, L. H. L. (2010). Incipient fault diagnosis and prognosis for hybrid systems with unknown mode changes. *Prognostics and Health Management Conference, 2010. PHM '10.* doi:10.1109/PHM.2010.5413418

Zaccarelli, N., Li, B. L., Petrosillo, I., & Zurlini, G. (2013). Order and disorder in ecological time-series: Introducing normalized spectral entropy. *Ecological Indicators*, *28*, 22–30. doi:10.1016/j.ecolind.2011.07.008

Zaffalon, M., & Hutter, M. (2002). Robust Feature Selection by Mutual Information Distributions. In *Proceedings of the 18th International Conference on Uncertainty in Artificial Intelligence (UAI-2002* (pp. 577–584). Morgan Kaufmann.

Zaidan, M. a., Harrison, R. F., Mills, A. R., & Fleming, P. J. (2014). Bayesian Hierarchical Models for Aerospace Gas Turbine Engine Prognostics. *Expert Systems with Applications*. doi:10.1016/j.eswa.2014.08.007

Zaidan, M. A., Mills, A. R., & Harrison, R. F. (2013). Bayesian framework for aerospace gas turbine engine prognostics. In *IEEE Aerospace Conference*. Big Sky, MT: IEEE.

Zaidan, M. A., Relan, R., Harrison, R. F., & Mills, A. R. (2014). *Integrated Bayesian Prognostics of Complex Systems*. Automatic Control and Systems Engineering, University of Sheffield.

Zaidan, M., Harrison, R. F., Mills, A. R., & Fleming, P. J. (2013). *Bayesian Hierarchical Models for Aerospace Gas Turbine Engine Prognostics*. Automatic Control and Systems Engineering, University of Sheffield.

Zaidan, M., Mills, A., & Harrison, R. (2013). Bayesian framework for aerospace gas turbine engine prognostics. *Aerospace Conference, …*, 0–7. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6496856

Zarrop, M. (1979). A Chebyshev system approach to optimal input design. *IEEE Transactions on Automatic Control*, *24*(5), 687–698. doi:10.1109/TAC.1979.1102160

Zarrop, M. B. (1974). Comments on "Optimal Inputs for System Identification." *Notes*, (June), 299–300.

Zarrop, M. B. (1979). *Optimal Experiment Design for Dynamic System Identification.* (Lecture No., p. x, 197 p). Berlin and New York: Springer-Verlag.

Zhang, P., & Ding, S. X. (2008). On Fault Detection in Linear Discrete-Time, Periodic, and Sampled-Data Systems. *Journal of Control Science and Engineering*, *2008*, 1–18. doi:10.1155/2008/849546

Zhang, X. J. (1989). *Auxiliary Signal Design in Fault Detection and Diagnosis (Lecture Notes in Control and Iinformation Sciences)* (p. 213). Springer-Verlag. Retrieved from http://www.amazon.com/Auxiliary-Detection-Diagnosis-Iinformation-Sciences/dp/0387515593

Zhang, Y., & Ji, Q. (2005). Sensor selection for active information fusion. In *Proceedings of the 20th national conference on Artificial intelligence - Volume 3* (pp. 1229–1234). AAAI Press. Retrieved from http://dl.acm.org/citation.cfm?id=1619499.1619531

Zhang, Y., & Ji, Q. (2010). Efficient sensor selection for active information fusion. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics : A Publication of the IEEE Systems, Man, and Cybernetics Society*, *40*(3), 719–28. doi:10.1109/TSMCB.2009.2021272

Zugasti, E., & Arrillaga, P. (2012). Sensor Fault Identification on Laboratory Tower. In *6th European Workshop on Structural Health Monitoring* (pp. 1–8). Dresden, Germany. Retrieved from http://www.ewshm2012.com/Portals/98/BB/tu2b1.pdf

Appendix A

Patent Review

A patent review has been performed on the fault detection / diagnosis of gas turbine engine and its subsystems. [Figure 42] provide a very broad overview of the results obtained by entering different search term in Google patents search engine.

Patent Review Statistics:

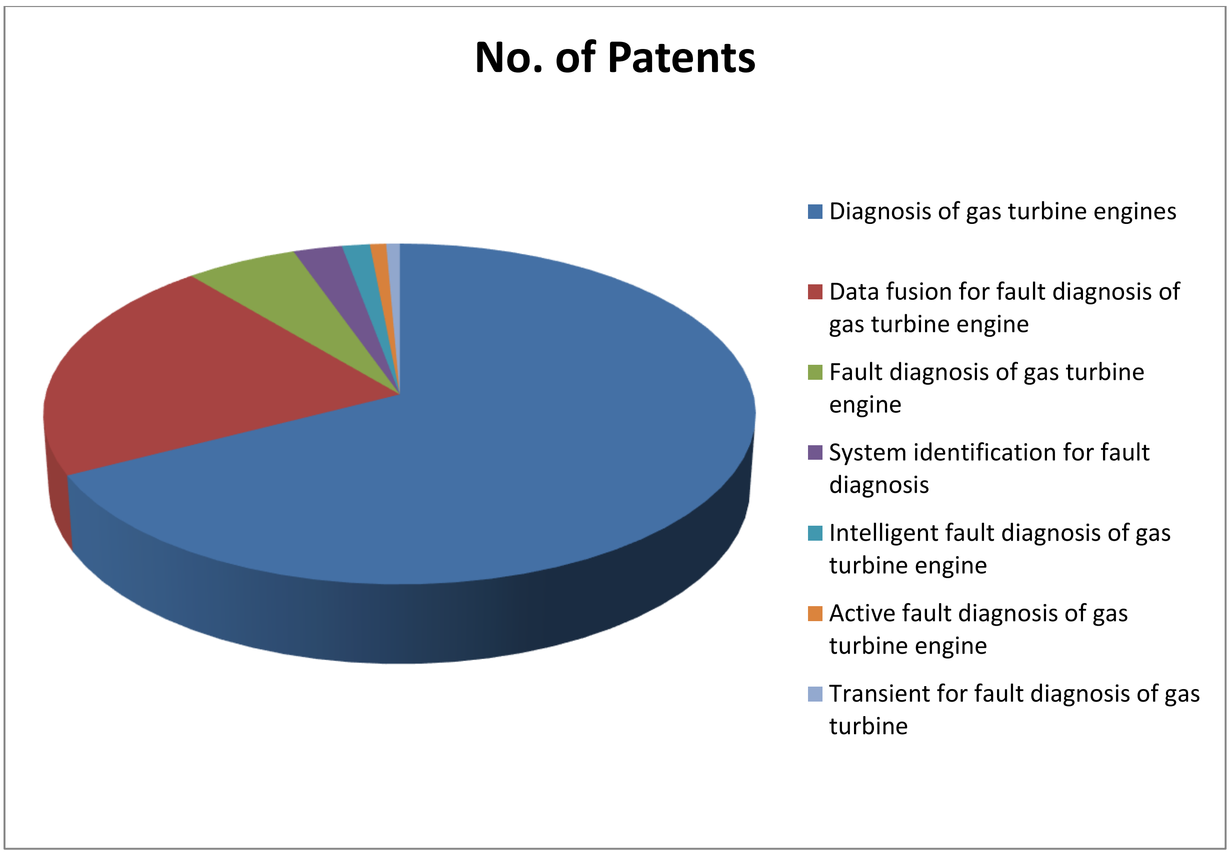


Figure 42: Patent Review Statistics

Review of the relevant patents:

[Table 9] shows few of the patents filed by various organizations that are relevant to the present thesis. The detailed description containing dates, assignee and abstracts of the patents can be found in the Appendix B.

No.	Assignee	PATENT NO	TITLE
1	Honeywell International Inc., Morristown, NJ(US)	**US 6,868,325 B2**	Transient fault detection system and method using hidden markov models
2	Honeywell International Inc., Morristown, NJ (US)	**US 7,043,348 B2**	Transient fault detection system and method
3	Honeywell International Inc., Morristown, NJ (US)	**US 7,693,643 B2**	Fault detection system and method for turbine engine fuel systems
4	United Technologies Corporation, Hartford, CT (US)	**US 7,769,507 B2**	System for gas turbine health monitoring data fusion
5	Honeywell International Inc.,Morristown, NJ (US)	**US 7,945,397 B2**	System and method for gearbox health monitoring
6	NIXON & VANDERHYE, PC 901, NORTH GLEBE ROAD, 11TH FLOOR ARLINGTON, VA 22203 (US)	**US 2010/0155634A1**	Performance monitoring and prognostics for aircraft pneumatic control valves

7	Honeywell International Inc., Morristown, NJ (US)	**US 2010/0303611 A1**	Methods and systems for turbine line replaceable unit fault detection and isolation during engine startup
8	Honeywell International Inc., Morristown, NJ (US)	**US 20080183311 A1**	Apparatus and method for automated closed-loop identification of an industrial process in a process control system.
9	Honeywell International Inc., Morristown, NJ (US)	**US 201110112659 Al**	System identification in automated process control

Table 9: Patent Review

Patent No. [1, 2 & 7] in [Table 9] discuss the use of transients or transient event e.g. start-up based information for fault detection whereas patent no. [3, 5 & 6] discuss various approaches for subsystem fault detection/health assessment. The limitations of these patents lie in the fact that most of these approaches deal with either building a passive anomaly detector consisting of a feature extractor and a reasoning module/classifier or a residual/baseline comparison based approach for fault detection and are most of the time applicable in easily identifiable fault modes. These inventions do not take in to consideration the incipient faults or do no talk about their prognosis. These inventions also do not use an integrated approach of using steady state and transient data for system health monitoring.

Patent No. [4] in [Table 9] discuss a framework for data fusion consisting of a data alignment module, an analysis module and a high level diagnostic feature information fusion module for gas turbine engine health monitoring. This invention takes in to consideration already available information from sensors of different modalities such as aircraft sensors, structural assessment sensors, vibration sensors, gas path sensors, lubrication and fuel system sensors and combine that information with FADEC fault codes, pilot observations as well as engine maintenance history at different levels of hierarchy for recommending maintenance action. The limitation of this approach is that it does not consider or include the information about the health of auxiliary subsystem for proper assessment of engine health. This approach is a generic approach which can easily be modified according to chosen data fusion architecture

Patent No. [8] describes a method for automated closed-loop identification of a multiple model structure-model order combination of an industrial process based on prediction metric or rank for in a process control system. Patent no. [9] further build on the above described approach and discusses an approach based on injecting an additional signal optimized for identification. This invention describes a method for system model identification by performing experiments on a system to be controlled, comprising:

- Selection and discriminating between the two models that enables improvement in model quality, wherein for a given input to the system the two selected models produce different outputs.
- Designing or determining an input having a control component and an identification component where the control component is used for control of the system and the identification component is used for identification experiments.
- Wherein the system identification experiments are based on modification of linear quadratic control to perform system identification in closed loop.
- The control component and identification component of the input is determined simultaneously.

Invention described in patent no. [9] deal with the problem of optimal input design for system identification by selecting two model of the same process and determining an input signal for discrimination in a control theoretic (linear quadratic control) way specifically used for automated process control and model predictive controllers whereas there are other approaches already discussed in the document above which can be used for the same purpose. There is no mention of the applicability of the method discussed in this approach to a real system.

Appendix B

Patent No. US 6,868,325 B2
Date: Mar. 15,2005
Title: Transient fault detection system and method using hidden markov Models
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract:
A transient fault detection system and method is provided that facilitates improved fault detection performance in transient conditions. The transient fault detection system provides the ability to detect symptoms of engine faults that occur in transient conditions. The transient fault detection system includes **a Hidden Markov Model detector** that receives sensor data during transient conditions and determines if a fault has occurred during the transient conditions. Detected faults can then be passed to a diagnostic, system where they can be passed as appropriate to maintenance personnel.

Patent No. US 7,043,348 B2
Date: May 9, 2006
Title: Transient fault detection system and method
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract:
A transient fault detection system and method is provided that facilitates improved fault detection performance in transient conditions. The transient fault detection system provides the ability to detect symptoms of fault in engine that occur in transient conditions. The transient fault detection system includes a **feature extractor (PCA)** that measures sensor data during transient conditions and extracts salient features from the measured sensor data. The extracted salient features are passed to a **classifier (ANN)** that analyzes the extracted salient features to determine if a fault has occurred during the transient conditions. Detected faults can then be passed to a diagnostic system where they can be passed as appropriate to maintenance personnel.

Patent No. US 7,693,643 B2
Date: Apr. 6,2010
Title: Fault detection system and method for turbine engine fuel systems
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract:
A system and method is provided that facilitates improved fault detection. The fault detection system provides the ability to detect symptoms of fault in the fuel system of a turbine engine. The fault detection system captures selected data from the turbine engines that is used to characterize the performance of the fuel system. The fault detection system includes a **feature extractor (PCA)** that extracts salient features from the selected sensor data. The extracted salient features are passed to a **classifier (fuzzy clustering system)** that analyzes the extracted salient features to determine if a fault is occurring or has occurred in the turbine engine fuel system. Detected faults can then be passed to a diagnostic system where they can be passed as appropriate to maintenance personnel.

Patent No. US 7,769,507 B2
Date: Aug. 3, 2010
Title: System for gas turbine health monitoring data fusion
Assignee: United Technologies Corporation, Hartford, CT (US)
Abstract:
An apparatus for assessing health of a device comprising **a data alignment module** for receiving a plurality of sensory outputs and outputting a synchronized data stream, **an analysis module** for receiving the synchronized data stream and outputting at least one device health feature, and a high level **diagnostic feature information fusion module** for receiving the at least one device health feature and outputting a device health assessment.

Patent No. US 7,945,397 B2
Date: May 17, 2011
Title: System and method for gearbox Health monitoring
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract:
A system includes a **plurality of sensors** configured to measure one or more characteristics of a gearbox. The system also includes **a gearbox condition indicator device**, which includes a plurality of sensor interfaces configured to receive input signals associated with at least one stage of the gearbox from the sensors. The gearbox condition indicator device also includes a processor configured to identify a fault in the gearbox using the input signals and an output interface configured to provide an indicator identifying the fault. The processor is configured to identify the fault by determining a family of frequencies related to at least one failure mode of the gearbox, **decomposing** the input signals using the family of frequencies, **reconstructing** a gear signal using the deconstructed input signals, and **comparing** the reconstructed gear signal to a **baseline** signal. The family of frequencies includes a gear mesh frequency and its harmonics.

Patent No. US 2010/0155634 Al
Date: Jun. 24, 2010
Title: Performance monitoring and prognostics for aircraft pneumatic control valves
Assignee: NIXON & VANDERHYE, PC 901, NORTH GLEBE ROAD, 11TH FLOOR ARLINGTON, VA 22203 (US)
Abstract:
A method estimates the health state of an **aircraft pneumatic control valve** through **indirect measurements of available sensors.** Measurements from identical valves operating under the same condition are compared. Residues are translated into estimates of individual valve degradation state. Historical degradation states can be used to predict expected time to failure.

Patent No. US 2010/0303611 Al
Date: Dec. 2, 2010
Title: Methods and systems for turbine Line replaceable unit fault Detection and isolation during Engine startup
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract: Systems and methods for isolating a performance anomaly within one or more line replaceable units (LRUs) on a gas turbine engine by monitoring the start-up transient are presented. The system comprises **a set of sensors**, an **anomaly detector** and a **fault isolation reasoner.** Each sensor of the set monitors at least one operating parameter of at least one engine component. The anomaly detector is configured to detect an anomaly in a **component by comparing a particular value** of an **operating parameter** to a **baseline value** of that parameter. The specific cause of the start-up anomaly is isolated utilizing a set of **component reasoners** that is based on the nature of the detected anomaly. The key events during the engine start-up are identified by the combination of monitoring physically relevant phases of a start-up and monitoring the engine control schedule. The values at these key events are used for comparing at the anomaly detector

Patent No. US 20080183311 A1
Date: July 31, 2008
Title: Apparatus and method for automated closed-loop identification of an industrial process in a process control system.
Assignee: Honeywell International Inc., Morristown, NJ (US)
Abstract:
An apparatus, method, and computer program are provided for **automated closed-loop identification** of an industrial process in a process control system. **Multiple models** (such as multiple model structure-model order combinations) can be identified, where the models are associated with a process to be controlled. One or more metrics (such as a **prediction metric** or **rank**) can be determined for each of the models. At least one of the models can be selected based on the one or more metrics. A final model for controlling the process can be provided (such as to a controller), where the final model is based on the at least one selected model. A **band pass filter** could be designed using some of the identified models. The band pass filter could be used to identify at least one other of the models or to determine at least one of the one or more metrics.

Patent No. US 201110112659 Al	
Date: May 12, 2011	
Title: System identification in automated Process control	
Assignee: Honeywell International Inc., Morristown, NJ (US)	
Abstract:	
The systems and methods described herein allow for **automatic identification experiments in a closed loop**, where the old control strategy, already tuned and tested, is utilized. The strategy is modified to **inject additional signal optimized for identification**. The experimenting time may be reduced by performing only those system manipulations which explore model uncertainties important to potential degradation of controller performance by discrepancy between the system and the model. The disruptions are reduced by keeping the control loop closed, which eliminates waiting for steady state before applying steps to the inputs and reduces the risk of process limits crossing. **The energy of additional signal can be set to meet the maximum allowable disruption requirements. The energy of additional signal is in a direct relation to the speed of identification related information gathering**. It can be varied in time to follow the needs of system operators.	