CRANFIELD UNIVERSITY

LUCA MARINAI

GAS-PATH DIAGNOSTICS AND PROGNOSTICS FOR AERO-ENGINES USING FUZZY LOGIC AND TIME SERIES ANALYSIS

School of Engineering

PhD THESIS

CRANFIELD UNIVERSITY

SCHOOL OF ENGINEERING

PhD Thesis

July 2004

Luca Marinai

Gas-path Diagnostics and Prognostics for Aero-engines Using Fuzzy Logic and Time Series Analysis

Supervisor: Professor Riti Singh

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

Reducing the direct operating-costs is now crucial in order to ensure competitive advantages for airlines and manufacturers, and so effective advanced engine-condition monitoring methodologies are necessary. Hence gas-path diagnostics and prognostics methods are reviewed and the specifications for such effective tools deduced, together with their pertinent future prospects.

First, the considerable value that a preliminary observability study adds to the diagnostics process was recognised. A secure procedure has been devised: it is capable of (i) the identification of the severity of correlations between any two of the available measurements, as well as the correlations between any two of the component changes, (ii) the identification of more complex correlations that involve more than two changes in performance parameters, and (iii) the quantification of the quality of the system observability through a pertinent parameter. This enables comparisons among a significant number of measurement set selections.
The core of the research is a novel gas-path diagnostics (GPD) method that uses fuzzy logic in order to provide secure quantification of the gas-path component faults. A fuzzy diagnostics system was set up for the Rolls-Royce Trent 800 engine that relies on an extensive statement of fuzzy rules generated using an engine model to achieve a quantitative solution through a non-linear approach, which is competent to achieve (i) SFI (single fault isolation) in the presence of noisy data, (ii) tuning over a known global deterioration level for all the performance parameters (baseline) computed for the previous flight, (iii) partial MFI (multiple fault isolation) with up to 2 degraded components (i.e. 4 performance parameters) considerably faulty at the same time, (iv) SFI while isolating systematic errors in the measurements (biasses). A bias-tolerant system was devised by means of the NOT logical operator and a new formulation of the fuzzy rules that includes the location of the bias.
An innovative prognostics framework was devised, which uses ARIMA models and regression models respectively for short and long term investigations, to compute forecasts and the associated prediction intervals, which are aimed at assisting the prognostics decision-making process. This is strictly related to the diverse business intentions: in this study safety and economic related applications are investigated. For example, the optimisation of the TBO (time between overhauls) considering maintenance cost and additional fuel cost due to the deterioration is studied and the potential cost savings for the operators highlighted.

HMP 1.1 for performance analysis was developed: it is a health-monitoring and-prognostics framework consisting of three modules that perform respectively observability study, gas-path diagnostics and prognostics. The substantial benefits that can be achieved with such a tool, relative to the enhanced maintenance planning and improved mission scheduling, are discussed in the thesis via applications to the Trent 800 engine.

AKNOWLEDGEMENTS

I wish to express my gratitude and admiration to Professor Riti Singh, who firstly gave me the opportunity to undertake this work, and then supported me through the past three years providing me with his invaluable guidance, advice and all the necessary means to achieve a successful outcome. He also encouraged and enabled me to participate in a number of international conferences, congresses and workshops which have been a fantastic experience.

Rolls Royce plc sponsored and supervised this research project within their University Technology Centre (UTC) at Cranfield University. I would like to thank Mr Phil Walsh, Mr Steve Brown, Mr John Martin and Dr Marco Zedda – of the Performance Department of Rolls-Royce Civil Aerospace, Derby – as well as Mr Barry Curnock, Mr Richard Dingley and Mr Paul Whatley – of the Performance Department of Rolls-Royce Defence Aerospace, Bristol – for their help, support and guidance during the development of the present work.

I am grateful to Professor Renzo Lazzeretti of the University of Pisa that introduced me to Cranfield and guided me with his experience and wisdom. Very special thanks also go to Professor Pericles Pilidis, Dr Ivan Li, Mr Richard Hales and Miss Rachel Smith who gave me their support when I needed it, and to all the rest of colleagues and friends of the UTC and of the Department of Power, Propulsion and Aerospace Engineering. I wish to thank Professor Douglas Probert, also editor of the Applied Energy Journal, for the valuable time spent in helping to develop my technical writing skills and for his extensive useful advice. Furthermore, I have been lucky while pursuing my doctorate in that I have had the help of a number of MSc and Socrates students. So, thank you very much to Amit, Giovanni, Jason, Priya, and Thierry.

Finally, I would like to take this opportunity to thank my parents and my sister Luisa, for their extraordinary unconditional support and encouragement, all the friends that made my stay at Cranfield enjoyable and memorable, and my girlfriend Sandra that supported me through this experience.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF TERMS

Artificial neural network (ANN)	A collection of mathematical models that emulates some of the observed properties of biological nervous systems and draws on the analogies of adaptive biological learning.
Baseline	A quantifiable physical condition of level of performance from which changes are measured.
Bayesian belief network (BBN)	BBNs use a probabilistic framework to model dependencies in a diagnostic framework. BBNs can be extended into influence diagrams through the addition of decision nodes which allow modelling of potential actions to take. The price is a considerable increase in computational burden.
Bias	A systematic error in a measurement.
Computational burden	negative property of excessively time consuming software
Condition monitoring	The gathering of data from a gas turbine in service, in order to understand its condition and optimise its operating costs
Data fusion	Information from several diagnostic sources will only in rare cases agree completely. This is a result of the various underlying models, the data, assumptions made, etc. It is the task of data fusion techniques to resolve potential conflict between the information, taking into account for example historical performance of individual tools. Besides, information from different analysis types such as oil, debris, vibration analyses can be fuse together to gain a higher level of diagnostics. Approaches can be solved for example using rule-based approaches.
Decision making	Decision-making techniques try to optimize a decision over some type of utility.
Deterioration	Reduction in the component's capacity to perform to its design value
Expert system (ES)	Expert Systems have been used for many years to capture heuristic knowledge about system behaviour. Diagnostic expert systems allow the identification of problems with user interaction. Uncertainty has successfully been integrated into expert systems but challenges remain regarding updating and maintenance.

Fuzzy logic (FL)	A rule-based approach, founded on the formulation of a novel algebra, typically used in the analysis of complex systems and to enable decision-making processes affected by uncertainty.
Gross domestic product (GDP)	The total output of goods and services produced within a country; the broadest measure of economic output with the exception of GNP (gross national product), which includes a country's nationals who work in other countries.
Gas-path analysis (GPA)	A commonly-used term for gas turbine performance diagnostics.
Gas-path diagnostics (GPD)	The process of isolating and assessing faults in the engine's gas-path.
Gas-path prognostics (GPP)	The process of predicting the evolution of the faults in the engine's gas-path.
Genetic algorithm (GA)	Genetic algorithms are search algorithms based on mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomised information exchange to form a search algorithm with some of the innovative flair of human search.
Health parameters	Efficiency and flow capacity of the gas-path components as indicative of their health condition
Hybrid systems	Hybrid systems try to combine the advantages of several approaches, such as neural network, fuzzy logic and genetic algorithms.
Kalman filter (KF)	An algorithm for producing best estimates of the component changes and sensor biases that produced an observed set of gas-path measurement differences from expectation.
Learning	In changing environments, it is essential for the diagnostic tools to adapt to different external stimuli, but also to different demand profiles, and perhaps to changed system status. This requires advanced learning techniques.
Noise	A random component of measurement error caused by numerous small effects, which cause disagreements between repeated measurements of the same parameter.
Observability	The ability to perceive changes in performance parameters from the observation of a given number of measurements in the gas-path.

Revenue passenger-kilometers (RPK)	The number of passengers multiplied by the number of kilometers they fly.
Search space	A collection of all possible solutions to the problem from which the best solution is chosen.
Smearing	Tendency to distribute a considerable faults, that interests few components, over a large number of the engine's components and sensors
Trend analysis	In trend analysis, system changes are tracked over time and extrapolations allow the prediction of impending failures. Step changes, increase in variance, etc., can indicate potential problems. Statistical techniques are traditionally found here.

ABBREVIATIONS

AC	auto-correlation
AGARD	Advisory Group for Aerospace Research and Development
AI	artificial intelligence
AIAA	American Institute of Aeronautics and Astronautics
ANN	artificial neural-network
AR	auto-regressive
ARIMA	autoregressive-integrated-moving-average
ARMA	auto-regressive-moving-average
ASEE	American Society of Engineering Education
ASME	American Society of Mechanical Engineers
BBN	Bayesian-belief network
BSI	British Standard Institution
CBM	condition based maintenance
CD	cumulative deviation
CI	confidence interval
COEHM	combined engine-health monitoring
COEHMP	combined engine-health monitoring and prognostics
COM	centre of maximum
COMPASS	condition-monitoring and performance analysis software system
d.o.a.	degree of activation
d.o.m.	degree of membership
DOC	direct operating-cost
ECM	engine condition monitoring
EHM	engine health monitoring

EKF	extended Kalman-filter
EMS	engine monitoring system
EPR	engine pressure ratio
ERT	exchange rate table
ES	expert system
FADEC	full authority digital engine-control
FF	fuel flow
FFBPNN	feed-forward back-propagation neural-network
FIS	fuzzy inference system
FL	fuzzy-logic
FMPT	fleet-management programme
FOD	foreign object damage
GA	genetic-algorithm
GDP	Gross domestic product
GPA	gas-path analysis
GPD	gas-path diagnostics
GPP	gas-path prognostics
GT	gas-turbine
GUI	graphical user interface
HMP	health-monitoring and prognostics
HP	high-pressure
HPC	high-pressure compressor
HPT	high-pressure turbine
HS	high severity
ICM	influence coefficient matrix
IEKF	integrated extended Kalman-filter
IFAC	International Federation of Automatic Control
IPC	intermediate-pressure compressor
IPT	intermediate-pressure turbine
ISABE	International Society for Air-Breathing Engines
KES	knowledge-based engineering systems
KF	Kalman-filter
LP	low-pressure
LPT	low-pressure turbine
LS	low severity
MA	moving-average
MAE	mean absolute error
MCPHT	maintenance cost per hour
MF	membership function
MFI	multiple-fault isolation
MLP	multi-layer perceptron
MOPA	multiple operating-point analysis
MS	medium severity
NASA	National Aeronautics and Space Administration

OEM	original equipment manufacturer
PAC	partial auto-correlation
PHM	prognostics and health management
Pi	total pressure at station i
PI	prediction interval
PNN	probabilistic neural-network
RMS	root-mean-square
ROI	return on investment
RPK	revenue passenger-kilometre
RPM	revolution per minute
RRAP	Rolls-Royce aerothermal performance
SAE	Society of Automotive Engineers
SAM	standard additive model
SFC	specific fuel consumption
SFI	single-fault isolation
SV	singular value
t	time
T.O.	take off
tamb	ambient temperature
TBO	time between overhauls
TEMPER	Turbine Engine Module Performance Estimation Routine
TET	turbine entry temperature
TGT	turbine gas temperature
Ti	total temperature at station i
TTFO	time to first overhaul
UTC	university technology centre
Wi	mass flow at station i
WLS	weighted least-squares

CHAPTER 1 - INTORDUCTION: PROSPECTS FOR GAS TURBINE HEALTH MONITORING

1.1 Set the scene: novel opportunities in the gas turbine aftermarket

1.1.1 Introduction

At present, there are deep financial uncertainties in the civil air-transportation market and therefore intense competition among airlines. Hence, the development of advanced maintenance techniques in order to reduce operating-costs (Cribbes, 1997, Singh, 2001). Engine-related costs contribute to a large fraction of the direct operating-costs (DOCs) of an aircraft, because the propulsion system requires a significant part of the overall maintenance effort that has to be expended for each aircraft.

The aftermarket for aero gas turbines has undergone significant changes and is set to see more profound changes in the future. Such developments are of great importance to this industry, where reliance on the aftermarket is an essential part of the business case (Singh, 2003).

This project has been focused on the study and development of methodologies underpinning advanced gas turbine performance analysis. This chapter considers the role of this technology in the changing business environment in which contracts based on an agreed rate per engine flying hour are increasingly in demand. As a matter of fact, to set the scene for this research, a discussion follows about the techno-economic reasons why this project was undertaken considering the gas turbine market and aftermarket in the context of the global civil air-transportation business.

The next sections provide a global update of the prospective in the gas turbine aftermarket. By directly surveying leaders' points of view in the field (i.e. major aircraft manufacturers, engine manufacturers and airlines), a thorough analysis of the key trends in the business are highlighted, considering the following elements:

- An expanding air transportation world market, despite the difficulties and changes since the shock of September 2001. Prominent trends in the sector and the business forecasting in the next 20 years.
- Increasing expectations regarding unit prices, delays and financing. Need for reducing DOC.
- The key role of the gas turbine aftermarket.
- High competition among airlines (need for fixed maintenance costs) and high competition among gas turbine manufacturers to enter the aftermarket sector.
- Changes in the gas turbine business model and role of the aftermarket, new long-term contracts for the major players based on an hourly-utilization basis payment.

- Emergence of new active strategies, services, total care solutions, and the need for new monitoring and analysis methodologies.

The sections from 1.1.2 to 1.1.5 provide the analysis of the above mentioned issues and show how a key competitive advantage for manufacturers will be their understanding of this market and one consideration within this will be engine performance analysis capabilities. Section 1.2 considers engine-condition monitoring in general and gas-path diagnostics and prognostics in particular and their role in the gas turbine industry. Section 1.3 is dedicated to the definition of the research aim and the preliminary objectives of the work; a guide to the thesis is then given in section 1.4.

1.1.2 Market condition: need to reduce DOC

This section is dedicated to analyse the features of the air-transportation world market, through the assessments for the future of air travel published by major players. Despite the current crisis in the air-transportation business, long-term demand will grow strongly probably in a highly competitive and complex market scenario.

Economic models show that air travel is highly correlated to economic growth as measured as gross domestic product (GDP) – see Figure 3. International trade and lower fares also drive air travel.

The current short-term cycle is severe. The airline industry is in the midst of the most serious short-term downturn in modern aviation history. World traffic measured in revenue passenger-kilometres (RPKs) experienced a reduction in 2001, and no growth in 2002. The outlook following the shock of September 2001 has been characterised by 'a deeper recession and a stronger recovery' compared with forecasts (Figure 1). DRI-WEFA, forecasting group whose projections of economic growth are used by Airbus (2002), believe that the impact of the 2001 crisis on world GDP growth will be confined to 2001 and 2002. In 2003, geopolitical conflict and the SARS virus have continued to negatively impact the vitality of the airline industry. The Boeing forecast, published in "the Boeing current market outlook" (Boeing, 2003), predicts 5% smaller world RPKs in 2022 than there would have been without the current short-term cycle.

The long-term forecast remains healthy. In the short term, air travel is influenced by business cycles, consumer confidence, and exogenous events. Over the long term, cycles smooth out and GDP, international trade, lower fares, and network service improvement trends become paramount. During the past thirty years, worldwide air travel measured in RPKs has been characterised by periods of strong growth interspersed with slowdowns caused by a series of economic or financial crises (Figure 2). The impact of each crisis has lasted only a short time, after which air travel demand has rebounded strongly. Despite the short-term severity of the current crisis, Airbus forecasters expect that, in the long-term, history will repeat itself, and worldwide air travel will resume its pattern of robust growth. While considerable growth has already

been achieved, great potential still remains for further growth, especially in markets serving the less-developed countries. Here, the use of air transport is currently only a fraction of that in the developed countries, and strong growth in demand can be expected as wealth increases and air travel becomes affordable by more and more of these nations' large populations.

During the next 20 years, it has been suggested that the average economies will grow annually by 3.2%, and air travel will continue its historic relationship with GDP by growing at an average annual rate of 5.1% (Airbus, 2002).

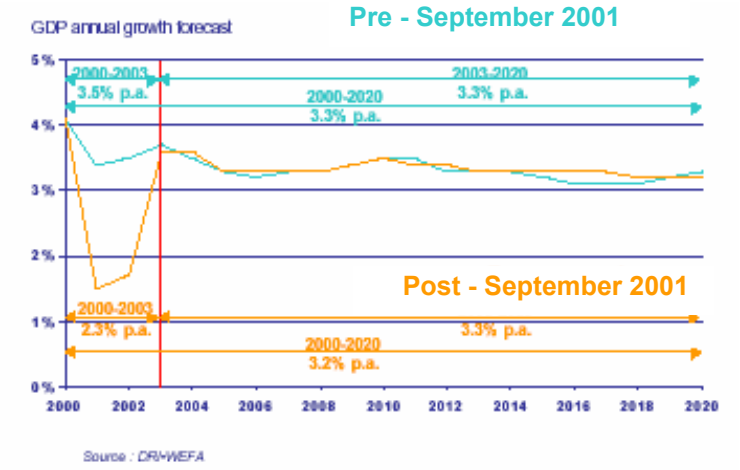


Figure 1: World GDP growth : deep recession and strong recovery , (Airbus, 2002)

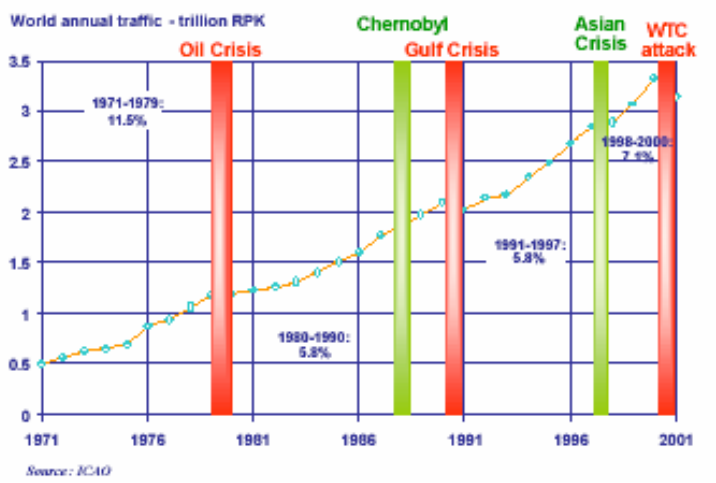


Figure 2: Air travel has grown more than six-fold during the past 30 years, (Boeing, 2003)

1.1.2.1 Economic growth is the prime driver of demand

Historically, RPKs and GDP have shown a close correlation (Figure 3) and Airbus remains confident that overall, as wealth increases and air travel becomes affordable by more and more people, RPKs will continue to be correlated with, and in the long term to grow more rapidly than, GDP.

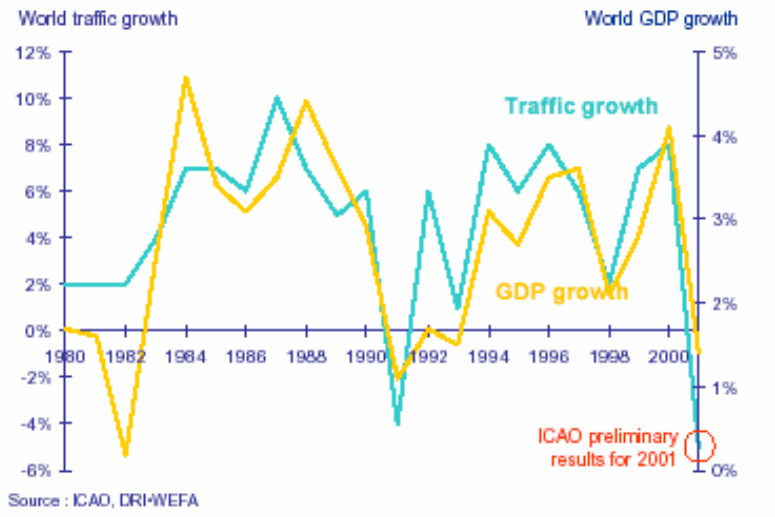


Figure 3: Growth in air travel is strongly correlated with the growth in GDP, (Airbus, 2002)

Overall, Airbus predicts that during the next twenty years RPKs will grow at an average annual rate of 4.7%. By 2020 world annual RPKs will reach 8.3 trillion compared with the 3.3 trillion carried in 2000 (Figure 4).

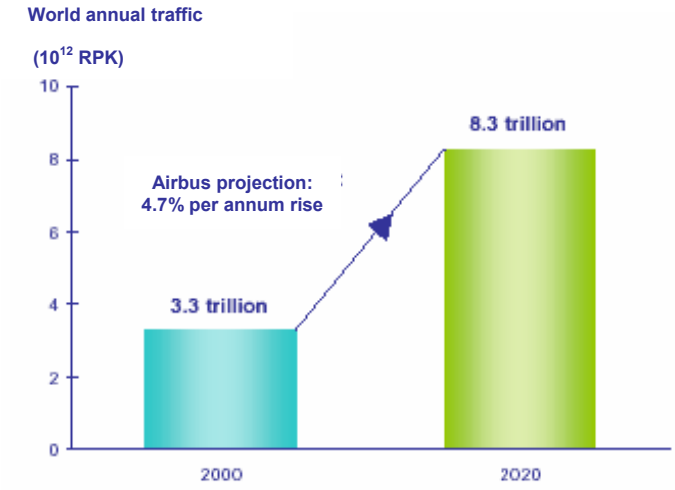


Figure 4: Predicted air-world traffic growth, (Airbus, 2002)

This average growth figure embraces a wide variety of growth rates in the different individual submarkets, which largely reflect their differing degrees of maturity. Airbus forecasters also have predicted that, in the current climate, a new generation of low-cost airlines will prosper, and that in response to increasingly severe cost pressures established airlines will be driven even further to improve the efficiencies of their route networks and to use low-unit-cost aircraft. Hence the present knock-on of severe competition among gas-turbine manufacturers, that is going to be discussed in the next section.

1.1.3 Gas turbine industry's point of view: the aftermarket

1.1.3.1 Role of gas turbine aftermarket in the global civil air transportation business

Considering the scenario of global civil air-transportation market, airlines are under increased pressure to reduce their operating costs because of heightened competition, eroding revenue and severe uncertainties. This environment is pushing towards the application of advanced fault-diagnostics and prognostics techniques to review maintenance philosophies in order to reduce operating costs (Singh, 2001). In this respect, the propulsion system calls for a significant portion of the overall aftermarket. Figure 5 shows that the maintenance cost together with the fuel bill represent 18% of the total costs. The profit may be seen as large in absolute terms. However, when compared with the revenue and costs, it may be a relatively small percentage. Thus, any change in either of the two elements – revenue or costs - could have detrimental effects on the total profits.

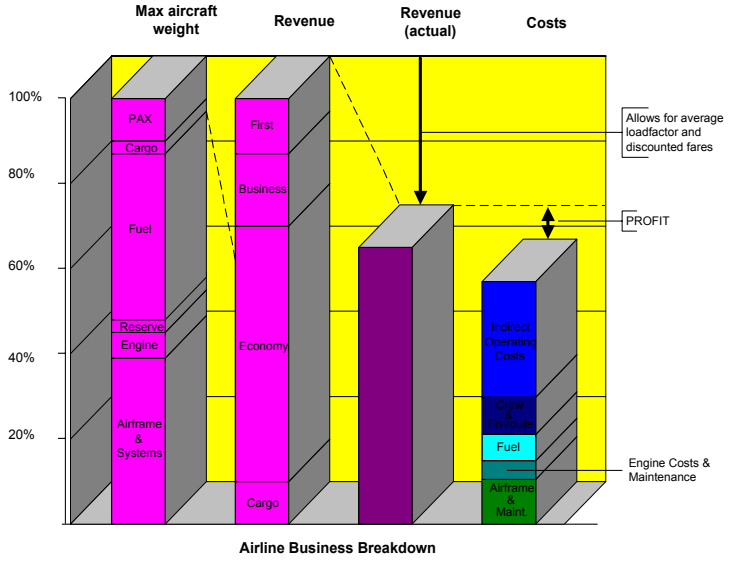


Figure 5: Airline business breakdown, (Singh, 2001)

1.1.3.2 Heightened competition and need for reducing DOC

For Lufthansa, engine related operating costs contribute more than 25% to the direct operating costs (DOC) of a short range aircraft such as an A320-200 (Rupp, 2002). Figure 6 shows that engine-related operating costs can be broken down into three main categories. Maintenance procedures have a direct effect on maintenance and overhaul costs, and also indirectly influence the fuel costs. A jet engine gradually degrades over time due to several mechanisms, which can be fully or partly reversed during an overhaul. The higher fuel consumption, resulting from engine deterioration, has become a serious economic problem since the sharp rise in unit fuel prices after the oil crisis in 1973.

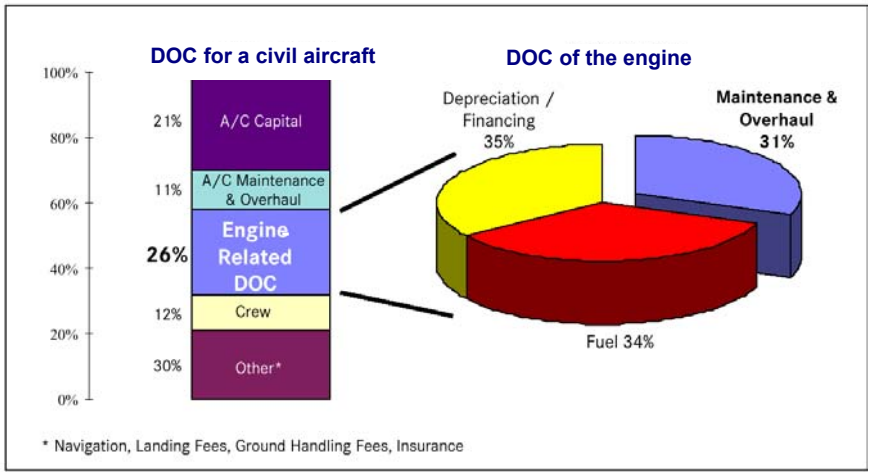


Figure 6: Typical costs as fraction of a civil aircraft's DOC, (Rupp, 2002)

Maintenance and fuel-related engine costs make up a large portion (17% for Luftansa, (Rupp, 2002)) of the overall aircraft DOC. Increased competition has reduced the profits from engine sales and so original equipment manufacturers (OEMs) rely more and more on the highly-profitable aftermarket business, especially in times of economic downturns. In fact, the in-service costs of engines are generally much larger than the corresponding initial selling prices. Nevertheless, within the present severe competition, airlines demand for a reduction in DOC and related reduction of maintenance costs. This has led OEMs to offering airlines full care packages discussed in the next few sections.

1.1.3.3 Opportunities in the aftermarket

Previously, advances in technology, such as cooled blades or the move from pure jets to turbofans, offered large advantages in functionality, making the engine business a research- and design technology-led industry. In the 1970s and 1980s, integration of engineering and manufacturing functions reduced costs. More recently, considerable interest has been devoted to lowering operating-costs by introducing improvements in engine reliability, achieving extended "life on wing" and upgrading maintenance schedules (Singh, 2003).
As the business emphasis shifts to the aftermarket, the changing paradigm will favour those business leaders who recognise the importance of the contribution that "intellectual adventure" can make in the aftermarket as it has, in the past, in other areas of gas turbine technology (Singh, 2003): certainly, gas turbine diagnostics and prognostics are part of this challenging research opportunity.

1.1.4 Strategic development: total care solutions

The role of the gas turbine aftermarket is changing by re-defining the business paradigm for the civil air transport market. The gas-turbine business model has been based on revenue strictly related to the aftercare. As a senior aero gas turbine industrialist once famously stated, 'the aero engine business at the end of the 21st century is like the razor business: you can give away the engines because the money is made on the blades!' (Singh, 2003). This business model, in which the engine manufacturer has revenue directly related on the aftercare, is nowadays changing towards a situation in which much of the

technical risk is transferred from the airline to the gas turbine maker. Fleet management and comprehensive engine aftercare service is provided by the manufacturer to the airline that, based on an agreed rate per engine flying hour, pays for when the aircraft is in the air producing revenue.

1.1.4.1 Airlines' point of view

Nowdays, it is becoming increasingly essential for airlines to be able to focus on their core business activities, whilst maximising operational reliability and minimizing financial risk and costs. Removing technical and financial uncertainty associated with engine aftercare is currently considered the key to address each of these issues. Therefore, the airlines demand is for high quality fleet-management and comprehensive engine-aftercare service, based on an agreed prior to the engine delivery rate per engine flying hour. The manufacturer provides the opportunity for the user to be released from the technical and financial unknown of engine maintenance and management. By transferring this traditional technical risk to the manufacturer, the engine operator is freed by problems of engine fleet management, product reliability and the uncertain cost of ownership.

1.1.4.2 Engine manufacturers' point of view

The next 20-year total value of the aeroengine market is approximately $750bn, of which 45% is judged to be for the aftermarket (Singh, 2003). In the previous business model, the impact of the aftermarket for engine manufacturers was critical and twofold. Firstly, the margins in the aftermarket are higher than in original equipment sales. Next, whilst new equipment sales can be deferred during economic downturns, the aftermarket may sustain the business until the next economic upturn. The business paradigm has been changing: from the gas turbine industry side, improvements concerning the in-service operations of engines have already had significant impacts on the business (Singh, 2001). Prolonging a gas-turbine's life, could lead to a scenario in which the engine will not require a major service during the aircraft's life (e.g. twenty-five years). In the previous after-sales maintenance scenario, this would have resulted in the partial or complete loss of the engine-manufacturers' aftermarket business. So, companies would have had to make compensating higher profits on the original equipment sale. More interestingly, the loss of the aftermarket revenues based on decades of prime incumbency eliminates a major market entry barrier. Perhaps this is when a new wave of companies will gain entry into the business. As matter of fact, not only the original equipment manufacturers, but also competitors, users and new specialist players are entering this business. In these circumstances, a key competitive skill for manufacturers will be their detailed understanding of this market (Singh, 2003). The improved engine-reliability, long engine-life and a lucrative market have led to engine manufacturers seeking long-term maintenance contracts based on an agreed rate per engine flying hour. This new business model could therefore be a win-win opportunity for airlines and manufacturers.

1.1.4.3 Competitive analysis: gas turbine manufacturers

GE Aircraft Engines, Pratt & Whitney, and Rolls-Royce will continue to maintain a traditional stronghold on the gas turbine engine markets worldwide. Through a myriad of strategic partnerships, each company maintains strengths in particular sub-segments of the overall market. Each is a leader in the sale of large turbofans, and plays a major role in the turboprop and turboshaft business. The market is concentrated in the large turbofan segments. However, dynamic growth in particular sub-segments such as regional/corporate jets and commercial helicopters will continue to stimulate competition from a wider variety of players than the traditional three. In addition, the advancement of gas turbine technology via several key development programmes will drive manufacturers to offer better deals with lower operational costs for airlines. The most important factor that will determine success is the after-market oriented competitive strategies.

1.1.4.4 The answer and solution

'Power by the Hour™' (trade mark held by Rolls-Royce) type of contracts, which includes the capital cost plus a blend of financing and maintenance after the engine's sale, are increasingly being demanded. Similarly, highly successful General Electric's 'Maintenance Cost per Hour™' (MCPH™) contracts and Pratt & Whitney 'Fleet Management Programme™' (FMP™) contracts offer long-term service agreements. These programmes provide engine maintenance on a flat rate per engine flight-hour basis, enabling airlines to accurately forecast operating costs, reduce cost of ownership and improve asset utilization (Marinai et al. 2003c). In these circumstances, a key competitive advantage for manufacturers will be concerned with engine-condition monitoring methodologies. Among them, at present, particular consideration should be given to gas-path diagnostics that plays a primary role in an aero-engine performance oriented business. Engine gas-path diagnostics has been recognised, for some time, as important means for making more informed decisions on the usage, maintenance, overhaul or replacement of the engine or one of its components. Deterioration can affect factors such as thrust (or power) and specific fuel-consumption (SFC). As a consequence of progressive performance losses, operation of the engine can become cost ineffective or even unsafe, hence monitoring techniques are employed and accordingly maintenance actions are undertaken. Besides, engine manufacturers have become more focused on deterioration modelling and prognostics capability in order to achieve greater confidence in their cash-flow projections. The main aims are to achieve significant benefits in mission scheduling and maintenance planning, as well as to reduce the costs of maintenance servicing (Marinai et al., 2004).

1.1.5 Summary

In section 1.1 the emergence of a renewed challenge of designing advanced health monitoring and prognostics methodologies is discussed in view of some techno-economic issues that are summarised below.

- At present, there are deep financial uncertainties in the civil air-transportation market but long-term forecasts are positive.
- The profit may be seen as large in absolute terms, however, the competition is high. The expectations regarding prices, delays and financing are increasing.
- Reducing the engine operating costs seems to be the keyword.
- The high competition among airlines reflects in a severe competition among gas turbine manufacturers. This context is going to be played in the aftermarket sector: recently considerable interest has been devoted to operating costs gaining improvements in engine reliability and "life on wing".
- From one side: the airlines demand for fleet management and comprehensive engine aftercare service based on an agreed rate per engine flying hour. The engine is paid for when the aircraft is in the air producing revenue.
- From the other side: the improvements in gas turbine life, could led to a scenario in which engines will not require a major service for the duration of the aircraft's life. The result could be the partial or complete loss of the engine manufacturers' aftermarket business.
- Win-win situation. A new business model seems therefore to provide a win-win situation for airlines and manufacturers: new type of contracts, which includes the capital cost plus a blend of financing and maintenance after the engine's sale and hourly-utilization basis payment, are taking place.

In these circumstances, a key competitive advantage for manufacturers will be their advanced engine health monitoring methodologies. Among them particular consideration is given to gas-path diagnostics and prognostics that play a primary role in an aero-thermal performance oriented business.

An overview of gas turbine condition monitoring techniques is considered in the next section emphasizing the crucial role of diagnostics and prognostics.

1.2 Engine-condition monitoring

1.2.1 Introduction

Engine-condition monitoring (ECM) and engine diagnostics have been recognised as important assets in making more informed decisions on the gas turbine usage. The importance of such techniques has been reemphasised by the changes in market positioning discussed earlier.

The objective of gas turbine maintenance is to guarantee a safe, punctual and economic aircraft operation. Monitoring and maintenance techniques must be used to ensure that the engine operates cost-effectively and safely. There are several types of maintenance strategies adopted according to the individual requirements. A broad classification is given in the next section.

1.2.2 Maintenance methods

When investigating the evolution of gas turbine maintenance strategies, three generations can be identified:

- First generation: run to failure maintenance is a strategy that maintains high output levels for as long as possible until a breakdown occurs.
- Second generation: preventive maintenance is a strategy that operates a planned servicing routine to prevent failure.
- Third generation: condition based maintenance is a strategy that detects component failures based on data derived from condition monitoring.

1.2.2.1 Run to failure

At first sight, the run to failure method might seem to be a good strategy as maintenance will only be performed when absolutely necessary. However, this is obviously not a viable maintenance policy for the safety-conscious airline business, where reliability is of prime importance. In addition, a failed component can cause damage in other parts of the engine, so resulting in very high maintenance costs.

1.2.2.2 Preventive maintenance

This was devised whereby an engine or component is withdrawn from service at regular intervals for inspection and repair. During the 1950s, maintenance intervals for reciprocating engines could be as low as 1000 hours. With growing experience and through engine modifications, the intervals increased rapidly. In addition, it was realised that the cold part of the jet engine justified much longer times between overhauls (TBO). However, the preventive maintenance strategy is not the optimal method for achieving maximum reliability and minimum costs. One of the reasons is that the removal of components before they fail results in losses of useful life. In addition, frequent maintenance can also reduce the reliability of the system. Moreover, mechanical components do not all fail at the same intervals (especially if repaired parts are used) and so unless an overly conservative approach is adopted, unplanned overhauls will still be experienced. These unscheduled maintenance activities are a heavy burden on

the airline's finances and for commercial jet aircraft can cost as much as one million pounds per aircraft per year (Dunn, 1997).

1.2.2.3 Condition-based maintenance

It follows that large cost savings can be achieved by doing the maintenance at the airline's convenience and also by minimising the loss of useful component life. This is the aim of condition-based maintenance (CBM), which determines the health of the powerplant by monitoring the condition of its components. Therefore, the maintenance can be carried out before a fault occurs while simultaneously using up as much as possible of the component's useful life. However, this is an ideal situation and some of the following issues have to be considered:

- It is usually neither practical nor economic to monitor every component, and so, a selection has to be made on the most appropriate and critical parts.
- Condition monitoring involves the measurement of the variation of some parameter and therefore it is only suitable for components where a fault occurs in a time-dependent fashion. Similarly, some failures cannot be directly monitored, such as fatigue or creep of structural components. These parts will generally still be hard lived, i.e. they are removed from service as soon as their age reaches a specified expected life.
- When a fault is highlighted in the system, the operator has to determine whether the fault is real or if it is the result of a different problem such as measurement error.

Engine-health monitoring (EHM) systems have been developed in order to deal with some of the above issues.

1.2.3 Diagnostics and engine-health monitoring systems

EHM systems generally combine engine condition monitoring (ECM) and diagnostics techniques by collecting, processing and displaying data to allow meaningful conclusions regarding engine health to be drawn in a cost-effective way (Escher, 1995). EHM systems have been developed significantly over the years from the use of simple hand recording and calculating procedures to the use of sophisticated on-board data-gathering systems, which provide data for subsequent processing via a ground-based computer program (Barwell, 1987).
The advent of full authority digital engine-control (FADEC), which interfaces with the aircraft systems, has greatly facilitated data acquisition. Data from the electronic engine control unit and other measured engine parameters are transmitted in digital form to the aircraft condition-monitoring system, which also records in-flight aircraft data. The information is then forwarded to the ground-based system, either during aircraft turnarounds or now increasingly in-flight. It follows that, at any point in time, the operator has information on the engine's health, which allows him to make decisions on the maintenance required for that particular engine. Advanced EHM systems are capable of isolating component failures and evaluating the degree of component deterioration (Li, 2002a).

1.2.3.1 Concurrent techniques

Although this thesis is focused on performance analysis, this document would not be complete without some comments on the other methodologies that constitute an ECM system. A key point in any ECM system is the concurrent utilisation of a number of techniques to keep track of the behaviour of various components and the health status of the subsystems. Some of the common methods employed in ECM are:

- Lubricating oil analysis.
- Vibration monitoring.
- High pressure turbine exit temperature spread.
- Visual inspection / Boroscope inspection.
- Transient monitoring.
- Life-cycle counting.
- Exhaust gas temperature monitoring.
- Gas-path performance monitoring.
- Gas-path debris monitoring.
- Eddy-current checks.
- Radiography.

Some or all the listed techniques are used depending on the complexity and aim of the monitoring system. Even though the present work will focus on aero-thermal performance analysis, no technique is more relevant than the others. Actually, these techniques are commonly used in parallel in order to monitor the engine by crosschecking their outputs. An interesting case is represented by gas-path debris and aero-thermal analysis (English, 1995). The former is particularly suitable to quantify certain faults at an early stage, while the latter is able to quantify them when they have reached such a level that the component performance is directly affected. From this viewpoint, the two techniques can be considered complementary.

1.2.3.2 Benefits and concerns

The success of any health-monitoring system depends largely on the accurate classification, identification and quantification of faults. However, depending on the required accuracy, the complexity and associated computational resources of the system can be considerable. Besides, the reliability of the EHM system reduces when its complexity increases, ant this results in an engine availability reduction, in the case that the EHM equipment becomes safety critical.

The main advantages in using EHM systems are a better control of engine-condition and operation, improved maintenance shedule, improved engine availability with enhanced airline reputation and passenger satisfaction, as well as reduced costs of operation (Barwell, 1987).

Costs can be reduced by (Barwell, 1987, Burnell, 1995):

- A better understanding and quantification of in service engine deterioration.
- Ordering spares ahead of engine appearance in the maintenance shop.
- Scheduling engine removal convenient to airline operation and avoiding in-flight shut-downs and aborted take-offs.

- Improving maintenance as the operator has more information on which to base decisions; identification of the component faults prior to stripping the engine is of great advantage because it is difficult and time consuming to inspect an engine thoroughly enough to identify the faulty parts.

However, these benefits are only achieved as a result of considerable investments and provided that the necessary knowledge and personnel to operate an effective EHM system are available. In addition, some of stated the benefits, such as the ability to schedule maintenance convenient to airline operation and ordering spares in advance, can only be partly achieved with a simple diagnostics-based EHM system. In order to make the optimal use of the engine-monitoring information, a forecasting capability is required.

1.2.4 Prognostics

Prognostics is the ability to assess and predict into the future health condition of the engine or one of its components for a fixed time horizon or predict the time to failure (Brotherton, 2000). Prognostics capability can be integrated into the previously discussed EHM systems, by introducing forecasting algorithms able to provide predictions with the required confidence level.
The goals of EHM with prognostic capability are similar to those stated above. However, as pointed out previously, the scale of the benefits resulting from the addition of a forecasting capability is expected to be far greatest. One reason for this is that the prediction tool enables a higher level of condition-based-maintenance for optimally managing total life-cycle costs (Byington et al., 2002). Therefore, this ability to detect and isolate impending faults or to predict the future condition of a component based on its current diagnostic state and available operating data is currently a high priority research topic (Byington et al., 2002). Diagnostics and Prognostics in general are methodologies that make use of the ECM technique mentioned in section 1.2.3 to provide statements and advice concerning the status of the engine at the time when the parameters are monitored (i.e. diagnosis) or in the future by predicting how this status will evolve (i.e. prognosis). As far as prognostics is concerned, a lot of research has been carried out with respect to predicting the component life under the planned usage pattern from a mechanical point of view (e.g. considering creep and thermal fatigue), but forecasting the gas-path performance deterioration is a rather new research topic. We refer to gas-path diagnostics and prognostics when this methodologies use gas-path monitored dependent parameters (e.g. gas-path temperatures, pressures, shaft speeds and fuel flow).

1.2.5 Gas-path diagnostics and prognostics: an introduction

The performance of an aero-engine deteriorates over time as a consequence of the degradation of its components. Significant benefits with respect to maintenance servicing can be achieved through the application of accurate and reliable gas-path diagnostics and prognostics. Deterioration can affect relevant factors such as thrust (or power) and specific fuel consumption. As a consequence of progressive performance loss, operation of the engine can become cost ineffective (e.g. leading to excessive SFC) or even unsafe (e.g. insufficient take off thrust). Therefore maintenance techniques must be used to

ensure that the engine is operated cost-effectively and safely. An engine-diagnostic process estimates changes in component performance parameters (i.e. efficiency and flow capacity) given a set of measurements taken along the engine's gas-path. This allows the identification of the component(s) responsible for the performance loss and facilitates the choice of the recovery action to be undertaken. Despite its importance, an accurate assessment of the actual health condition of the engine components is difficult to achieve due to (i) the small number of measurements typically available, (ii) the measurement uncertainties and (iii) the simultaneous presence of engine and sensor faults. A preliminary observability study can be undertaken to investigate the suitability of the given set of measurements to diagnose the given faults.

Once the diagnosis is performed, the prognosis provides a prediction into the future of the engine's or the component's health condition. Gas-path prognostics is emerging as a novel research area of interest for the gas turbine industry. Maintenance support, flight operations, fleet management reliability engineering and quality-assurance teams would benefit from the use of a gas-path prognostics process that provides forecasts and advices based on the expected short and long term behaviours of the engine suffering from the diagnosed condition.

Industry has yet to obtain the full extent of the added value that advances in gas-path diagnostics and prognostics can offer. Among the quantifiable benefits from the use of appropriate gas-path diagnostics and prognostics are (Singh, 2003):

- Life extension, based on individual engine/component condition and usage profile. Application of "life expired" engines for other uses to absorb residual creep or fatigue lives.
- Reduced need for the holding of spares.
- Availability management to limit the need for unplanned maintenance. Improved "departure" statistics. Reduced in-flight shutdown rates and maintenance away from base. Enhanced airline reputation.
- Clarity in defining cost-effective aftermarket agreement objectives.
- Instrumentation selection against usage objectives.

1.2.6 Defence, marine and industrial applications

Relatively recent advances in computing, data collection and general modelling capability have enabled the development of maintenance strategies significantly based on condition monitoring. The greatest effort to develop a comprehensive and cost-effective monitoring system has been made for civil aero engines. These techniques are now being adopted for military, industrial, shore based as well as marine gas turbines (Figure 7).

Although ECM was first introduced with the civil air application in mind, it was recognised that also military engines can benefit by the new maintenance approach: operation and life-cycle costs can remarkably be improved by introducing ECM techniques. The importance of ECM for military engines can be properly appreciated by considering the following example: the eight-engine

strategic bomber B52 is planned to remain in service for a long time (2040 is a deadline proposed for the bomber's withdrawal). In the light of the extremely long life on wing in general required for military engines, health monitoring will play a key role in the operability and overall life-cycle cost of the military aircraft as well.

As far as the future marine gas turbine industry is concerned, demand for ever-increasing power from the propulsion engine, with the order-winning criteria being economy and reliability, faces the engine manufacturers. The changes in the role of the aftermarket, which are re-defining the business paradigm for the civil air-transport business, also have a resonance in the defence and marine sectors. The concept of contracts based on hourly-utilization basis payment, which are being introduced for the airlines, is becoming a reality for the marine and military propulsion as well.

Gas turbines play a important role in the energy-supply business and also here changes in the aftermarket are happening. Major users are setting up as competitors to the engine manufacturers in providing engine maintenance services, changing industrial relationships and economics. Substantial benefits from EHM systems are available for industrial gas turbines as well, although the interest of users in condition monitoring is relatively recent. At first, the use of vibration-based condition monitoring techniques for aero-derivatives gas turbines has made industrial users more sensitive to maintenance issues. Maintenance systems partially based on condition monitoring have recently been introduced on a wide range of machineries and are showing significant economic advantages.

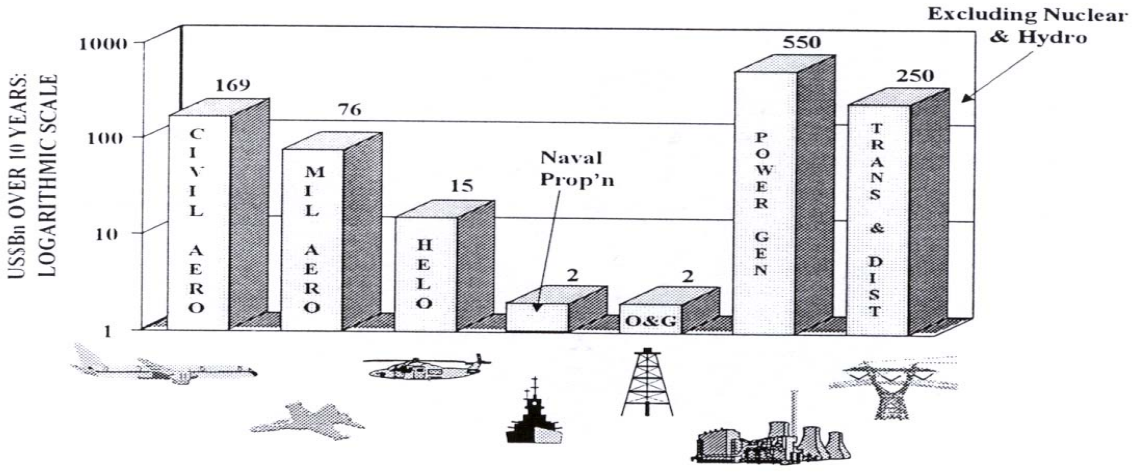


Figure 7: The market for power systems (Singh, 2001)

1.3 Research aim

1.3.1 Introduction

In the light of the considerations of the previous sections, engine-monitoring and analysis techniques emerge as a key area of investigation with a considerable importance for the gas turbines' business in general, and for the civil air-transportation one in particular. Moreover the availability of highly increased on-board computational power and data-storage capabilities are changing the engine-monitoring scenario. Hence, a new opportunity for enhancing the capacity of extracting information regarding the engines' health from available on-line and historical gas-path data is facing airlines and engine manufacturers.

Because of this growing importance of engine-monitoring systems, a lot of research (notably at Cranfield University) has been done on advanced gas-path diagnostics and prognostics methodologies to maximise on-wing time and improve maintenance operations. A University Technology Centre (UTC), funded by Rolls-Royce plc, was established at Cranfield University in 1998 to carry out research regarding gas turbine performance. The study described in this thesis has been undertaken, since October 2001, within the Theme 1 of the UTC, studying novel methods for gas turbine performance analysis, diagnostics and prognostics.

1.3.2 Research aim statement

The aim of this research was to develop a novel and versatile framework able to offer a high level of gas-path diagnostics competences, and prognostics capabilities for civil aero-engine on-board applications.

1.3.3 Preliminary objectives of this research

The research aim naturally separates into two elements: diagnostics and prognostics processes. As far as the **diagnostics** process is concerned, the research strategy (i.e. qualitative objectives) included the following steps:

- To conduct a review of the pertinent literature in order to identify research opportunities in terms of novel promising approaches.
- To define the requirements for an advanced gas-path fault diagnostic process.The importance of observability studies in the diagnostics process was considered.
- To develop a conceptual model in view of these requirements.
- To implement a novel approach based on that conceptual model.
- To evaluate the effectiveness of the developed diagnostics framework with simulated case studies based on real data recorded on-board.

The research strategy (i.e. qualitative objectives) for the **prognostics** process comprised:

- To conduct a review of literature in order to identify the problem and research opportunities in terms of potential effective methodologies.
- To define the requirements for an advanced gas-path fault prognostic process.

- To chose a systematic prognostics approach and to develop a conceptual model in view of these requirements.
- To implement a novel approach based on that conceptual model.
- To investigate different applications, including consideration of safety and economic interactions.
- To evaluate the potentialities and the effectiveness of the developed prognostics framework with simulated case studies based on real data recorded on-board.

1.4 Subject matter covered: a guide to the thesis

Chapter-1 is dedicated to set the scene for the project. The techno-economic reasons why this study was carried out are investigated. The emergence, in a strongly competitive environment, of new active aftermarket strategies to provide total care solutions make the engine manufacturers highly challenged to design advanced health monitoring and prognostics systems. Hence, the research aim is stated and the work-strategy undertaken is defined.

In Chapter-2, three fundamental elements of a performance analysis tool for gas turbine in-service health monitoring have been surveyed, they are: (i) observability analysis methods, (ii) gas-path diagnostics techniques, and (iii) gas-path prognostics techniques. A literature review from the early works up to the state of art of those methodologies was performed with the twofold aim of identifying pros and cons of the present methodologies and specifying the desirable requirements for novel diagnostics and prognostics systems.

Chapter-3 highlights the gaps in contemporary research that provide the guidelines for defining the requirements for a new potential framework. Hence, the broad objectives of the project are derived. Gas-path diagnostics and prognostics processes are considered separately. Contributions and benefits from the research project described in this thesis are summarised.

Chapter-4 describes a fuzzy-logic based diagnostics process. A brief introduction to fuzzy-logic techniques is given, together with a simplified example that is provided for a thorough understanding of all the features of a fuzzy-logic based process for gas turbine diagnostics. Then the set-up procedure of such a process for a Rolls-Royce Trent 800 engine is described and the choices made during its implementation are justified. A preliminary observability analysis of that engine's behaviour is performed and discussed. The results from case studies are analysed in order to evaluate the capabilities and the accuracy of the fuzzy diagnostics approach in the presence of noise in the measurements. The system is tested to prove that it is competent to achieve (i) SFI (single fault isolation) capability in the presence of noisy data, (ii) tuning over a known global deterioration level for all the performance parameters (baseline) computed for the previous flight, (iii) partial MFI (multiple fault isolation) capability with up to 2 degraded components (i.e. 4 performance parameters) considerably faulty at the same time, and (iv) SFI while isolating systematic errors in the measurements (biasses). A bias-tolerant system was devised by means of the NOT logical operator and a new formulation of the

fuzzy rules that includes the location of the bias. The method and its results are discussed in the light of the requirements stated in Chapter-3.

Chapter-5 presents an innovative prognostics framework which has been devised to meet the previously considered requirements. This makes use of two different statistic-based approaches for short and long term investigations that characterise different prognostic problems. The first technique is based on ARIMA models and some applications to the problem of performance rejection are considered. The second technique is based on regression analysis: some case studies regarding the optimal gas turbine maintenance cost to benefit ratio are investigated. A discussion of the proposed framework is given.

Chapter-6 presents the final conclusions, comments and recommendations of this investigation as to the accuracy and applicability of the techniques proposed, and summarizes the future studies whose importance emerged during the development of the current project.

Finally, appendices complement the main text.

CHAPTER 2 - METHODOLOGIES: STATE OF THE ART

2.1 Problem definition: gas-path diagnostics and prognostics

2.1.1 Introduction

As stated in the previous section, a considerable interest is devoted to reduce gas turbine operating costs. In this scenario, enhanced maintenance practice based on condition monitoring is conceived to improve engine availability and reduce costs of servicing. Hence, the need for a tool able to diagnose and prognose faulty engine-conditions emerges. In this contest, aero-thermal gas-path diagnostics and prognostics play a primary role in an industry in which the core of the businesses is based on aero-thermal performance. It has been recognised (Provost,1995) the importance of a preliminary observability study to determine the gas-path measurements that are needed to enable the required analysis of component changes and/or sensor biasses to be performed, as well as to assess both possible measurement redundancy and the ability of a set of measurements to successfully differentiate between all the component changes and sensor biasses being sought. This investigation is crucial for the interpretation of the gas-path analysis results in order to provide correct fault diagnoses, by determining which particular combinations of faults are not detectable. Once a fault is assessed, by means of a diagnostic process, a prognostic process provides the forecast regarding the evolution of the fault over time and its probability. Three critical elements of a performance analysis tool aimed at investigating the gas turbine health during its operation here arise:

- Observability pre-processing.
- Gas-path diagnostics process.
- Gas-path prognostics process.

The purpose of this chapter is to identify the fundamental difficulties with these three processes, to find solutions to the issues in the literature and eventually to highlight the limitations of those solutions. To achieve this, the following aspects are investigated in the next sections:

- Performance analysis role within the gas turbine design procedure: production engines, pass off tests, after service tests and on wing monitoring.
- Difficulties in devising performance analysis techniques.
- Causes of deterioration.
- Evolution of the deterioration.
- Observability problem definition.
- Gas-path diagnostics problem definition.
- Gas-path prognostics problem definition.
- Survey of methodologies for observability analysis.
- Survey of methodologies for gas-path diagnostics.
- Survey of methodologies for gas-path prognostics.

In view of the outcome of the investigation of the listed aspects, Chapter-3 is dedicated to highlight gaps in the contemporary research that provide the guideline to define the objectives of this project.

2.1.2 Gas turbine performance analysis: introduction

The synthesis and analysis processes in a typical Gas Turbine design procedure are illustrated in Figure 8.

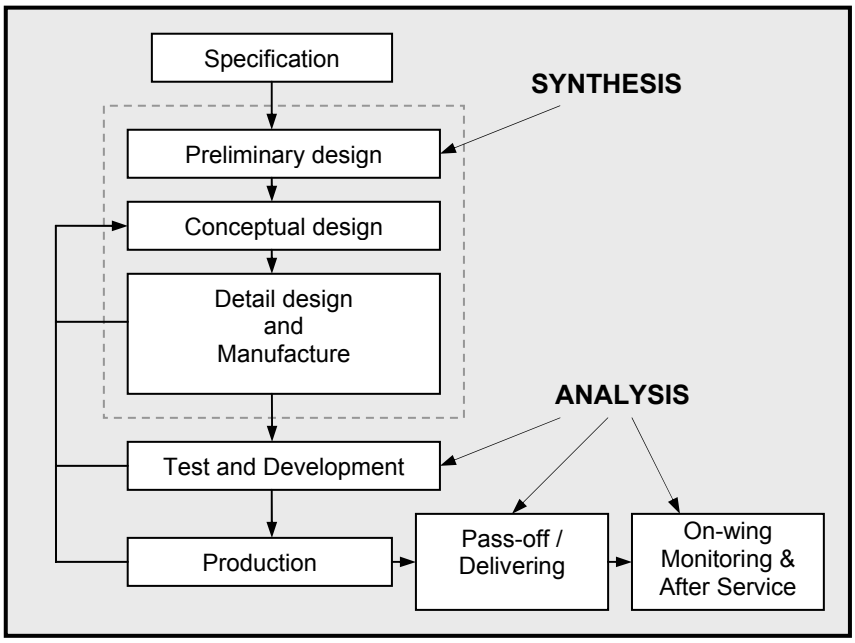


Figure 8: Typical gas turbine design procedure - Synthesis and Analysis processes

Gas turbine analysis can have different objectives depending on what the results will be used for. As far as the analysis of performance is concerned the main purposes are three. First, throughout the development phase, test data analysis provides a detailed understanding of the performance of all components useful for the design process. Second, before the engine is delivered to the customer or after a major overhaul, a pass off test is done to assess the prescribed performance. At this stage a thorough performance analysis is required only if the test failed. The third application of analysis is in condition monitoring (or after service) to achieve cost-effective maintenance.

2.1.2.1 Performance modelling and analysis

The simplicity of calculation of gas turbine cycles has allowed the implementation of many simulation programs that model gas turbine using characteristics of their components and thermodynamic relationships. The steady state performance modelling problem of calculating the values of dependent parameters knowing the operating condition and some independent parameters, can be analytically expressed by means of equation (2.1).

$$z=h(x,w) \qquad (2.1)$$

where:

$z \in R^M$ is a vector of M measurements (dependent parameters); typically they are temperatures, pressures, spool speeds, airflows, fuel flows and thrust.
$x \in R^N$ is a vector of N performance parameters (independent parameters); they are efficiencies and flow capacities for rotating components and discharge coefficient for propelling nozzles. They are also called health parameters.
$w \in R^P$ is a vector of P environmental and power setting parameters.
h is a vector-valued function, typically non-linear, it represents the gas turbine mathematical model.
The inverse process, namely performance analysis, consists of calculating performance parameters and therefore the performance of the gas turbine components, using measurements as input. In particular, the analyst is interested in identifying parameters changes from a presumed nominal state (rather than their absolute values).
Figure 9 based on a famous diagram by L.A. Urban, shows the various relationships in the analysis process: the analyst works from right to left on this diagram, in order to isolate and assess changes in engine module performance from knowledge of measurements taken along the engine's gas-path. Discernable shifts in measurements from a baseline level are necessary for determining the shift in engine operation from a prescribed level.

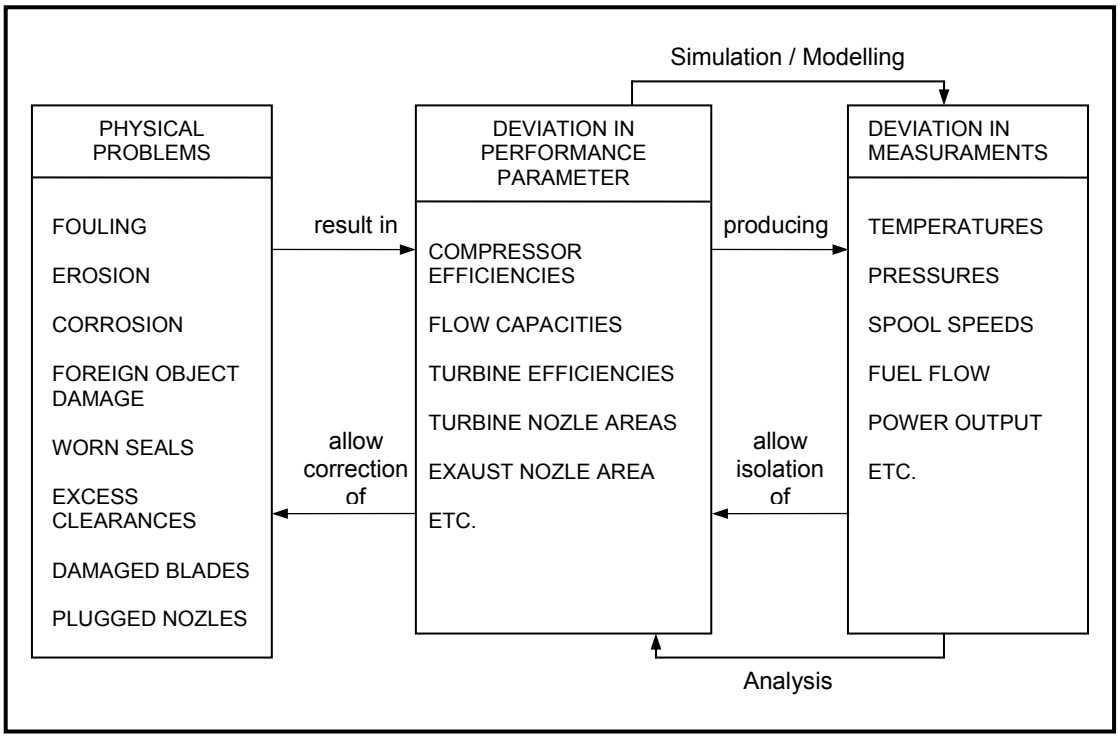


Figure 9: Performance Analysis, (Urban,1969)

2.1.2.2 Complexity of the analysis problem

Gas-path diagnostics methodologies are devised in order to assess changes in performance and not absolute values. For a given operating point, we are interested in calculating the deviation between the nominal parameter value and the observed value. A thorough approach must recognise that the differences

between expected and observed measurements can be due to the following reasons:

- Component performance parameter changes.
- Random error and bias or systematic error in power setting parameters of the gas turbine.
- Random error and bias in the measurements.

Unknown performance parameter changes and measurement biasses affect the gas turbine resulting in a set of measurement deviations that are different from the observed measurement deviations because corrupted by random noise (Figure 10). A correct approach requires to modify equation (2.1) as follows:

$$\Delta z = h(\Delta x, w) + v + b \qquad (2.\ 2)$$

where:

Δ is the delta operator that calculates the difference between observed and expected values.

$v \in R^M$ is the zero-mean measurement noise vector.

$b \in R^M$ is the sensor bias.

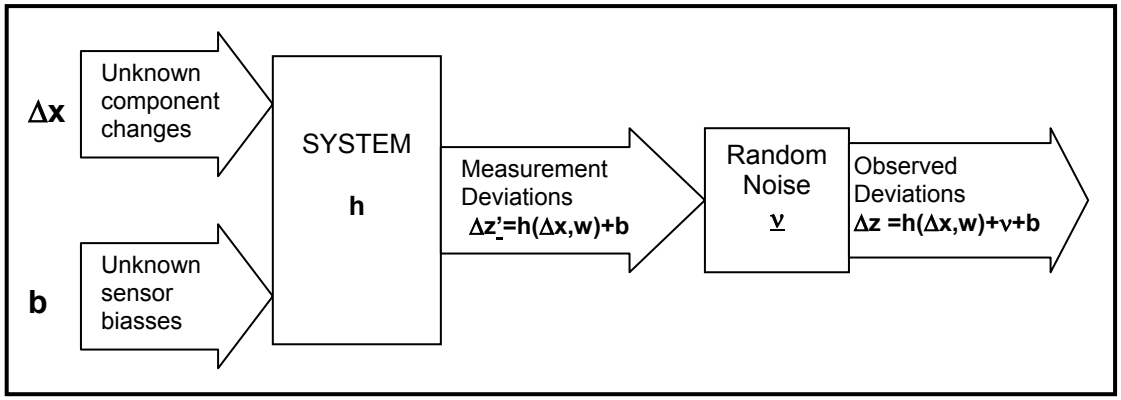


Figure 10: The analysis process

Besides the corruption of the measurements due to sensor errors, a lack of system observability typically complicates the analysis problem. System observability would require a large number of sensors to be installed on the engine. Nevertheless, as far as in-flight data analysis is concerned, this number is seriously limited by consideration of weight, bulk, cost and reliability as discussed in detail in section 2.2.

An ulterior issue that needs to be taken into account is that the data must typically be corrected for factors such as installation losses (e.g. associated with test cells), instrumentation calibration errors, effects of bleed and effects of Reynold's number. After the correction a gross delta value for each parameter is obtained including the net delta due to the degradation which is the portion of the diagnostic interest, the portion due to instrumentation non-repeatability and typically the presence of other unsought elements as shown in Figure 11.

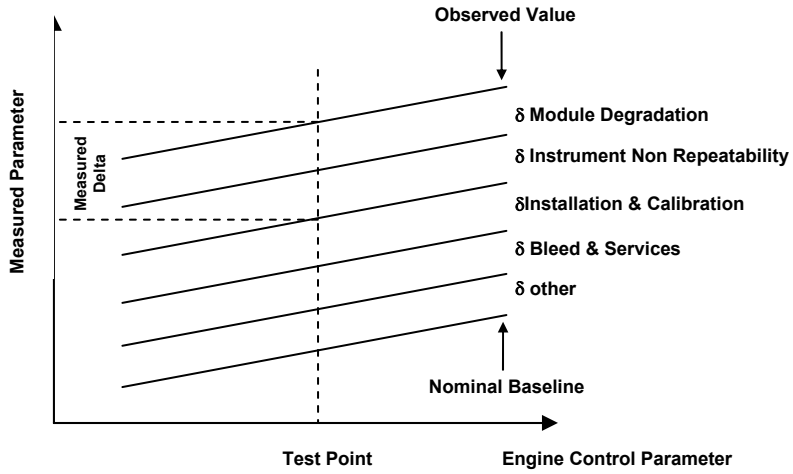


Figure 11: Calculation of measurement of deltas

2.1.2.3 Measurements uncertainty

Measurements are always subject to errors. In the 70's, the lack of a standard method for estimating the errors associated with gas turbine performance data had made impossible to compare measurement systems between facilities, and there had been uncertainty over the interpretation of error analysis. A lot of work has been done to define standard method of treating measurement error for gas turbine engine performance parameters (Abernethy et al, 1973a; Abernethy et al., 1973b). Performance data errors propagate from the basic measurement through functional relationships, and therefore there is the need for introducing methods for modelling measurement error and handling error traceability (La Grandeur, 1986). The error is the difference between what we measure and what is true (Figure 12). In this work a brief introduction to uncertainty models is going to be given. We distinguish two components of measurement error: the bias error and the precision or random error (Figure 13). The uncertainty estimate is the interval about the measurement that is expected to encompass the true value.

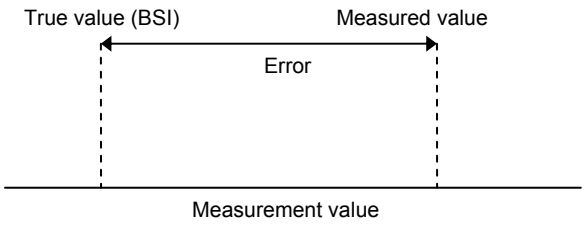


Figure 12: Measurement error

Random error or precision error is seen in repeated measurements (Figure 13). There are always numerous small effects which cause disagreements. The standard deviation is used as a measure of precision error (Abernethy et al., 1973a; Abernethy et al., 1973b).

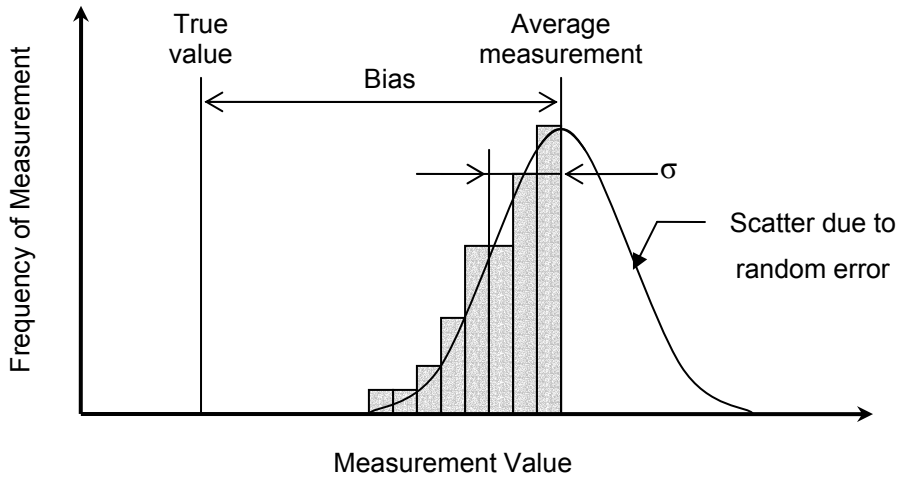


Figure 13: Measurement Bias and Precision error

The second component, bias, is the constant or systematic error (Figure 13). In repeated measurements, each measurement has the same bias. To determine the magnitude of bias in a given measurement situation, we must define the true value of the quantity being measured. Therefore, the bias is difficult to be identified. We must, instead, rely on the best information available. Usually we rely on the engineering judgment of instrumentation and measurement engineers to provide an upper limit or bound on the bias.

Abernethy et al. (1973a) categorize bias into five classes (Table 1): large known biasses, small known biasses, large unknown biasses and small unknown biasses that may have unknown sign (±) or known sign. The large known biasses are eliminated by comparing the instrument to a standard instrument and obtaining a correction. This process is called calibration. Small known biasses may or may not be corrected depending on the difficulty of the correction and the magnitude of the bias. The unknown biasses are not correctable. That is, we know that they may exist but we do not know the sign or magnitude of the bias. Small unknown biasses stem from errors introduced from the hierarchy of calibrations that relate the BSI (British Standard Institution) standard to the working instrument. Every effort must be made to eliminate all large unknown biasses. The introduction of such errors converts the controlled measurement process into an uncontrolled worthless effort. Large unknown biasses usually come from human errors in data processing, incorrect handling and installation of instrumentation, and unexpected environmental disturbances such as shock and bad flow profiles. We must assume that in a well controlled measurement process there are no large unknown biasses. To ensure that a controlled measurement process exists, all measurements should he monitored with statistical quality control charts. Drifts, trends, and movements leading to out-of-control situations should be identified and investigated. Histories of data from calibrations are required for effective control (Abernethy et al., 1973a; Abernethy et al., 1973b).

Table 1: Five types of bias error, (Abernethy et al. 1973a)

	KNOWN SIGN AND MAGNITUDE	***UNKNOWN MAGNITUDE***	
LARGE	(1) CALIBRATED OUT	(3) ASSUMED TO BE ELIMINATED	
SMALL	(2) NEGLIGIBLE	(4)UNKNOWN SIGN	(5) KNOWN SIGN
	CONTRIBUTES TO BIAS LIMIT	CONTRIBUTES TO BIAS LIMIT	

2.1.3 Deterioration causes

As far as in-service gas turbine monitoring is concerned, its application is aimed at identifying changes in health parameters caused by physical degradations of the engine's components. The following section will discuss the physical causes that generate mechanical degradations of the engine.

Fouling can be defined as "the degradation of flow capacity and efficiency caused by adherence of particulate contaminants to the gas turbine airfoil and annulus surfaces" (Diakunchak, 1992). Gas turbines are particularly susceptible to fouling because of the large quantities of air they ingest. It can normally be eliminated by washing and affects both the compressor and turbine. While compressor fouling is caused by particles that enter the engine with the inlet air, turbine fouling is also caused by fuel contaminants. The incoming air consists of hard and soft particles. Hard particles such as dust, dirt, sand, rust, ash and carbon particles and soft particles such as oil, unburned hydrocarbons, soot, airborne industrial chemicals, fertilizers, herbicides etc. can provide a source for fouling. Many of the contaminants are smaller than 2μm (Kurz and Brun, 2001). In the case of compressor fouling, the change in blade shape causes a reduction in compressor flow capacity and a reduction in compressor isentropic efficiency. The effect of fouling on compressor flow capacity is more significant than the effect on efficiency. The reduction in mass flow capacity varies with operating speed, ambient temperature and altitudes (Saravanamuttoo, 1985). Furthermore, compressor fouling not only reduces the flow capacity and efficiency, but also reduces the compressor surge margin and this may result in compressor surge (Diakunchak, 1992).

Erosion can be defined as "the abrasive removal of material from the flow path components by hard particles in the air or gas stream" (Diakunchak, 1992). The contaminants causing erosion, are particulates such as dirt, sand, rust ash, carbon particles and dust. They are bigger than those causing fouling and typically have a diameter of 20m or more (Diakunchak, 1992). Erosion increases the surface roughness, changes the aerofoil profile and throat opening, and enlarges blade tip and seal clearances. Unlike fouling, erosion can

not be recovered by washing but can only be restored through the repair or replacement of components (Escher, 1995).

Corrosion can be defined as “the loss of material from flow path components caused by the chemical reaction between these components and contaminants that enter the gas turbine with the inlet air, fuel, or injected water/steam”. Salts, mineral acids, and reactive gases such as chlorine and sulphur oxides, in combination with water, can cause wet corrosion, especially of the compressor airfoils. Elements like sodium, vanadium and lead in metallic or compound form can also cause high temperature corrosion of the turbine airfoils. Hot end surface oxidation is another form of corrosion (Diakunchak, 1992). Similar to erosion, corrosion can result in the loss of material and increase in surface roughness. In addition, corrosion results in a loss of performance and service life of the component affected. Typically, compressor corrosion results in a reduction in compressor flow capacity and isentropic efficiency, whilst turbine corrosion results in an increase in turbine effective area/flow capacity and a reduction in isentropic efficiency. Coatings are usually applied on turbine and compressor aerofoils to protect them against corrosion. Besides, corrosion diminishes the in-service life of the affected component (Escher, 1995).

Foreign object damage (FOD) is caused when “larger objects such as hailstones, runway gravel or birds are ingested into the engine” (Lakshminarasimha et al., 1994). Alternatively FOD can be the result of engine internal pieces (including ice from the inlet) breaking off and being carried downstream. The effect of FOD can range significantly from non-recoverable (with washing) performance deterioration to catastrophic engine failure (Diakunchak, 1992). The effect of FOD on performance degradation varies significantly with the severity of the damage. FOD results in a large reduction of the component isentropic efficiency and in some cases can change the flow capacity of the damaged component; the value is very much dependent on the severity of the damage. An increase in flow capacity can be the result of lost blades. A decrease of flow capacity can be the result of foreign particles blocked in the gas-path. A blockage can be caused by desert sand that has been virtually glued to the turbine blades because of the heat.

Thermal distortion is a fault that normally occurs at combustor exit/turbine entry where temperatures are highest. Distortion is caused by problems such as faulty fuel nozzle spray patterns and warped combustor components which cause changes in the radial and circumferential temperature traverse pattern at the combustor exit. This can result in temporary or permanent deformation of downstream components such as cracked, bowed, warped, burned, lost or damaged turbine nozzle guide vanes, area changes, increased leakage, and relative thermal growth between the static and rotating members (English, 1995). High temperature can cause the first stage turbine blades to untwist. These blades untwist as a result of creep damage during sustained high temperature operation (MacLeod et al., 1992). Bowed, burned, warped, untwisted or damaged blades can cause a reduction in turbine isentropic efficiency due to increased air leakage and reduced airfoil performance. The

damage of the blades can also result in changes to the effective flow area. However, the most significant effect will usually be on turbine isentropic efficiency (MacLeod et al., 1992).

Rubbing wear is the removal of material from the rotor blade tips and knife edge seals due to contact between static and rotating parts (Zaita, 1998). Many engines use abradable surfaces where a certain amount of rubbing is allowed during the run-in of the engine in order to establish proper clearances. The material removal will typically increase seal or tip gaps (Kurz et al., 2001). Contact between rotating and stationary parts can be induced by several features (Crosby, 1986):

- Relative thermal growth between the static and rotating members.
- Centrifugal growth of the rotating member.
- Axial movement of rotating parts either due to rotor end loads or "flapping" of the rotor blades due to some induced vibration.
- Distortion of casing relative to the rotors due to heavy externally induced loads, e.g. in-flight turbulence, heavy landing.
- Engine operating procedure.

Rubbing wear accounts for the major part of deterioration, especially in the early stages of an engine's service life (Crosby, 1986).

2.1.4 Evolution of the deterioration

When the gas turbine is in-service, performance losses resulting from the deteriorations of its components get progressively worse, unless appropriate maintenance is undertaken. Little quantitative information is available in the public literature about the evolution of the deterioration during engine operation. Detailed studies were carried out during the late 1970s and early 80s under the heading of the NASA JT9D and CF6 engine diagnostics program (Sallee, 1978, Sasahara, 1986 and Wulf, 1980). The results from the analysis suggest that deterioration can be divided into two time frames (Sallee, 1978):

- Short deterioration that occurs rapidly in the first few hundred flights after entry into commercial service.
- Long term deterioration that occurs more gradually as service usage accumulates.

In Figure 14 short term losses are defined as those occurring at the aircraft manufacturer during aircraft acceptance flights prior to initiation of revenue service (Wulf, 1980). The acceptance test allows the engine to slowly adjust its clearances between rotating and stationary components without damaging them (Naeem et al., 2001). Thus, the engine is delivered to the customer with a declared performance and mechanical condition compatible with minimum in-service deterioration (Crosby, 1986). The initial installation losses shown in Figure 14 are losses experienced during revenue service but prior to the first removal for repair.

After the first installation the engine continues to degrade but most of the deterioration losses can be recovered during the refurbishment, so that there is only a slight increase in unrestored performance. The rate of deterioration does generally not follow a linear path as shown in Figure 14 but exhibits a rate that diminishes with time. During the early life of the engine the main cause of

deterioration is the increase in clearances. There obviously comes a point where the clearances no longer increase except under rare instances of violent manoeuvring and hence the rate of deterioration decreases. Nevertheless, the performance of the engine will continue to deteriorate due to erosion corrosion and fouling, however at a much slower rate. Figure 15 shows a typical deterioration rate versus cycles, for the RB211.

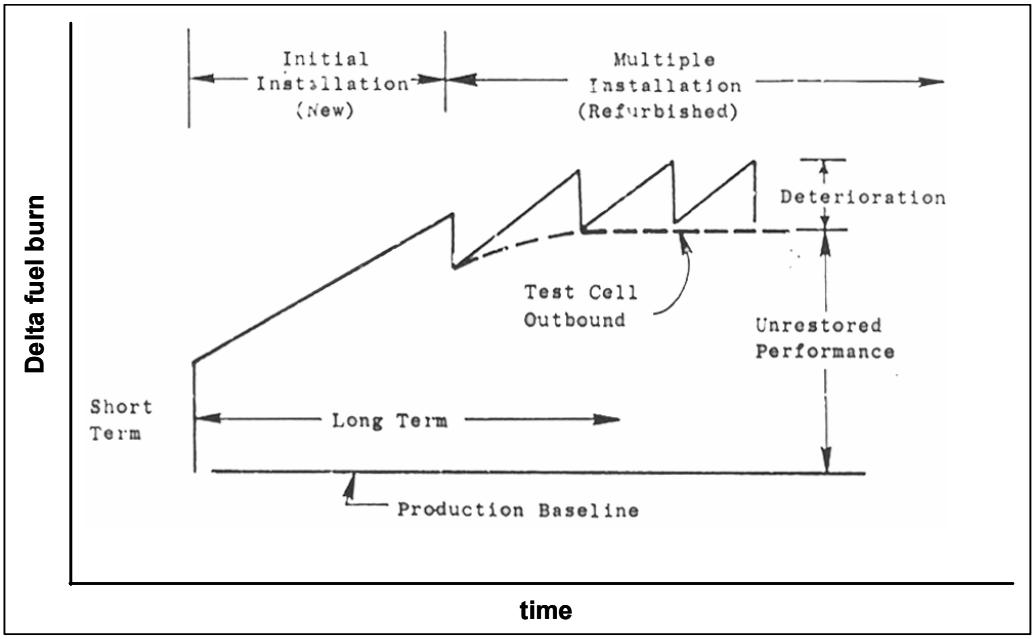


Figure 14: Engine deterioration schematic, (Wulf, 1980)

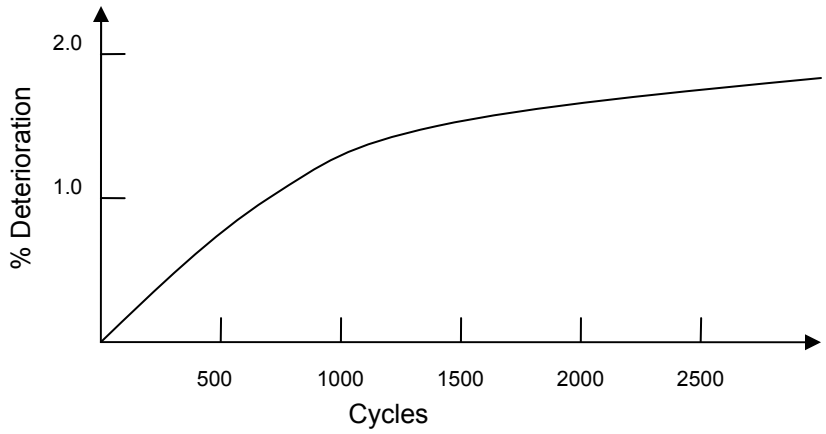


Figure 15: RB211 deterioration, (Crosby, 1986)

Three types of engine deteriorations can normally be identified (Diakunchak, 1992):

- Recoverable, as a result of cleaning, washing or general overhaul. This comprises all the deteriorations that can be eliminated by simply washing the engine and hence refers mainly to fouling.
- Non-recoverable, even after cleaning and washing. This includes all the surface deposits that can not be removed even with regular washing of

the engine. Additionally, any flow path damage, surface erosion/corrosion, tip and seal clearance increase, etc., will not be affected by cleaning and is therefore also referred to as non-recoverable deterioration. After a maintenance action, the performance loss due to recoverable and non-recoverable deteriorations is restored, while unrestored performance is accountable for permanent component degradation Figure 14.

- Permanent, this is not recoverable, despite an overhaul including for instance the re-establishment of all clearances and replacement of damaged parts. This is because some of the deteriorations that have occurred are too difficult and/or expensive to remove.

Short term deterioration resulting from the acceptance test is generally permanent while long term deterioration can usually be classified as a recoverable or non-recoverable performance loss.
Engine deterioration can be correlated against hours of service life or cycles, where one cycle is a flight consisting of a take off, a cruise period, a descent and a landing. During his experience in the JT9D diagnostics program, Sallee (1978) found that, as far as engines for civil application are concerned, many of the deterioration mechanisms correlated better with the number of flights rather than hours. The reason for this is that a civil aircraft spends the majority of its time during cruise where the engine operates at lower rotational speeds, pressure and temperatures than at take off. In addition a significant amount of the deterioration takes place when the aircraft flies through polluted atmosphere (including sand, birds…) near the ground. Furthermore, rubbing wear typically occurs during transient phases, such as take-off and climb, where the potential mismatch in thermal and centrifugal growth is largest. Therefore, the rate of deterioration is generally correlated against engine cycles rather than hours (Crosby, 1986).

2.1.5 Gradual and rapid deteriorations

Fouling, blade erosion and corrosion, worn seals, excessive blade tip clearance and their synergic effects induce gradual changes in the thermodynamic performance of the engine and its components. This results in gradual changes in the set of measurements. Foreign object damage, system failures and sensor faults result in rapid changes in the set of measurements.
Therefore gradual and rapid deteriorations can be distinguished and treated separately (Figure 16). The former implies that all the engine components are deteriorating slowly, whereas the latter may be the result of a single event. The presence of two different fault-mechanisms and the difficulties in solving simultaneously the two problems with the same diagnostics algorithm has led to the necessity of implementing two complementary diagnostics methodologies. There are several techniques available to address the problem of estimating gradual as well as rapid deteriorations, namely MFI (multiple fault isolation) and SFI (single fault isolation) methods respectively (Volponi, 2003). Traditionally a diagnosis has been mostly performed by inspecting a single-point observation leading to a snap-shot calculation. MFI and SFI methods, in order to operate simultaneously, require an algorithm for event detection (e.g. based on a time-series analysis). In a similar way, a distinction between gradual and rapid

degradations is considered necessary in modelling the evolution of the engine deterioration for prognostics studies. As far as gradual deterioration is concerned forecasting algorithms can be devised that make prediction based on historical data; whereas rapid deterioration, once detected and estimated, can be taken into account in the prognosis. Besides, hazard plots, specific to the type of engine and the mission's route, can be used to predict the probability that such events occur (Marinai et al., 2003c).

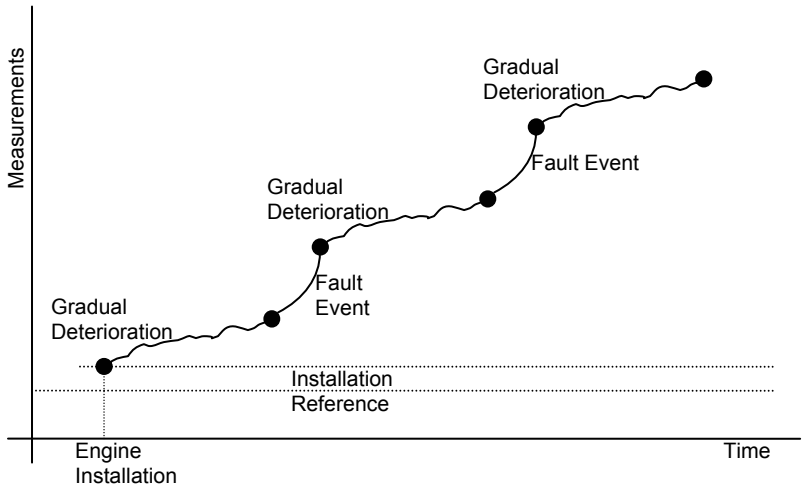


Figure 16: Gradual vs. Rapid deterioration

Table 2: Deteriorating causes

TYPES OF FLIGHT-ENVIRONMENTS	CAUSES OF DEGRADATION
Rainy	Ground water ingestion / Acid rains
Icy	Ice ingestion
Industry	Pollutant
Presence of birds	Birds ingestion
Sea	Salt
Desert / dusty environment	Sand
Volcano	Smoke
Military Missions	Gun residue / smoke/fire ingestion
Fire fighter	Smoke / Soot

It can be obsedved that the information for predicting an event and the one for the prediction of the evolution of deteriorations have different nature. As a metter of fact a relatively short time series of recorded in-flight data for one specific engine contains only information concerning the evolution of gradual deterioration. Whereas, if the interest is in the statistics of a possible event that can happen to a given engine during a given operation we should be looking at a blend of different data. These includes hazard plots - with the probability that a given fault happens over time for a given family of engines, and information regarding the type of flight in particular the environment conditions. It is well known that engine's components degradations have different temporal evolutions in different operating conditions depending on the environment

severity and the operating point(s) of the engine. The list of types of flight-environments given in Table 2 is by no means exhaustive but gives the flavour of how diversely they can affect the degradations.

2.1.6 Summary: definition of the gas-path diagnostics and prognostics problems

The main goals of a Gas-Path Health Monitoring System (Figure 17) are on the one hand the assessment of the component's current health (combining all the information available) by means of the diagnostic process and, on the other hand, the prediction of the engine health for a fixed time horizon or the prediction of the time to failure via the prognostic one. The importance of a preliminary observability analysis was recognised. In this section (2.1), the observability, diagnostics and prognostics problems were defined highlighting their complexity and the associated difficulties.

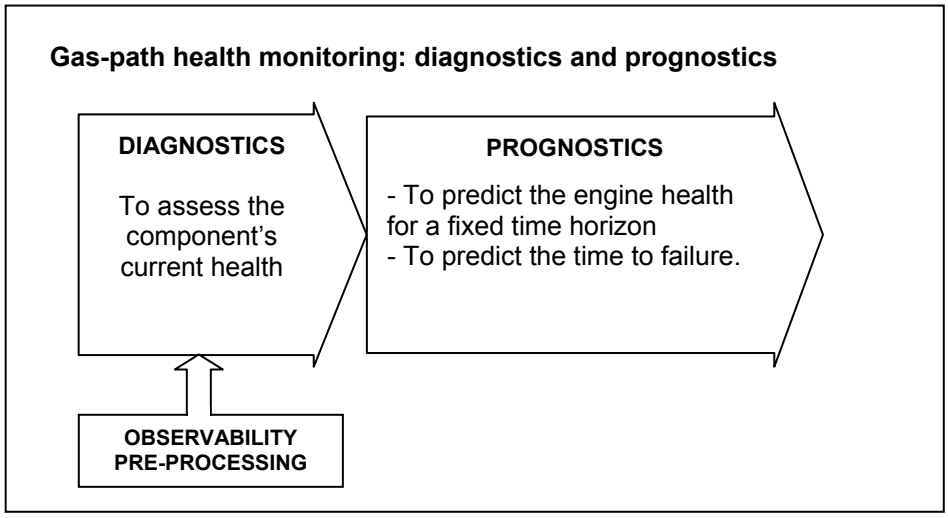


Figure 17: Gas-Path Health Monitoring System

An accurate diagnosis of the actual health condition of the engine components is difficult to achieve due to (i) the small number of measurements available (lack of observability), (ii) the measurement uncertainty, i.e. noise and biasses, and (iii) the simultaneous presence of engine and sensor faults. In the past three decades many methods were deviced to deal with these problems: techniques developed to undertake observability investigations are considered in section 2.2, while a range of gas-path diagnostics approaches are reviewed in section 2.3.
The outcome of the diagnostics process, namely the estimated performance parameters, as well as the monitored gas-path variables can be collected in time-series and can be used to make forecasts and prognoses concerning the health of the engine and/or its components. Pertinent works are reviewed in section 2.4.

The techniques for health monitoring developed in the early '70s were conceived to suit requirements of severe computational limitations. The availability of highly increased on board computational power and data-storage capability is going to change the engine health monitoring scenario. A new

opportunity of enhancing the capability of extracting statistical information regarding engines' health from available historical gas-path data is facing airlines and engine manufacturers. In this contest, the ability of modelling and predicting deteriorations, by means of forecasting algorithms, has become a new potential research area, in order to achieve substantial economic- and safety-related gains. Gas-path parameter forecasts can be used in the decision-making process for scheduling missions and planning maintenance. The maintenance as well as spare-part orders can be effectively planned by means of an accurate prognostics capability for a fleet of engines. The development of a prognostics competence is critical for:

- Estimating the risk of failure for a given lead time horizon (e.g. for risk analysis, performance rejection).
- Identifying the optimal maintenance costs to benefit ratio.
- The optimal management of the fleet.
- A competent combination between the engine-health monitoring process and the engine parts life-tracking process.

2.2 Observability study: procedures

This section discusses methods of determining the suitability of different sets of measurements, taken through the gas-path, to provide the information needed to estimate changes in performance parameters.

2.2.1 Introduction

A thorough evaluation of the health of the various gas-path components requires a large number of performance parameters to be assessed and a large number of sensors to be installed. On the other hand, the number of sensors is seriously limited by weight, bulk and cost concerns. Besides, the reliability of the monitoring system may even be reduced if the number of sensors increases, as the probability of a sensor fault increases. As a consequence, in-service monitoring has to be carried out with a limited number of sensors.

To diagnose specific faults the instrumentation set should be properly chosen. In fact given a set of sensors some faults will be more observable than others. Nevertheless, diagnostics usually has to be performed using the sensors available, which may not be the most suitable, because often the choice of sensors is not dictated by gas-path diagnostics requirements. The lack of observability that affects most of the diagnostics processes is a very important issue (Gelb, 1974), even though the development of a quantitative approach to gas turbine engine faults observability has long been neglected (Provost, 1995).

Increasing the number of sensors can enhance the diagnostics capability. Nevertheless, a number of issues have to be taken into account. Considerable cost is related to sensors installation and maintenance. This suggests that an excessive use of on-board instrumentation can result in a loss rather that a gain in terms of Return on Investment of the Engine Monitoring System (Figure 18).

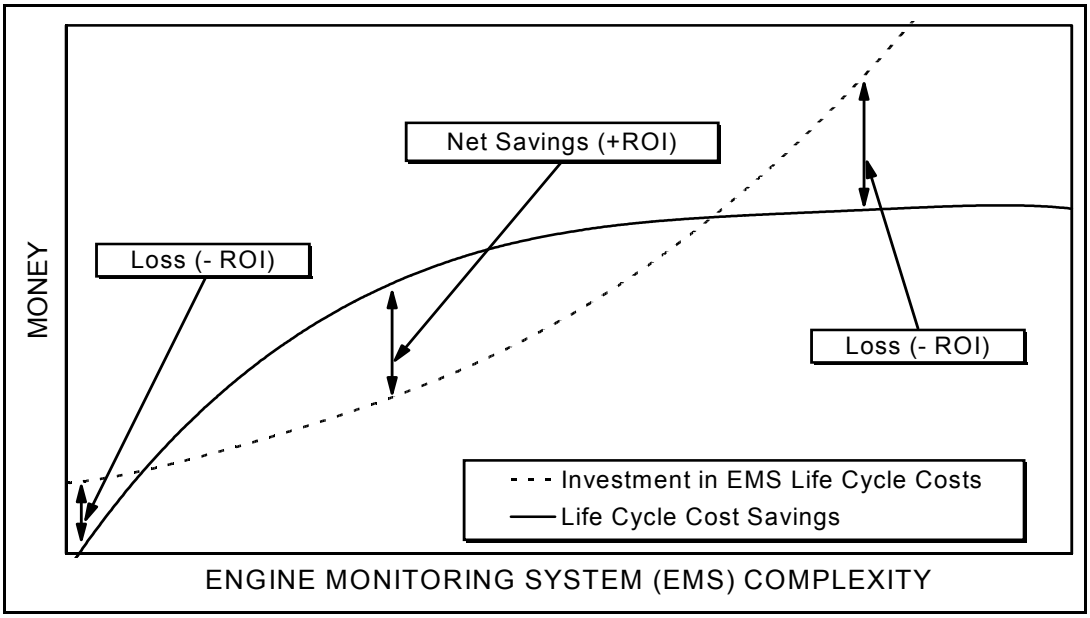


Figure 18: Return on investment of an engine monitoring system.

2.2.2 Available sets of measurements

The list of gas-path measurements available to diagnose component performance changes can vary significantly depending on the engine being analysed: production engines especially when installed in aircraft have very limited set of measurements available. Development engines have anything from a limited set, tailored to the monitoring of endurance, running to full suite of pressure/temperature rakes, in many planes designed specifically for detailed component performance determination. In Table 3 the gas-path measurements that are usually available in a modern turbofan are listed (Provost, 1995).

Table 3: List of the gas-path measurements that are usually available

Shaft speed
Fuel flow
Thrust (on a test bed)
Total inlet airflow (derived from total and static pressure and total temperature)
Total and/or static pressures in the engine gas-path (either single or multi-tapping)
Total and/or static temperatures in the engine gas-path (either single or multi-tapping; 'static temperature' = wall-mounted transducer)
Shaft torque
Power setting parameter, e.g. EPR

In addition, measurements defining the flight conditions (air inlet condition and engine forward speed) must be available (Table 4), e.g. altitude, Mach number, total air temperature (flight), cell pressure and air intake temperature (test bed), wet and dry bulb temperature (for humidity determination).

Table 4 Flight condition

Flight condition: -air inlet condition -engine forward speed	Altitude
	Mach Number
	Total air temperature (flight)
	Cell pressure and air intake temperature (test bed)
	Wet and dry bulb temperature (for humidity determination)

2.2.3 Test for correlation between measurements

This section discusses a test described by Provost (1995) that should be undertaken to check whether any of the chosen gas-path measurements respond to all (or nearly all) the gas-path parameter changes in similar way to other measurements. The result of this test indicates whether there is any redundancy between measurements. For example if two measurements are present purely for performance analysis reasons and the test shows analytical redundancy, deletion of one of them may be considered if cost of providing such instruments is high.

According to Provost (1995), a first linearization of equation (2.1) is required. Having generated two lists, one containing possible component performance parameter changes and the other containing available measurements, the method is based on the approximation that the set of M measurement deviations Δz are expressed as a linear combination of N elements of a constant vector Δx.

$$\Delta z = H \cdot \Delta x \qquad (2.3)$$

The Exchange Rate Table H is determined calculating the percentage change in each gas-path measurement for a 1% change in each component performance parameter (or unit change if the datum value is zero) in turn, at the flight conditions and power level at which the analysis is to be done. An example of the format of a component Exchange Rate Table is shown in Table 5.

Table 5: 2D Component Space.

vector: % change in measur.A for a 1% in comp.1, comp2

	%ΔZ for 1% Component Change 1	%ΔZ for 1% Component Change 2
measurement A	**	**
measurement B	**	**
measurement C	**	**
measurement D	**	**
…..	…..	…..

The test is based on the evaluation of the angle between pairs of vectors; this can be done performing the following steps:

1. Normalization: divide each element in each row of the Component Exchange Rate Table by the square root of the sum of the squares of the elements in that row.
2. Dot product: multiply this matrix by its transpose.
   $\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + ... = |a||b|\cos\vartheta$, the dot product of every pair of normalized vectors produces the cosine of the angle between them.
3. Check correlation: the resulting symmetric matrix of cosines can be inspected for possible correlation between the measurements.

Experience indicates that cosines with magnitude greater than 0.7 are worthy of note (Provost, 1995).

- 0.7-0.8 some correlation
- 0.8-0.9 high degree of correlation
- 0.9-1 very high degree of correlation. The measurements are redundant.

2.2.4 Test for correlation between performance parameter changes

This section discusses a second test discussed by (Provost, 1995), that enables us to check whether any of the component changes and/or sensor biasses produce similar change in all (or nearly all) of the gas-path measurements. This test indicates whether any component changes or sensor biasses cannot be distinguished from on another, which may be unacceptable. This means that a

mathematical solution of the problem does not exist because the vectors are linearly dependent, being the number of measurements considered not sufficient. Classic cases of such situations are:

- HP compressor efficiency changes and HP turbine efficiency changes cannot be distinguished from one another if HP compressor delivery temperature is not available.
- Generally, HP turbine capacity changes, cooling air-flow changes and HP compressor delivery pressure sensor biasses cannot be distinguished from one another.

According to Provost (1995) three tables can be generated to create the System Matrix (Figure 19):

1. The Component Exchange Rate Table that contains the changes in the gas-path measurements due to 1% change in each component performance parameter.
2. The Environmental Exchange Rate Table (or Slope Table) that is obtained by calculating the changes in the gas-path measurements due to environmental measurement biasses. The environmental sensor biasses are taken into account by determining the effect that 1% biass in each in-flight condition and power setting measurements has in the measurement set. This is done by changing the flight condition parameters and the power setting parameter by -1%, and calculating the resulting percentage changes in the gas-path measurements.
3. The Identity Matrix is introduced in order to include the gas-path measurement sensor biasses. Such an identity matrix merely states that a 1% sensor bias produces a change of 1% in that measurement while leaving the others unaffected.

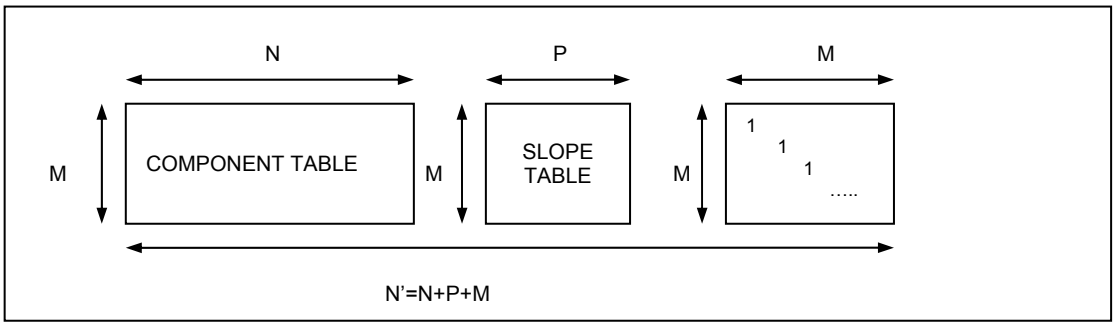


Figure 19: System Matrix

A 'Measurements Space' is defined as a coordinate system that contains component/bias vectors that are the columns in the System Matrix (Table 5), parallel vectors indicate similar effect on the measurements.

Table 6: 2D Measurement Space

vector:	%ΔZ for 1% Component Change 1	%ΔZ for 1% Component Change 2	%ΔZ for 1% Component Change 3	……	%ΔZ for 1% Meas.A *(bias)*	……
measurement A	**	**	**	……..	1	……..
measurement B	**	**	**	……	0	……

The test is based on the evaluation of the angle between pairs of vectors; this can be done as follow:

1. Normalization: divide each element in each column of the System Matrix by the square root of the sum of the squares of the elements in that column.
2. Dot product: multiply the transpose of this matrix by the matrix. $\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + ... = |a||b|\cos\vartheta$, the dot product of every pair of normalized vectors produces the cosine of the angle between them.
3. Check correlation: the resulting symmetric matrix of cosines can be inspected for possible correlation between component changes and sensor biasses.

Experience indicates that cosines with magnitude greater than 0.7 are worthy of note (Provost, 1995).

- 0.7-0.8 some correlation
- 0.8-0.9 high degree of correlation
- 0.9-1 very high degree of correlation. The component changes and/or the sensor biasses are effectively indistinguishable.

2.2.5 Observability analysis

A system is said to be completely observable if every transition of the state Δx – see equation (2.3), can be determined from the observations Δz. The two tests described before are limited to the study of the correlations between either pairs of component changes and/or sensor biasses – see first test, or pairs of measurements –see second test. A more effective observability study should be aimed at determining all the particular combinations of component changes and/or sensor biasses that are not detectable with a given set of measurements. Several methods to test observability are available. The most common method is the invertability of the System Matrix but this produces a 'yes/not' type of answer that is not useful for our purpose. A more useful approach was proposed by Provost (1995), which is based on a thorough study of the eigenvalues and eigenvectors of the System Matrix. Section 4.4.5 provides a description of this method together with an application to the Trent 800 engine. Besides, a development of this theory is proposed: a parameter is introduced that quantifies the observability of the System Matrix. This can be used in order to provide a quick comparison among a vast number of different sets of measurements.

2.2.6 Summary

A preliminary observability analysis should always be undertaken before any diagnostics attempt. The development of a quantitative approach to gas turbine engine faults observability is necessary to (i) chose the most suitable measurements set (for diagnostics purposes), and (ii) to provide significant information to interpret the result of the diagnostics process. In this section two observability tests have been reviewed and a method based on eigenvectors and eigenvalues has been mentioned.

2.3 Gas-path diagnostics: a comparative review of procedures

2.3.1 Introduction

Gas turbine performance diagnostics is aimed at detecting, isolating and assessing deviations in components' performance, faults in engine systems, and instrumentation errors from knowledge of shifts in a set of measurements (often limited in number) taken along the gas-path. However, it has been recognised above, that measurement deviations may be due to the following reasons:

- Component performance parameter changes.
- Random errors and biasses in power setting parameters.
- Random errors and biasses in the measurements.

Besides, the limited number of measurements available introduces a lack of observability in the problem.

This section is aimed at identifying the state-of-the-art in gas-path diagnostics methodologies through a survey of the techniques developed since the pioneer Louis A. Urban (1969) first explored these problems in 1967.

A remarkable research effort has been devoted towards the development of engine gas-path diagnostics (GPD) processes. Different techniques emerged with their advantages and limitations. An extensive review of methods was provided by Li (2002b). A most recent update of gas-path diagnostics methodologies is contained in the Von Karman Institute lecture series 2003-01 on gas turbine condition monitoring and fault diagnosis edited by Mathioudakis and Sieverding (2003). Many pertinent tools have been devised during the last three decades and a critical review of the most used techniques and their applications now follows – see also Table 7, highlighting similarities, differences, advantages and limitations. The literature survey is structured according to a classification in ten key procedures:

- Linear with ICM inversion.
- Non-linear with ICM inversion.
- Linear Kalman filter.
- Linear WLS.
- Non-linear Kalman filter.
- Non-linear model-based with GA.
- Artificial Neural Networks.
- Bayesian Belief Networks.
- Expert systems.
- Fuzzy Logic.

A summary of the pros and cons, discussed in the next ten sections, is presented in Table 7 with the aim of guiding through the review. Specific comments to this Table are summarised in section 2.3.13.

Table 7: Summary of GPD methodologies, X means that the methodology involves that feature, (Marinai et al. , 2004)

METHODOLOGY / *STRATEGY INVOLVED*	LINEAR GPA WITH ICM INVERSION	NON-LINEAR GPA WITH ICM INVERSION	LINEAR KALMAN-FILTER	LINEAR WLS	NON-LINEAR KALMAN-FILTER	NON-LINEAR MODEL-BASED WITH GA	ARTIFICIAL NEURAL-NETWORKS	BAYESIAN-BELIEF NETWORKS	EXPERT SYSTEMS	FUZZY-LOGIC
LINEAR/ NON-LINEAR MODEL	Linear	Non-linear	Linear	Linear	Non-linear	Non-linear	Non-linear	Non-linear		Non-linear
SMALL CHANGES OF HEALTH PARAMETERS	X		X	X						
MEASUREMENTS RANDOM NOISE			X	X	X	X	X	X	X	X
BIAS			X	X	X	X	X		X	X
N PARAMETERS M MEASUREMENTS	M=>N	M=>N	M<N	M<N	M<N	M<N	M<N	M<N	M<N	M<N
SINGLE/MULTIPLE FAULT(S)	MFI	MFI	MFI	MFI	MFI	SFI/limited MFI	SFI/limited MFI	SFI/limited MFI	SFI/limited MFI	SFI/limited MFI
SMEARING VERSUS CONCENTRATION			Smearing	Smearing	Smearing	Concentration	Concentration	Concentration	Concentration	Concentration
DIFFICULTY AND DEPENDENCE ON TRAINING / TUNING			prior knowledge	prior knowledge	prior knowledge	Number of string assignment	Long training and data selection	Effort in gathering info for setting-up		Rules explosion
ARTIFICIAL-INTELLIGENCE BASED						X	X	X	X	X
COMPUTATIONAL BURDEN (SLOW IN RECALL MODE)						X		X		
MODEL-FREE							X	X	X	X
HIGH DIMENSIONALITY PROBLEMS						X	X	X	X	X
DATA-FUSION CAPABILITY							X	X	X	X
BLACK-BOX NOT OBSERVABLE							X			
GOOD ACCURACY IN PREDEFINED RANGES ONLY						X	X	X	X	X
EXPERT KNOWLEDGE CAPABILITY								X	X	X
ON-WING			X	X	X		X			X

2.3.2 Linear gas-path analysis with ICM inversion

Gas turbines dependent and independent parameters, as defined in the section 2.1.2.1, are characterised by a highly non-linear relationship. A first approach to the problem is based on the assumption that the changes in the (independent) health parameters are relatively small and the set of governing equations can be linearized by a Taylor series expansion around a given steady-state operating point. These linearized equations are expressed in matrix form. The set of M measurements changes Δz are expressed as a linear combination of N performance parameters changes Δx.

$$\Delta z = H \cdot \Delta x \qquad (2.4)$$

The matrix H has been referred to by several names, Exchange Rate Table and Influence Coefficient Matrix (ICM) being two of the most popular.
This formulation of the problem leads to a simple solution:

$$\Delta x = H^{-1} \cdot \Delta z \qquad (2.5)$$

The matrix H^{-1} is referred to as Fault Coefficient Matrix (FCM). This method is based on the assumption that the ICM is invertible and the measurements are free of noise. The inversion of the ICM requires that the number of performance parameters is less than or equal to the number of measurements ($N \leq M$). Otherwise estimation techniques, discussed later on, should be used.

A method, based on the assumptions above, was introduced by Urban (1969), and called Gas-Path Analysis (GPA). In this work the term gas-path diagnostics (GPD) is used in general, while Gas-Path Analysis is reserved for the approaches based or derived from Urban's work. Several authors have described applications of this approach (e.g. Passalacque, 1974 and Escher, 1995) and highlighted the following advantages:

- Faulty component isolation: the method isolates the degraded or faulty component responsible for the deterioration in performance.
- Multiple fault isolation (MFI) capability, the linearization allows superimposition and therefore the capacity of isolating more than one faulty component.
- Quantification of performance loss for each component.

and the following limitations:

- The method requires a large number of pertinent measurements for the analysis.
- It does not deal with sensor noise and bias.
- It relies on the assumption of linearity that is only acceptable in very small ranges of values of influential parameters about the operating condition.

The assumption of linearity becomes a severe limitation when deteriorations cause the engine to operate further away from the operating point for which the matrix was calculated. This consideration let to the development of non-linear GPA.

2.3.3 Non-linear gas-path analysis with ICM inversion

The linear approach discussed earlier has the severe limitation that in many circumstances the level of error introduced by the assumptions of the linear model can be of the same order of magnitude as the fault being sought. One way of improving the accuracies of the predictions is to try to solve the non-linear relationship between the considered health parameter and measurements using an iterative method as described by Escher (1995). This theoretical relationship between N independent parameters x and M dependent parameters z can be expressed in mathematical terms as in equation (2. 2). In the neighbourhood of x, each of the functions h_i can be expanded in a Taylor series. The linear equations can be obtained by neglecting the 2nd and Higher Order Terms (HOT). This linear approximation is then employed recursively and an exact solution is obtained by Escher (1995) using the Newton-Raphson technique. Essentially, via this approach, an ICM is generated taking into account a small deterioration in the engine-component's performance. The ICM is then inverted to get the FCM. The FCM is then multiplied with the vector of engine measurements (obtained from a deteriorated engine). This gives a vector of changes in the engine-component's performance parameters. From the results calculated, a new ICM is generated and this process is recursively repeated until the solution converges to a fixed limit. Figure 20 shows the benefit of using a non-linear approach over the linear one.

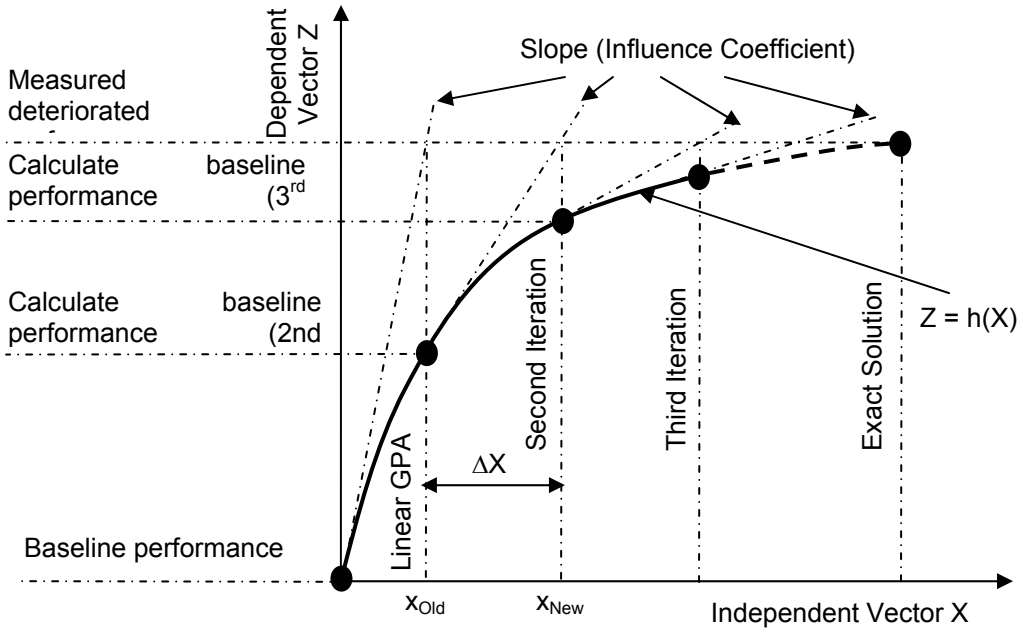


Figure 20: Simplified illustration of linear and non-linear GPA, (Escher, 1995)

The list of advantages and limitations given in the previous section still holds for the non-linear approach discussed, with the exception that this method theoretically achieves a more accurate solution because is based on a non-linear and therefore more realistic model of the problem. To overcome the limitations that face the two techniques, when only a small number of

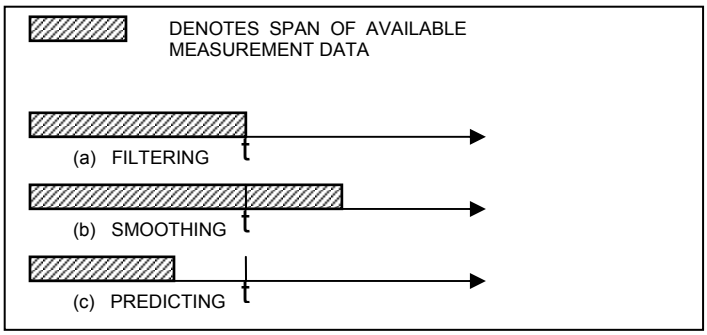
measurements is available (i.e. lack of observability) and in the presence of measurement uncertainties (e.g. noise and biasses), estimation techniques like the Kalman filter (KF) method, Weighted Least Squares (WLSs) approach, and their variants, have been used. These have been adopted by the major OEMs such as Rolls-Royce, Pratt & Whitney and General Electric.

2.3.4 Kalman filter based GPA: linear approach

The theory of Estimation Methods (Bryson et al, 1975,Gelb, 1974), allows to find the 'best' estimate of the solution of the diagnostics problem as defined before. An optimal estimator is a computational algorithm that processes measurements to deduce a minimum error estimate of the state of a system. Among the presumed advantages of this type of data processor, the most relevant are that it minimizes the estimation error in a well defined statistical sense and that it utilizes all measurement data plus prior knowledge about the system. The corresponding potential disadvantages are its sensitivity to erroneous a-priori models and statistics and the inherent computational burden.

Three types of estimation problems of interest are depicted in Figure 21. When the time at which an estimate is desired coincides with the last measurement point, the problem is referred to as *filtering*; when the time of interest falls within the span of available measurement data, the problem is termed *smoothing*; and when the time of interest occurs after the last available measurement, the problem is called *prediction* or *forecast* (Gelb, 1974). In this thesis, pertinent smoothing and forecasting methods are discussed in Chapter-5 as part of a gas-path prognostics framework. In this section, an application of the Kalman filter to GPD is revised.

Figure 21: Three types of estimation problems (estimation desired at time t)



A block diagram of the optimal filtering technique, developed by Kalman for estimating the state of a linear system, is shown in Figure 22.

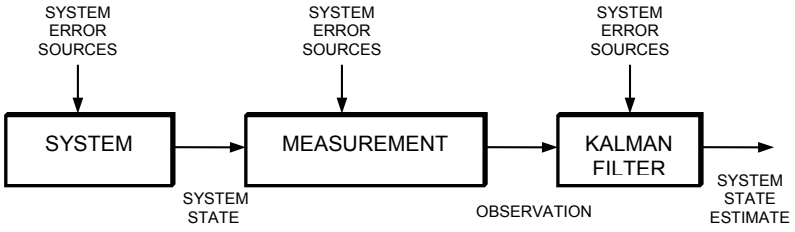


Figure 22: Block diagram typically depicting system, measurement and estimator.

The method is based on the following assumptions:

- Linear approximation. Though there are means of extending to the non-linear case.
- Noise is independent from one sampling time to the next.
- Noise is assumed to be Gaussian in terms of amplitude and it is assumed that at any given point of time, the probability density of Gaussian noise amplitude takes on the shape of a normal bell-shaped curve.

The problem is defined by the following sets of equations**:**

$$z_k = H_k x_k + v_k \qquad \text{Measurement equation} \qquad (2.6)$$

$$x_k = \Phi_{k-1} + w_{k-1} \quad k = 1,2 \qquad \text{System equation} \qquad (2.7)$$

where:

- $z_k \in R^M$: Measurement vector
- $x_k \in R^N$: State Vector
- $H_k \in R^{MxN}$: Model Matrix
- $v_k \in R^M$: Measurement noise, assumed to be Gaussian, white (i.e. uncorrelated), zero mean with covariance matrix R_k
- $\Phi_k \in R^{NxN}$: Transition matrix
- $w_k \in R^N$: Process noise, assumed to be Gaussian, white, zero mean, with covariance matrix Q_k

Matrices H_k and Φ_k, measurement and process noise statistics are assumed to be known. The following initial conditions are assumed:

$$E[x(0)] = \hat{x}_0 \qquad (2.8)$$

$$E[(x(0) - \hat{x}_0) \cdot (x(0) - \hat{x}_0)^T] = P_0 \qquad (2.9)$$

Where the operator E[.] is the mean value. Another assumption is that process and measurement noises are uncorrelated:

$$E[w_i \cdot v_j] = 0 \text{ for all } i \text{ and } j. \qquad (2.10)$$

The Kalman filter produces a recursive estimation *x* of the state vector at time *k* based on the current measurement vector *z* and the previous state vector estimation x_{k-i}. This feature can be exploited for real time applications provided the above hypothesis are all correct: the Kalman filter provides the minimum variance, unbiased and consistent estimate of the state vector, given a set of measurement vectors. The estimate is minimum variance as it minimizes the following quantity:

$$J_k = E[\tilde{x}_k^T(+) \cdot \tilde{x}_k(+)] \qquad (2.11)$$

Where,

- $\tilde{x}_k$ is the estimation error : $\tilde{x}_k = \hat{x}_k - x_k$.
- (+) means that the quantity has been updated with the measurement vector z_k .

- (-) means that the quantity has been evaluated just before the measurement.

An unbiased estimation is one whose expected value is the same as the quantity to be evaluated

$$E[\hat{x}_k] = x_k \qquad (2.12)$$

A consistent estimate is one which converges to the true value of *x* as the number of the measurement increases.

The following equations make up the Kalman filter (Gelb, 1974):

$\hat{x}_k(-) = \Phi_{k-1} \cdot \hat{x}_{k-1}(+)$ State estimate extrapolation (2.13)

$P_k(-) = \Phi_{k-1} P_{k-1}(+) \Phi_{k-1}^T + Q_{k-1}$ Error covariance extrapolation (2.14)

$\hat{x}_k(+) = \hat{x}_k(-) + K_k(z_k - H_k \cdot \hat{x}_k(-))$ State estimate update (2.15)

$P_k(+) = (I - K_k H_k) P_k(-)$ Error covariance update (2.16)

$K_k = P_k(-) H_k^T (H_k P_k(-) H_k^T + R_k)^{-1}$ Kalman gain matrix (2.17)

As shown by equation (2.11) the KF minimizes a quadratic cost function step by step i.e. after each measurement. It can be shown that if the system is linear as given by equations (2.6) and (2.7), then minimization over the time steps is the same as the minimization of the complete cost function evaluated over the whole set of measurements. It can be shown (Bryson et al., 1975) that for a linear system described by equations (2.6) and (2.7), and subject to the above assumptions the solution provided by the Kalman filter is optimal with respect to every common criterion (minimum variance, maximum likelihood , minimum error). Moreover the technique is recursive. Though when used with linear GPA, Kalman filters improve the diagnostics accuracy, the issues of non-linearity has still not been addressed. The analogy between KF and WLS is remarkable; the details of a WLS based approach are considered in the next section.

Pratt & Whitney, through Hamilton Standards, have been pioneering this field and implementing a Kalman filter based method. Several adaptations have been designed to cope with some of the filter's limitations (Urban et al, 1992) and they are integrated into the now available software (MAPIII, TEAMIII, EHM, ADEM). Similarly, the gas-path diagnostic tool currently used in Rolls-Royce is based on a modified version of the Kalman filter technique so called the 'Concentrator' applied to the linear GPA (Provost, 1987,1988, and 1995). The diagnostic system (COMPASS) developed for test-cell diagnostics of aero engines has then been used for on-board applications. As far as General Electric's approach is concerned, the WLS estimation has been applied (Doel, 1994) and implemented within a diagnostic tool (TEMPER). The previously mentioned applications are able to estimate both component performance changes and sensor biasses.

The KF and WLS based methods present the following advantages:

- Optimality: the cost functional is minimized.
- Recursively: limited commuting requirements.
- Prior knowledge: knowledge about the statistics of components deterioration can be introduced.

- Measurement noise: can be assumed to be white and Gaussian.
- Sensor errors: can be included in the state vector as bias.

Limitations in the application of KF techniques to the linear GPA are according to Zedda (1999a):

- Requirement for prior knowledge and tuning: the choice of the process noise covariance matrix is often arbitrary. There usually exists no statistically significant population of faulty engines to base the performance parameter's standard deviations assigned on it. Sensitivity studies can be helpful but the deviations of the state vector elements may be strong.
- The "smearing" effect: often only a limited number of components and sensor are fault affected, while the KF tends to smear the faults over many engine's components and sensors.
- System model and divergence: the Kalman filter produces an optimal solution provided the hypotheses about the system are correct. In the case of gas turbine diagnostics, even though we might assume equation (2.6) to be sufficiently precise, almost nothing is known about equation (2.7), which describes the temporal evolution of the fault. As the method should be able to detect deterioration due to various kinds of fault, both slowly varying (erosion, corrosion, fouling) and abruptly varying (FOD), equation (2.7) is not available. Therefore, it should be somehow estimated and this can impair the final diagnostics accuracy. In fact the use of techniques to completely estimate equation (2.7) introduces errors and as measurements are collected and used by the algorithm the system learns the wrong state too well. The consequence is divergence, i.e. the estimated solution becomes and more distant from the actual solution.
- Non-linearity and optimality: the errors due to the assumed approximations of non-linear systems with a linear one may not be negligible even if no estimation technique is employed (Escher, 1995). Therefore a non-linear estimation technique seems more suitable. The application of a non-linear version of Kalman filter though is no easy task. Many problems are associated with the use of the common Extended Kalman Filter (EKF) and Iterated Extended Kalman Filter(IEKF), as pointed out by Jazwinski (1970). The main drawbacks are that the estimates are often biased and suboptimal (i.e. the cost function is not minimised).

2.3.5 Weighted least squares based GPA: linear approach

The method here described is based on a linear formulation of the problem as previously discussed. To evaluate the component's performance correctly, sensor error must be considered in the analysis. WLS facilitates the determination of engine state in the presence of sensor error. WLS is based on the linearized model of the measurement process given by:

$$z = H \cdot x + v + b \qquad (2.18)$$

where:

$z \in R^M$ is a vector of M measurements;
$x \in R^N$ is a vector of N performance parameters;
H is the ICM discussed before;
$\nu \in R^M$ is the zero-mean measurement noise vector;
$b \in R^M$ is the sensor bias.
The method relies on the minimization of the error between the observed measurement deviations and the measurement deviations that would be generated by the deviations in x. This can be done by minimizing the scalar quantity:

$$J = 1/2\left\{x^T P_0^{-1} x + (z - Hx)^T R^{-1} (z - Hx)\right\} \quad (2.19)$$

where P_0 and R were defined in the previous section. Solving the minimization of *J* with respect to *x* will give the state vector estimate with the highest conditional probability. The optimal solution x_o is:

$$x_o = \left(P_0^{-1} H^T R^{-1} H\right)^{-1} H^T R^{-1} z \quad (2.20)$$

The true measurements z_o for the engine state x_o is:

$$z_o = H.x_o \quad (2.21)$$

Hence, the weighted least-squares algorithm provides best results when measurement deviations are small. In order to overtake this limitation GE has included the 'fault logic' capability in their program TEMPER. The fault-logic is used to search for large deviations in component performance, or for large measurement errors, when the solution residual is large. The solution residual J_o provides the mechanism for recognising that a specific case is far from nominal conditions. TEMPER assumes that J_o follows the Chi-squared distribution and the fault logic in invoked whenever the residual exceeds the 95 percent limit. A summary of the characteristic of the diagnostic system follows, according to Doel (1994) :

- Weighted least square algorithm provide better result for small deviations, and in order to address this problem a feature called "fault logic" is incorporated to determine large deviations.
- The technique is affected by 'smearing' (underestimation of the actual fault and attribution of the remainder fault to other engine components and to measurement error).
- The algorithm is linear for the measurement vector, whereas the gas turbine performance is highly non-linear.
- A problem is the input requirements for the algorithm. Comprehensive data analysis and careful judgment is required for the statistical and baseline inputs as accurate baselines are critical for augmentation strategies for implementation of the fault logic.
- TEMPER does not provide any aid in interpreting the results and there have been a number of cases where it has identified the same component repeatedly even though it has been overhauled.

The reason why KF and WLS based estimation technique have been applied to linear GPA as well as their common limitations have been discussed in the previous section.

2.3.6 Kalman filter based GPA: non-linear approach

This type of diagnostic methods are based on accurate modelling of non-linear steady state gas turbine performance. At steady state conditions, the dependent and independent parameters of gas turbines can be expressed with a non-linear relationship:

$$z = h(x,w) + v \qquad (2.\ 22)$$

with the same meaning of the symbols. The idea of the non-linear model-based methods is shown in Figure 2.4. The real engine component parameter vector x determines engine performance represented by the measurement vector z. With an initial guessed parameter vector $\hat{x}$, the engine model provides a predicted performance measurement vector $\hat{z}$.

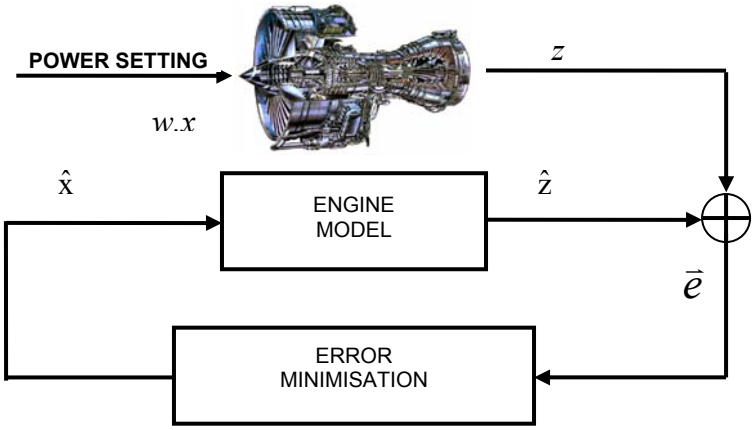


Figure 23: Non-Linear Engine Fault Diagnostic Model (Li, 2002)

In the last section the limitations of the linear KF approach have been highlighted. Additional comments follow (Zedda et al. 1999c):

- If an actually linear system has to be processed, then the most sensible choice is a common Kalman filter, as it is optimal with respect to any reasonable minimization criterion (especially minimum variance and maximum likelihood). Moreover, the solution is achieved through recursive technique, which can be very useful due to the limited computing power required.
- If the system is actually non-linear, a linearization can be done and the common-linear KF can be applied in a straightforward way, in this case, though the difference between the actual and the simulated behaviour of the system may be large and may lead to divergence, i.e. ever increasing distance between the state and the estimate. In practices, the onset of divergence manifests itself by inconsistency of the residuals with their predicted statistics. Residuals become biased and larger as more measurements are collected and processed. A similar behaviour of the estimator is observed when a measurement becomes biased (Healy et al. 1997). If divergence effects are added to the “smearing” effect, typical of KF, the estimation accuracy may become unacceptable. As far as gas turbine diagnostics is concerned, the linearization of a process characterised by such large non-linearity as a gas turbine engine is probably responsible for inaccuracies of the estimation, especially when

time varying multiple faults are present. In some instances, divergence does occur (Urban et al., 1992).

- If the effect of non-linearity of the system of the estimation accuracy is ascertained, a non-linear version of the KF can be used to try to model more accurately the engine behaviour. The most commonly used filtering techniques are the Extended Kalman Filter (EKF) and the Iterated Extended Kalman Filter (IEKF).

The non-linear formulation is here presented and compared with the non-linear one; it can be shown (Bryson and Ho, 1975) that the linear KF minimizes the cost function given by:

$$J = \frac{1}{2}(x(0) - \hat{x}_0)^T \cdot P_0^{-1} \cdot (x(0) - x_0) + \frac{1}{2}\sum_{i=1}^{k-1} w_i^T Q_i w_i + \frac{1}{2}\sum_{i=1}^{k-1}(z_i - H_i x_i)^T \cdot R_i^{-1} \cdot (z_i - H_i x_i) \quad (2.23)$$

For linear problems, the minimum variance and the maximum variance and the maximum likelihood coincide and therefore the KF can be considered as the best choice, provided the modelling of the process is sufficiently accurate. If the process is non-linear then the cost function to be minimized can be rewritten as:

$$J = \frac{1}{2}(x(0) - \hat{x}_0)^T \cdot P_0^{-1} \cdot (x(0) - x_0) + \frac{1}{2}\sum_{i=1}^{k-1} w_i^T Q_i w_i + \frac{1}{2}\sum_{i=1}^{k-1}(z_i - h_i(x_i)^T \cdot R_i^{-1} \cdot (z_i - h_i(x_i)) \quad (2.24)$$

A solution minimising the cost function (2.24) ensures maximum likelihood and approaches minimum variance asymptotically as the number of measurements increases. Therefore, the aim of a non-linear filter would be minimize a cost function, possibly in a recursive way. It can be shown that both the EKF and the IEKF approaches produce biased and sub optimal estimates due to the linearization of the cost functions. From a practical point of view, this leads to a low accuracy estimation. The most non-linear squares estimation algorithms require a choice between an optimal solution and a recursive formulation. If recursivity is a paramount requirement, then optimality is compromised. As third possible solution has been discussed by Zedda (1999a): the proposed estimation-technique splits the problem of cost function minimization into a linear first step and a non-linear second step by defining new first step states that are non-linear combinations of the unknown states.

2.3.7 Non-linear GPD with genetic algorithms

A genetic algorithm (GA) based diagnostics method was developed at Cranfield (Zedda, 1999a). It is a model-based approach, which is theoretically similar to those of non-linear model-based methods described earlier. The GA is applied as an effective optimization tool to obtain a set of engine component parameters that produce a set of predicted engine-dependent component parameters through a non-linear gas turbine model that best match the measurements. The solution is obtained when an objective function (or cost function), which is a measure of the difference between the predicted and measured engine-dependent parameters, achieves its minimum value, a schematic diagram of diagnostics strategy is shown in Figure 24.

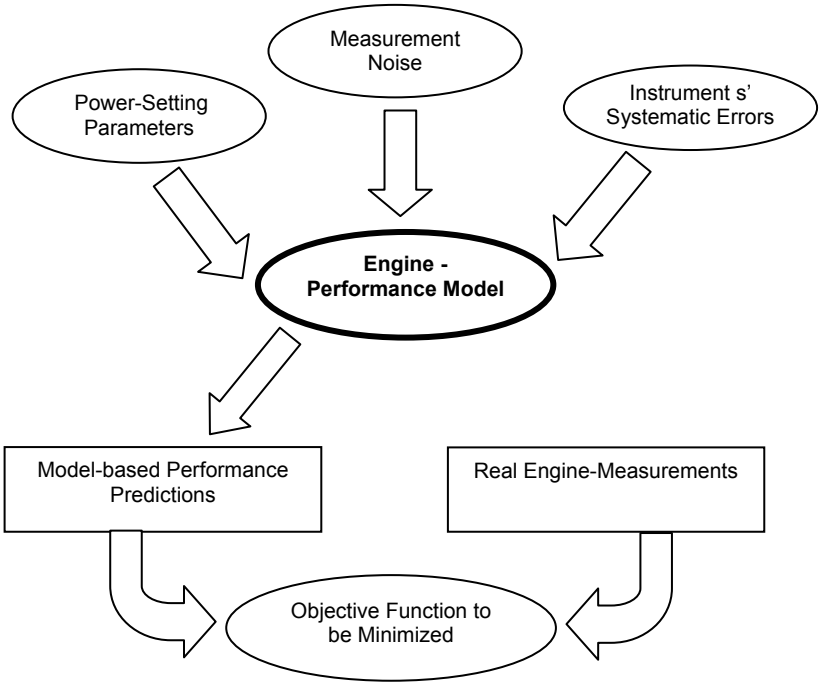


Figure 24: The diagnostics strategy, (Zedda, 1999)

The methodology is based on non-linear modelling and optimal-estimation theory. The procedure estimates the performance parameters expressing the fault condition of the engine components in the presence of measurement noise and biasses. The measurement uncertainty is supposed to affect even the power setting parameters. Estimation is performed through an 'ad hoc' Genetic Algorithm (GA). The GA uses an accurate non-linear steady-state model of the engine's behaviour. The only statistical assumption required by the technique concerns the measurement noise and the maximum allowed number of faulty sensors and engine components, which is enclosed as constraint.

2.3.7.1 The objective function

In this method the performance parameters are estimated by minimising an objective function by means of GAs. The basic requirements for the objective function are the following (Zedda et al., 1999b):

- It should be a measure of the consistency between actual and predicted measurements.
- Measurement noise should be accounted for.
- Measurement biasses should be accounted for.
- Its minimisation should reduce 'smearing' effect.
- Evaluation of the function should not be too burdensome from a computational point of view.

One of the objective functions used by Zedda (1999a) is the following:

$$J(x,w) = \sum_{j=1}^{M} \frac{\left| z_j - h_j(x,w) \right|}{z_{odj}(w)\,\sigma_j} \qquad (2.25)$$

where:

$z \in R^M$ is a vector of M measurements (dependent parameters);
$x \in R^N$ is a vector of N performance parameters (independent parameters);
$w \in R^P$ is a vector of P environmental and power setting parameters.
$H(\cdot)$ is a vector-valued function, typically non-linear, it represents the gas turbine mathematical model;
z_{odj} value of the j-th measurement in the off design undeteriorated condition;
σ_j standard deviation of the j-th element of the measurement noise vector v assumed to be Gaussian.

A modified version of the function (2.25) has been defined to deal with measurement biasses as well:

$$J_{kl}(x,w) = \sum_{\substack{j=1 \\ j \neq k,l}}^{M} \frac{\left| z_j - h_j(x,w) \right|}{z_{odj}(w)\sigma_j} \qquad (2.26)$$

If no biasses affected the measurements, then the minimisation would estimate (x,w) so that the equations used in terms of the summation would be mutually consistent. The inconsistency due to biased measurement would manifest itself with larger values of the objective function, since no (x,w) can be found to correspond to predicted measurements fitting sufficiently well the real ones. The problem can be overcome by eliminating in the summation of function of the M_{bias} terms corresponding to the biased measurements. The terms k and l are biased and therefore neglected for the calculation of the objective function. Typical procedure proceeds step by step as described in Zedda (1999a).

The technique largely relies on the concept of relative redundancy, which is usually available in development engine test bed. When the instrumentation is more limited, like in pass off test, the problem can be overcome by application of Multiple Operating Point Analysis (Gulati et al. 2000a, Gulati et al. 2001).

2.3.7.2 A typical search-space

Several thousand combinations of faults could be generated depending on a number of influential factors, such as number of engine components, number of performance parameters associated with each component, size of steps chosen between lower and upper limits of deterioration. Predictions can be obtained by simulating the engine performance for each fault conditions, even allowing for measurement noise. An objective function search-space generated by varying the LPC flow capacity from –3.5% to +3.5% and the LPC efficiency from 0 to 3.5% and comparing it with data generated by introducing 2% efficiency deterioration in LPC is shown in Figure 25. Each point on the surface plot is a potential solution and the best solution is the point having the lowest objective function (Zedda, 1999a).

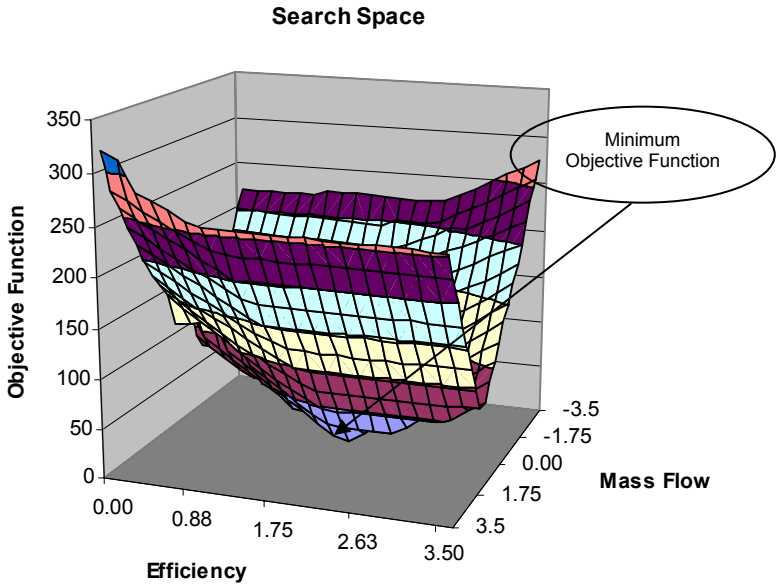


Figure 25: A typical search-space (Zedda, 1999a)

2.3.7.3 Optimization technique: genetic algorithms

This method, as said above, requires an algorithm capable of searching through the vast search-space effectively and in a reasonably short period. In order to arrive at the global minimum, the algorithm has to overcome several local minima. Several optimizing techniques like calculus based techniques (requiring the calculation of gradient) and other methods like the downhill simplex method and the Powell's direction set method could be used and has been considered by Zedda (1999a). However, the optimization technique based on the genetic algorithm (GA) was found to be the most suitable for this kind of problem. GA has emerged as a powerful optimization tool recently and finds a wide range of applications in different fields.

The metaphor underlying the genetic algorithm is that of natural evolution. In evolution, the problem that each species faces is the one of searching for beneficial adaptations to a complicated and changing environment. The GAs follow step-by-step the procedure followed by the natural principle of survival of the fittest. The nomenclature essential with GAs is also borrowed from the vocabulary of natural genetics (Michalewicz, 1996).

In the context of this technique, a *string* refers to a possible solution and a collection of possible solutions or *strings* is called a *population*. The *fitness* of the *string* is a function of the objective function and is inversely proportional to it. The best *string* would therefore have the highest *fitness,* which means that the value of objective function would be minimized (Goldberg, 1989).

Typically, the diagnostics algorithm based on a GA starts off with a *population* that is randomly created and the objective function value for each of these *strings* from the population is calculated. The objective function to be minimised

is mapped onto a fitness function and the larger the *fitness* the higher the probability of *survival*. The mapping of the objective value onto the fitness function could be linear or non-linear. The aim of the algorithm is to reach the global minimum by successfully avoiding the local minima. The GA then works over a number of iterations or *generations* and each of these *generations* is based on three fundamental operators in the GA: s*election, crossover and mutation*. The *selection* operator chooses the *strings* to be used in the next *generation* according to a "survival of the fittest" criterion. The *crossover* operator allows information exchange between *strings*, in the form of swapping of parts of the parameter vector in an attempt to generate fitter *strings*. *Mutation* is used to introduce new or prematurely-lost information in the form of random perturbations to the values without exceeding the predecided upper and lower thresholds (Goldberg, 1989). Figure 26 shows a schematic diagram of a generic generation.

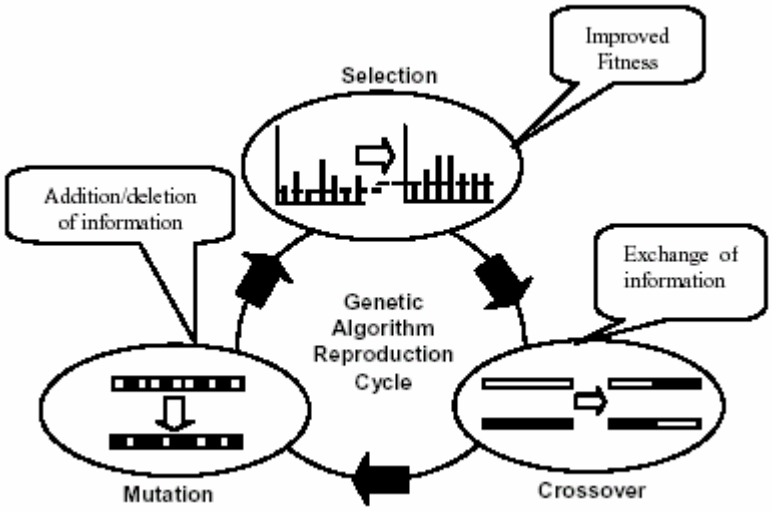


Figure 26: Single generation of GA reproduction cycle.

Compared with typical calculus-based optimization methods, GA has several distinctive features (Zedda et al. 1999b):

- The GA based diagnostic method does not require calculation of complex functions or complex derivatives and therefore any non-smooth function can be optimized.
- It uses probabilistic rules rather than deterministic rule to create strings for the subsequent generations. This makes it possible for the algorithm to proceed in different paths every time it runs.
- Constraints can be dealt with in a very different way, such as by means of penalty functions or design of specific operations.
- The method uses a global search procedure and therefore avoids getting stuck in local minimum.

The diagnostics method presents the following advantages (Zedda, 1999a):

- The minimization of the above discussed objective function provides a result that is inherently more robust to measurement bias than the other

methodologies considered before. In fact it has the ability to concentrate on a limited number of fault-affected performance parameters using bias-free measurements. This reduces the degree of undetermination of the estimation problem.

- The method relies on a fully non-linear model. Even cases of relatively large deterioration of up to four performance parameters can be estimated.
- The method performs concentration on the faulty engine components, whereas smearing effect affects most KF-based approaches.
- The method is robust with respect to small modelling and measurements errors, while KF-based approaches intrinsically non-robust to discrepancies from the assumed model.
- Minimal statistical information is demanded.

and suffers from the following limitations (Zedda, 1999a):

- The methodology is more computationally burdensome than classic estimation techniques.
- Although multiple faults can be detected, the application described by Zedda (1999a) is limited to four parameters experiencing simultaneous deterioration.
- Care must be taken when using the GA in assigning the number of strings. Even though the rule for the assignment of the number of strings for different fault classes can easily be established by trial-and-error and the achieved accuracy is not a strong function of the rule of assignment itself, active awareness of these issues is necessary for the correct utilization of the technique. This makes the method difficult to use, and requires a trained person for its worthwhile operation.

Some of these limitations have been overtaken by further developments of the technique. In fact, the theory previously introduced was implemented for development engines where many measurements are available. It was applied to a three-spool military turbofan engine RB199 (Zedda et al., 1999b) and a two-spool low by-pass military turbofan engine EJ200 (Zedda et al., 1999c): it reached a high level of accuracy. Gulati et al. (2000 and 2001) combined the concepts proposed by Stamatis et al. (1991) concerning multiple point diagnostics together with a Genetic Algorithm approach to produce a GA based Multiple Operating Point Analysis (MOPA) method for gas turbine fault diagnostics. This method is suitable for problems where only limited instrumentation is available. It was applied to a RB199 engine and showed interesting results. Further developments of the method led to the study discussed by Sampath (2003a), in which a hybrid approach wasadopted for the intercooled recuperated WR21 engine, thereby gaining by reducing computational-burden. The application of evolution strategy was also investigated by Sampath et al., (2003c). Sampath (2003a) described a diagnostics model that operates in three distinct stages. The first stage uses response surfaces for computing objective functions to increase the exploration potential of the search-space while reducing the computational load. The second stage uses the heuristics modification of genetic algorithm parameters through a master-slave type configuration. The third stage uses the elitist model

concept genetic algorithm to preserve the accuracy of the solution in the face of randomness. This fault-diagnostics model was integrated with a nested neural-network to form a hybrid diagnostics model. The nested neural-network is employed as a pre-processor or filter to reduce the number of fault classes to be explored by the GA-based diagnostics model. The hybrid model improves the accuracy, reliability and consistency of the results obtained.

2.3.8 Artificial neural-network based GPD

Artificial Neural-Networks (ANNs) have been extensively investigated for use in fault diagnoses. ANNs consist of parallel distributed processors able to store knowledge as experience and make it available for use. They are trained to map inputs to outputs via non-linear relationships in a framework that mimics the learning process performed by the human brain. Generally, a Neural-Network (NN) operates in two phases - a learning phase and an operating phase. The purpose of the learning phase is to determine the NN parameters, which enable the network to function properly in the operating phase. A typical NN shows the following features (Sampath, 2003a):

- The network is made up of units called neurons, each performing a weighted sum of its own inputs. The sum is then passed through a function to the output.
- The knowledge is stored in inter neuron connections called weights. This knowledge for the weights is acquired through training (also called learning phase). The purpose of the learning phase is to determine the NN parameters, which will enable the network to function properly in operating phase. The NN can be classified according to the kind of training. If the training algorithm uses different input and output patterns the learning is called *supervised*, if the training algorithm has to extract information just from the input patterns the learning is named *unsupervised*. If input and output patterns are the same, the training is *self supervised* and the net performs auto association.

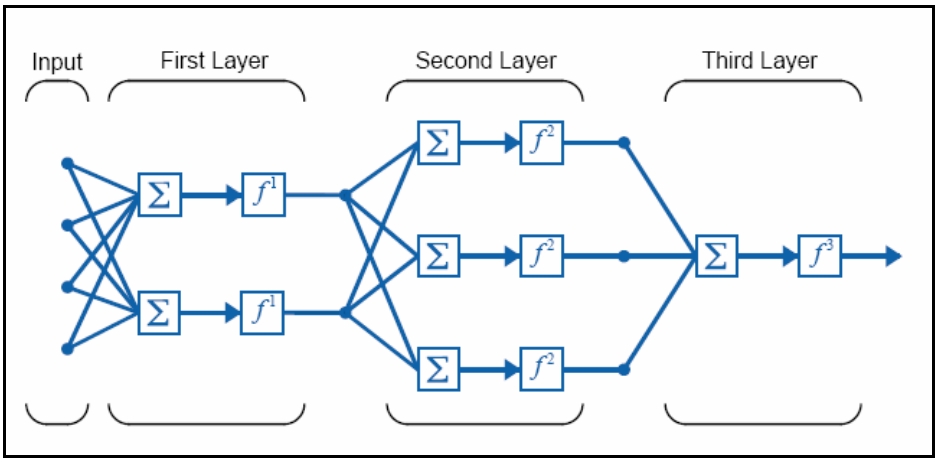


Figure 27: A typical Multilayer Perceptron

The Multi-Layer Perceptron (MLP) with back-propagation training is the most common architecture used for GPD purposes. It is also called the Feed-Forward

Back-Propagation Neural-Network (FFBPNN).The configuration of a typical MLP is shown in Figure 27.

The use of ANNs in GPD experiences the following advantages:

- ANNs are trained by means of information extracted from data. This makes them particularly suited for finding solutions to problems for which there are no exact solution, but only a large number of examples. In general ANN proved to be advantageous in problems that are difficult to describe mathematically.
- Solutions can be provided for cases, which have never been encountered in the training phase, i.e. they generalize from examples.
- ANNs are able to deal with the non-linear nature of the GPD problem.
- ANNs can deal with the level of noise that typically affects the gas-path measurements.
- ANNs can be used for on-line GPD applications due to the fast computational speed in recall mode. Once the net structure/architecture is chosen, trained and verified, the performance simulation code is not called anymore during the NN calculations, and this reduces the computational burden in recall mode.

and limitations:

- Like other AI tools, ANNs are unable to perform creditably outside the range of data to which they have been exposed during the training: this implies that a massive amount of data from encountered and foreseeable faults/conditions of operation would be required in each ANN development.
- Training times are long though each depends on the network type, size and the amount of training data. ANNs require retraining when machine operating conditions change. This could mean, for instance, retraining after a machine overhaul.
- Its deficiency in providing descriptive results: there is no way of accessing the neural network's "reasoning": it is only possible to inspect the predictions it makes.
- It is sometimes difficult to provide the confidence level associated with the output result.
- Another major drawback of neural networks as reported by Zedda et al. (1998) is the requirement of large amount of training data and the long times required to train the networks. Trained networks with frozen network weights would require retraining the networks if the engine undergoes overhaul.
- As the number of engine operating points, that need to be diagnosed, increases, the diagnostics error is bound to increase unless an alternative means of data correction is devised (Ogaji, 2003).

A wide range of different neural networks has been implemented to perform GPD. An application of Feed-Forward Back-Propagation neural networks to gas turbine diagnostics is discussed by Torella et al. (1995).

Zedda et al. (1998) introduced a modular neural-network system (Figure 28) to tackle large-scale diagnostic problems and applied it to predict the behaviour of a Garrett TFE 1042 engine. The module classifies the input measurement

vectors according to the fault category (A or B). This kind of classification overcomes the problem of low diagnostics accuracy usually obtained when the two categories are mixed. Moreover, since the characteristics of the fault categories are different, separate diagnostics can be developed. Such a system has the drawback that requires a large number of nets and long training time.

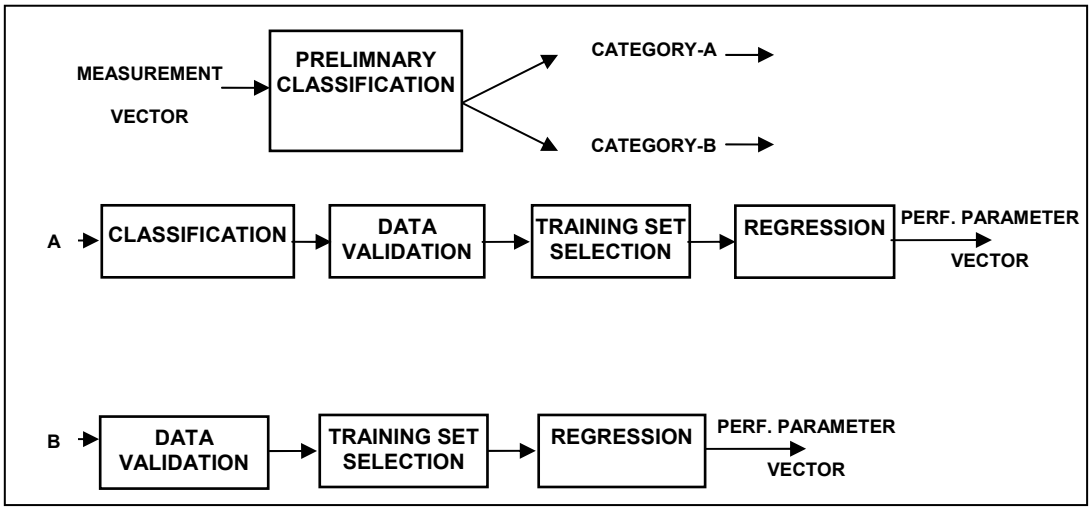


Figure 28: Modular ANN system for fault diagnostics (Zedda & Singh, 1998)

A multiple neural network architecture that performs sensor and component fault diagnosis is described by Kanelopoulos et al. (1997). The authors suggested the use of different networks to isolate the sensor and component faults and showed that this layout provides better results than a single network. This concept was further developed by Ogaji (2003) that generated a cascaded network to isolate component and sensor faults. Volponi et al. (2000) introduced a hybrid neural-network where part of the model was replaced by influence coefficients: they reported that the accuracy of such a network was favourable compared with a back-propagation net and Kalman Filter approach. Green (1997) discussed the need to incorporate an ANN with other AI techniques to perform the estimations of active life, diagnostics and prognostics capabilities for engines.

Eustace (1993) applied a Probabilistic Neural-Network (PNN) to diagnose faults in any engine within a fleet of 130 GE low by-pass F404 military engines. The authors used a statistical correlation technique to select five out of eight available engine monitoring parameters as inputs to the network. This approach is interesting, considering the fact that, even for healthy engines, the values of the parameters vary from engine to engine. Results from the network show an accuracy of 87.6% which is acceptable considering the variability in the baseline.
In many situations, the information on which the diagnosis has to be based on is not sufficient to obtain a unique solution for unknown health parameters. Probabilistic arguments are then very useful for deriving solutions that are close to the existing experience (Mathioudakis et al., 2003). PNNs are used to provide solution to pattern classification problems through Bayesian classifiers. PNN uses a supervised training set to develop distribution functions within a pattern

(middle) layer. In the recall mode, the developed functions are used to determine the likelihood of a given pattern being a member of a class or category with the criteria solely based on the closeness of the input feature vector to the distribution function of a class.

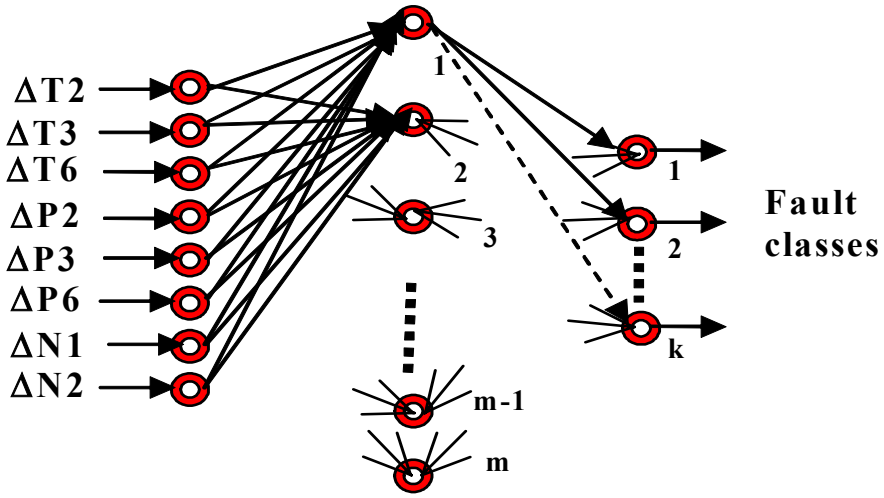


Figure 29: Structure of a PNN

PNNs have three layers. The input layer has as many elements as there are separable parameters needed to describe the objects to be classified. The middle layer organizes the training set such that each input vector is represented by an individual processing element. And finally, the output layer, also called the summation layer, has as many processing elements as there are classes to be recognised. Figure 29 shows a typical PNN with *m* hidden layer neurons and *k* output classes. PNNs are simple on design and with sufficient data are guaranteed to generalize well in classification tasks. Training of the PNN is much simpler than with back-propagation. However, the pattern layer can be quite huge if the distinction between categories is varied and at the same time quite similar in special areas. In addition, PNNs are slower to operate in the recall mode as more computations are done each time they are called.

Furthermore, a desirable property of a diagnostics method is the ability to detect measurement biasses due to improper calibration or faulty sensors. An incorrect diagnosis due to engine-sensor bias may either detect a non-existent engine fault or ignore a real fault. Hence, sensor failure detection, identification and accommodation (SFDIA) became a critical issue especially when measurements from failed sensors are used in the feedback. The use of NN technology can provide an alternative approach to the SFDIA problem (Zedda, 1999a). One variant of the ANN is the Auto Associative Neural-Network (AANN), see Figure 30. This has attracted much attention recently. It has the advantage of fast computing and can perform SFDIA simultaneously. The major distinctive feature of the AANN is that it contains a bottleneck layer (Figure 30). The number of bottlenecks should be less than the number of input nodes, but larger than the degree of freedom for which the system is being diagnosed. The input and output layers have identical parameters.

Theoretically, the ANN maps the input parameter to provide identical outputs or a corrected output. All the noise filtering is undertaken in the mapping section, namely from input to the bottleneck layer. Once the intrinsic states are retrieved from the mapping process, decoding the bottleneck nodes back to the output nodes is carried out by the de-mapping section (Ogaji et al., 2002).

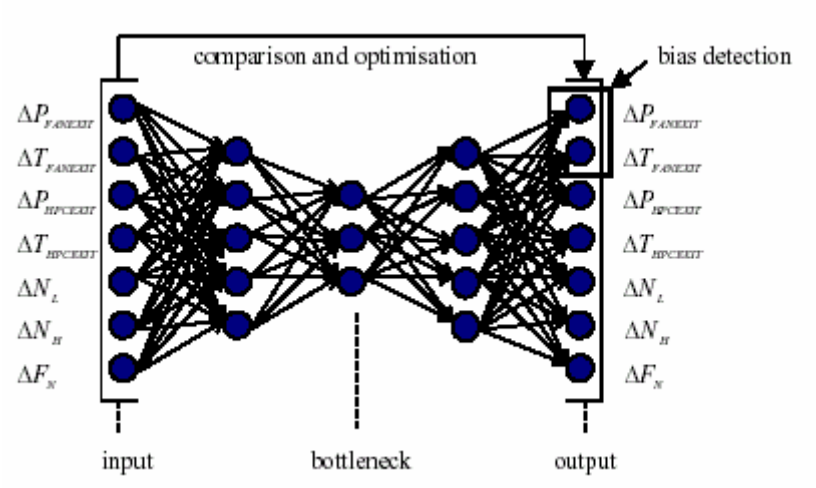


Figure 30: Auto Associative Network

2.3.9 Bayesian belief networks based GPD

Recently, Bayesian belief networks (BBNs), have been introduced as a new type of probabilistic expert system part of the artificial intelligence techniques. BBNs based GPD systems are reviewed in this section.

By means of the Bayes' theorem, the diagnostics problem can be expressed through the following conditional probability density function or posterior density function definition:

$$p(x/z) = \frac{p(x)p(z/x)}{p(z)} \qquad (2.27)$$

It can be noted that the knowledge of the maximum $p(x/z)$ provides the solution of the estimation problem. This is due to the meaning of $p(x/z)$: it is the PDF of x (set of performance parameters) given a realisation of the measurement vector z. Differently, from other heuristic methods, a statistical approach has the distinct advantage of yielding probability density functions instead of simply point-values. Therefore, it can provide also the confidence levels of its results (Law et al., 1991), which makes it especially suitable for engineering applications (i.e. in scheduling maintenance activities).
A belief network uses a graphical representation of probabilities by representing the cause and effect relationships among the elements of the network. It is a form of graph, consisting of nodes representing variables and arcs for the probabilistic dependencies between these variables. Each node contains a conditional probability distribution that describes the relationship between the node and the parents of that node. In fact, the nodes can represent anything, an

observation, a fault or some intermediate value. The list of possible states for each node must be mutually exclusive and collectively exhaustive.

Palmer (1998)) presents a BBN based diagnostic system developed for the CF6 family, nevertheless no details were provided about the topology of the BBN. Romessis et al., (2001) and Mathioudakis et al. (2003) provided a thorough description of an application of a BBN to turbofan engine diagnostics. A development of that method was carried out by Kadamb (2003) that set-up a GPD system, for a three spool engine, based on BBNs. The system was able to diagnose relatively large deteriorations in performance parameters relying on a fully non-linear model. The performance parameters (efficiency and flow capacity of turbines and compressors of a three shaft engine) were designated as the parent nodes and the gas-path measurements as the child nodes. The relationships between these were defined through the links and each link had a probability associated with it. The links between a parent node and a child node were established only if that particular child node (weremeasurement) was affected by the fault. A typical layout of a BBN is shown in Figure 31.

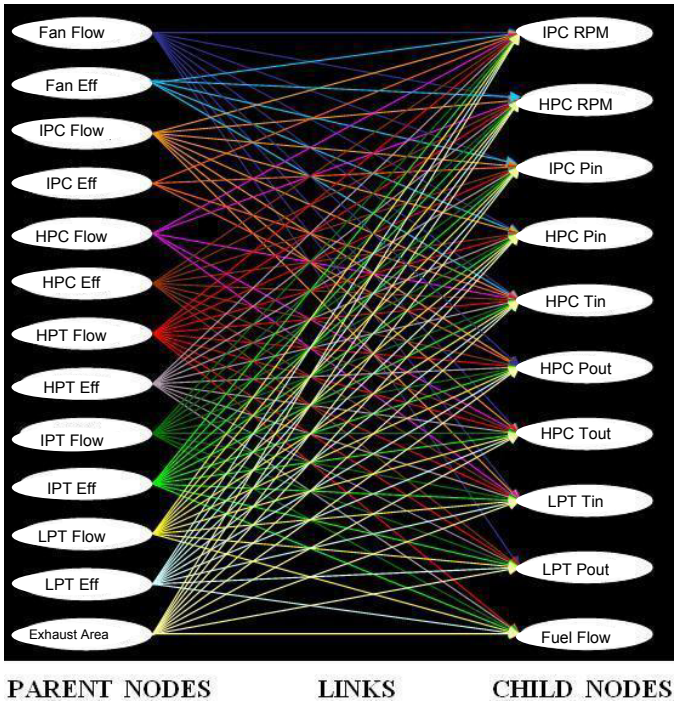


Figure 31: Typical BBN layout (Kadamb, 2003)

The use of BBNs in gas-path diagnostics shows the following advantages:

- It allows the introduction of many types of data. The data can be qualitative, continuous numbers or discrete numbers.
- It supports multiple simultaneous faults.
- Engine model hardware changes can be easily entered into the network. For techniques such as neural nets it takes a long time to retrain the system.
- Generic faults can be included in the system to catch problems areas not covered by any of the more specific faults whilst, systems such as neural nets would need enough information and test cases for training.

- Although a belief network can be set-up by using an engine simulation model, once it is set-up it does not require to run the engine model gaining in computational speed.

and imitations:

- Although a system that can isolate more than 4 deteriorated parameters (2 degraded components) is theoretically feasible, the price is a considerable increase in computational burden (Kadamb, 2003).
- Sensor biasses have not been dealt with in the BBN-based studies reviewed here.
- Substantial time and effort is required to gather the information needed for setting it up.
- Belief network based tools require a trained user to set it up.

2.3.10 Expert systems

The development of Expert Systems (ESs) has been one of the key break through of artificial intelligence (AI). The typical form of an ES encompasses expert knowledge in a pattern matching process, so as to be able to provide an interpretation of a new situation (i.e. the input) via an inference engine. This is often achieved by means of rules. Most ES are developed on shells which could be broad based such as CLIPS (C language integrated production systems) or tailored to meet specific objectives such as General Electric's GEN-X (GENeric eXpert system), which is their platform for developing diagnostic packages. Initial attempts to integrate ES – incorporating procedural maintenance for diagnosed faults – with GPD failed predominantly from insufficient instrumentation set to permit a meaningful modular diagnosis and also from insufficient PC power at that time (Doel, 1990). Recent studies were dedicated to the implementation of ES for fault diagnostics (Doel, 1990; Ali et al., 1990; Torella, 1993; Torella, 1997; DePold et al., 1999, and Spina et al., 2002). in cases in which only a qualitative answer can be sufficient. An example of ES application, discussed by Torella et al. (1999), regards diagnosis and troubleshooting of GTs from a probabilistic relationship between failures (causes) and symptoms (behaviour) because an observed symptom could be the result of several different causes having various degrees of probabilities and vice versa. The process was coded by means of a belief network.

2.3.11 Fuzzy-logic based GPD

Recently fuzzy-logic-based diagnostic methodologies have been devised by taking advantage of a convenient way of mapping an input space to an output space (pattern recognition) in the presence of uncertainty. The subject of the application of fuzzy-logic to engine fault diagnosis is dealt in much greater detail in chapter-4. This section is dedicated to review how different authors have applied this technology for GPD purposes in various ways, using different features of fuzzy systems. The input and output are discretized and this enables complex mathematical problems to be simplified (Zadeh, 1969). A model based fuzzy-logic for sensor fault accommodation was presented by Healy et al. (1997) and a fuzzy rule-based and case-based expert system for turbo-machinery diagnosis was developed by Siu et al. (1997). Tang (1999) described

a process based on fuzzy-logic and neural-network for engine diagnosis, taking into account faults in gas-path components, instrumentation, and rotor or oil subsystems. Ganguli (2001a and 2001b) presented a linearized fuzzy-logic intelligent process for gas turbine fault isolation. The method uses rules developed from a model of performance influence coefficients (via a linearized approach) to isolate five specified engine faults while accounting for measurement uncertainties along the gas-path. The goal of the process is to identify the faulty component without quantifying the deterioration, the output fuzzy sets are not further decomposed (defuzzified) using linguistic variable. The results showed FL to have good success rates except for the IPC and HPC where fault isolation accuracies were given as 97% and 94% respectively.

An ulterior fuzzy-logic process was developed (Ganguli, 2002) for helicopter rotor system fault isolation. More recently Karvounis et al. (2003) described a model-based process for assessing the overall condition for the T700 engine and determining when to service it. A fuzzy-logic formulation is implemented to automate the detection and provide an end-of-flight estimate for future prognostics.

Fuster et al. (1997) introduced a gas turbine diagnostics model where the uncertainty of component parameters was expressed by fuzzy-logic likelihood value and the fault symptoms were described by *True* or *Fals*e. An adaptive model for accurate simulation of gas turbine performance with the capability of adapting to engine particularities was developed and described for the first time by Stamatis et al. (1990a). In this method, modification factors (MF) which are the ratio of parameter values of reference performance maps and the values of the actual maps were introduced. The modification factors for every component were obtained through a Non-linear Generalized Minimum Residual method. Observation of the changes of modification factors between nominal and deteriorated engine can lead to detection of the location and the kind of fault of the engine, (Stamatis et al., 1990b). Proper selection of modification factors with optimization can also be used for fault detection of gas turbine components and sensors.

Most recently, a non-linear model based process was devised by Marinai et al (2003a): it uses FL technology for GPD of a three-spool turbofan engine. The system was first designed to isolate and quantify single component faults. Results, from tests carried out on the IPC, showed good fault estimation capabilities. A sensitivity analysis of the influence of the various fuzzy parameters on the diagnosis accuracy was provided by Alexander (2003). Marinai et al. (2003b) implemented a process capable of multiple fault isolation.

The use of fuzzy-logic technology in gas-path diagnostics experiences the following advantages:

- Similarly to the ANN, fuzzy rules are generated by means of information provided from training data. This makes them particularly suited for finding solutions to problems for which there are no exact solution, but only a large number of examples. In general fuzzy-logic has been shown

advantageous in approximating processes difficult to describe and solve mathematically.

- Fuzzy-logic systems are universal approximators (if enough rules are used), therefore solutions can be provided for cases, which have never been encountered before in the statement of the rules. i.e. they generalize from examples.
- Fuzzy-logic is able to deal with the non-linear nature of gas turbine performance and as ANN, FL could be used to perform data-fusion for gas turbine diagnostics: GPD, vibration analysis, gas-path debris analysis can provide a comprehensive input to generate a rule-based system able of higher diagnostics competences.
- Fuzzy algebra is particularly suitable to deal with the uncertainty that characterises diagnostics problems, a fuzzy-logic based system can deal with the gas-path measurements uncertainty.
- A fuzzy-logic system is fast and easy to set-up and can be used for on-line application due to the fast computational capability in inference mode.
- Although the rules are generated by means of an engine model, the final system is said to be 'model free'. The engine model is not run during the inference calculations.
- The generation of the rules, differently from the ANN training phase, is fast and observable. Fuzzy systems and neural networks are both black-box approximators and suffer all the pros and cons that come with that model-free status. As far as ANN are concerned, we cannot be sure of what functions or patterns they have learned without checking a large number of input-output test cases. The fuzzy black box contains rules that are observable and can be tuned or removed as independent modules. ANN forget part of what they have learned each time they are learning something new, new patterns can distort old patterns (Kosko, 1997).

and limitations:

- The model-free feature, that allows data-fusion and gains in computational burden, comes with the restriction that a fuzzy system does not admit model-based proofs of stability and robustness.
- Like other AI tools, fuzzy systems are unable to approximate credibly outside the range of data to which they have been exposed: this implies that a massive amount of data from encountered and foreseeable faults/conditions of operation would be required in their development.
- Fuzzy systems face the problem that the number of rules increases according to the complexity of the process that is being modelled. Nevertheless rules reduction strategies can be adopted.
- The accuracy of a fuzzy diagnostics system needs to be traded-off with the computational speed and burden. Nevertheless, in fault diagnostics, precision in isolating the faulty component is more demended than the actual accuracy in the estimate.

It is being suggested in the field of research (Ganguli, 2001a and 2001b, Karvounis, 2003, Marinai et al. 2003a and 2003b, and Marinai, 2004) that fuzzy-logic could provide accurate answers to gas-path diagnostics problems in gas turbines. Amongst other things, this research will seek to identify ways in which fuzzy-logic can be applied to improve gas-path diagnostics.

2.3.12 Diagnosis with transient data

Gas turbine diagnostics is commonly carried out monitoring and analysing steady state measurement data. Nevertheless, in some cases good quality steady state data are difficult to obtain or even not available. For example, some combat aircraft can operate for up to 70% of the total mission time with their engines in non-steady-state conditions. In addition, some gas turbine faults phenomena only appear during transient processes but could seriously degrade the operability of the engine especially at altitude, during aircraft manoeuvres and following missile release, such as miss-scheduled nozzle and compressor blade movement due to control system faults. Therefore, gas turbine fault diagnostics may be achieved using transient measurement data (Li, 2002b).

Merrington (1989) considered a Least Square Estimate (LSE) approach. Gas turbine engine transient process is simulated from consistent non-linear idle/max or max/idle transient data and fault diagnosis is achieved by means of LSE. Two fault cases are discussed: (i) a biased exhaust gas temperature sensor error and (ii) a changed final nozzle schedule. Recently, the development of a diagnostics model based on genetic algorithm using transient data of a turbofan engine was shown by Sampath et al. (2003b). Figure 32 shows a comparison of acceleration of HP spool (N_H) of a clean engine and a faulty engine. The specific approach used in this method is to compare model-based information with measured data obtained from the engine during slam acceleration. The measured transient data is compared with a set of simulated data from the engine transient model, under similar operating conditions and known faults, through a Cumulative Deviation (CD). The CD is the deviation between the parameters obtained from the transients of clean engine (Baseline) and the engine with a fault and is the difference between the areas subtended by the curves in Figure 32. The CDs obtained from the comparisons are minimized for the best match using Genetic Algorithm (GA). A diagnostics model based on ANN has been investigated by Sampath et al. (2003b).

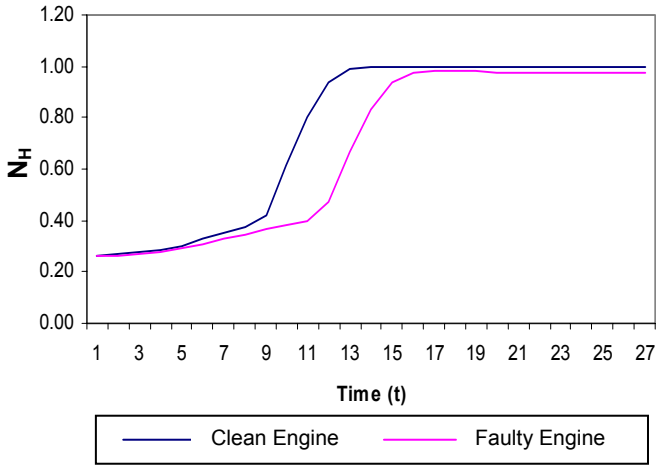


Figure 32: Typical measurement deviation during transients

Caution is required when such a diagnostics approach is to be implemented, the following aspect have to be considered (Curnock, 2000).

- For civil engines the amount of transient data from a single flight is quite small consisting one of the major acceleration (idle to take-off) and one major deceleration (cruise to idle for descent). A further acceleration and deceleration are available if reverse thrust is used.
- Instrumentation accuracy is poorer for transient than for steady-state. Several parameters are worse than 1% accurate at low speed. It is therefore not possible to determine values of health parameters within 1%.
- Transient manoeuvres cover regions of the HP compressor characteristic not normally covered in steady-state running. Slightly different values of the health parameters can be expected in these regions compared with those obtained from diagnosis of steady-state data. As the differences will be small it is unlikely that they can be used to any advantage, other than as approximate confirmation of the steady-state valued.

2.3.13 Summary and conclusions

In section 2.3 an extensive survey of advanced diagnostics methods has been provided and a summary of it is in Table 7. Ten of the key techniques, used in a wide range of applications, are described together with their advantages and limitations. Some of the approaches are based on the assumption that the changes in the health parameters are relatively small and the set of governing equations can be linearized. The inadequacy of this model has led to the development of non-linear methods being devised. Other techniques, such as WLE and Fuzzy-Logic approaches are particularly suitable for dealing with measurement uncertainties. Algorithms based on ICM inversions are suitable only if the number of measurements is more than the number of health parameters. Besides, they are not able to deal with measurement uncertainties. Estimation techniques as well as AI-based methods can deal with diagnostics with only a few measurements. A distinction can be made here (Volponi, 2003) between techniques more suitable for estimating (i) gradual deteriorations and others for (ii) rapid deteriorations as discussed in section 2.1.5. We will refer to

such methods as MFI (multiple-fault isolation) and SFI (single-fault isolation) ones respectively. The former implies that all the engine components (whose shifts in performance we are estimating) deteriorate slowly whereas the latter implies a rapid trend shift probably due to a single entity (or perhaps two) going awry. AI-based methods are more suitable for SFI problems, because they are based on an approximation of all the possible solutions for the limited number of cases used to train the system. The extension to all the possible combinations (even in a limited search-space) is theoretically possible, but extremely burdensome computationally and highly inconvenient. AI-based techniques do not experience the "smearing" problem (i.e. the tendency to smear the faults over a large number of the engine's components and sensors) that estimation techniques suffer from, but on the contrary have good "concentration" capabilities to identify the faulty component. Estimation techniques require prior information and the solution can be dramatically affected by this choice. Similarly, AI-based methods require particular care during the set-up phase. Moreover, AI-based algorithms can be excessively time-consuming, either in the actual calculation as in the case of a GA, or in the training phase as for an ANN. As far as fuzzy systems are concerned, they need a large number of rules for systems with many inputs and outputs such as a diagnostic system. An appropriate learning algorithm can restrict the number of rules required (Kosko, 1997). ANN, ES, BBN and FL are referred to as model-free systems. This model-free feature is responsible for data-fusion capability and gains in computational burden, but comes with the limitation that no model-based proofs of stability and robustness are possible. Besides, they cannot be accurate out of the range of variability for which they have been trained and set-up. Numerous tests must be conducted to validate these techniques. ES, BBN and FL approaches have the additional quality that they can be used to encompass expert knowledge in the system. Eventually, it has been recognised that the capability of making diagnoses on-line for on-board applications requires a blend of fast calculation and the ability to achieve a worthwhile solution with only a limited number of measurements which are affected by a considerable level of noise. The techniques that are better suitable for these circumstances are indicated in Table 7.

Particular interest, in this work, has matured regarding the potentialities of fuzzy-logic in implementing diagnostic methodologies by taking advantage of its convenient way of non-linear mapping an input space into an output space (pattern recognition) and dealing with uncertainty. A new interpretation of how this technology can be employed and weather it is able to performe effective diagnoses is estensively discussed in Chapter-4.

2.4 Trending and prognostics

2.4.1 Introduction

As discussed in chapter 1, many airlines nowadays demand payment for their engine maintenance costs on an hourly utilization basis. Thus, engine manufacturers have become more focused on performance-deterioration modelling and prognostics capability in order to achieve greater confidence in their cash-flow projections. Hence, a survey of methods for predicting the performance deterioration of civil aero-engines is undertaken. The main aims of these techniques are to achieve significant benefits in mission scheduling and maintenance planning, as well as to reduce the costs of maintenance servicing.

2.4.2 Survey

We have described in section 2.1.3 and 2.1.4 how jet engines' rotating-components degrade during operation and how this affects their performances (Saravanamuttoo, 1985). Gas-path diagnostics is aimed at detecting, isolating and assessing these changes in performances, even in the presence of possible instrumentation faults and system failures. A comprehensive overview of performance-diagnostics techniques can be found in section 2.3. Related to diagnostics, the goal of prognostics is to predict the engine's or component's health condition studying how gas-path measurements and performance parameters evolve over the operating time.
Recently much attention has been devoted to devising prognostics algorithms for the analyses of engine performances within both the civil and military sectors. The literature survey shows that prognostics is not a mature technology and some authors give more a description of the desirable features and the objectives of a prototypic prognostic system rather then a description of implemented algorithms.

Prognostics outcomes influence directly the problems of scheduling missions and planning maintenance. With an accurate prognosis-capability for a fleet of engines, the maintenance as well as spare-part orders can be effectively planned.

Sheuren (1998) presented an artificial-intelligence based prognostics and health management process (PHM) for the Joint Strike Fighter. Thereby the military services aim at eliminating traditional inspections and calendar-based maintenance: remedial actions are preferably based solely on existing condition. Jaw et al. (2001) and (Green, 1997) described how the U. S Air Force included artificial-intelligence based prognostics algorithms in the health management software. A knowledge–based expert system for prognosis, and its integration with the diagnostics results according to Pratt and Whitney experience was presented by DePold et al. (1999). Ghiocel et al. (1999) approached life prediction by focusing on uncertainty propagation using space-time variant stochastic process models. This concept has been applied: for example, a prognostics algorithm for a US navy's ship propulsion systems was discussed by Kacprzynski et al. (2001). The importance of data-fusion was

highlighted by Roemer et al. (2001) and the prognostics value for risk assessment in decision-making was discussed by Roemer (2000). A review of prognostics approach applications to gas-turbine health monitoring can be found by Byington et al. (2002). Ghiocel (2001) illustrated the application of a hybrid stochastic-neuro-fuzzy-inference process to fault diagnostics and prognostics. Brotherton (2000) proposed an integrated prognosis process that uses a dynamically-linked ellipsoidal basis function neural-network. A review of artificial-intelligence based methods was presented by Brotherton (2000). A prototype health-monitoring-and-prognostic process applied for the gas-turbine engine on the US Army M1 Abrams tank was discussed by Greitzer et al. (1999). Greitzer et al. (2003) investigated regression analysis for predicting the remaining useful life of engine components as part of the development of a diagnostics/prognostics system for the gas turbine installed on the US Army M1 Abraham tanks. Sasahara (1986) and Bakken et al. (1996) proposed a method based on regression analysis for forecasting the deteriorations to compute the optimum maintenance interval for engine components. The latter application of prognostics was also investigated by Moes (2003). The developed maintenance manager uses past and forecasted fuel flow data to compute the most cost-effective maintenance interval for the engine. Similarly, Beschorner (2003) investigated the employment of a time series forecasting technique based on regression analysis on simulated transient take-off data to predict maintenance requirements.

The capability of smoothing and detecting trends in time series has a relevant impact in the ability do provide a prognosis. Provost (1987,1995 and 1988) described an application of the Kalman filter to estimate underlying level and trend in a process generating noisy observations. The methodology was patented and is currently used in Rolls-Royce. A comparison with other often-used smoothing algorithms such as Moving Average and Exponential Smoothing was provided and the conclusions are discussed below. One of the problems pointed out regarding the Moving Average method was that it is not recursive, i.e. the calculations do not just depend on the current observation and immediately previous level estimates. The required storage of a history of data makes it computational inefficient. The Exponential Smoothing method can be regarded as a limiting case of Moving average method: it is recursive, and easy to implement. However, both of the above methods contains no trend estimate terms which means that (Provost, 1995):

- No estimates of trends are available.
- Time lags between the observations and the estimates are introduced, particularly when analysing a series in which a ramp is present.
- Both methods also do not make proper allowance for different intervals between the observations.

The Kalman filter based approach encompassed in COMPASS (Rolls-Royce) suffers from none of these drawbacks; nevertheless the limitations of this method are:

- The increase in computational load.
- A lack of versatility when changes in trends complicate the forecasting problems.

- No prediction intervals and therefore no levels of confidence in the results are provided.
- Short and long term investigations are approached in the same way although they have different requirements (O'Donovan, 1983).
- No allowances for rapid deterioration and trend changes are provided.

Marinai et al. (2003c) focused on forecasting algorithms, based on time-series models, which were applied to solve prognostics problems in engine-performance analysis: the probability of deterioration and its magnitude during the next time period of interest were calculated. Two techniques to handle gradual deteriorations in different prognostics problems were presented. Box-Jenkins ARIMA method was used to provide accurate forecasts for immediate and short-term forecasting. Regression analysis was designed to handle prognoses that require medium and long-term predictions, focusing on physics-based mathematical models of the degradation.

2.4.3 Prognostics strategies in engineering

A lot of work has been done in predicting the component life under the planned usage from a mechanical point of view considering creep, thermal fatigue, etc. However, forecasting the gas-path performance deterioration is a fairly new research subject. Various approaches to prognostics have been developed, mainly to study mechanical problem. These techniques range from simple historical failure rate models to high-fidelity physics-based models. For practical purposes, these approaches can be generalized into three basic forms; Figure 33 illustrates the hierarchy of potential prognostic approaches in relation to their applicability and relative accuracy (Byington et al., 2002).

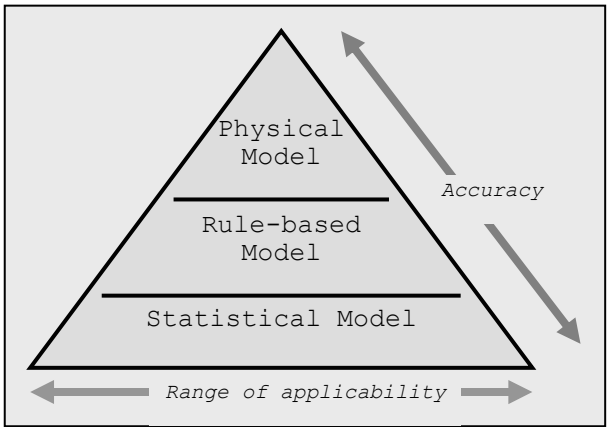


Figure 33: Hierarchy of prognostic approaches.

Physics-based prognostics:
These are models that have been developed by experts and validated on data. A physics-based stochastic model is a technically comprehensive modelling approach that has been traditionally used for component failure mode prognostics. For example, it can be used to evaluate the distribution of remaining useful component life (Kacprzynski et al., 2002).

Rule-based prognostics:
These are systems that embody rules that have been developed and refined by human engineering and maintenance experts. Examples of these systems are rule-based expert systems and fuzzy-logic systems. For example, Kacprzynski et al. (2001) proposed a probabilistic-based technique that utilizes the known information on how measured parameters degrade over time, based on a specific test, to assess the current severity of parameter distribution shifts and project their future state.

Statistics-based prognostics:
In the case where physics- or rule-based models are absent but historical observations of the phenomenon are available, a statistics-based prognostic model may be a suitable alternative. This class of prognostic models can be implemented by investigating historical data with time series analysis.

2.4.4 Summary and conclusions

Recently, several attempts were made to device tools with prognostics capability, but a mature technology is not available. Different authors identified diverse features that such a tool should have and addressed them with various approaches. The main purpose of the above literature survey was to identify the requirements and the specifications for designing an advanced prognostic process. These will be commented and put in the contest of the present work in the next section 3.2. Most of the studies carried out showed qualitative nature or lack of statistical rigour. Some of them (e.g. Sasahara, 1986, Provost, 1995, and Bakken et al., 1996) were able to identify the trends but no confidence level for the results is provided. Generally, no distinction is made between short and long term investigation, that notoriously have different requirements (O'Donovan, 1983), neither between gradual and rapid deterioration phenomena which are investigated. Three typical prognostics approaches employed in engineering were reported in section 2.4.3. It is our opinion that the statistic-based approach, at the present time, is the most suitable for predicting the aero-thermal performance of a gas turbine given its simplicity and wide applicability. Future studies, if more sophisticated models of the deterioration will be available, could be dedicated to investigate other approaches.

2.5 Summary

In the light of the techno-economic considerations discussed in Chapter-1, this work focussed on designing advanced health monitoring and prognostics methodologies. Hence, the aim of this research, stated in section 1.3.2 is to develop a novel and versatile model able to offer a high level of gas-path diagnostic competence, and prognostic capability to be used for civil aero-engines on-board applications. In view of this research aim, this chapter provided a break-down of the research aim in three main tasks: (i) observability pre-processing, (ii) diagnostics process, and (iii) prognostics process. The three problems listed above were thoroughly defined discussing their complexity and the associated difficulties, in section 2.1. Hence, a literature survey provides an overview from the early works up to the state of art of the methodologies developed to assist these types of investigations. Two tests and a method based on eigenvectors and eigenvalues were reviewed to undertake observability studies. Table 7 provided a summary of gas-path diagnostics methodologies and it was commented in section 2.3.13 . Section 2.4.3 discussed different prognostics approaches typically used in engineering – see also Figure 33.

Pertinent publications were reviewed in order to point out advantages and limitations of the current methodologies as well as to provide a guideline for the next chapter which is dedicated to identify the gaps in the contemporary studies. Hence, the objectives of this research project will be stated according to the desirable requirements for novel diagnostics and prognostics systems. The research method to achieve those objectives will be then discussed.

CHAPTER 3 - RESEARCH OBJECTIVES AND STRATEGY FOR THE INVESTIGATION

3.1 Introduction

This study was sponsored by Rolls-Royce plc. The major techno-economic factors for undertaking this work are discussed in Chapter-1: these are concerned with the emergence for engine manufacturers of new active aftermarket strategies to provide total-care solutions, which make them highly challenged to design advanced health monitoring and prognostics methodologies. The aim of this research, as stated in Chapter-1, was to develop a novel and versatile framework able to offer a high level of gas-path diagnostic competence, and prognostic capability to be used for civil aero-engines on-wing applications.

In Chapter-2 a literature survey provided an overview from the early studies up to the state of the art of those methodologies. Techniques developed to undertake observability investigations were reviewed in section 2.2. Table 7 (see section 2.3) provided a critical summary of the gas-path diagnostics methodologies. Different prognostics approaches employed in engineering were reported in section 2.4.3 and were shown in Figure 33. Authoritative publications were reviewed with the twofold aim of identifying the pros and cons of the present methodologies as well as specifying the desirable requirements for novel diagnostics and prognostics systems as it follows in the next section.

In this Chapter, the research objectives will be outlined and the research method to achieve those objectives discussed. The contribution to science and the benefits resulting from undertaking this investigation are anticipated.

3.2 Gaps in contemporary methodologies: research objectives

In Chapter-2, three fundamental elements of a performance analysis tool aimed at monitoring the gas turbine's health during its operation have been surveyed:

- Observability pre-processing
- Fault diagnostics process
- Fault prognostics process

The importance of a good integration among the three aspects was pointed out. As far as the **observability** pre-processing is concerned, techniques that range from simple tests to more sophisticated approaches based on eigenvalues and eigenvectors analysis were presented (section 2.2). The result of the observability study enables an effective interpretation of the GPD results to be achieved irrespective of which diagnostics technique is employed. The outcome of a reliable diagnosis must not only indicate the quantifiable faults, but must also warn about the faults that are not observable with the given measurement set.

The fact that gas-path **diagnostics** is a mature technology, from some extents, and that a lot of work was carried out in this area is not in doubt. Nevertheless, new issues have been encountered through the history of its development and new techniques have been identified for their particular strengths towards some of these problems. However, no technique provides a satisfactory and complete answer to all of them: the limitations of the most popular approaches were critically analysed in section 2.3. A distinction was made (Volponi, 2003) between those techniques more suitable for estimating (i) gradual deteriorations (MFI) and those for (ii) rapid significant deteriorations (SFI). The need for a framework that enables the two processes to operate automatically in concert with one another, and without corruptive interaction, was indicated by Volponi, (2003).

In consideration of the primary role that gas-path diagnostics plays, in an industry in which, the core of the businesses is based on aero-thermal performance, the need for an effective approach cannot be ignored. A thorough analysis of the published literature, summarised in Table 7, suggests that the requirements for an advanced diagnostics process should be:

- Based on a non-linear model.
- Designed specifically for SFI and/or MFI.
- Capable of detecting with reasonable accuracy significant changes in performance.
- Able to provide a 'concentration' capability on the actual fault.
- Competent to make a worthwhile diagnosis using only few measurements (N>M).
- Able to deal with random noise in the measurements and sensor bias.
- Light in computational requirements.
- Fast in undertaking diagnoses for on-wing applications.
- Readily adaptable to similar systems in a reasonably short time: exempt from training and tuning uncertainties, difficulties and dependences for setting-up parameters.
- Free from a lack of comprehensibility due to black-box behaviour.
- Capable of data-fusion.
- Able to incorporate expert knowledge.

Although fuzzy systems have only recently appeared as advanced gas-path diagnostics methodologies with quantitative capability (Marinai et al., 2003a), some authors (Siu C. et al., 1997, Ganguli, 2001a and 2001b, and Ogaji et al., 2003) consider them to have good prospects in solving gas-path fault diagnostic problems due to their advantageous rule-based pattern recognition model and their inherent capability of dealing with uncertainty. The first objective of this research is to identify a novel effective way of applying fuzzy-logic technology in a quantitative sense, aimed at fulfilling the above listed requirements and at contributing to gas-path diagnostics development (Chapter-4).

As far as the **prognostics** process is concerned, dissimilarly from the diagnostics one, it cannot be considered a mature technology. Most of the pertinent research carried out has a qualitative nature or lacks statistical rigour. First of all, it was here recognised that gradual and rapid deteriorations have to be distinguished and treated separately. The information for predicting fault events and for predicting the deterioration's evolution are located in different types of data as discussed in section 2.1.6 and therefore require different approaches. The requirements that emerge for an advanced process that performs effective prognostics are the following:

- Able to include the confidence level in the projections (see prediction intervals as discussed more thoroughly in Chapter-5).
- Designed to cope with short and long term investigations.
- Able to cope with slow and/or rapid deterioration scenarios.
- Designed with different strategies for dealing with performance parameters and/or measurements.
- Possibly, easy to understand/explain in management meetings, where cash-flow projections are discussed.
- Light in computational requirements.

Hence, a second objective of this research is to identify a framework, based on statistical models, that predicts the performance deterioration of civil aero-engines and fulfils the above listed requirements. The main aims are to achieve significant benefits in mission scheduling and maintenance planning, as well as to reduce both fuel consumption and the cost of maintenance servicing (Chapter-5).

3.2.1 Contribution from current research

The aim of this project was to develop a novel and versatile framework able to offer a high level of gas-path diagnostics competences, and prognostics capabilities for civil aero-engine on-board applications. The research aim naturally separates into two elements: diagnostic and prognostic processes. Moreover to achieve effective diagnostics, an observability study module was devised.

The outcomes of this study were significant components of a performance health-monitoring-and-prognostics (HMP) framework. The general structure consists of three sub-processes – observability study, fault assessment or diagnostics, as well as forecast and prognostics – that are carried out via modules that compute the actual calculations and communicate with each other. The three modules constitute the software named HMP 1.1 for performance analysis (see the Main window of the software in Figure 34), which is the outcome of this research – see Appendix C.

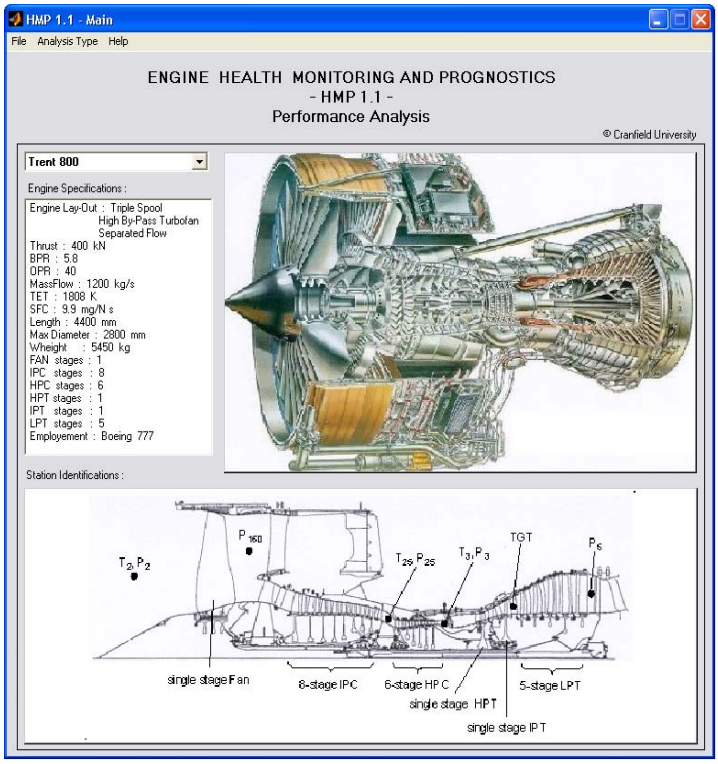


Figure 34: Main window of HMP 1.1 for performance analysis, image courtesy of Rolls-Royce

The core of this work is the diagnostics procedure. Nevertheless, to achieve a greater confidence in the diagnostics results an observability pre-processing study is required. Chapetr-4 is dedicated to the procedures developed to perform the diagnostics and observability study. After having diagnosed a fault, the prognostics module provides prospects and decision-making competences based on them. The prognostics process is dealt with in Chapter-5.

It will be demonstrated later, in the document, that the outcome of this PhD research makes the following novel contributions:

- The use of a fuzzy logic based process, that employs rules generated by means of an engine model, and defuzzifies the output to achieve a quantitative solution is a new concept in engine fault diagnostics as, so far, only qualitative types of outputs were attempted. The statement of a high number of rules to approximate, in a predefined search-space, the non-linear relationship between measurements and performance parameters, taking into account the measurement uncertainties, is also a new concept. First, a SFI procedure based on fuzzy logic was implemented. An additional feature was devised to tune the rules of the fuzzy diagnostics system over a known deterioration level for all the performance parameters (baseline). This baseline is assumed to be calculated at the previous flight with a MFI method and represents the global deterioration level. This feature enables SFI and MFI processes to operate in concert with one another. Moreover, a fuzzy logic based procedure, capable of a partial MFI with up to 2 degraded components (4 performance parameters) which are significantly faulty simultaneously was implemented. Another innovative outcome of this research is the

idea implemented to use fuzzy logic to deal with systematic errors in the measurements (biasses). A bias-tolerant system was devised by means of the NOT logical operator and a new formulation of the fuzzy rules that includes the location of the bias.

- A novel idea was proposed to quantify the quality of the system observability through a pertinent parameter. This enables extensive comparisons among a significant number of systems characterised by different measurement set selections.
- An innovative prognostics framework was devised, that uses two different statistic-based approaches for short and long term investigations (specific to different prognostic problems), and treats differently gradual and rapid deteriorations as well. The idea of using prediction intervals that take into account the errors in deterioration modelling as well as the forecasting errors in the prognoses of the performance of the engine (or of one of its components) is a new concept. A procedure that uses forecasted measurements in the input to the diagnostics process, and gains statistical information on the predicted performance parameters in the output is proposed here as a novel logic scheme.

3.2.2 Benefits from the current research

Industry has yet to obtain the full extent of the added value that advances in gas turbine health monitoring systems offer, particularly in the new business scenario based on an agreed payment rate per engine flying hour from the airlines. Overall, the benefits that are expected to arise from the use of the diagnostics and prognostics techniques presented in this thesis can be summarised as follows:

- Definition of service work packages, based on actual diagnosed condition, instead of those for an “average” engine.
- Improved safety in operating gas turbine engines.
- Reduced overall life-cycle cost.
- Optimisation of the maintenance intervals and prioritisation of the related tasks to enhance operational availability.
- Engineering justification for scheduling maintenance while identifying corresponding economic benefits.
- Reduced need for spares holding.
- Clarity in defining cost-effective aftermarket agreement objectives.
- Pertinent instrumentation selection against usage objectives.
- Enhanced availability management to limit the need for unplanned servicing: improved “departure” statistics, reduced in-flight shutdown rates and maintenance actions away from base.
- Enhanced airline reputation.

3.3 Research strategy

The preliminary strategy proposed in Chapter-1, in view of the detailed objectives defined above, can now be developed as follows. For the diagnostic

process, the main steps that will be undertaken – see Chapter-4, are listed below:

- Analysis of fuzzy algebra, fuzzy systems and their opportunities in GPD.
- Development of a conceptual model for a fuzzy-logic based diagnostic process.
- Considerations about the choice and modelling of an engine, to carry out the research.
- Observability study for the specific engine.
- Implementation and set-up of a fuzzy-logic based SFI system for that engine from the conceptual model. In the light of the objectives, first take into account the problem of fault diagnosis in the presence of sensor noise, neglecting the possibility of biasses. Identification of the most significant parameters in the process.
- Sensitivity study to optimize the functional and system parameters.
- Evaluation of the effectiveness of the developed diagnostic framework with simulated case studies.
- Enhancement to a partial MFI capability considering up to 2 components being faulty simultaneously.
- Extension of the diagnostic capability to the sensor bias problem and test it with simulated data.
- Discussion.

Chapter-4 is entirely dedicated to the description of how this strategy was carried out. As far as the prognostic process is concerned, the research strategy includes:

- Development of a conceptual prognostic model in view of those requirements that takes into account the capability of dealing with (i) short-term and (ii) long-term investigations.
- Implementation of two novel approaches, based on that conceptual model to cope with the two different scenarios.
- Investigation of the two developed methodologies in safety and economic applications.
- Evaluation of the potentialities and the effectiveness of the developed prognostics framework with simulated case studies.
- Discussion

3.4 Summary and conclusions

Research objectives and strategy have been discussed in this chapter. Overall, the aim of this research is to develop relevant elements within a health-monitoring and prognostics framework for performance analysis capable of advanced diagnostics and prognostics. The research aim logically was broken down into two main research objectives concerning the development of diagnostic and prognostic competences. A requirement concerning the observability investigation capability embedded in the diagnostic process emerged.

Chapter-2 provided a review of pertinent publications aimed at identifying the pros and cons of the current methodologies. Hence, the desirable requirements for novel diagnostics and prognostics systems were derived and stated in section 3.2.

It can be concluded that there is no single technique which addresses all the issues. Some of the techniques are complementary and each has its own advantages and limitations.

As far as the diagnostic methodologies are concerned, particular interest matured towards the potentialities of fuzzy-logic technology, namely: (i) its convenient way of non-linear mapping an input space into an output space (pattern recognition) by using a rule-based approach, (ii) its attitude to deal with uncertainties by means of its fuzzy approach, and (iii) its suitability to performe data fusion and propensity to be fused with other techniques. Chapter-4 proposes a novel way of using this technology to fulfil most of the requirements stated in this chapter. Moreover, the importance of a preliminary observability study was recognised: Chapter-4 presents how this issue is dealt with in this research.

Future studies should be dedicated to the attempt to combine more than one technique to offset the limitations of one with the advantages of another within a combined engine-health monitoring (COEHM) scheme (Singh, 2003). This task can be thought of in two ways. One is to use critically the results from different methods to work in concert with one another, exploiting a potential synergic effect. The second way is to design hybrid systems (Sampath, 2003a).

On the other hand, industry is showing a major interest in engine health monitoring and prognostics (HMP) schemes. Any future solution is expected to offer a high level of diagnostics competence, integrated to prognostics capability including life-cycle management and economic implications (Singh, 2003). The availability of large quantities of operational data from individual engines and fleets will provide statistical databases which, when taken together with advanced diagnostic methods, will allow important advances to prognostics, so further increasing the market value of these technologies. The future challenge is to design effective combined engine health monitoring and prognostics (COEHMP) procedures. This study does not intend to design a complete COEHMP but to contribute to its future realization. In Chapter-5 a prognostics module is proposed in the attempt to fulfil the requirements stated above.

CHAPTER 4 - GAS-PATH DIAGNOSTICS USING FUZZY LOGIC

4.1 Introduction

This chapter presents a new gas-path diagnostics (GPD) method. The novelty of this technique lies in the use of fuzzy logic to provide secure isolation and quantification of gas-path component faults. Fuzzy logic is introduced because of its inherent capability of dealing with GPD problems due to its rule-based nature and its fuzzy approach. The rule-based architecture is used to perform pattern recognition of measurement fault signatures, while the fuzzy approach is advantageous to deal with the uncertainties that typically affect the GPD problem, namely the measurement errors and the undetermined mathematical formulation. These features created a research opportunity; and an application of the method to the Rolls-Royce Trent 800 engine and its encouraging results will show, in this chapter, that the promises of fuzzy logic were justified. A graphical user interface (GUI) was devised as the diagnostics module part of the Heath Monitoring and Prognostics (HMP) framework presented in this thesis – see HMP 1.1 interface in Appendix C. Firstly, its SFI (single fault isolation) capability was proven – see section 4.6. Then a partial MFI (multiple fault isolation) capability, with up to 2 gas-path components being considerably faulty simultaneously, was tested – see section 4.7. Eventually a novel bias-tolerant process to cope with systematic errors in the measurements (bias) using fuzzy logic is presented and tested – see section 4.8. It was recognised that, before any diagnosis is computed, an observability analysis must be performed to study the correlations between the measurements used in the input to the diagnostics process and the correlations between the performance parameters sought. The results of two preliminary observability tests performed for the Trent 800 are discussed in section 4.4. Besides, section 4.4.5 describes the procedure implemented within this project to identify a suitable measurement set (i.e. an enhanced observability study) and to advise the diagnostics process with the quality of the observability that will affect the diagnostics capability. The role that these outcomes play in the diagnostics process is highlighted.

4.1.1 A guide through the chapter

Section 4.2 is aimed at guiding the reader through the fuzzy logic process step-by-step from an introduction to the theory to the application to gas-path diagnostics. Section 4.3 introduces the Rolls Royce Trent 800 engine involved in the project and the instrumentation set used in the diagnostics process. The observability study is described in section 4.4. Section 4.5 is then dedicated to the development of the fuzzy diagnostics system for the three spool Trent 800 engine and to the sensitivity studies carried out for a pertinent set-up of the methodology. The GUI devised for this purpose is also introduced. The accuracy of the SFI capability of the system in the presence of noisy measurements and a method used to enhance such a capability is discussed in section 4.6. This section also describes an additional feature of the system whose rules can be tuned over a global deterioration baseline to enhance the SFI role in GPD. A fuzzy diagnostics system able to perform a partial MFI and

its accuracy are discussed in section 4.7, while, section 4.8 presents a bias-tolerant SFI fuzzy diagnostics system and its results. A second GUI (part of the HMP 1.1 diagnostics module) was devised to make use of the fuzzy diagnostics model in order to compute the diagnoses and plot the results; this is presented in section 4.9. A discussion of the outcomes of the procedures presented in this Chapter-4 in the light of the requirements specified in Chapter-3 is reported in section 4.10, whereas the recommendations that ensued from this research concerning observability study and GPD are presented in section 4.11.

4.2 Fuzzy logic systems

4.2.1 Introduction

Fuzzy logic is a new rule-based approach, founded on the formulation of a novel algebra, typically used in the analysis of complex systems and to enable decision-making processes (Zadeh, 1969) to be performed.

Fuzzy engineering is the specific research area investigated in order to model engineering processes with fuzzy systems. These are able to provide appropriate approximations of various phenomena if enough rules are defined. The quality of the approximation is strictly related to the quality of the rules. This is not a standard view of fuzzy systems but is the view taken in this thesis according to the definition of fuzzy engineering given by Kosko (1997). A different view is that fuzzy logic is a linguistic theory that models human reasoning with vague rules of thumb and common sense. This, without any doubt, holds in many applications. Fuzzy systems, as described in the next section, rely on the formulation of fuzzy algebra. This is a generalization of the abstract set theory, based on new definitions concerning fuzzy sets and logical operators (Zadeh, 1969).

Fuzzy logic is used in this research to provide the capability of approximating the relationships between the M-dimensional input space of the gas-path measurements and the N-dimensional output space of the performance parameters by using a number of fuzzy rules. The rules in turn depend on fuzzy sets able to deal with uncertain or vague estimations of the process variables.

Fuzzy logic is all about the relative importance of precision. It is a convenient way to map inputs into outputs (Zadeh, 1995) and the primary mechanism for doing this is a list of if-then statements called fuzzy rules. All the rules are evaluated in parallel and the order of the rules is unimportant. To set up a system that interprets rules, we first have to define all the elements of a fuzzy system (i.e. fuzzy sets, membership functions, logical operators and architecture of the rules) and then the elements of the inference process, namely the algorithms for the implication, the aggregation and the defuzzification phases. The fuzzy inference process interprets the values in the input vector and, based on a set of fuzzy rules, assigns values to the output vector.

This section is aimed at describing the fuzzy logic process step by step by providing an introduction to the theory and an application to gas turbine performance diagnostics. The next three sections will go from general to specific, first introducing the underlining ideas and then discussing the implementation details according to the specific problem of GPD. The following three subjects shall be discussed:

- **Fuzzy Algebra,** which provides an introduction to the general concepts and the basic elements of the architecture of a fuzzy system – see section 4.2.2.
- **Fuzzy inference systems**, which explains the specific fuzzy inference methods used in this research project – see 4.2.3.
- **Functionality of a fuzzy logic diagnostic system**, this goes into the detail of how fuzzy logic is used, in this work, to devise a novel gas turbine diagnostics process. This section, for simplicity refers to the example of a single shaft jet engine with three measurements and four performance parameters. This system is not optimized to perform diagnostics but is only used to introduce the main concepts – see section 4.2.4. In section 4.5, all these concepts are then used to set up an effective diagnostics system for the Trent 800 engine.

4.2.2 Fuzzy algebra: basic elements of a fuzzy system architecture

Engineering science typically deals with uncertain variables and approximations to a fixed number of decimals places that depend on the accuracy capability but also on the necessity and costs of being accurate. When a decision has to be made based on uncertain values of a set of variables, a binary logic based on either-or laws can become a limitation.
A fuzzy system based on multivalue logic can help in modelling a process when a mathematical model of how the system's outputs depend on the inputs is not available or is not accurate or when it is necessary to deal with the uncertainty present in the inputs. Besides, a fuzzy model is beneficial in order to introduce different sources of information in the decision making process (data fusion) and when it is advantageous to include expert knowledge or statistical inputs.
Fuzzy logic systems rely on the formulation of a novel abstract set theory and algebra: a generalization of the set theory, based on fuzzy sets as well as logical operators will be considered below. The four main elements of a fuzzy logic inference process are listed in Figure 35 and discussed in the following sections.

It will be proved that fuzzy set theory, introduced by Zadeh in 1965, is a generalization of abstract set theory. In other words, the former always includes the latter as a special case; definition theorems, and proofs of fuzzy set theory always hold for non fuzzy sets. Because of this generalization, fuzzy set theory has a wider scope of applicability than traditional set theory in solving engineering problems that involve high degrees of uncertainty and, to some degree, subjective evaluation (Kandel, 1986).

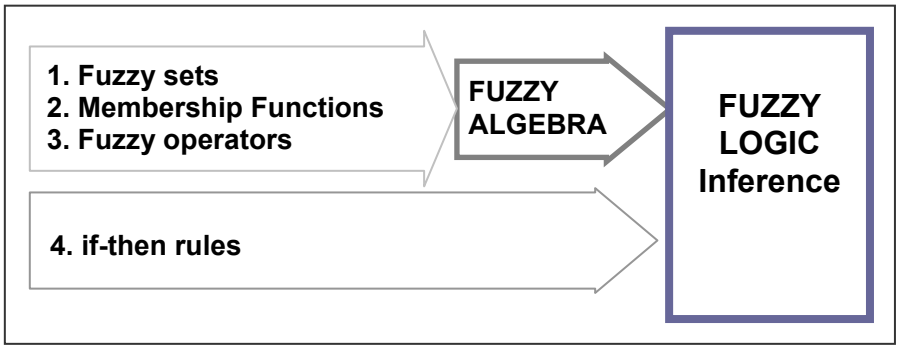


Figure 35: Fuzzy algebra and fuzzy logic inference.

4.2.2.1 Fuzzy sets

The basic concept behind fuzzy algebra and fuzzy logic systems is the definition of fuzzy sets. A fuzzy set does not have distinctly delineated boundaries and contains elements with a partial degree of membership.

In standard algebra a traditional set includes elements with a boolean or two-value logic. This means that an element belongs or does not belong to the set. The degree of membership of an element can only be either 0 or 1 (or alternatively either 0% or 100%). If we consider the example in Figure 36, the numbers A=51, B=60 and D=69 are elements of the set S, while the number D=71 is not. This concept is graphically described in Figure 37. The numbers included in the range between 50 and 70 belong to the set of cool air temperature.

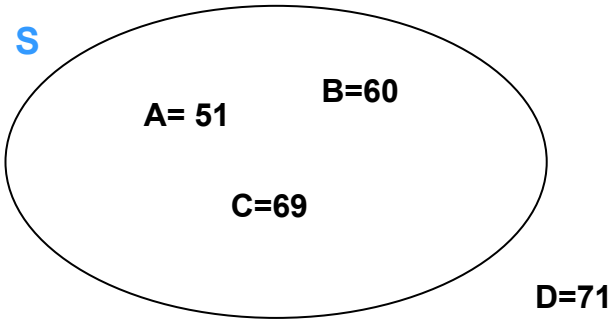


Figure 36: Standard Set

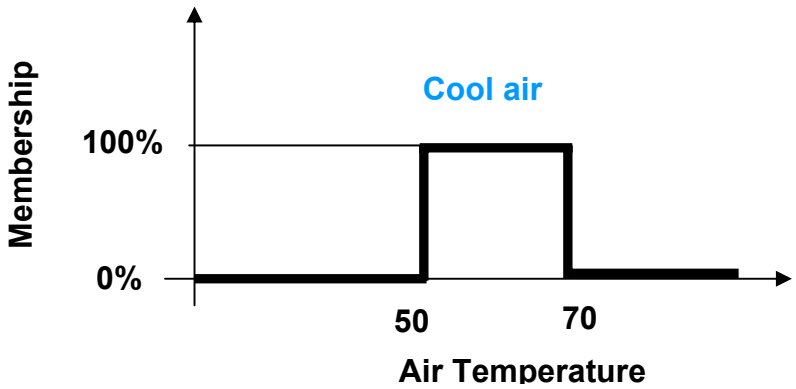


Figure 37: Diagram of a Standard Set

On the other hand, a fuzzy set admits elements with a partial degree of membership according to a defined membership function (MF). In the example of Figure 38, the membership function is triangular, therefore the degree of membership decreases going towards the margins of the set.

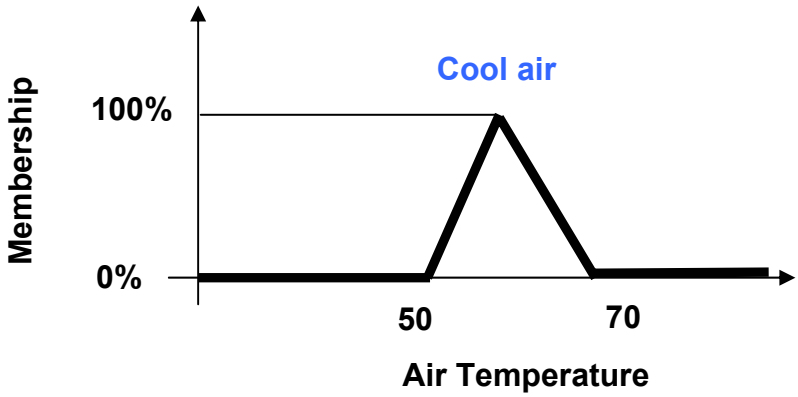


Figure 38: Diagram of a Fuzzy Set

In Figure 39 the two overlapping fuzzy sets of cool and right air temperature are considered. A value of temperature such as 68 degrees has distinct values of degree of membership to the two sets and consequently activates the two MFs with two different degrees of activation.

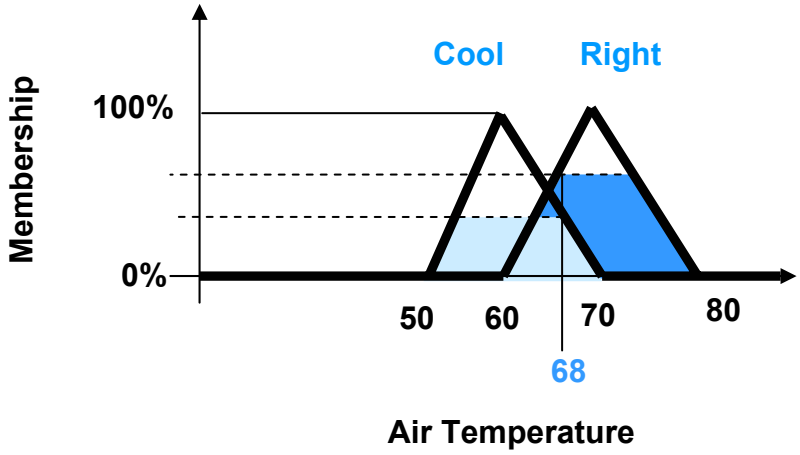


Figure 39: Diagram of two overlapping Fuzzy Sets

Going from the graphical representation to the analytical form, let X={x} denote the space of objects. Then, a fuzzy set A in X is a set of ordered pairs

$$A = \{x, \mu_A(x)\}, x \in X \qquad (4.\ 28)$$

where $\mu_A(x)$ is the degree of membership of x in A and the function μ_A is called membership function (MF). Usually $\mu_A(x)$ is a number in the interval [0,1], with the grades 1 and 0 representing, respectively, full membership and non membership in a fuzzy set. It maps each element of the input space X to a membership value. The input space is sometimes referred to as the universe of discourse. The membership function itself can be an arbitrary curve whose shape is defined as a function that suits the problem from the point of view of simplicity, convenience, speed, and efficiency.

Summarizing, the following concepts have been introduced so far:

- Fuzzy set.
- Degree of membership.
- Membership function (MF).
- Degree of activation (d.o.a.)

The next section will consider the logical operators, the third element of the fuzzy inference process – see Figure 35.

4.2.2.2 Logical operators

Fuzzy logic is a generalization of standard boolean logic, this means that the logical operations, as defined in this section, will hold in standard algebra, as well. As far as the logical operators AND, OR, and NOT are concerned, Figure 40 shows the truth tables according to traditional logic.

A	B	A and B
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	A or B
0	0	0
0	1	1
1	0	1
1	1	1

OR

A	not A
0	1
1	0

NOT

Figure 40 : standard logical operations

Figure 41.a shows a graphical representation of the logical operators in a two value logic. Many methods are available in the literature for their implementation in a multivalue logic or fuzzy logic. In this work the, following algorithms are considered:

- AND using minimum or product (a·b).
- OR using maximum or algebraic sum (a+b-a·b).
- NOT using the complement.

An example of fuzzy operators using the first options in the list above is shown in Figure 41, where we replace *A* AND *B*, where *A* and *B* are limited to the range (0,1), by using the function *min*(*A*,*B*). Using the same reasoning, we can replace the OR operation with the *max* function, so that *A* OR *B* becomes equivalent to *max*(*A,B*). Finally, the operation NOT *A* becomes equivalent to the operation (1 – *A*). Once the logical operators are defined, any construction using AND, OR, and NOT applied to fuzzy sets can be resolved.

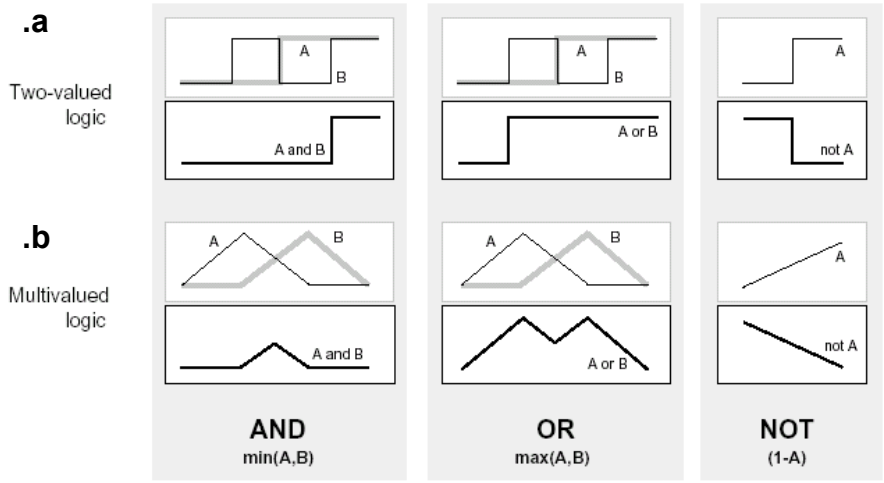
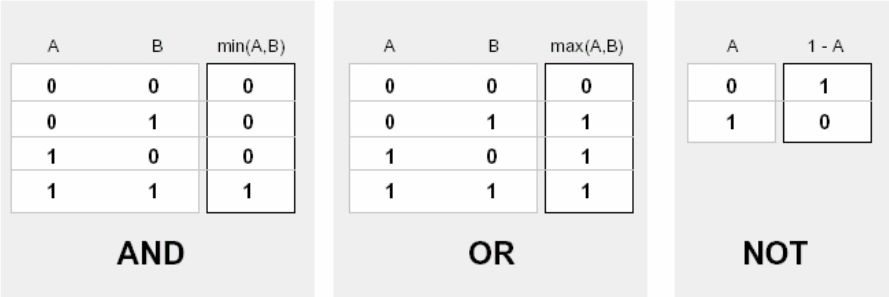


Figure 41: Two-valued and multi-valued logic.

It can be proved that these definitions still hold in traditional algebra, considering Figure 42. As an example, looking at the AND operator in the table we can see that: the min(0,0)=0, min(0,1)=0, min(1,0)=0 and min(1,1)=1. These outcomes coincide with the traditional logic truth tables of Figure 40. Similarly, we can reason for the second options in the list of possible algorithms provided above (e.g. change min with product to implement the AND operator).



A	B	min(A,B)
0	0	0
0	1	0
1	0	0
1	1	1

AND

A	B	max(A,B)
0	0	0
0	1	1
1	0	1
1	1	1

OR

A	1 - A
0	1
1	0

NOT

Figure 42: Example of logical operator, fuzzy algebra

In fuzzy algebra AND, OR, and NOT are known as the fuzzy intersection or conjunction (AND), fuzzy union or disjunction (OR), and fuzzy complement (NOT), but as said before the definition of them is by no means unique.

4.2.2.3 Fuzzy rules

Fuzzy rules play a key role in the fuzzy inference process – see Figure 35. Fuzzy systems are universal approximators if enough rules are stated. Fuzzy sets and fuzzy operators, that constitute the fuzzy algebra, are the elements of if-then rule statements. A single fuzzy if-then rule assumes the form 'if *z* is in the fuzzy set A then *x* is in the fuzzy set B'. The if-part of the rule "*z* is in *A*" is called the antecedent, while the then-part of the rule "x is in *B*" is called the consequent.

With reference to the Figure 43, an M-dimensional input space (in performance diagnostics, the measurements) is mapped into a N-dimensional output space (performance parameters) by means of m rules. Each input vector partially

activates all the rules in parallel. The rule can be associated with different rule-weights w_i. Eventually, a defuzzifier calculates the outcome solution based on the activation of the MFs. It will be proved that an additive fuzzy system compute a conditional expectation E(X|Z) and therefore an optimal non-linear estimation.

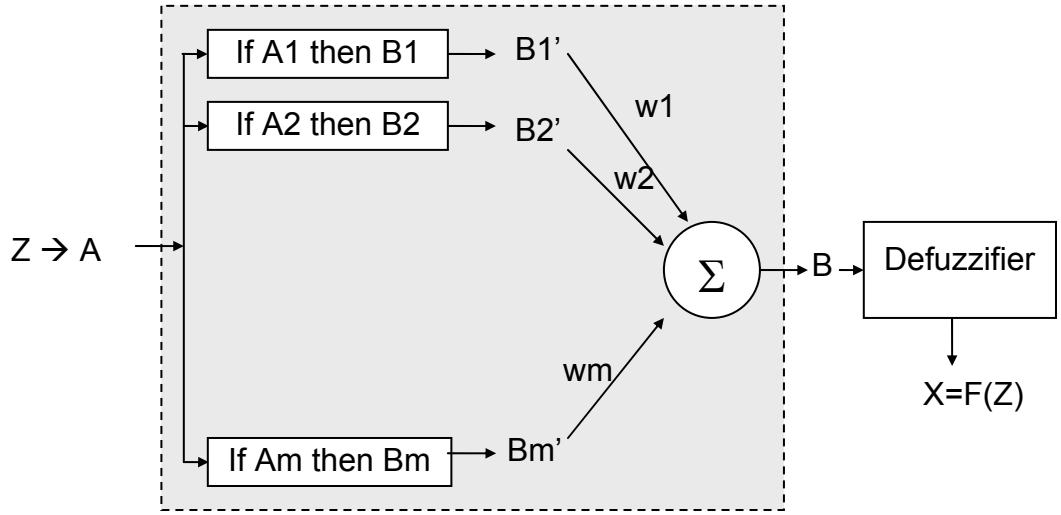


Figure 43: Additive fuzzy system architecture.

Interpreting an if-then rule (i.e. evaluating the antecedent) involves the following phases: (i) fuzzifying the input and (ii) applying any necessary fuzzy operators. Successively, its outcome is applied to the consequent (known as implication). In the case of two-valued or binary logic, when the if-part of the rule is true, the then-part is true. In a multivalue logic the antecedent is a fuzzy statement, so if the antecedent is true to some degree of activation, then the consequent is also true to that same degree. This is a very important concept that will be thoroughly described in section 4.2.4 referring to a simplified graphical example.

Therefore, interpreting one if-then rule is a three-part process:

- Fuzzify inputs: resolve all fuzzy statements in the antecedent to a degree of membership between 0 and 1.
- Apply fuzzy operator (AND, OR, NOT) to multiple part antecedents: If there are multiple parts to the antecedent, apply fuzzy logic operators and resolve the antecedent to a single number between 0 and 1. This is the degree of support for the rule.
- Apply implication method: Use the degree of support for the entire rule to shape the output fuzzy set. The consequent of a fuzzy rule assigns an entire fuzzy set to the output. This fuzzy set is represented by a membership function that is chosen to indicate the qualities of the consequent. If the antecedent is only partially true, (i.e., is assigned a value less than 1), then the output fuzzy set is truncated according to the implication method.

In general, one rule by itself does not do much good. What is needed are a number of rules that can play off one another. The outcome of each rule is a fuzzy set partially activated for each output of the fuzzy system. The output

fuzzy sets for each rule are then aggregated into a single fuzzy set for each output. Finally, the resulting set is defuzzified, or resolved to a single number (Zadeh, 1995).

4.2.3 Fuzzy inference systems

Fuzzy engineering can be implemented according to a three step procedure aimed at defining the system architecture. The first step is the identification of the input and output variables Z and X. In a diagnostics system the input variables are the elements of the set of measurements and the outputs the performance parameters representative of the health of the engine. Nevertheless, we will discuss how either input or output can include different information (e.g. vibration information for the input or bias for the output). The second step is aimed at selecting the right membership functions for these variables. The third step relates the output sets to the input sets through fuzzy rules. The way in which the rules are stated depends on the learning algorithm. Rules in this work are generated running an engine model. The choice of the right learning algorithm has a big impact on the accuracy of the fuzzy system.

Once the system architecture is defined, fuzzy inference is the process that computes the outcome provided an input to the system. There are two main types of inference methods known in the literature as Mamdani and Sugeno. A Mamdani-type inference is based on the fact that fuzzy sets are defined for inputs and outputs. Therefore, after the aggregation process there is a fuzzy set for each output variable that needs to be defuzzified.
On the other hand, a Sugeno-type system is based on the definition of the output MFs as single spike rather than a distributed fuzzy set, this is also known as singleton output membership function and can be considered as a pre-defuzzified fuzzy set. This improves the efficiency of the process simplifying the computation. The outcome is just the weighted average of a few data points. The GPD method developed in this work uses the Mamdani inference strategy.

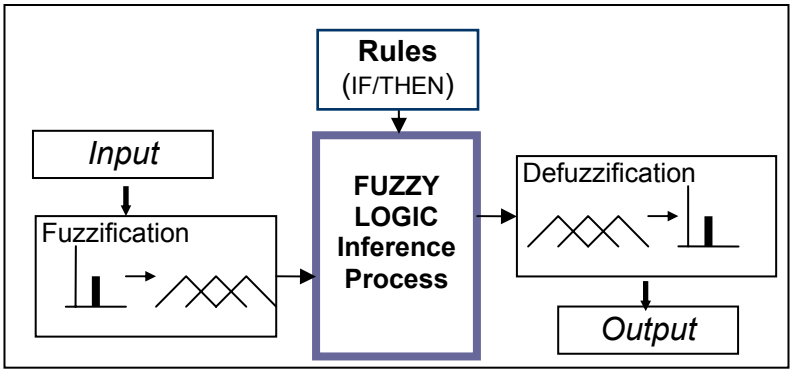


Figure 44: Configuration of a rule-based fuzzy logic system

A typical fuzzy logic system, as shown in Figure 44, involves fuzzification, rules evaluation and defuzzification:

- A fuzzifier turns numeric values (input measurements) into degree of activation of input MFs.

- An inference engine accumulates the effects of each rule on the output MFs, it includes logical operations, implication and aggregation phases.
- A defuzzifier calculates the outcome based on the activation of the output MFs.

The functionality of the fuzzy diagnostics system considering fuzzification, logical operations, implication, aggregation and defuzzification phases will be discussed in the next section through an example.

4.2.4 Fuzzy diagnostics system functionality via a graphical example

In this section a basic example of fuzzy logic system for engine diagnostics is provided according to the novel idea formulated in this work. It is not meant to be the description of a complete diagnostic system but only a simplified case to illustrate all the features of a fuzzy diagnostics system for a thorough understanding. A complete fuzzy logic diagnostic system developed and tested for the Trent 800 engine is then studied in section 4.5.

The example refers to a single shaft turbojet with 3 measurements (shaft speed N, fuel flow FF, and exhaust gas temperature EGT) and 4 gas-path performance parameters (compressor efficiency and flow capacity, and turbine efficiency and flow capacity) – see Figure 45.

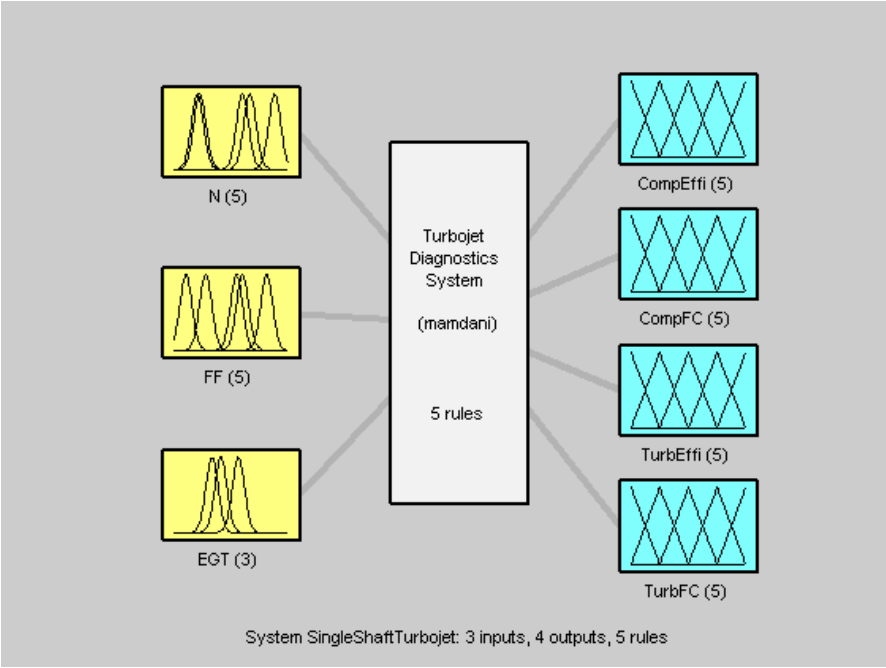


Figure 45: A simplified fuzzy-logic based diagnostic system.

We want to set-up a system that has all the features of a real fuzzy logic based diagnostic system but is limited in number of inputs, outputs and rules; in order to create an example suitable for the purpose of this section. To do that, we simply state 5 rules that relate the fault signature in the measurements to the changes in performance parameters at the following 5 conditions:

1. No changes in performance parameters, clean engine.

2. 1% drop in compressor efficiency
3. 1% drop in compressor flow capacity
4. 1% drop in turbine efficiency
5. 1% drop in turbine flow capacity

Where, the % delta changes are calculated from a clean engine baseline, computed at the same power setting and environmental condition. When the goal is to implement a complete diagnostic system, rules have to be created to approximate the whole defined search space. In other words, they have to cover a vast number of combinations of deteriorated conditions in terms of changes in efficiency and flow capacity that represent a sub-set of all the possible combinations. By no means, these 5 rules define such a sub-set; nevertheless, they are helpful to show the functionality of such a system.

In Figure 45, the information flows from left to right, from the 3 inputs to the 4 outputs. The parallel nature of the rules is one of the most important aspects of fuzzy logic systems. If enough rules are defined, the search space is approximated in regions where the system's behaviour is dominated by either a rule or another. In the case of the example the system is only able to choose among something close to the regions where the 5 rules are created. Hence, the task of this system is simple: being able to distinguish, in the presence of uncertainty, which of the 5 faulty conditions correspond to the fault signature in the input measurements. Obviously, only a limited number of sets of measurements are allowed as inputs.

4.2.4.1 System set-up and fuzzy rules' learning algorithm

The novelty of this work stands in the application of fuzzy logic to devise a gas-path diagnostics system. The procedure, according to which this was achieved, as well as its key concepts are described in this section.

Once the inputs and outputs are selected the set up procedure must define the system's architecture in all its elements and the fuzzy rules' learning algorithm. The rules, theoretically, can have many forms and each form is characterised by a learning algorithm used to tune the rules and therefore to make the fuzzy system's approximation capability effective.

Firstly, the range of variability of the outputs must be defined. This constitutes the constrained search space where the solution is looked for. In the case of engine diagnostics, assumptions will be made later regarding how the performance parameters vary and the type of ranges it is sensible to expect them to vary within. Once the search space is defined by the minimum and the maximum of each range, an increment value is defined to divide each range in a finite number of constant variations. For example, for the four performance parameters we define:

1. Compressor efficiency range: from -2% to 0 with steps of 0.5%.
2. Compressor flow capacity range: from -2% to 0 with steps of 0.5%.
3. Turbine efficiency range: from -2% to 0 with steps of 0.5%.
4. Turbine flow capacity range: from -2% to 0 with steps of 0.5%.

A number of combinations of variations of the performance parameter changes within the search space can be identified by varying them in their ranges with the increment values, first one at a time and then simultaneously. A system with SFI capability has up to 2 parameters changed at the same time.

For each output performance parameter, a number of levels of variation (i.e. fault levels) are therefore defined. In this example the fault levels for each parameters are -2, -1.5, -1, -0.5, 0. Hence, a number of output MFs can be defined centred in the values of each of these fault levels. For example, in Figure 46, 5 output MFs are defined for the compressor efficiency, in the range from -2 to 0%.

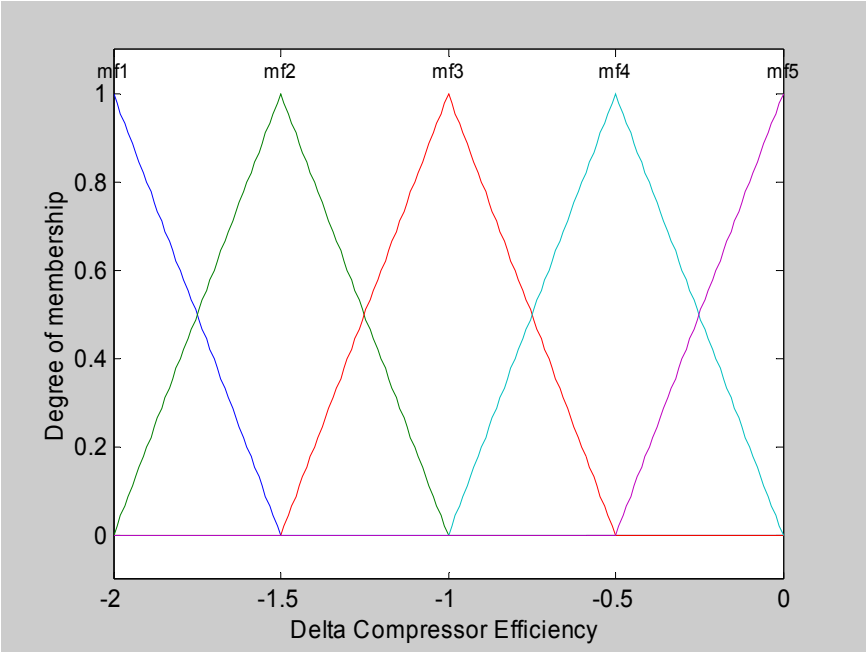


Figure 46: Example of 5 output MFs for the output variable Compressor Efficiency.

For each combination of variation of the performance parameters in the search space a fuzzy rule is created. A typical rule states that if the changes in the measurements are given values, than the corresponding changes in the performance parameters are other values that identify the deterioration.
Each combination of values of the search space defines a condition the engine model can be run at. The use of data obtained from the engine model to generate the rules preserves the non-linearity of the problem. Table 8 contains the output of the engine model in terms of % changes from the baseline condition for the three measurements N, FF and EGT corresponding to the 5 conditions considered in this example. As an example, the second rule can be read from left to right in the table: if ΔN=-0.983% and ΔFF=-1.234% and ΔEGT= 0% than $\Delta\eta_C$ = -1%, $\Delta\Gamma_C$= 0%, $\Delta\eta_C$= 0%, $\Delta\Gamma_C$=0%. The fault signature in the measurements corresponds to a deterioration level in the performance parameters.

Table 8: Engine model output for the 5 rules

Changes / Rules	% ΔN	% ΔFF	% ΔEGT	% $\Delta\eta_C$	% $\Delta\Gamma_C$	% $\Delta\eta_C$	% $\Delta\Gamma_C$
1	0	0	0	0	0	0	0
2	-0.983	-1.234	0.000	-1	0	0	0
3	0.530	-0.131	0.000	0	-1	0	0
4	-0.975	-0.805	0.347	0	0	-1	0
5	0.160	0.525	0.116	0	0	0	-1

In this example, we decided for simplicity to write only five rules that involve only individual changes in the output parameters equal to 0 or -1% and therefore only the output MFs centred in 0 and in -1% are used. Nevertheless, in a real system all the rules that involve all the combinations of variations of parameters in the search space should be stated.

Using the engine model, all the measurement variations, from a clean baseline, that correspond to all the combinations in the search space, are computed. For each measurement a MF is centred in the deviation values calculated (see Table 8). These values (centres of the MFs) are sorted from the minimum to the maximum. At this stage, a number of input membership functions is defined in the range of variability of each measurement and this number depends on the number m of rules generated – see Figure 47. In this example Gaussian MFs are used. The proecedure described below, was implemented within this work, to overcome the fact that the number of input MFs may be required to be less than the number of rules (expecially in the case of a MFI system). A maximum number of input MFs, indicated with K, is then chosen (based on the experience), and its value can be different from measurement to measurement. For the i-th input measurement, the sorted values of deviations (outcomes of the engine model for a number m of rules) are considered and if two values overlap one of them is discharged. The remaining values are counted, if they are, in number, less or equal to K (the maximum number of MFs required) one MF is centred in each of these values that at the most are m (Figure 47 shows the 5 input MFs for the variable N). Otherwise, the difference between each value and its consequent value, in the sorted list, is computed. The smallest value of difference between two measurement deviations is identified and these two values are substituted with their average value. A MF is then centred in this average value. This is repeated until the number of values that are centres of the input MFs is equal to K. Figure 50 shows how the first and second MFs are substituted with only one MF to meet the requirement of 4 input MFs for the variable N. The five input MFs for the variable FF are represented in Figure 48. In the case of EGT there are three overlapping outcomes of the engine model equal to zero, therefore according to the procedure only three MFs are defined shown in Figure 49.

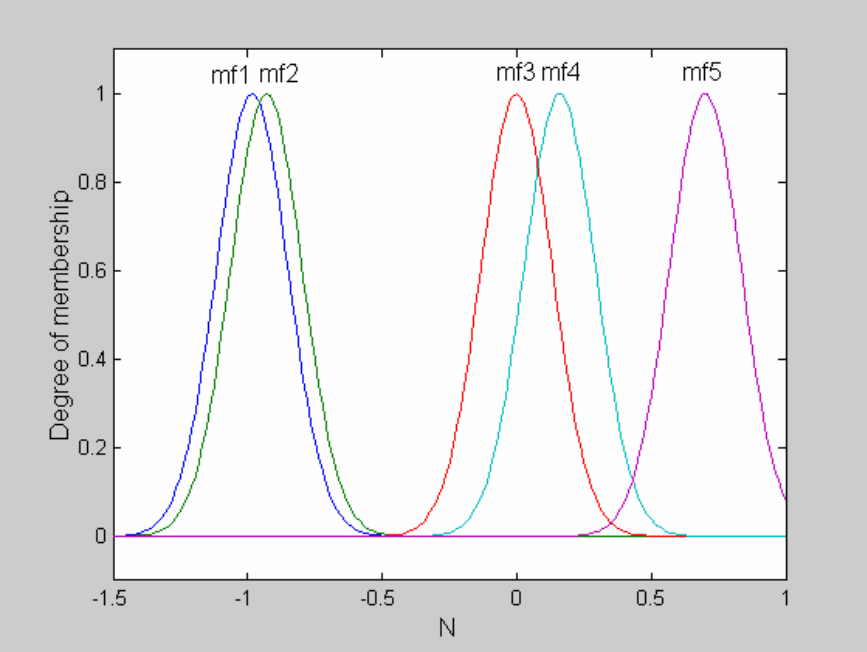


Figure 47: Example of 5 MFs for the input variable N.

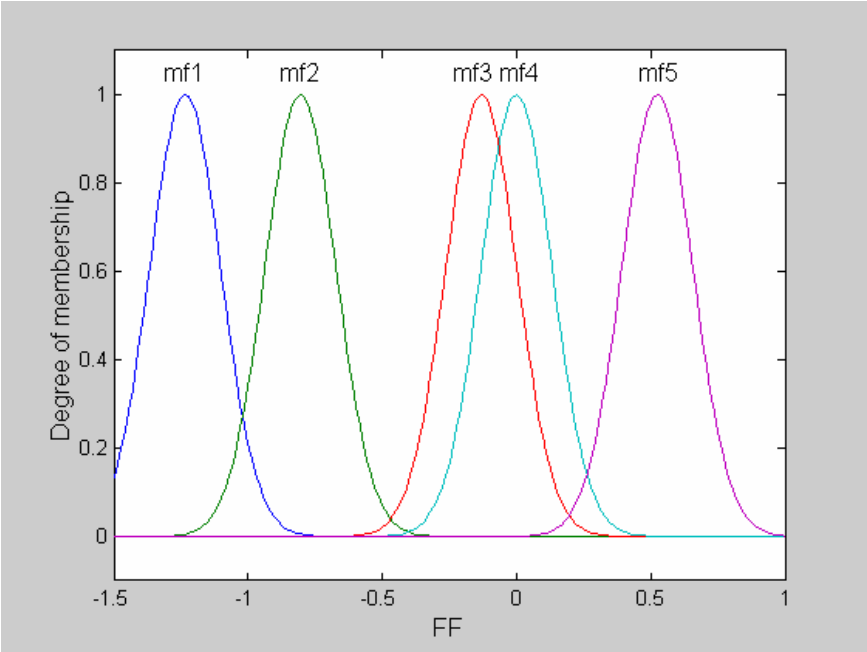


Figure 48: Example of 5 MFs for the input variable FF.

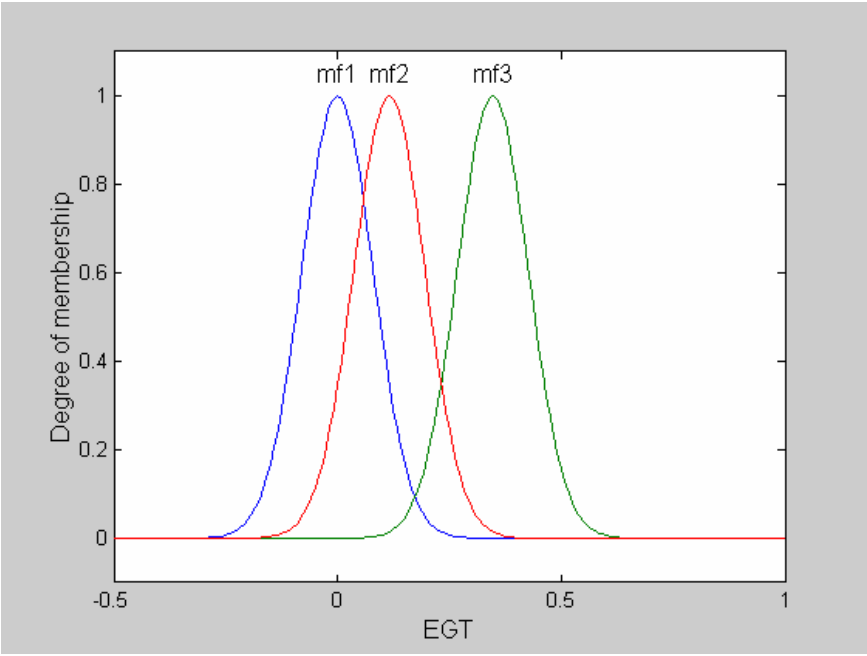


Figure 49: Example of 3 MFs for the input variable EGT.

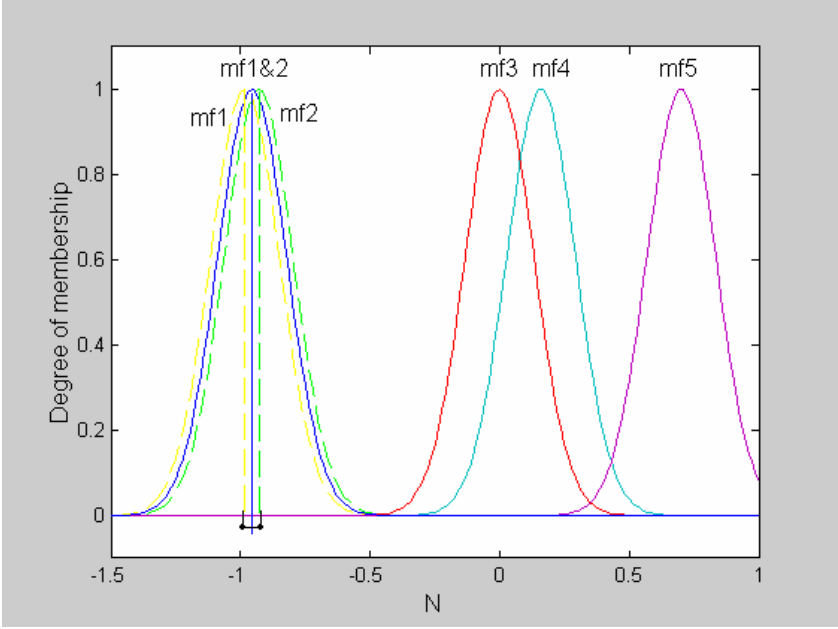


Figure 50: Example of 4 MFs for the input variable N.

Once the input and output MFs have been defined, the fuzzy rules can be stated. To state the fuzzy rules the procedure requires that each value in Table 8 is substituted with its corresponding MF. To achieve this, the measurements deviations of the table are associated (in the procedure used to define input MFs) with an input MF that is either centred in that value or centred in the mean value of a cluster of values grouped according to the procedure. On the other hand, the deviations in performance parameters of the table are always associated with a MF centred in that value.

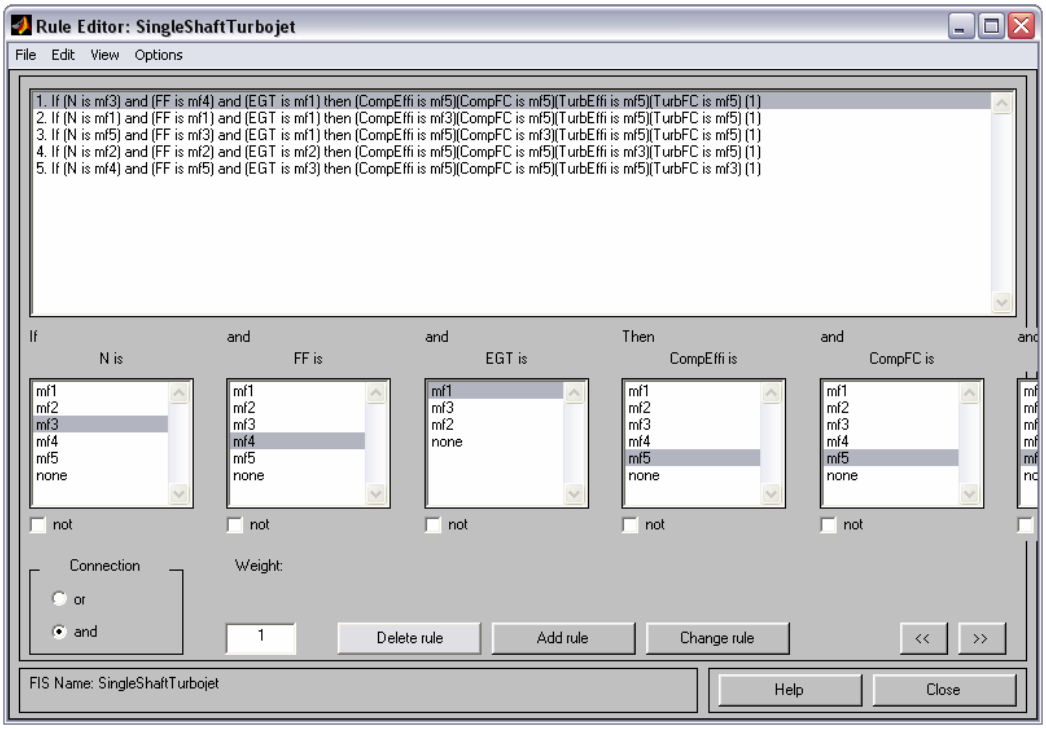


Figure 51: Example of fuzzy rules.

The fuzzy rules corresponding to Table 8 are stated in Figure 51. Note that a rule is divided in two parts: (i) the if-part that contains the fault signature in the measurements represented with MFs linked with the AND operator, and (ii) the then part that contains the MFs of the output performance parameters that characterise the fault condition.

4.2.4.2 Fuzzy inference

In the previous section the procedure to set-up a diagnostics system was described. When the system is set up, it can be used in inference mode. This means that a set of three measurements (N, FF and EGT) is taken along the gas-path of the engine or in this case is simulated with the engine model with an implanted fault. The deviations (%Δ) in the measurements are calculated from a clean engine baseline, computed at the same power setting and environmental conditions. These %Δ are given as input to the fuzzy logic system that performs fuzzy inference to calculate the output performance parameter % changes. The fuzzy inference process is performed through the following five phases: (i) fuzzification of the input variable, (ii) application of the fuzzy operator (i.e. AND) in the antecedent, (iii) implication from the antecedent to the consequent, (iv) aggregation of the consequents across the rules, and (v) defuzzification. These elements are defined in detail below as we step through each of them.

Figure 52 shows a graphical representation of the 5 rules. The left part illustrates the if-part with the pattern in the measurements characterised with MFs. On the right, the then-part part of the rules contains the output MFs that correspond to the faults. In the figure the input vector (to the bottom of the figure) is [ΔN,ΔFF,ΔEGT]=[0,0,0] , as expected the first rule is completely activated and the answer is no fault.

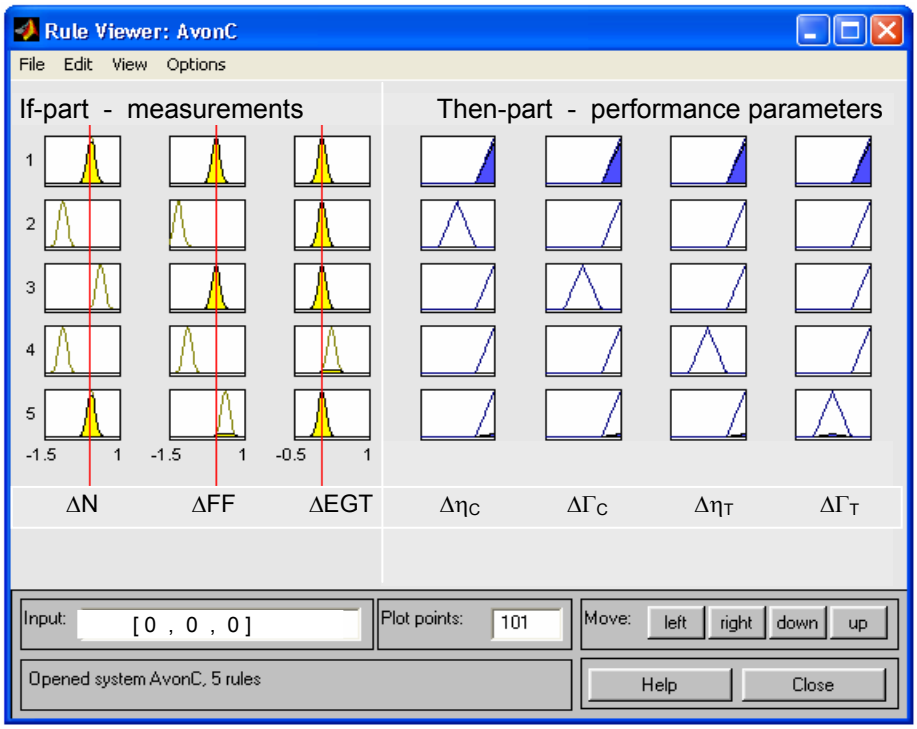


Figure 52: Graphical representation of the 5 rules

(i) Fuzzification

The fuzzification process turns numeric values (input measurements) into degrees of activation of the input MFs. The inputs are always numeric values and the outputs are fuzzy degrees of membership. Figure 53 shows how an input vector (in the example [ΔN,ΔFF,ΔEGT]=[-0.782, -1.05, 0.122]) has a degree of membership in the input MFs of the if-part of all the rules, always in the range 0,1. This is also refered to as degree of activation of the input MF.

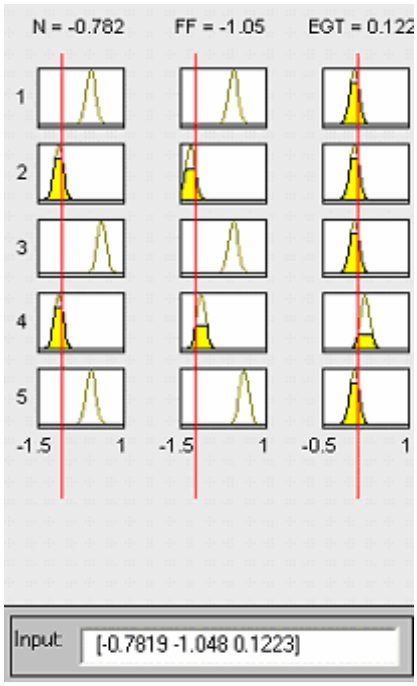


Figure 53: Fuzzification

(ii) Logical operators

With the fuzzification phase we assess the degree to which each part of the antecedent has been satisfied for each rule. In general the antecedent of the rule is constituted by more than one MF connected with a fuzzy logical operator, that in a diagnostics system is the AND operator. Therefore, in the example, the AND operator is applied to obtain one value that represents the result of the antecedent for each rule. The input to the logical operator is, in general, two or more degrees of activation from fuzzified inputs. The output is a single degree of activation.

In the example in Figure 54, for rule number 4, the AND operator can be implemented as minimum or as product of the values of d.o.a. resulting from the fuzzification. A comment that can be made is that using the minimum does not take into account all the three values, while the product does. On the other hand, the product tends to reduce the outcome d.o.a to pass to the consequent, and this affects defuzzification methods based on integration such as the centroid. In fact:

- If the values of d.o.a. are for example: [0.4, 0.5, 0.5] the min is 0.4 and the product is 0.1.
- If the values of d.o.a. are for example: [0.1, 0.9, 0.9] the min is 0.1 and the product is 0.081.

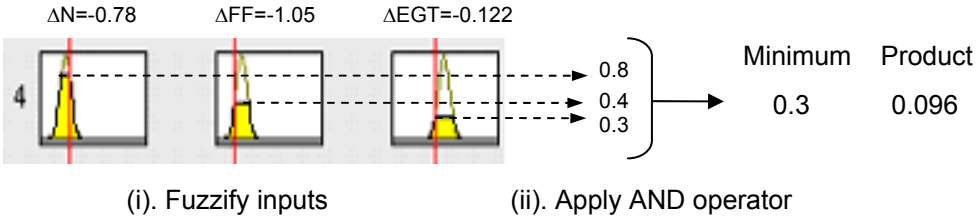


Figure 54: Apply AND operator

(iii) Implication

The output of the antecedent can be multiplied times a weight generally called rule's weight. This is used when there is the necessity of giving more weight to a rule than another. Although in the opinion of the author this feature could have high potentials in diagnostics, it was not used in this work; nevertheless some comments will be made.

In the implication phase, the MFs in the consequent are reshaped using the output form the antecedent. The input for the implication process is a number and the output is a fuzzy set. In this work two methods are used: (i) the minimum that truncates the MFs in the consequent to the level given by the antecedent, and (ii) the product which scales the output MFs according to that value. These two concepts are illustrated in Figure 55.

(iv) Aggregation

In a fuzzy system decisions are based on the activation of all the rules in parallel, therefore the rules must be combined in some manner in order to make a decision. Aggregation is the process by which the fuzzy sets that represent the outputs of each rule are combined into a single fuzzy set. Aggregation only occurs once for each output variable just prior to the fifth and final step, the defuzzification. The input of the aggregation process is the list of truncated

output functions returned by the implication process for each rule. The output of the aggregation process is one fuzzy set for each output variable. Notice that as long as the aggregation method is cumulative (which it always should be), then the order in which the rules are executed is unimportant. In this work two methods have been used: maximum and summation (this is simply the sum of each rule's output set). The two methods are graphically illustrated in Figure 56.

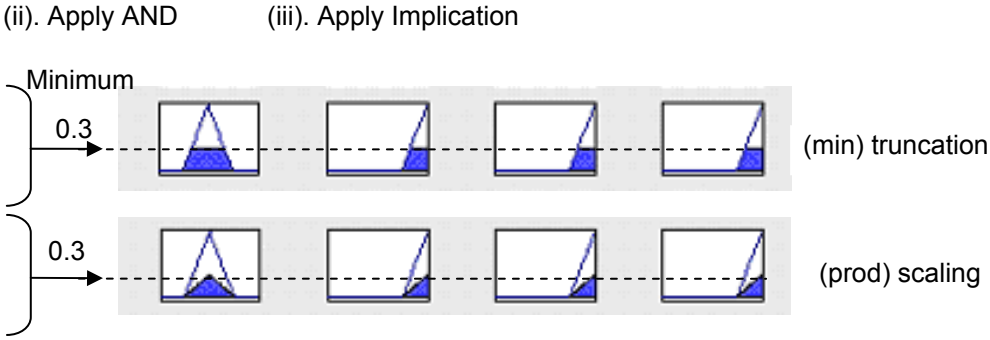


Figure 55: Implication

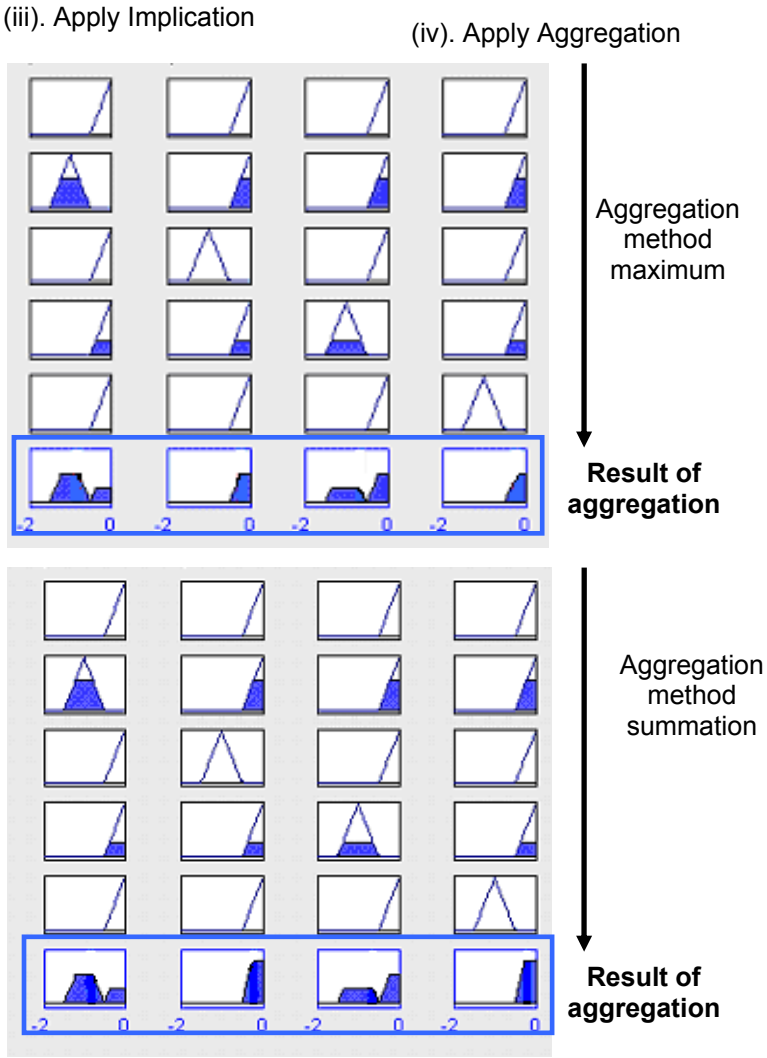


Figure 56: Aggregation

(v) Defuzzification

The defuzzification process takes in input the aggregated fuzzy set and computes a single number per output parameter. Defuzzification is the process which carries out arbitration between the possible outcomes of the inference process, based on the strength of activation of each of those outcomes. The process is carried out individually for each of the output variables, and combines the separate activations of the membership functions belonging to that variable into a single value. Among the different techniques available in the literature, two methods were implemented: (i) centroid function, and (ii) a method known as 'centre of maximum'.

With the centroid function an average of the centres of the membership functions weighted by their respective activations is calculated – see Figure 57 (a).

On the other hand, the centre of maximum scheme computes the numeric output value as the centre of the membership function that is most strongly activated. In the diagram in Figure 57 (b), the output value is shown as coinciding with the centre of the left-most membership function. Figure 58 illustrates the defuzzification phase for the example under study computed with the two techniques.

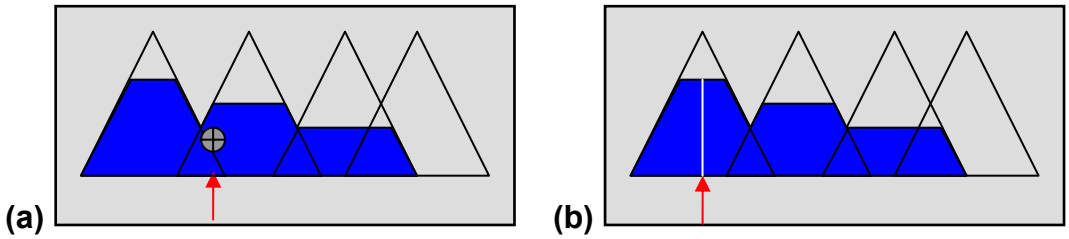


Figure 57: defuzzification with (a) centroid function, and (b) centre of maximum (COM) function

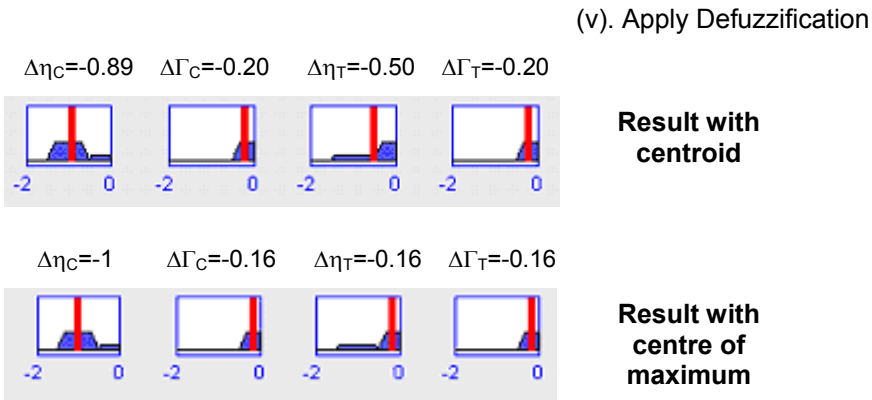


Figure 58: Defuzzification

4.2.4.3 Functionality considering the fuzzy inference diagram

In this section a fuzzy logic system was devised. Summarizing the system has 3 inputs (for each of them Gaussian MFs have been designed), 4 outputs (for each of them 5 Triangular MFs have been designed), and 5 rules. As shown in Figure 59, the capability of performing the 4 phases of the fuzzy inference process, that includes (i) AND operator, (ii) implication, (iii) aggregation, and (iv) defuzzification, by using different optional algorithms was implemented. The following options were made available:

- AND implemented as: minimum or product.
- Implication implemented as: minimum or product.
- Aggregation implemented as: maximum or summation.
- Defuzzification implemented as: centroid or centre of maximum functions.

Note that the OR operator has not been used. The most relevant combinations of choices for the 4 considered phases of the inference process have been implemented to realise the fuzzy inference diagrams shown between Figure 60 and Figure 66. For all the diagrams the input is kept fixed to [ΔN,ΔFF,ΔEGT]=[-0.782, -1.05, 0.122, but different outcomes are obtained because of the different system layouts.

A fuzzy inference diagram is the composite of all the partial diagrams that we have considered above. It simultaneously displays all parts of the fuzzy inference process that we have examined.

In the diagrams of Figure 60 and Figure 61 the defuzzification is obtained with centre of maximum function and the difference between the two is the AND operator that is implemented as product in Figure 60 and as minimum in Figure 61. Going to Figure 62 the AND is implemented as product but the centroid function is used for the defuzzification process. Note how the solution is also affected by less activated rules. This diagram can also be compared with the one in Figure 63 where the only parameter changed is the AND operator implemented as minimum. Although the product for the AND algorithm tends to reduce the outcome d.o.a., it takes into account all the input MFs' activation. Figure 63 can then be compared with Figure 64 because, in the first the aggregation method is performed with the maximum fuction and in the second with the summation. With the aggregation phase implemented as summation, one of the output MF that constitute the singular output fuzzy set can be activated by more than one rule, all the various activations are added up and therefore taken into account. On the other hand, if the maximum is used, only the most activated rule contributes to activate the various MFs in the singular output fuzzy set. Figure 65 and Figure 66 differ for the implication method. In the first diagram it is implemented by using the minimum function, truncating the MFs in the consequent to the level given by the antecedent, and in the second by using the product, scaling the output MFs according to that value.

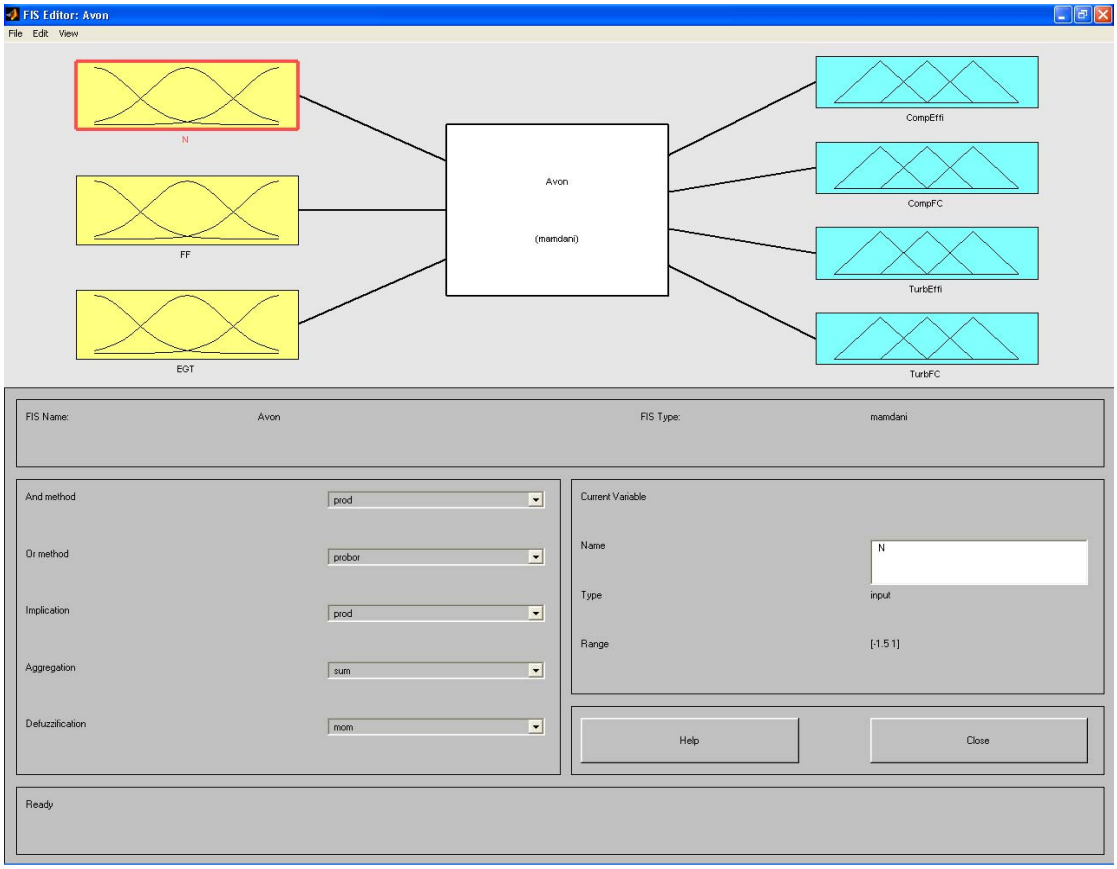


Figure 59: Example of system and the operators

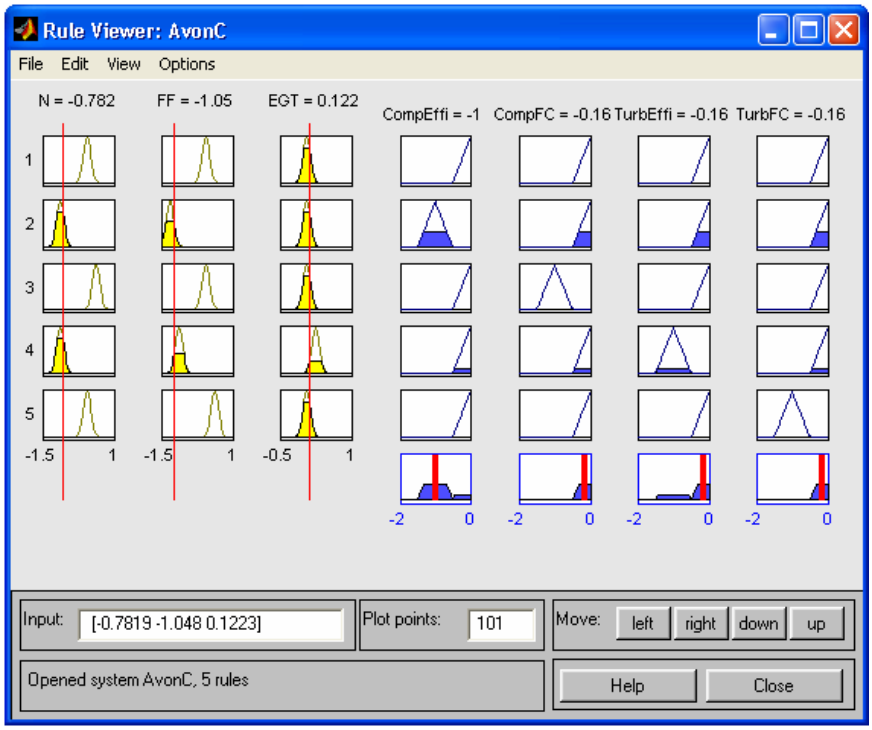


Figure 60: Fuzzy inference diagram when AND=product, implication=minimum, aggregation=maximum, defuzzification=centre of maximum

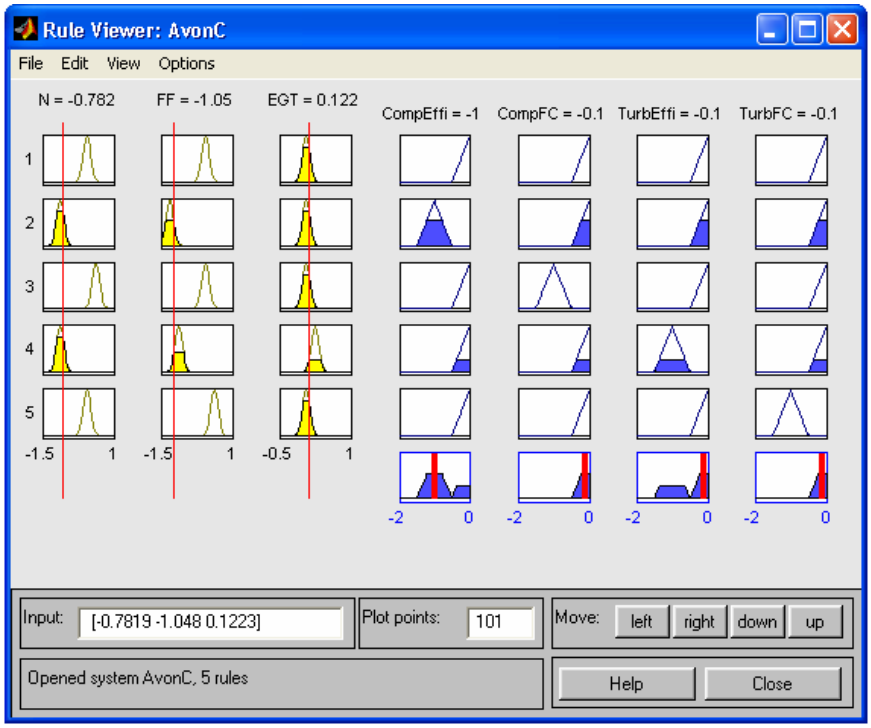


Figure 61: Fuzzy inference diagram when AND=minimum , implication=minimum, aggregation=maximum, defuzzification=centre of maximum

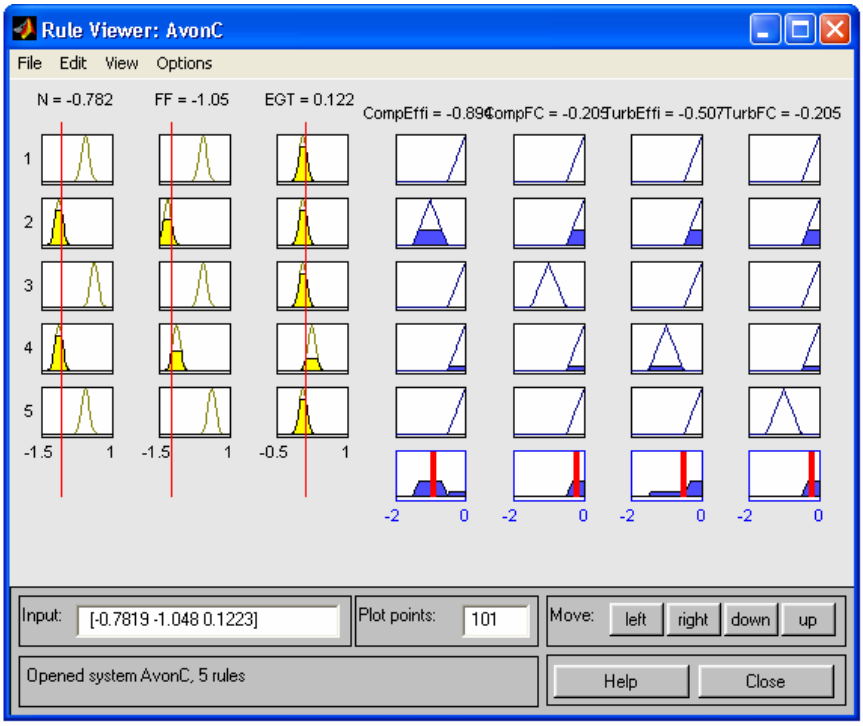


Figure 62: Fuzzy inference diagram when AND=product , implication=minimum, aggregation=maximum, defuzzification=centroid

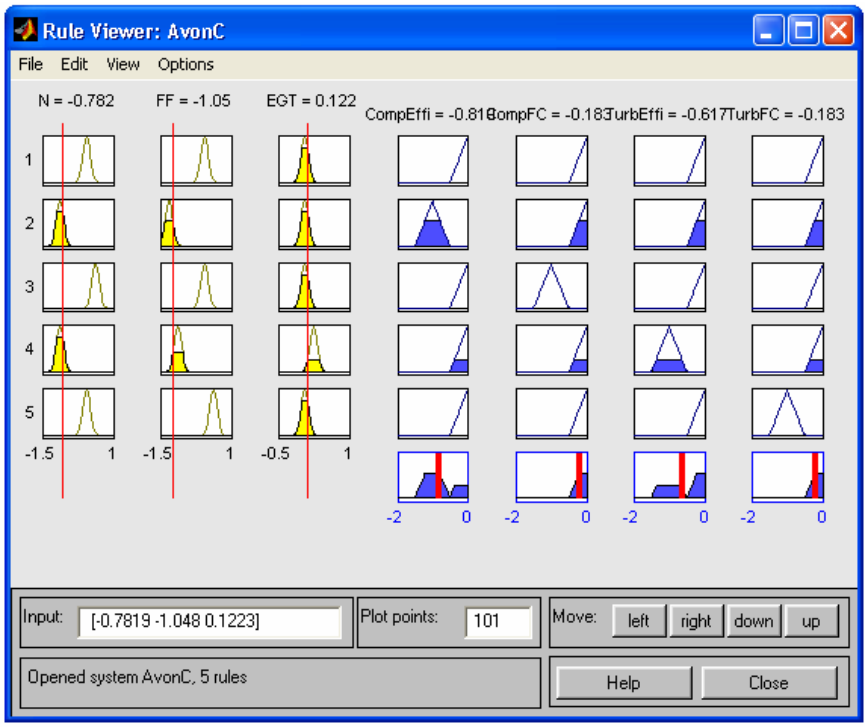


Figure 63: Fuzzy inference diagram when AND=minimum , implication=minimum, aggregation=maximum, defuzzification=centroid

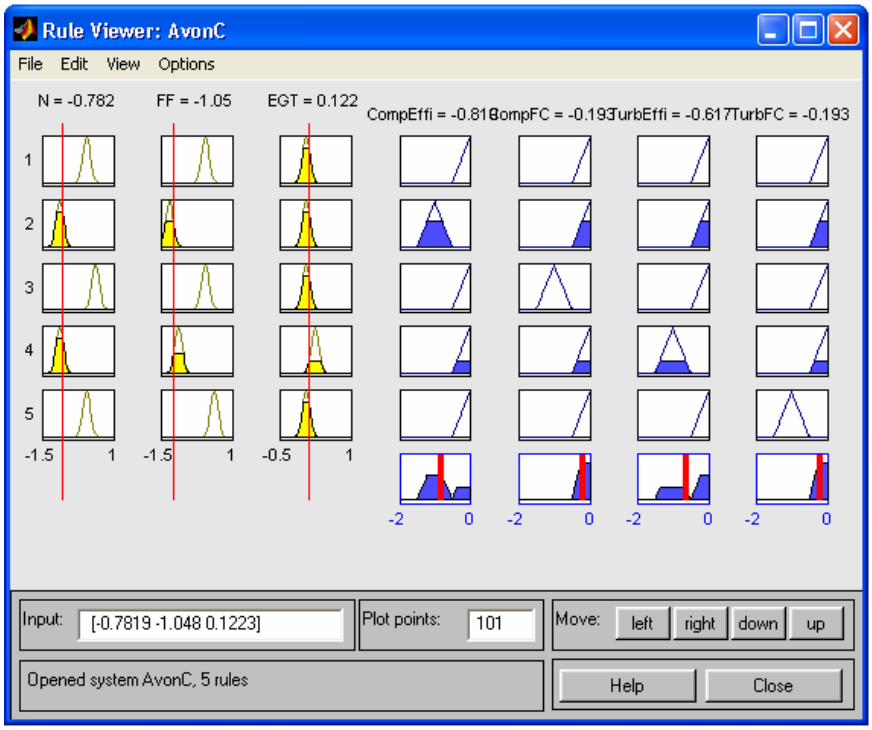


Figure 64: Fuzzy inference diagram when AND=minimum , implication=minimum, aggregation=sum, defuzzification=centroid

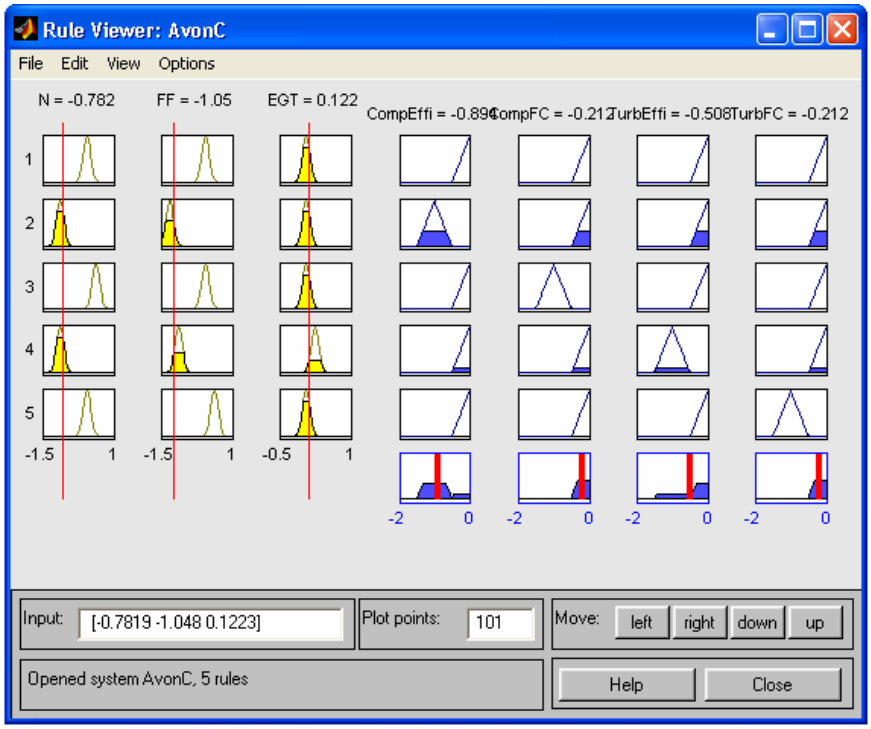


Figure 65: Fuzzy inference diagram when AND=product , implication=minimum, aggregation=sum, defuzzification=centroid

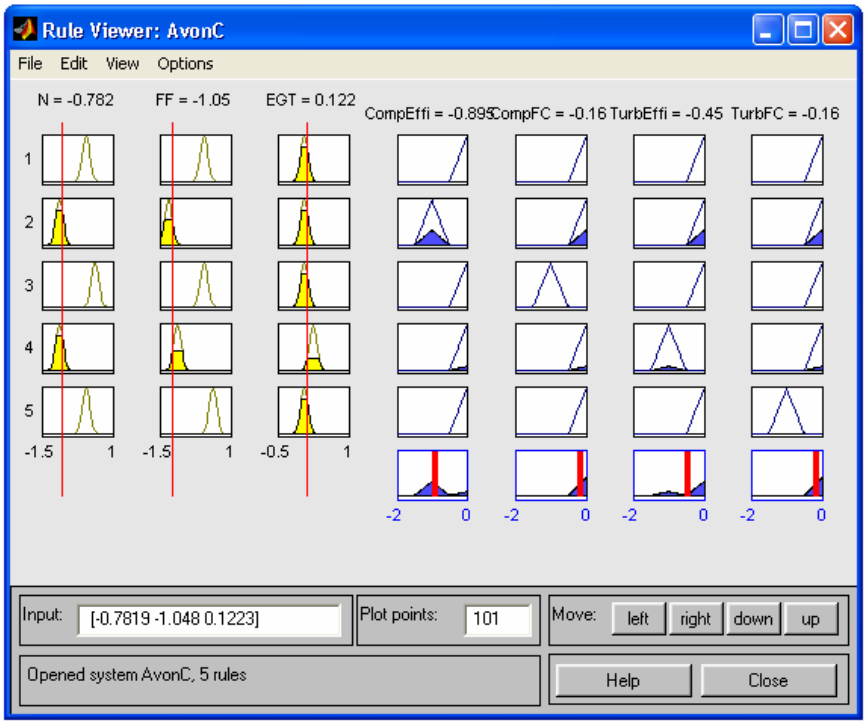


Figure 66: Fuzzy inference diagram when AND=product , implication=product, aggregation=sum, defuzzification=centroid

4.2.5 Algebraic form and conditional mean interpretation

This section is dedicated to the derivation and discussion of the algebraic formulation for a specific form of additive fuzzy systems known as standard additive model (SAM). A SAM has the following features:

- Based on Mandami-type inference.
- AND operator implemented as product.
- Implication method implemented as product.
- Aggregation method implemented as summation.
- Defuzzification as centroid function.

Parameters like number, type, width, density for input and output MFs as well as rules' definitions (number, type, density) and rules' weights are left as variables that can be used in order to optimize the system. Formulations similar to the one that follows for the SAM can be derived for all the systems described in the previous section characterised by different layouts – see Kosko (1997) .

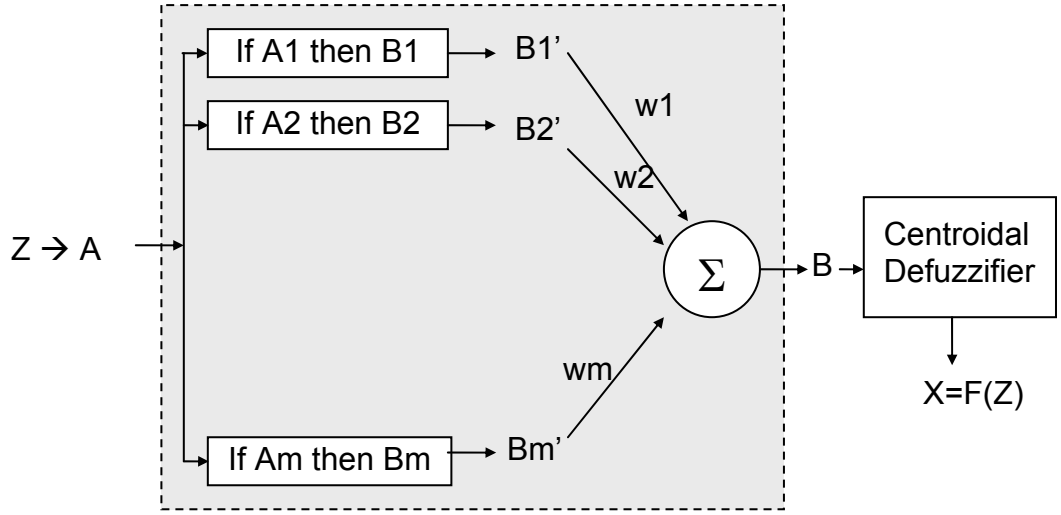


Figure 67: Additive fuzzy system architecture

A fuzzy system $F=R^n \rightarrow R^p$ uses m rules to map input vectors z into output vectors F(z) – see Figure 43. The if-part fuzzy set $A_j \subset R^n$, of the j-th rule, has a degree of activation a_j: $R^n \rightarrow [0,1]$. A diagnostic system, as discussed in the previous section, can calculate the if-part degree of activation a_j from n scalar degrees of activation $a_j^1, \ldots a_j^n$, by computing the AND operator as the minimum function, but at the expense of ignoring the correlations among the various elements of the input vector. Alternatively, product is considered one of the most powerful and light in computation ways to factor a joint if-part set function and therefore to implement the AND operator:

$$a_j(z) = \prod_{i=1}^{n} a_j^i(z_i) \qquad (4.\ 29)$$

It does not ignore n-1 of the fit values for each x as does the minimum combiner. The product of fit values gives a small number, nevertheless only the relative values matter in the SAM ratio below in equation (4. 34). The then-part fuzzy set $B_j \subset R^p$ has set function b_j: $R^p \rightarrow [0,1]$.

The “standard” part of a SAM is represented by the product $a_j x b_j(z,x) = a_j(z) b_j(x)$ to perform the implication in the inference procedure. The input z activates the j-th part of the rule to the degree $a_j(z)$ and, in the implication phase, this scales the then-part to $a_j(z)B_j$. In the aggregation phase the system sums the scaled then-parts, which represent the additive part in the SAM:

$$B(z) = \sum_{j=1}^{m} w_j a_j(z) B_j \subset R^p \tag{4. 30}$$

for scalar rule weights w_j. In the previous section, it was discussed how, on the other hand, the maximum can be computed instead of the summation in the aggregation phase.
Once the aggregation is performed, the output F(z) is computed by defuzzifying the result of the generalized fuzzy set B(z): F(z)=defuzzified (B(z)). The centroid and the centre of maximum have been considered in this work as defuzzification methods. The centroid uses all the information in the output set B(z) in a Bayesian sense, as it will be discussed by formulating the additive statistic theorem in the next section. This leads to the SAM theorem:

$$F(z) = \frac{\sum_{j=1}^{m} w_j a_j(z) V_j c_j}{\sum_{j=1}^{m} w_j a_j(z) V_j} \tag{4. 31}$$

with m bounded then-part volumes or areas:

$$V_j = \int_{R^p} b_j(x_1 \ldots x_p) dx_1 \ldots x_p > 0 \tag{4. 32}$$

and m then-part set centroids

$$c_j = \frac{\int_{R^p} x b_j(x_1 \ldots x_p) dx_1 \ldots x_p}{\int_{R^p} b_j(x_1 \ldots x_p) dx_1 \ldots x_p} \tag{4. 33}$$

In the case of weights all equal to one, the SAM theorem becomes:

$$F(z) = \frac{\sum_{j=1}^{m} a_j(z) V_j c_j}{\sum_{j=1}^{m} a_j(z) V_j} \tag{4. 34}$$

4.2.5.1 Fuzzy system as conditional mean

This section discusses a key concept in the application of fuzzy algebra to gas-path diagnostics: all centroidal fuzzy systems F compute a conditional mean, $F(z) = E[X|Z=z]$. An additive fuzzy system splits this global conditional mean into a convex sum of local (rule) conditional means.

The fuzzy system F approximates a process with the centroidal mean E[X|Z]. Nevertheless, to do so, it does not use a mathematical model of the process involved; it uses fuzzy rules: F is a model-free statistical estimator. This section discusses how this model freedom accounts for most of a fuzzy system's modeling power. The rules can have many forms and each form is characterized by a learning algorithm used to tune the rules. We have discussed the learning algorithm devised in this work in section 4.2.4.1.

Before proving that a centroidal fuzzy system computes a conditional means it is necessary to clarify the role of a fuzzy logic approach compared to a statistic one. The February 1994 issue of IEEE transaction on Fuzzy Systems (Kosko, 1994 and Laviolette et al., 1994) covers the subject of whether fuzziness differs from randomness in a statistical sense both at the set level and at the system level.

According to Kosko (1997), the set-level debate turns on how we use binary or multivalue sets to model events. Fuzziness deals with the degree to which an event occurs, while randomness deals with whether an event occurs. Standard probability theory maps only binary sets to real numbers.
We can view a fuzzy set as a random set (Nguyen, 1978) or locus of two-point conditional probability densities. Then, the set degree a(z) = degree($z \in A$) becomes the local conditional probability prob{Z=A|Z=z}. The complement fit value 1-a(z) = degree($z \notin A$) becomes the dual probability prob{ Z≠A|Z=z }. Suppose A is the subset of cool air temperatures – see Figure 38. On the global set view a(z) is the degree to which air temperature value z is a cool value. On the local random-set view a(z) is the probability that the temperature is cool given that the temperature value is z. We can equally view A as a locus of multivalued set values or as a locus of two-point conditional probabilities.
As far as the system-level debate is concerned, Kosko (1997) shows with simple arguments that all the centroidal fuzzy systems are probabilistic systems. This concept is illustrated below. A centroidal fuzzy system F computes a conditional expectation E[X|Z] and therefore a mean-squared optimal non-linear estimation (Papoulis, 1994). The power of fuzzy systems lies in both the optimality of their results and its model-free structure. Most popular conditional-mean systems use a mathematical model of the process. The linear Kalman filter uses a Gauss-Markov state model and assumes that all variables are jointly Gaussian. A fuzzy system does its mean-squared best to model a system or approximate a function with its rules or paired sets or densities.

The proof that centroidal fuzzy systems are conditional means follows from the ratio structure of the centroid and the boundedness and non-negativity of the set values b(z,x) of the combined set B of the activated then-part sets B_j in equation (4. 30). Each input z gives its own B(z) and thus its own output F(z):

$$F(z) = Centroid(B(z)) \tag{4. 35}$$

$$= \frac{\int_{R^p} xb(z,x)dx}{\int_{R^p} b(z,x)dx} \quad (4.\ 36)$$

$$= \int_{R^p} xp(x|z)dx \quad (4.\ 37)$$

$$= E[X|Z=z] \quad (4.\ 38)$$

for each $z \in R^n$. This holds because the ratio in equation (4. 36) of the joint distribution to the marginal defines a proper conditional probability density:

$$p(x|z) = \frac{b(z,x)}{\int_{R^p} b(z,x)dx} \quad (4.\ 39)$$

even though b(z,x)>0 may hold.

According to the SAM theorem (4. 31), if F is additive and SAM in structure, then F(z) is a convex sum of centroids or local conditional (then-parts set) means. We now show that this convexity property holds for all additive systems and that the same convex weights decompose the conditional variance of the system. The conditional variance gives a confidence measure for each output F(z) and shows that the structure of the then-part sets B_j matters after all.

In practice we can compute the local conditional means or centroids in advance but not so with the local conditional variances. The result holds for general additive mapping F: $R^n \rightarrow R^p$ that combines activated then-part sets with the sum combiner. We provide here the formulation and proof of the theorem for the scalar case F: $R^n \rightarrow R$ for simplicity and to avoid the matrix notation for the covariances. This formulation is known as the Additive Statistic Theorem.

Suppose F: $R^n \rightarrow R$ is an additive fuzzy system such that F(z) =centroid(B) and $B(z) = \sum_{j=1}^{m} w_j B_j'(z) = \sum_{j=1}^{m} w_j a_j(z) B_j$. Then:

$$F(z) = E[X|Z=z] = \sum_{j=1}^{m} p_j(z) E_{B_j'}[X|Z=z] \quad (4.\ 40)$$

$$V[X|Z=z] = \sum_{j=1}^{m} p_j(z) V[X|Z=z, B_j'] \quad (4.\ 41)$$

The convex coefficients $p_j(z)$ are weighted volume ratios of the activated sets B_j':

$$p_j(z) = \frac{w_j V_j^{'}(z)}{\sum w_k V_k^{'}(z)} \tag{4. 42}$$

$$V_j^{'}(z) = \int b_j^{'}(z,x)dx \tag{4. 43}$$

Proof. The chain of equalities, repeated below, shows that the system F computes a conditional expectation for each input z. Besides it shows that F(z) is a convex sum of local conditional mean realizations or centroids.

$$F(z) = Centroid(B(z)) \tag{4. 44}$$

$$= \frac{\int_{-\infty}^{\infty} xb(z,x)dx}{\int_{-\infty}^{\infty} b(z,x)dx} \tag{4. 45}$$

$$= \int_{-\infty}^{\infty} x p_B(x|z)dx \tag{4. 46}$$

$$= E[X|Z=z] \tag{4. 47}$$

$$= \frac{\sum_{j=1}^{m} w_j \int_{-\infty}^{\infty} xb_j^{'}(z,x)dx}{\sum_{j=1}^{m} w_j \int_{-\infty}^{\infty} b_j^{'}(z,x)dx} \tag{4. 48}$$

Multiplying and dividing by the same quantity:

$$= \frac{\sum_{j=1}^{m} w_j \int_{-\infty}^{\infty} b_j^{'}(z,x)dx \frac{\int_{-\infty}^{\infty} xb_j^{'}(z,x)dx}{\int_{-\infty}^{\infty} b_j^{'}(z,x)dx}}{\sum_{j=1}^{m} w_j \int_{-\infty}^{\infty} b_j^{'}(z,x)dx} \tag{4. 49}$$

$$= \frac{\sum_{j=1}^{m} w_j V_j^{'} \int_{-\infty}^{\infty} xb_j^{'}(z|x)dx}{\sum_{j=1}^{m} w_j V_j^{'}} \tag{4. 50}$$

$$= \sum_{j=1}^{m} p_j(z) E_{B_j'}[X|Z=z] \tag{4. 51}$$

$$= \sum_{j=1}^{m} p_j(z) c_j'(z) \tag{4. 52}$$

which proves (4. 40). As far as the variance (covariance) is concerned, it can be proved in the same way:

$$V[X|Z=z] = \frac{\int_{-\infty}^{\infty} (x - E[X|Z=z])^2 b(z,x) dx}{\int_{-\infty}^{\infty} b(z,x) dx} \tag{4. 53}$$

$$= \sum_{j=1}^{m} p_j(z) \int_{-\infty}^{\infty} (x - E[X|Z=z])^2 p_{B_j'}(x|z) dx \tag{4. 54}$$

$$V[X|Z=z] = \sum_{j=1}^{m} p_j(z) V[X|Z=z, B_j'] \tag{4. 55}$$

As we can see, the additive statistic theorem shows that not all additive outputs F(z) are the same. Some outputs are more certain than others. The most certain output F(z) = c_j' occurs when $p_j(z)=1$ and thus when the system does not interpolate between then part centroids. The spread or dispersion about these centroids leads to system uncertainty. The fuzzy system pays for its interpolation in increased uncertainty.

Concluding, we want to remark the impact of the fact that all fuzzy system outputs do not have the same uncertainty status. Consider two systems with m rules each. They have the same sets and rules except for one difference. The then-part of the first system has less variance than the then-part of the second even though they have the same centroids and unity volumes V_j. In other words, both systems define the same mapping and have the same first order statistics, but they have different variance and therefore accuracy. This shows that the choice of the MFs can have a heavy impact on the accuracy of the system. Besides, this uncertainty view comports with the view that the tendency to reduce the number of rules in approximating a process or function decreases the accuracy and gives a rougher function approximation.

4.2.6 Comments on fuzzy rules for a diagnostics system

A distinction was made (Volponi, 2003) between techniques more suitable for estimating (i) gradual deteriorations and others for (ii) rapid deteriorations (i.e. faults) as discussed in section 2.1.5. We referred to such methods as MFI (multiple-fault isolation) and SFI (single-fault isolation) respectively. The former implies that all the engine components (whose shifts in performance we are

estimating) deteriorate slowly whereas the latter implies a rapid trend shift probably due to a single entity (or perhaps two) going awry. AI-based methods such as fuzzy logic systems are more suitable for SFI problems, because they are based on an approximation of all the possible solutions for the limited number of combinations used to train the system. The extension to all the possible combinations (even in a limited search-space) is theoretically possible, but extremely burdensome from a computational point of view. In this project a fuzzy logic diagnostic system was firstly set up to secure an effective SFI capability. Then a partial MFI capability was tested considering the simultaneous deterioration of up to 4 health parameters (two components).

The number of necessary fuzzy rules grows exponentially with the number of system variables. Any attempt to reduce the number of rules is inevitably associated with less precise approximation capability. In general, we must trade some accuracy for ease of computation.

In this work a diagnostic system for the three-shaft Trent 800 turbofan was developed – see section 4.5. The 6 gas-path components investigated are: FAN, intermediate pressure compressor (IPC), high pressure compressor (HPC), high pressure turbine (HPT), intermediate pressure turbine (IPT) and low pressure turbine (LPT) – see second column of Table 9. When these 6 components are considered for GPD, the number of possible combinations C of components degraded can be calculated with the following equations:

$$C = \frac{n!}{k!(n-k)!} \tag{4. 56}$$

this gives the number of combinations of n=6 components taken k at a time. According to equation (4. 56), all the possible combinations are listed in Table 9.

Table 9: Combinations C of 6 gas-path components taken k at a time

k / C	1 at a time	2 at a time	3 at a time	4 at a time	5 at a time	6 at a time
1	FAN	FAN - IPC	FAN - IPC - HPC	FAN - IPC - HPC- HPT	FAN - IPC - HPC- HPT-IPT	FAN-IPC-HPC-HPT-IPT-LPT
2	IPC	FAN - HPC	FAN - IPC - HPT	FAN - IPC - HPC- IPT	FAN - IPC - HPC- HPT-LPT	
3	HPC	FAN - HPT	FAN - IPC - IPT	FAN - IPC - HPC - LPT	FAN - IPC - HPC- IPT-LPT	
4	HPT	FAN - IPT	FAN - IPC - LPT	FAN - IPC - HPT- IPT	FAN - IPC - HPT- IPT-LPT	
5	IPT	FAN - LPT	FAN - HPC- HPT	FAN - IPC - HPT - LPT	FAN - HPC – HPT- IPT-LPT	
6	LPT	IPC - HPC	FAN - HPC - IPT	FAN - IPC - IPT - LPT	IPC - HPC- HPT-IPT-LPT	
7		IPC - HPT	FAN - HPC - LPT	FAN - HPC - HPT - IPT		
8		IPC - IPT	FAN - HPT - IPT	FAN - HPC - HPT- LPT		
9		IPC - LPT	FAN - HPT - LPT	FAN - HPC - IPT - LPT		
10		HPC - HPT	FAN - IPT - LPT	FAN - HPT - IPT - LPT		
11		HPC - IPT	IPC - HPC - HPT	IPC - HPC- HPT - IPT		
12		HPC - LPT	IPC - HPC- IPT	IPC - HPC - HPT - LPT		
13		HPT - IPT	IPC - HPC - LPT	IPC - HPC- IPT - LPT		
14		HPT - LPT	IPC - HPT- IPT	IPC - HPT - IPT - LPT		
15		IPT - LPT	IPC - HPT - LPT	HPC - HPT - IPT - LPT		
16			IPC - IPT - LPT			
17			HPC - HPT - IPT			
18			HPC - HPT - LPT			
19			HPC - IPT - LPT			
20						

Considering that the number of parameters representative of the health of the each component are always 2, 2k is the number of parameters deteriorated simultaneously in each rule (each run of the engine model) when we simulate k degraded components at a time.

For example if 2 degraded components at a time are simulated, 4 parameters are changed in the generation of each rule.

On the other hand, the equation:

$$N = f^{g} = f^{2k} \tag{4. 57}$$

computes the number of permutations of f (=3 in the example of Table 10) fault levels (e.g. 0, 1, 2 % change in performance parameters) taken g=2k (=4 in Table 10) at a time with repletion. The parameter g=2k represents the number of parameters changed at a time. In the case of Table 10 the number of permutations with repetitions are $N=f^{2k}=3^4=81$. As we have six components, we have C=15 combinations of 2 components (and 4 parameters) taken a time: the final number of rules to generate in this example would be the product TotalCombinations=C·N=15·81=1215.

Table 10: Example of 4 deteriorated parameters at a time

η_i	Γ_i	η_j	Γ_j
0	0	0	0
1	0	0	0
2	0	0	0
0	1	0	0
1	1	0	0
2	1	0	0
..	..	..	..

Summarizing, the number of TotalCombinations for a three spool engine, with six gas-path components, and so the number of rules that are accordingly generated, is given by:

$$TotalCombinations = C \cdot N = f^{2 \cdot k} \cdot \frac{6!}{k!(6-k)!} \tag{4. 58}$$

where k is the number of degraded components simulated at a time and f is the number of fault levels, as performance parameters percentage changes from the clean engine.

Given six components and two health parameters per component, we have 12 performance parameters (η and Γ of the components). Therefore, we define the search space as the 12-dimensional space of the ranges of variability of the 12 parameters in % changes from the clean value. The solution of the diagnostic problem will be looked for within the constrained search space.
The learning algorithm devised in this work builds the fuzzy-logic-based diagnostic system with a number of rules equal to TotalCombinations as defined above, noting that there is no justification to omit some combinations if the

purpose is to approximate the dependency between measurements and performance parameters when the latter vary in a given search space. Nevertheless, the values of the f fault levels can either be chosen as uniformly distributed in the ranges of the search space or not. This work is dedicated to the study of fuzzy systems with uniformly distributed fuzzy rules, so the density of the fuzzy rules is left unchanged through a given search space, though it is varied from system to system to trade accuracy towards computational burden as discussed before.

4.2.6.1 Fuzzy systems and neural networks

A last comment can be made about the strong analogy that exists between fuzzy systems and neural networks. Neural networks, as fuzzy systems, can approximate a function or process that represents a relation of cause and effect and can act as universal approximators. A neural network, instead of stating rules, trains its synapses. The numerical synaptic values change when input data make the neurons fire. This enables nets to learn to recognise patterns and therefore to map inputs into outputs. The major difference is that, in the case of a neural network, a user has no way to know what the net has learnt or forgotten during the learning process. When the network is trained with new information there is an inevitable tendency to forget the old ones. On the other hand, fuzzy rules are modular and the user can always put them in or take them out at will.

4.3 The Trent 800 engine and Instrumentation

The engine involved with this research project is the Rolls-Royce Trent 800, a detailed description of the engine is provided in Appendix A. Figure 68 shows its three shaft configuration highlighting the typical sensor locations.

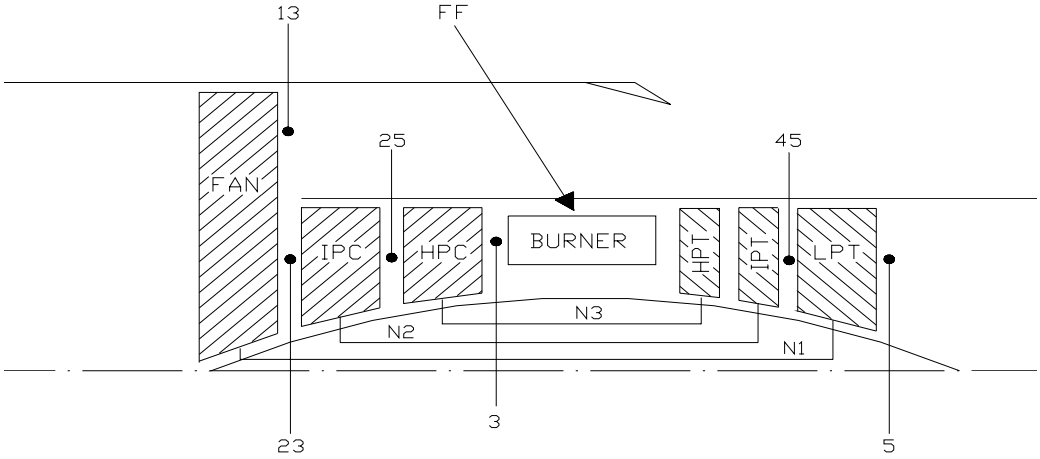


Figure 68: Three Shaft Turbofan engine Configuration.

The set of measurements available for the diagnostics process within this project is listed in Table 11, while the power setting and environmental parameters are reported in Table 12. Note that from now onwards the simbles in Table 11 will be used to indicate percentage changes in the measurements in order to simplify the nomenclature avoiding the simble Δ for the measurements. Besides, it will be clear where the text referes to changes in measurements and not absolute values from the fact that the simble Δ will be used to indicate the corresponding percentage changes in performance parameters.

Table 11: Measurement set

1	**N2 :**	IP Shaft Speed
2	**N3 :**	HP Shaft Speed
3	**FF :**	Fuel Flow
4	**P13 :**	FAN tip exit Total Pressure
5	**P25 :**	HPC entry Total Pressure
6	**P3 :**	HPC exit Total Pressure
7	**T25 :**	HPC entry Total Temperature
8	**T3 :**	HPC exit Total Temperature
9	**T45 :**	IPT exit Total Temperature
10	**T5 :**	LPT exit Total Temperature

Table 12: Power setting and environmental parameters

1	**N1 :**	LP Shaft Speed
2	**M :**	Mach Number
3	**Z :**	Altitude

In the simulation undertaken in this work, sensor noise is assumed to follow a normal distribution. The normal distribution standard deviation, given in terms of percentage deviation from the nominal value of each measurement, has ben

used to represent the noise level. Accurate values of standard deviations are provided by the sensor manufacturers but for the scope of this project the sensor noise standard deviations listed in Table 13 are considered sufficiently accurate and realistic. The performance simulations are carried out mainly using Turbomatch a steady-state performance simulation code developed at Cranfield University and described in Appendix B. The simulations are performed at the flight condition of 10000 m of altitude, 0.85 Mach and 0.8%PCN1 (which identifies the percentage of N1 respect the design point condition).

Table 13: Sensor noise standard deviations in % of the measured value

SENSOR TYPE	$STDV_i$
Temperature	0.4%
Pressure	0.25%
Fuel Flow	0.5%
Shaft Speed	0.05%

4.4 Observability study

4.4.1 Introduction

Section 4.3 provided a description of the engine involved within this research: the Rolls-Royce Trent 800. According to the requirements of this project a diagnostics process was developed to assess the changes in the 12 performance parameters listed in Table 14 from the knowledge of 10 measured parameters listed in Table 11 and the power setting and environmental parameters listed in Table 12.

Table 14: Changes in performance parameters sought with the diagnostics process

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$

Independently from the method used to perform diagnostics, the outcome is inevitably affected by the undetermined mathematical formulation due to the choice of number and type of inputs and outputs. An observability study was carried out in order to advise the diagnostics process with the inevitable correlations that will affect the diagnostics capability (lack of observability). The outcomes of this study are reported below.

The concept of observability in the study of linear systems has a very precise meaning, as first introduced by Kalman in 1960. A system Z=h(X) is said to be completely observable if every state X (vector) can be determined from the observation of Z (vector). The concept of observability is useful in gas-path diagnostics, which is aimed at reconstructing unmeasurable state variables from measurable ones. Several methods are available to test the observability of a linear system. The most common method studies the invertibility of the System Matrix (see section 2.2) but this produces a “yes/no” type of answer that is not useful for our purpose. A more effective approach was proposed by Provost (1995) aimed at quantifying the system observability. A development of that approach is under investigation at Cranfield as discussed in section 4.4.5.

The next sections describe the outcome of the observability study performed within this project. Two tests were performed to check whether the System Matrix is observable (with a yes/no type of answer) and to highlight the correlations that affect the diagnostics results – see sections 4.4.3 and 4.4.4. The main aim of the tests was to evaluate the correlations that exist among the measurements and among the performance parameters, in the specific diagnostics problem formulated above – see section 4.3 for number and type of inputs and outputs. A different approach is described in section 4.4.5: two procedures are described that provide a more complete evaluation of the quality of the observability (strong or weak observability). The HMP 1.1 observability module (GUI) that includes all the procedures described in this thesis is shown in Figure 69.

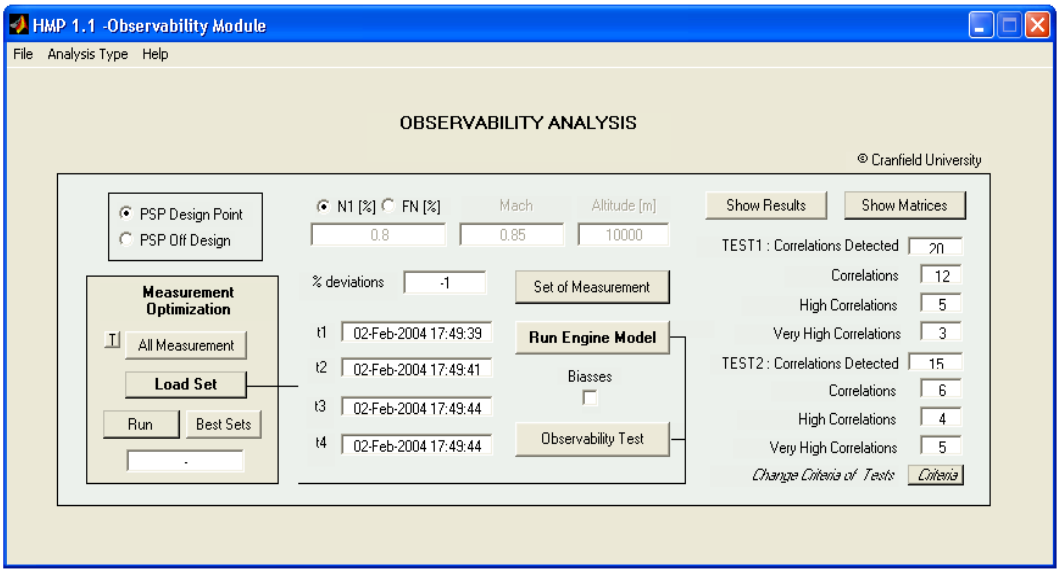


Figure 69: Graphical User Interface of the Observability Module of the software HMP 1.1.

4.4.2 Exchange rate table and system matrix

The observability study methods described below are all based on a first linearization of the functional relationship between measurements and performance parameters as discussed in section 2.2. Having chosen an engine operating condition – see Table 15 – the values of the available measurements are calculated with the engine model and they are listed in Table 16.

Table 15: Operating condition

N1	Mach	Altitude
0.8%	0.85	10000 m

Table 16: Clean engine condition (N1=0.8% , Mach= 0.85, Altitude=10000 m)

N2 [%rpm]	N3 [%rpm]	FF [kg/s]	P13 [atm]	P25 [atm]	P3 [atm]	T25 [K]	T3 [K]	T45 [K]	T5 [K]
0.8155	0.8585	0.8101	0.6209	2.8961	10.964	470.23	702.24	860.48	602.65

The Exchange Rate Table, according to the procedure described in section 2.2 is determined calculating the % change in each gas-path measurement for a 1% change in each component performance parameter (or unit change if the datum value is zero) in turn, at the flight conditions and power level at which the analysis is to be done. The Exchange Rate Table calculated at the operating condition of Table 15 is shown in Table 17. The associated Environmental ERT and the Identity matrices (see section 2.2) are listed respectively in Table 18 and Table 19. These three matrices constitute the 10 by 25 System Matrix.

Table 17: ERT Matrix (N1=0.8%, Mach= 0.85, Altitude=10000 m), 10 by 12

	1% $\Delta\eta_{FAN}$	1% $\Delta\Gamma_{FAN}$	1% $\Delta\eta_{IPC}$	1% $\Delta\Gamma_{IPC}$	1% $\Delta\eta_{HPC}$	1% $\Delta\Gamma_{HPC}$	1% $\Delta\eta_{HPT}$	1% $\Delta\Gamma_{HPT}$	1% $\Delta\eta_{IPT}$	1% $\Delta\Gamma_{IPT}$	1% $\Delta\eta_{LPT}$	1% $\Delta\Gamma_{LPT}$
N2	0.3801	-0.5297	-0.3262	0.5175	-0.2146	-0.0773	-0.2869	-0.1557	-0.3850	-0.2011	0.1803	0.5028
N3	0.0800	-0.4321	0.1130	-0.1351	-0.6593	0.4123	-0.8666	-0.5253	-0.2586	0.4345	0.0082	0.0850
FF	1.1838	-2.0516	0.5234	-0.0049	0.4148	0.1691	0.5456	0.3160	0.6086	0.2296	1.6183	0.2123
P13	0.0000	-0.5749	0.0467	0.0000	0.0387	0.0145	0.0515	0.0290	0.0564	0.0193	-0.0258	-0.0709
P25	0.7769	-1.0255	-0.3660	0.2624	0.7286	0.3522	0.9565	0.5663	-0.4040	-1.2603	0.4661	1.1118
P3	0.8209	-1.4137	-0.1459	0.1095	-0.0365	0.1277	0.0730	-0.7297	-0.2007	-0.1186	0.7661	0.8026
T25	0.3552	-0.2658	0.2914	0.0808	0.2339	0.1106	0.3105	0.1808	-0.0936	-0.3700	0.1425	0.3360
T3	0.3460	-0.3774	0.3346	0.0256	0.3019	0.0627	0.0085	-0.2150	-0.0441	-0.0185	0.2122	0.2307
T45	0.4591	-0.5125	0.6578	-0.0012	0.5358	0.2127	0.7101	0.3998	0.7717	0.2882	0.7903	-0.3893
T5	0.4646	-0.4978	0.7285	-0.0100	0.5874	0.2356	0.7766	0.4430	0.8529	0.3203	1.3623	-0.1809

Table 18: Environmental Exchange Rate Table (or Slope Table), 10 by 3

	N1	Z	M
N2	-0.8584	-0.0331	-0.0147
N3	-0.6954	-0.1771	-0.0711
FF	-3.4724	-0.4197	1.2665
P13	-0.8921	-0.7617	1.4139
P25	-1.7057	-0.4178	1.4192
P3	-2.3258	-0.4378	1.2222
T25	-0.4830	-0.1042	0.2552
T3	-0.6560	-0.1125	0.1794
T45	-0.9650	0.0105	0.2220
T5	-0.9240	0.0332	0.2589

Table 19: Identity Matrix, 10 by 10

	N2	N3	FF	P13	P25	P3	T25	T3	T45	T5
N2	1	0	0	0	0	0	0	0	0	0
N3	0	1	0	0	0	0	0	0	0	0
FF	0	0	1	0	0	0	0	0	0	0
P13	0	0	0	1	0	0	0	0	0	0
P25	0	0	0	0	1	0	0	0	0	0
P3	0	0	0	0	0	1	0	0	0	0
T25	0	0	0	0	0	0	1	0	0	0
T3	0	0	0	0	0	0	0	1	0	0
T45	0	0	0	0	0	0	0	0	1	0
T5	0	0	0	0	0	0	0	0	0	1

4.4.3 Measurement correlations study

This section describes the outcome of the study undertaken to assess if any two of the available gas-path measurements respond in a similar way to the component changes according to the procedure presented in section 2.2.3 (Provost, 1995).

The ERT matrix defined above is first normalised by rows, each element in each row is divided by the square root of the sum of the squares of the elements in that row, and then the obtained matrix is multiplied by its transpose. The result is the Measurement Correlation Matrix reported in Table 20. The element i,j of the matrix represents the cosine of the angle between the vectors of the rows i and j of the ERT. According to Provost (1995), cosines with magnitude greater that 0.7 are worthy of note.

Table 20: Measurement Correlation Matrix

	N2	N3	FF	P13	P25	P3	T25	T3	T45	T5
N2	1.0000									
N3	0.5236	1.0000								
FF	0.5822	0.3272	1.0000							
P13	0.3395	0.2606	0.6910	1.0000						
P25	0.5977	-0.0387	0.7290	0.6540	1.0000					
P3	0.7743	0.4917	0.8893	0.7345	0.7775	1.0000				
T25	0.5020	-0.0464	0.7313	0.5152	0.9028	0.6940	1.0000			
T3	0.6082	0.3805	0.8595	0.5652	0.6853	0.8637	0.7921	1.0000		
T45	0.0813	-0.0609	0.8009	0.4032	0.4466	0.4767	0.5827	0.6330	1.0000	
T5	0.0876	-0.0833	0.7771	0.3425	0.4439	0.4684	0.5746	0.6134	0.9716	1.0000

Table 21 lists the 5 correlations that were detected among which: none are very high correlations (0.9<cosine<1), 2 are high correlations (0.8<cosine<0.7), and 3 are simple correlations (0.7<cosine<0.8).

Table 21: Report of the Measurement Correlation Matrix analysis

Correlated	Highly Correlated	Very highly correlated
P3 and P25	P3 and FF	
T45 and FF	T5 and T45	
T5 and FF		

4.4.4 Performance parameter correlations study

This section describes the outcome of the second test undertaken to assess whether any of the component changes and/or sensor biasses produce similar change in all (or nearly all) of the gas-path measurements, according to the procedure presented in section 2.2.4 (Provost, 1995). The System Matrix defined above is first normalised by columns, each element in each column is divided by the square root of the sum of the squares of the elements in that column, and then the obtained matrix is multiplied by its transpose. The result is the System Correlation Matrix reported in Table 22. The element i,j of the matrix represents the cosine of the angle between the vectors of the columns i and j of the System Matrix, and also in this case values greater that 0.7 are worthy of note (Provost, 1995).

Table 23 reports the 20 correlations detected, among which: 5 are very high correlations (0.9<cosine<1), 8 are high correlations (0.8<cosine<0.7), and 7 are simple correlations (0.7<cosine<0.8). The top part of the table lists the results when biasses and errors in the environmental and power setting parameters are excluded form the analysis. The number of total correlations is reduced to 8 with only 2 very high correlations. This is a relevant outcome to consider when evaluating a diagnostics system which does not have the capability of dealing with biasses as the system described in this section 4.6 and 4.7. Differently, the results from the whole table are useful in the study of the bias-tolerant system described in section 4.8.

Table 22: System Correlation Matrix

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}	*Nl*	*Z*	*M*	**N2**	**N3**	**FF**	**P13**	**P25**	**P3**	**T25**	**T3**	**T45**	**T5**
η_{FAN}	1.0000																								
Γ_{FAN}	-0.9582	1.0000																							
η_{IPC}	0.3923	-0.3350	1.0000																						
Γ_{IPC}	0.4421	-0.3614	-0.3432	1.0000																					
η_{HPC}	0.6096	-0.4607	0.4923	0.2139	1.0000																				
Γ_{HPC}	0.6686	-0.6521	0.4253	0.0429	0.3826	1.0000																			
η_{HPT}	0.6129	-0.4754	0.4444	0.2189	0.9753	0.3854	1.0000																		
Γ_{HPT}	0.2026	-0.0771	0.3670	0.0777	0.7702	0.1715	0.7863	1.0000																	
η_{IPT}	0.2847	-0.2421	0.8818	-0.3514	0.5383	0.2128	0.5508	0.5205	1.0000																
Γ_{IPT}	-0.2528	0.1578	0.5265	-0.5958	-0.3894	-0.1293	-0.3984	-0.2883	0.5400	1.0000															
η_{LPT}	0.8933	-0.8494	0.6648	0.1954	0.6578	0.6262	0.6761	0.3401	0.6492	0.0735	1.0000														
Γ_{LPT}	0.6397	-0.6054	-0.3757	0.6795	0.2397	0.4166	0.2453	-0.1152	-0.5150	-0.7650	0.2849	1.0000													
Nl	-0.9647	0.9994	-0.3623	-0.3545	-0.4811	-0.6599	-0.4948	-0.0979	-0.2691	0.1483	-0.8651	-0.5915	1.0000												
M	-0.6215	0.7732	-0.0595	-0.2413	-0.2637	-0.4931	-0.2739	0.0358	0.0429	0.2797	-0.4444	-0.5509	0.7579	1.0000											
Z	0.7833	-0.8643	0.1403	0.3166	0.5282	0.5489	0.5557	0.1716	0.1054	-0.3963	0.6348	0.6275	-0.8563	-0.9384	1.0000										
N2	0.2031	-0.1788	-0.2494	0.8459	-0.1531	-0.1136	-0.1590	-0.1198	-0.2641	-0.1353	0.0736	0.3154	-0.1722	-0.0305	-0.0054	1.0000									
N3	0.0429	-0.1458	0.0864	-0.2209	-0.4704	0.6066	-0.4802	-0.4042	-0.1774	0.2924	0.0033	0.0533	-0.1395	-0.1630	-0.0262	0.0000	1.0000								
FF	0.6326	-0.6924	0.4002	-0.0081	0.2959	0.2488	0.3023	0.2432	0.4174	0.1545	0.6606	0.1332	-0.6967	-0.3864	0.4678	0.0000	0.0000	1.0000							
P13	0.0000	-0.1940	0.0357	0.0000	0.0276	0.0213	0.0286	0.0223	0.0387	0.0130	-0.0105	-0.0445	-0.1790	-0.7013	0.5222	0.0000	0.0000	0.0000	1.0000						
P25	0.4151	-0.3461	-0.2798	0.4290	0.5198	0.5181	0.5299	0.4357	-0.2771	-0.8482	0.1903	0.6975	-0.3422	-0.3847	0.5242	0.0000	0.0000	0.0000	0.0000	1.0000					
P3	0.4386	-0.4771	-0.1116	0.1789	-0.0260	0.1879	0.0404	-0.5614	-0.1376	-0.0798	0.3127	0.5035	-0.4666	-0.4031	0.4514	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000				
T25	0.1898	-0.0897	0.2228	0.1321	0.1669	0.1627	0.1720	0.1391	-0.0642	-0.2490	0.0582	0.2108	-0.0969	-0.0959	0.0943	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000			
T3	0.1849	-0.1273	0.2559	0.0419	0.2154	0.0922	0.0047	-0.1655	-0.0303	-0.0125	0.0866	0.1447	-0.1317	-0.1036	0.0663	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000		
T45	0.2453	-0.1730	0.5029	-0.0019	0.3823	0.3129	0.3934	0.3076	0.5293	0.1940	0.3226	-0.2442	-0.1935	0.0096	0.0820	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	
T5	0.2483	-0.1680	0.5569	-0.0163	0.4191	0.3466	0.4303	0.3409	0.5850	0.2155	0.5561	-0.1135	-0.1854	0.0306	0.0956	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 23: Report of the System Correlation Matrix analysis

Correlated	Highgly Correlated	Very highgly correlated
$\Delta\Gamma_{HPT}$ and $\Delta\eta_{HPC}$	$\Delta\eta_{IPT}$ and $\Delta\eta_{IPC}$	$\Delta\eta_{FAN}$ and $\Delta\Gamma_{FAN}$
$\Delta\Gamma_{HPT}$ and $\Delta\eta_{HPT}$	$\Delta\eta_{LPT}$ and $\Delta\eta_{FAN}$	$\Delta\eta_{HPC}$ and $\Delta\eta_{HPT}$
$\Delta\Gamma_{LPT}$ and $\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$ and $\Delta\Gamma_{FAN}$	
$\Delta\Gamma_{FAN}$ and M	$\Delta\eta_{LPT}$ and N1	M and Z
N1 and M	$\Delta\Gamma_{FAN}$ and Z	
$\Delta\eta_{FAN}$ and Z	N1 and Z	
P13 and M		
	$\Delta\Gamma_{IPC}$ and N2	$\Delta\eta_{FAN}$ and N1
	$\Delta\Gamma_{IPT}$ and P25	$\Delta\Gamma_{FAN}$ and N1

4.4.5 Current studies: enhanced observability study

The preceding section described a method to investigate correlations among pairs of component changes and/or sensor biasses which may affect the measurements in a similar way. However, a weak observability of the System Matrix may result in the following two situations:

1. A given fault that manifests itself as a combination (two or more) of component changes and/or sensor biasses, produces measurement changes similar to another fault: two different faults cannot be distinguished.
2. A given fault (combination of component changes and/or sensor biasses) produces no changes in the measurements: the fault cannot be observed with the available set of measurement.

The theory of observability studies these circumstances in detail. As mentioned above, the most common test of observability is the study of the invertibility of the System Matrix, by determining, for example, its rank. If the rank equals the number of system outputs it means that a unique solution is possible: it is always possible to calculate the component changes and/or sensor biasses from the knowledge of the measurements.

Provost (1995) and Brown (1966) proposed a alternative way of testing for observability which is revised in this section. Even in the case of a completely observable system it can be shown that a system can be more or less observable as discussed below. The presence of uncertainty in the measurable parameters can considerably affect the solution of a system with poor observability.

4.4.5.1 A geometric interpretation

As an example, Table 24 shows the ERT for a 2 dimensional case. The ERT has a rank equal to 2, therefore a unique solution can be found. Nevertheless, the quality of the observability can be investigated.

Table 24: 2D example ERT

	1% $\Delta\eta_{HPT}$	1% $\Delta\Gamma_{HPT}$
N3	0.0728	-0.7296
T5	0.7765	0.4430

For the specific case in which the changes N3 and P5 (note that N3 and P5 represent percentage change and not absolute values) are equal to 0.3, the following two equations can be obtained:

$$N3 = 0.0728\,\Delta\eta_{HPT} - 0.729\,\Delta\Gamma_{HPT} = 0.3 \text{ (say)} \quad (4.\ 59)$$

$$T5 = 0.7765\,\Delta\eta_{HPT} + 0.443\,\Delta\Gamma_{HPT} = 0.3 \text{ (say)} \quad (4.\ 60)$$

And the solution is:

$$\Delta\eta_{HPT} = 0.5874 \quad (4.\ 61)$$

$$\Delta\Gamma_{HPT} = -0.35 \quad (4.\ 62)$$

Figure 70 shows the equations (4. 59) and (4. 60) plotted in the space of the performance parameters. The exact solution is indicated as the intersection of the two solid lines. The acute angle of intersection between the lines becomes important when measurement uncertainty is taken into account. The dashed lines in the diagram represent shifts in measurements of $\pm$ 0.05. The exact solution point becomes an area of uncertainty that is larger when the acute angle of intersection between the two equations is smaller. Small measurement errors may result in a significant error in the unknown performance parameters. It can be seen that if we have fuzziness in the measurements, the corresponding fuzziness in the solution is held to a minimum if the lines intersect to a right angle; conversely, the fuzziness becomes more pronounced as degree of obliqueness of the intersecting angle increases.

The geometric interpretation, just described, becomes cumbersome for higher-order systems, so it is more appropriate to view this problem in terms of relationships among the row vectors of the System Matrix – see section 4.4.5.2.

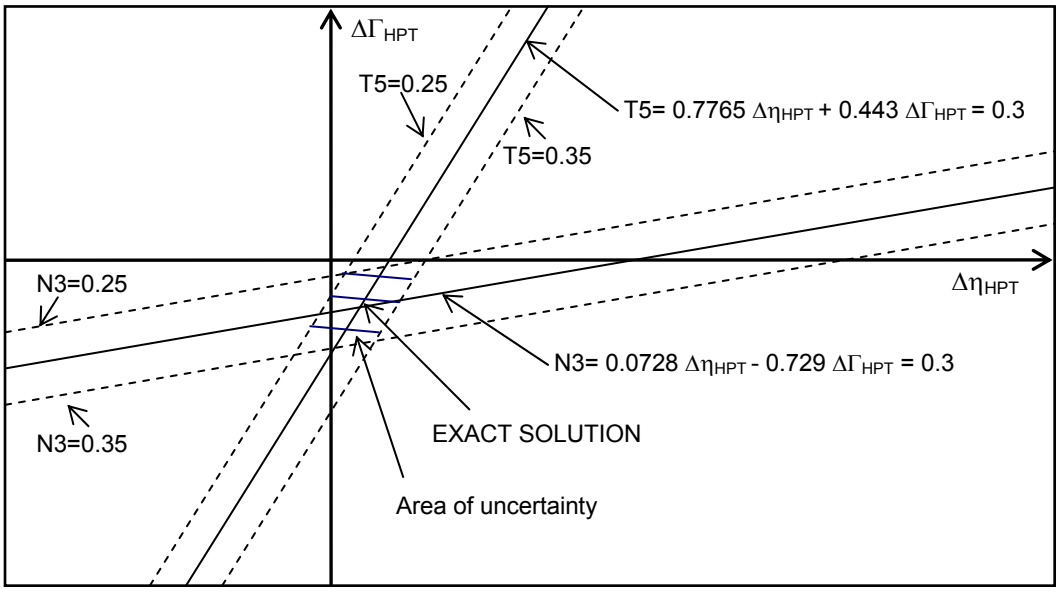


Figure 70: graphical representation, 2D example

Therefore it can be concluded that the rank of the System Matrix enables us to answer to whether the system is observable or not, while the magnitude of the acute angle among the rows of the matrix provides some information about how observable the system is (Brown, 1966): the higher the angle, the higher the observability. The next section will show that if there is a vector that is as closely normal as possible to all the rows of the ERT (plotted in Figure 70 for a 2D example), that implies that the rows are all closely parallel to each other and hence a high area of uncertainty exists.

4.4.5.2 The mathematical formulation

This section provides a method that can be used to assess the quality of the observability of a given system: two procedures are proposed. Even when a system is observable according to yes/no type of tests, it can be characterised by a strong or a weak observability.

For simplicity, let us limit the study to the analysis of the ERT, i.e. without sensor biasses included. The same formulation can be extended to the System Matrix. Firstly, the ERT matrix is normalised by row to produce a matrix called ERT_{norm} with rows which are of unit length (necessary since the angles between the equations denoted by ERT depend on ratios of elements by row):

$$ERT_{norm} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ .... \\ R_n \end{bmatrix} \qquad (4.\ 63)$$

The problem now is the identification of the column vector **U** (unit length), which is approximately normal to all the vectors R1, R2, …Rn. If we find that in our n-dimensional space there exists a vector that is as closely normal as possible to all the vectors of the matrix ERT_{norm}, we will have a maximum fuzziness situation, because that implies that the rows are all closely parallel to each other and that there is a certain degree of linear dependency among them. Therefore the answer to the question about the quality of the system's observability lies in the study of this matrix.

The vector **U** (unit length) is as closely normal as possible to R1, R2, …Rn when the following expression is a minimum:

$$L = (R_1\ \underline{U})^2 + (R_2\ \underline{U})^2 + (R_3\ \underline{U})^2 + \ldots + (R_n\ \underline{U})^2 \qquad (4.\ 64)$$

This can be written as:

$$\begin{aligned} L &= \underline{U}^T (R_1^T\ R_1\ \underline{U} + R_2^T\ R_2\ \underline{U} + \ldots)\ \underline{U} = \\ &= \underline{U}^T\ ERT^T_{norrm}\ ERT_{norrm}\ \underline{U} + \lambda\ (\underline{U}^T\ \underline{U} - 1) \end{aligned} \qquad (4.\ 65)$$

where λ is a scalar and $\underline{U}^T\ \underline{U} = 1$. To minimise L we compute:

$$\frac{\delta L}{\delta U} = 0$$

$$= 2\, ERT^{T}_{norrm}\, ERT_{norrm}\, \underline{U} - 2\lambda\, \underline{I}\, \underline{U} = 0 \qquad (4.66)$$

$$= (ERT^{T}_{norrm}\, ERT_{norrm} - \lambda\, \underline{I})\, \underline{U} = 0$$

This has a non-trivial solution only if:

$$\det (ERT^{T}_{norrm}\, ERT_{norrm} - \lambda\, \underline{I}) = 0 \qquad (4.67)$$

We define therefore the Observability Matrix as follows:

$$CORROBS = ERT^{T}_{norrm}\, ERT_{norm} \qquad (4.68)$$

the values of λ that satisfy the characteristic equation of the CORROBS matrix [nxn], are the eigenvalues. According to the traditional linear algebra theory, the eigenvalues of CORROBS are equal to the square of the SVs (singular values) of ERT_{norm}. For each eigenvalue the associated eigenvector, which satisfies the equations above, is calculated.

The magnitude of each eigenvalue is related to the angle between the rows of the ERT_{norm} matrix and the eigenvector associated with that eigenvalue: the smallest eigenvalue corresponds to the most orthogonal eigenvector (weak observability), while the highest eigenvalue corresponds to the less orthogonal eigenvector.

Besides, considering that, for a systems with ERTs with rank m, the sum of the eigenvalues of the matrix CORROBS is equal to m (as justified below) it can be noticed that if the system is highly observable the eigenvalues have all similar magnitudes: no clear directions closely orthogonal to all the rows of the ERT are identifiable, this would be associated with one eigenvalue significantly smaller than the others. On the other hand, a poorly observable system is characterised by some eigenvalues that are considerably small and others considerably high (being their sum constant): vectors approximately orthogonal to all the rows of the ERT are identifiable.

The following considerations can be made relative to the observability analysis of a system where m observable quantities are used to search for a total of n gas-path component changes and sensor biasses (Provost, 1995):

- n eigenvectors and associated eigenvalues are calculated (the eigenvectors are representative of component changes and/or sensor biasses).
- The product between the ERT matrix and the eigenvectors provides the Measurement Differences matrix whose elements are the measurement deviations projected in the directions of the eigenvectors.
- The eigenvectors are perpendicular to one another except where the eigenvalues are equal; in that case the eigenvectors are identical.

- The sum of all the eigenvalues is m (because the sum of the diagonal elements – i.e. trace – of any m by m sub-matrix with rank m extracted from the CORROBS is always equal to m, being this matrix obtained through the normalization procedure described above).
- If m<n, there are at least n-m eigenvectors representing combinations of component changes and sensor biasses that are not observable. These eigenvectors are characterised by :
    - Eigenvalues that are theoretically zero (in practice very small).
    - The columns of the Measurements Differences matrix associated with those n-m eigenvectors have elements that are all zero (hence non-observability, no measurement devotions can be read in the directions of those eigenvectors).
    - Being the eigenvectors representative of component changes and/or sensor biasses, it can be concluded that, the specific combination of component changes that corresponds to the elements of each of the n-m eigenvectors (and the combinations obtained by multiplying them for any factor) are not observable combinations.
- Each element of the eigenvectors can be multiplied by the same factors to give a new vector with the same proprieties, i.e. the direction of the eigenvector, not its magnitude, is important.

The following two procedures are therefore introduced to assess the quality of the observability of a given system.

Procedure 1: According to the considerations above (Provost, 1995), this procedure relies on the identification of the smallest eigenvalues of CORROBS and the corresponding eigenvectors: these eigenvectors provide the combinations of changes in performance parameters and/or sensor biasses that are not observable because they are associated with no changes in the measurements (see application in sction 4.4.5.3).

Procedure 2: The need for a parameter able to quantify the quality of the observability of a system (either completely observable or not) was recognised. Such a parameter would enable the comparison of a large number – calculated by means of equation (4. 69) – of systems and the choice of the most suitable measurement set for diagnostics purposes. A criterion is proposed in this work to compare two systems with the same number of inputs, outputs and rank from the point of view of the quality of the observability. The following statements can be made:

- The sum of all the eigenvalues of CORROBS is m.
- If the ERT is a square matrix n by n (being m=n) with rank n, since the sum of all the eigenvalues of CORROBS is constant and equal to m=n, and since the determinant of CORROBS is equal to the product of the eigenvalues (all different from zero), we can state that for a given set of measurements the higher its determinant, the better the observability of the system, because the smallest eigenvalue will be the highest possible, and the highest eigenvalue will be the smallest possible.

- If the ERT is an m by n rectangular matrix with rank m, the determinant of the CORROBS being theoretically zero, the study focuses on the m eigenvalues of CORROBS that are significantly different from zero. Their summation is approximately equal to m, being the remaining n-m eigenvalues theoretically zero (in practice approximately zero). We can state that for a given set of measurements the higher the product of the m eigenvalues significantly different from zero, the better the observability of the system.

It can be observed that in practice it is very unlikely that the rank of the System Matrix is smaller than m and therefore two systems with the same number of inputs and outputs will most likely have an associated System Matrix with the same rank. The reason why this is very unlikely lies in the fact that it would imply that 1% changes in performance parameters and/or sensor biasses produce exactly the same changes in all the measurements or a factor of these changes. This would generate two columns of the System Matrix that are mathematically linearly dependent. On the other hand, it is likely (as discussed above) that the two column vectors are approximately parallel and have a high degree of dependency.

In the generic case of an m by n ERT, Procedure 2 can therefore be used to choose the most suitable set of m measurements from a list of p (p>m) measurements that is characterised by the highest quality of observability: the best set is the one associated with the maximum value of the product of the m eigenvalues of CORROBS that are significantly different from zero. Note that if p=15 and m=10 the problem involves the identification of the most suitable set of 10 measurements among 15 given measurements and the possible combinations to study are 3003. The number of combinations C_m of p=15 measurements taken m=10 at a time is calculated according to equation (4. 69).

$$C_m = \frac{p!}{m!(p-m)!} \tag{4. 69}$$

The following limitations and comments, when applying this theory to gas-path diagnostics must be noted:

- Procedures 1 and 2 rely on the linearization of the problem around the condition of 1% changes in performance parameters and sensor biasses. Besides, the study is undertaken with an ERT matrix computed at a given operating condition: for example, the best measurement set identified at a given flight condition may not result optimal at another. Nevertheless, diagnoses, most of the times, are required at cruise condition and therefore the importance of the observability study at that operating condition still holds.
- Procedure 1 identifies the combinations of component changes (and sensor biasses) that cannot be observed because they are associated with no changes in the measurements. If a comparison among a considerable number of combinations of measurement sets is required this procedure would not be practical.

- The criterion proposed in Procedure 2 identifies a parameter respectively in the cases of an observable System Matrix (n=m) with rank m=n, and of a non observable System Matrix (m<n) with rank m<n, that quantifies the quality of the observability.
- When Procedure 2 is used to compare different types of measurement sets (with the same number of measurements) in the case of m<n, it must be remembered that a number of n-m eigenvectors that represent combinations of component and/or sensor that are not observable are expected. These eigenvectors have associated eigenvalues that are theoretically equal to zero (in practice less than 10^{-5}). These are associated with n-m directions with no observability.

4.4.5.3 A possible application

An example is now considered with reference to the ERT matrix reported in Table 17. The study is limited to the case without sensor biasses or errors in the environmental and power setting parameters. The data from Table 25 to Table 28 report respectively the eigenvalues of CORROBS, the eigenvectors, the Measurement Deltas matrix and the norms of the columns of the Measurements Deltas matrix, calculated according to the theory described above. The norms of the columns of the Measurements Deltas matrix is a measure of the changes in the measurements that would correspond to changes in performance parameters equal to the elements of the associated eigenvector. The smaller is the norm the less observable are the changes.

Table 25: Eigenvalues

4.9E-17	**5.6E-16**	1.1E-02	1.5E-02	2.4E-02	1.1E-01	1.5E-01	3.5E-01	5.9E-01	1.4E+00	2.3E+00	5.0E+00

Table 26: Eigenvectors

η_{FAN}	**0.00595**	**-0.00828**	0.79118	-0.05403	-0.19304	-0.13820	-0.02819	-0.06948	-0.35373	0.09107	0.10338	0.40611
Γ_{FAN}	**0.00262**	**-0.00017**	0.06751	0.04880	-0.00703	0.00157	-0.02136	-0.12605	-0.58609	0.31820	-0.33626	-0.64717
η_{IPC}	**-0.00495**	**-0.04605**	-0.30225	-0.03867	-0.30051	-0.38807	-0.16505	0.59147	-0.35289	-0.22059	-0.26606	0.20748
Γ_{IPC}	**0.00078**	**0.02114**	-0.15960	0.54934	0.08190	-0.69955	0.14965	-0.30956	0.01468	0.17897	0.13915	0.07995
η_{HPC}	**-0.01167**	**0.01115**	0.06713	-0.06859	0.65297	-0.00696	0.49500	0.31054	-0.05207	0.21973	-0.33323	0.24618
Γ_{HPC}	**0.00789**	**0.00352**	0.07781	0.55610	0.43410	0.25935	-0.60492	0.15682	-0.13642	-0.09731	0.02390	0.09871
η_{HPT}	**-0.41040**	**-0.01161**	-0.03728	0.37190	-0.41843	0.30987	0.07990	-0.04735	0.19640	0.31314	-0.46129	0.25016
Γ_{HPT}	**-0.14859**	**0.17679**	0.06762	-0.40560	0.18230	-0.35089	-0.54663	-0.17274	0.31949	0.23064	-0.36838	0.04016
η_{IPT}	**0.64103**	**0.49562**	0.06393	0.14834	-0.10545	0.05809	0.08168	-0.18111	0.07138	-0.27718	-0.41995	0.07146
Γ_{IPT}	**-0.62780**	**0.39471**	0.09828	0.01741	0.12420	-0.06509	0.14740	-0.14832	-0.12377	-0.58707	-0.07088	-0.10829
η_{LPT}	**0.02591**	**-0.34310**	-0.35300	-0.19045	0.12376	0.14521	-0.04784	-0.57044	-0.37938	-0.15745	-0.15216	0.40591
Γ_{LPT}	**-0.05926**	**0.66835**	-0.31269	-0.12678	-0.02598	0.16053	-0.04773	0.01758	-0.28342	0.39305	0.34849	0.22507

Table 27: Measurement Deltas matrix

N2	-2.0E-16	8.3E-16	-4.5E-03	2.0E-02	1.4E-02	-2.4E-01	4.1E-02	-3.5E-01	8.4E-04	2.6E-01
N3	3.4E-17	4.4E-16	3.1E-03	1.8E-02	3.8E-02	3.8E-02	-3.5E-01	1.3E-01	-2.8E-01	-9.1E-01
FF	-7.7E-16	-1.6E-15	1.1E-01	-2.9E-01	-3.9E-02	2.2E-02	-1.1E-01	-5.0E-01	1.0E-01	-8.2E-01
P13	8.0E-17	-3.5E-16	-1.2E-02	5.5E-03	3.7E-04	-2.3E-02	1.5E-02	1.1E-01	3.7E-01	-2.1E-01
P25	-1.3E-15	1.7E-15	3.1E-02	2.7E-02	2.4E-01	3.4E-01	-2.8E-01	-7.0E-02	3.8E-01	1.6E+00
P3	1.2E-15	5.6E-16	-9.8E-03	7.0E-02	-1.5E-01	4.1E-01	2.5E-01	-2.4E-01	-1.6E-01	-1.7E-01
T25	1.8E-16	-2.5E-16	-1.1E-02	-1.5E-03	-7.1E-02	-6.6E-02	-1.5E-01	2.0E-01	-9.0E-02	4.3E-01
T3	-2.2E-16	-1.7E-16	1.7E-03	-8.0E-03	4.2E-02	-3.7E-02	1.5E-01	2.5E-01	-2.6E-01	-6.4E-02
T45	-4.3E-16	3.8E-15	1.0E-01	1.3E-01	-2.5E-03	-1.1E-01	-5.0E-02	-1.2E-01	-5.4E-02	-5.2E-01
T5	1.7E-15	-2.9E-15	-1.6E-01	1.2E-02	5.9E-02	-1.0E-03	-9.6E-02	-4.1E-01	-3.4E-01	-5.4E-01

Table 28: Norms of the columns of the Measurements Deltas matrix

6.2E-15	1.3E-14	0.4476	0.5751	0.6475	1.2912	1.4874	2.3667	2.0297	5.5589

According to Procedure 1:

- 12 eigenvalues and 12 eigenvectors are calculated.
- 10 large eigenvalues (all greater than 0.01) are identified corresponding to 10 eigenvectors associated with different degrees of observability.
- 2 eigenvalues theoretically zero (less than 10^{-5}) that correspond to the 2 eigenvectors associated with no observability are recognised (note that the corresponding measurement differences associated are all smaller than 10^{-5}). The specific combinations of component changes that correspond to the elements of these two eigenvectors (and the combinations obtained by multiplying them by any factor) are not observable – see the corresponding norms of the columns of the Measurement Deltas that are approximately zero.

A second type of study was carried out, according to Procedure 2, to choose set of m=10 measurements from a list of p=15 (p>m) measurements (reported in Table 29) that is characterised by the highest quality of observability: 3003 combinations were investigated. The p=15 measurements include the set of 10 measurements available (see section 4.3, Table 11), within this project, for the engine investigated. Therefore, a comparison between the best set and the available set can be undertaken in Table 30, which details the measurements of the two sets and the associated product D of the the m eigenvalues of CORROBS significantly different from zero. The best set of 10 measurements is the one associated with the maximum value of this product. A further improvement observability-wise could be introduced by increasing the number of measurements up to 12 in order to make the System Matrix observable with rank m=n. The best set of 12 measurements with the best observability could be found by applying Procedure 2 in a similar way.

Table 29: 15 possible measurements

N2	N3	FF	FN	P13	P23	P25	P3	P5	T13	T23	T25	T3	T45	T5

Table 30: The available measurement set and a better choice of 10 measurements set with their values of D

Measurement set										D	Set
N2	**N3**	**FF**	**P13**	**P25**	**P3**	**T25**	**T3**	**T45**	**T5**	*2.2E-07*	*Available*
N2	N3	P23	P25	P3	P5	T23	T25	T3	T5	*2.5E-05*	*Best (Procedure 2)*

Concluding, a thorough evaluation of the health of the various gas-path components requires a large number of performance parameters to be assessed and therefore a large number of sensors to be installed. On the other hand, the number of sensors is seriously limited by weight, bulk, cost and reliability concerns as discussed in section 2.2. Diagnostics typically has to be performed using the sensors available, which may not be the most suitable because often the choice of sensors is not dictated by gas-path diagnostics requirements. Nevertheless, a discussion of two procedures to assess these requirements was presented. In the remaining part of this work the available set of 10 measurements (Table 11) is considered with the aim of investigating a realistic scenario.

4.5 A fuzzy logic based diagnostics system for a three spool engine

4.5.1 Objectives and scope

In the light of the specifications for a novel diagnostics methodology stated in Chapter-3, and considering the potentialities of fuzzy-logic theory in this field illustrated in section 4.2 , the research objectives are precisely to develop a fuzzy-logic based procedure that is:

- Based on a non-linear model.
- Designed specifically for SFI and/or MFI.
- Capable of detecting with reasonable accuracy significant changes in performance.
- Able to provide a 'concentration' capability on the actual fault.
- Competent to make a worthwhile diagnosis using only few measurements (N>M).
- Able to deal with random noise in the measurements and sensor bias.
- Light in computational requirements.
- Fast in undertaking diagnoses for on-wing applications.
- Readily adaptable to similar system in a reasonably short time: exempt from training and tuning uncertainties, difficulties and dependences for setting-up parameters.
- Free from a lack of comprehensibility due to black-box behaviour.
- Capable of data-fusion.
- Able to incorporate expert knowledge.

This section is entirely dedicated to the description of a novel way of developing a SFI capability in the presence of sensor noise neglecting the problem of sensor bias. Sections 4.7 and 4.8, then, respectively extend the method to the cases of a partial MFI system and of a bias-tolerant system.

The scope of this section is to illustrate an application of the devised fuzzy logic based gas-path diagnostics method to the Trent 800 engine. The most important parameters in the process are identified and optimized through a sensitivity study. The accuracy of the methodology in this specific application is then assessed with simulated case studies in section 4.6.

4.5.2 The methodology and identification of the key parameters

Gas-path analysis is formulated here as a problem of recognition of deteriorated measurement patterns by using a rule-based method that has its foundation on fuzzy algebra (Marinai et al., 2003b).

The inherent capability of fuzzy systems, previously pointed out in section 4.1, to deal with GPD problems is twofold. Firstly, a fuzzy diagnostics system takes into account the uncertainty in the measurements that affects the fault patterns characterization, at a set-level. Secondly, at a system-level, the learning algorithm devised in this project states fuzzy rules to map input sets of measurements into output sets of performance parameters, in a constrained

search space. This enables diagnoses even though the formulation of the diagnostics problem is analytically undetermined (as discussed in detail in Chapter-2).

In this section, the fuzzy logic system's architecture illustrated in section 4.2 is revisited with reference to gas-path diagnostics for a three spool Trent 800 engine. The diagnostic process, as shown in Figure 71, is designed to assess performance parameter percentage changes from a clean engine condition (12 outputs) given the knowledge of the measurement changes (10 inputs) calculated as percentage deviations with respect to a baseline determined by means of an engine model run at a specific power setting and environmental conditions. The fuzzy system $F=R^{10} \rightarrow R^{12}$ uses m rules to map the vector of input delta measurements z to a vector of output delta performance parameters x=F(z). The analysis is undertaken at the operating condition characterised by the following parameters: N1=0.8% , Mach= 0.85, Altitude=10000 m.

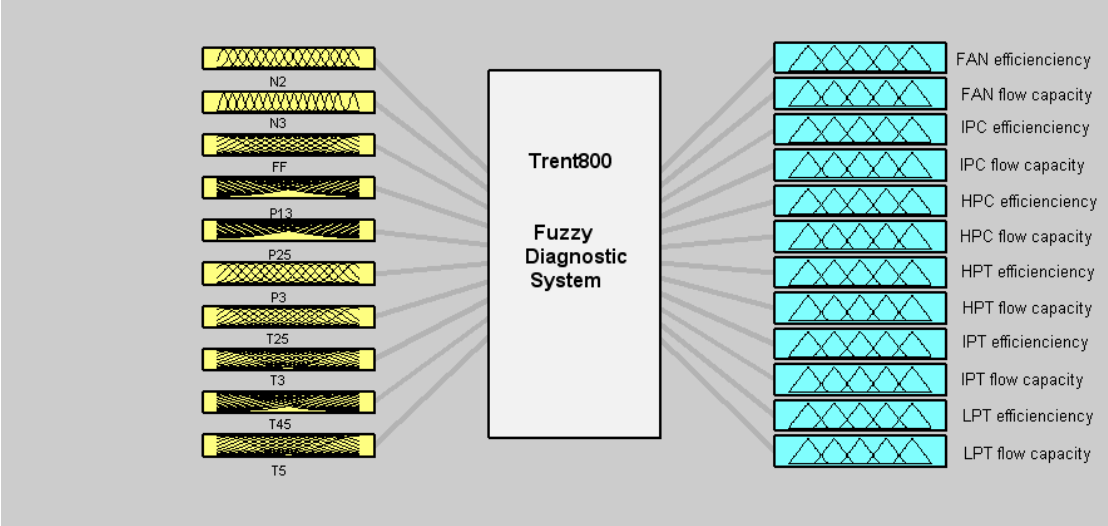


Figure 71: layout of the fuzzy logic diagnostic system

Diagnostics is made through a Mandami-type fuzzy inference process. The ranges of variability of the outputs - $\Delta\eta$ and $\Delta\Gamma$ for the 6 components – define the search space, where the solution is sought. A sensible choice of these ranges for a real life application would be between -5% and 0 for all the efficiency deltas and for the flow capacity deltas of the compressors, while they can cover positive values for the turbine flow capacity deltas going for example between -5% and +3%. The range of variability of each input variable is evaluated according to the sought output ranges through the engine model.

4.5.2.1 Fault levels combinations and if-then rules

The learning algorithm proposed in this thesisprovides a way of stating if-then rules that relies on the use of an engine model. Therefore, the rules are strictly related to the aero-thermal equations. The use of data obtained from the engine model to generate the rules preserves the non-linearity of the problem.

The rules have the general form *IF condition-1 AND condition-2 …THEN statement.* As discussed in section 4.2.4.1 the if-part of the rule refers to the fault signature in the measurements, represented through input MFs, evaluated by running the engine model at a defined deteriorated condition within the search space. On the other hand, the then-part of the rule refers to this condition characterised with output MFs.

Section 4.2.4.1 illustrated how to state fuzzy rules for a simple case of three measurements, four parameters and 5 rules. The procedure starts with the definition of the search space for the performance parameters. According to section 4.2.6 the search space includes all the combinations of changes in efficiency and flow capacity of the 6 components that the system is meant to deal with. The parameters that characterise the search space are: (i) the number of components that are considered deteriorated simultaneously (1 at a time for SFI), (ii) the maximum and minimum values of the ranges of variability of the performance parameters, and (iii) the increment value that divides each range in a finite number of constant variations (fault levels). For the purpose of illustrating the methodology, we consider the following search space:

- Number of components simultaneously deteriorated = 1 (SFI)
- Maximum variation in compressors' efficiencies = 0%
- Minimum variation in compressors' efficiencies = -3%
- Maximum variation in compressors' flow capacities = 0%
- Minimum variation in compressors' flow capacities = -3%
- Maximum variation in turbines' efficiencies = 0%
- Minimum variation in turbines' efficiencies = -3%
- Maximum variation in turbines' flow capacities = 1%
- Minimum variation in turbines' flow capacities = -3%
- Increment Value= 0.5%

The features of this search space are the following:

- It defines the 12 dimensional space of the ranges of variability of the 12 parameters in % changes from the clean condition values.
- It takes into account positive variation for turbines' flow capacity.
- It considers C=6 combinations of 1 gas-path components deteriorated at a time – see section 4.2.6.
- The increment value in the search space is 0.5%. This implies that the engine model is run for all the combinations of variations of the performance parameters within the ranges defined above, obtained going from the minima to the maxima with 0.5% steps. For example going from 0 to 3% of FAN efficiency the following 7 conditions of deterioration are generated: 0, -0.5, -1, -1.5, -2, -2.5, -3 %.
- We note that with 0.5 % steps all the ranges are divided in 7 combinations except for turbine flow capacity ranges, which are divided into 9 fault levels (being these ranges between 1% and -3%).
- The number of if-then statements generated is equal to 331.

The solution of the diagnostic problem will be looked for within the constrained search space. Note that the use of a constant increment value implies that the values of the f fault levels are chosen uniformly distributed within the ranges.

4.5.2.2 Input and output MFs

Fuzzy sets are defined for the inputs and the outputs. Each input range is spanned with a number M_i of MFs where the index i=1…n identifies the i-th measurement. These MFs according to the procedure described in 4.2.4.1 are centred, for each measurement, in the output of the engine model run at all the combinations identified in the search space, or in the mean value of a cluster of values grouped according to the procedure. On the other hand, the deviations in performance parameters of the table are always associated with a MF centred in the fault level specified in the search space.: N_j MFs for j=1…p are designed for the i-th performance parameter.

Two types of MFs were considered: triangular with amplitude s or Gaussian according to equation (4. 70) where m is the midpoint of the function and RMS=σ. The two types of MFs are shown respectively in Figure 72 and Figure 73.

$$MF(x) = e^{-0.5\cdot\left(\frac{x-m}{\sigma}\right)^2} \qquad (4.\ 70)$$

The optimal type of output MFs is not known a priori and therefore a sensitivity study (section 4.5.5) was undertaken to identifying the choice that contributes to an optimal accuracy of the diagnostics system. An example of 7 Gaussian MFs spanning the range for FAN $\Delta\eta$ is shown in Figure 74.

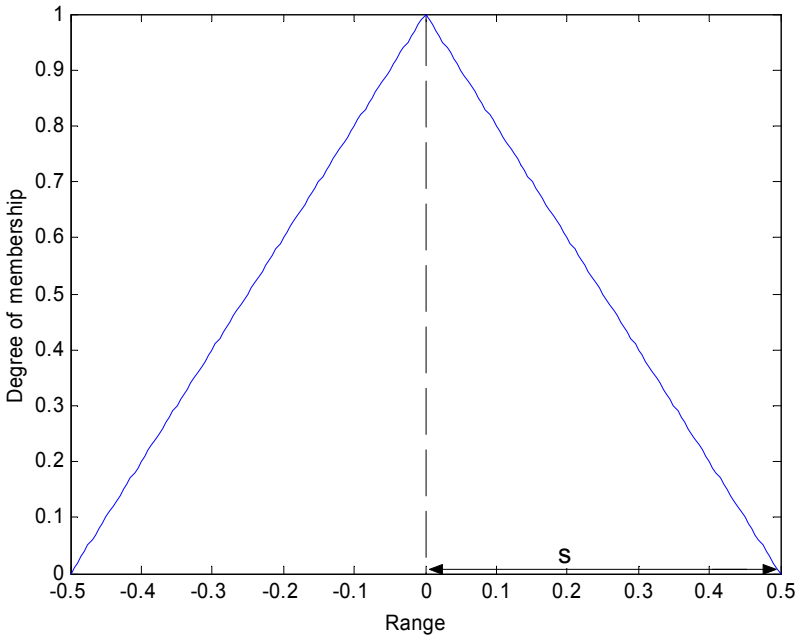


Figure 72: Triangular membership function

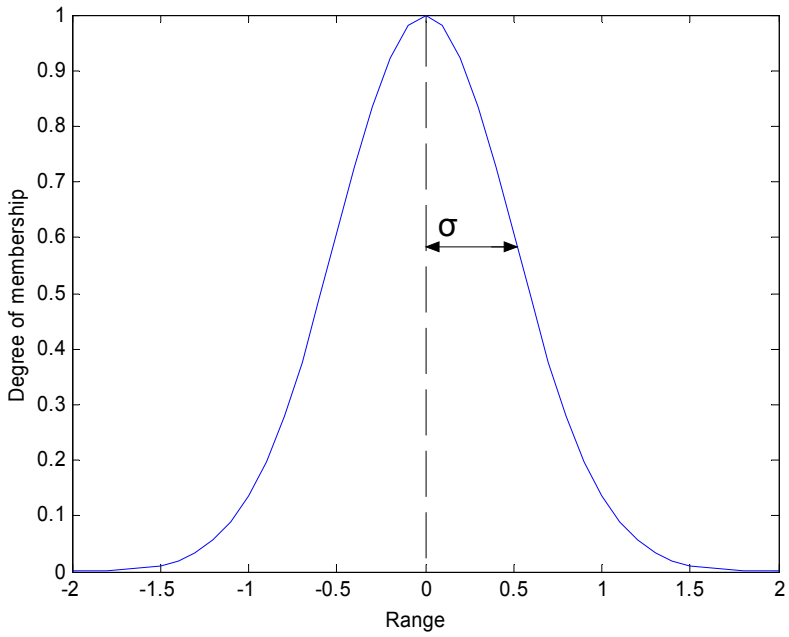


Figure 73: Gaussian membership function

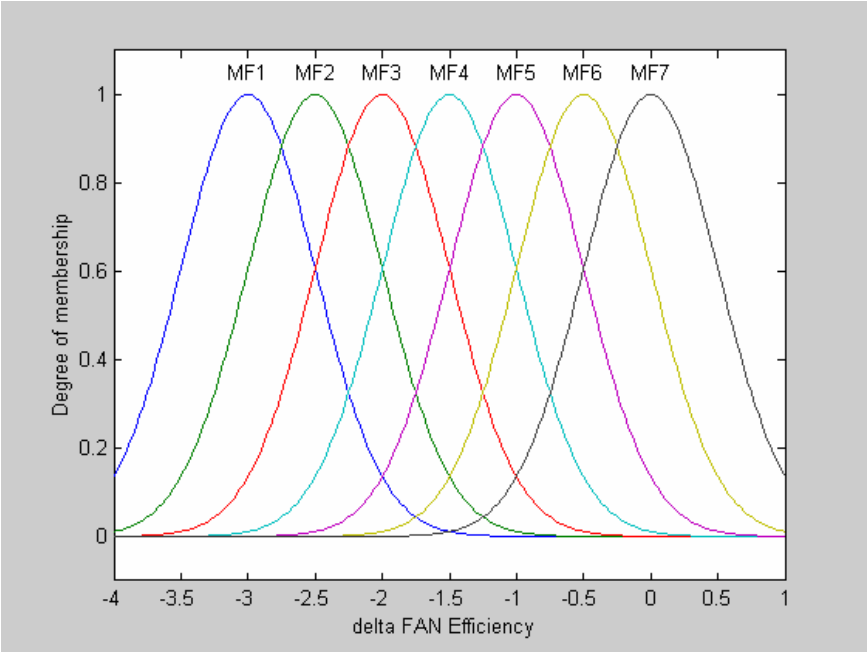


Figure 74: Example of 7 Gaussian MFs in a fixed performance parameter range for the output FAN $\Delta\eta$

On the other hand, a preliminary comment can be made here regarding the input MFs. The degree to which the measurement value z belongs to a given MF, in fuzzy algebra, was named a(z). Alternatively, a(z) can be interpreted as the probability that the measurement is the midpoint of the MF given that the measurement value is z. Therefore, we can view the input fuzzy set as a random set of two-point conditional probability densities, where the set degree

a(z) = degree($z \in A$) becomes the local conditional probability prob{$Z=A \mid Z=z$}. In this sense we can use Gaussian MFs for the input measurements with values of RMS equal to realistic values of sensor noise RMSs. This choice is in the opinion of the author an effective and consistent way of designing measurement MFs oriented to tackle the measurement uncertainty problem. However, at this level of the investigation, the option of using triangular MFs, generally considered very effective to design highly dimensional fuzzy systems, is left also for the input variables. This leaves open the opportunity to compare two input MF types – see section 4.5.5 – to identify the optimal system layout.

4.5.2.3 Fuzzy rules generation

The method developed in order to state fuzzy rules discussed in section 4.2.4.1, can here be applied to the case of a three spool engine. Each rule is composed of two parts: (i) the if-part that contains the fault signature in the measurements represented with MFs linked with the AND operator, and (ii) the then part that contains the MFs of the output performance parameters that characterise the fault condition. Table 31 and Table32 contain an example of data necessary to set up a rule generated by running the engine model. The use of data obtained from the engine model to generate the rules preserves the linearity of the problem. A rule states in terms of MFs, what in terms of numerical values can be read as follows: if the pattern in the measurements shows the deviations from a baseline listed in Table 31 then the combination of deterioration levels is in Table 32.

Table 31: Example of % changes in measurements from the baseline

ΔN2	ΔN3	ΔFF	ΔP13	ΔP25	ΔP3	ΔP5	ΔT25	ΔT3	ΔT45
0.460	-0.008	-0.949	-0.907	-1.117	-1.115	-0.804	0.169	0.111	0.182

Table 32: Example of % deltas in performance parameters from the clean engine

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
-2	-1.5	0	0	0	0	0	0	0	0	0	0

In general a rule will be formulated according to Table 33 and Table 34 created from Table 31 and Table 32 following the procedure of section 4.2.4.1. Table 33 shows the formulation of the if-part of the rule where the mf_i is the MF of the i-th input that is either centred in the value of Table 31 or centred in the mean value of a cluster of values grouped according to the procedure.
The algorithm that generates the input MFs for a number m of rules starts with the choice of K, the maximum number of input MFs (based on the experience). Then, for the i-th input measurement, the values of deviations (outcomes of the engine model for a number m of rules) are sorted and if two values are overlapped one of them is discharged. Then, the remaining values are counted, if the number of values is less or equal to K (the maximum number of MFs required) one MF is centred in each of these values that at the most are m (number of rules). Otherwise, the difference between each value and its consequent value, in the sorted list, is computed. The smallest value of difference between two measurement deviations is identified and these two values are substituted with their average value. A MF is then centred in this average value. This is repeated until the number of values that are centres of the input MFs is equal to K.

The symbol + represents the AND operator. Similarly Table 33 contains the then-part of the rule with the output MFs that identify the deteriorated condition.

Table 33: If-part of the fuzzy rule

If- part – Δmeasurements MFs
mf1 + mf2 + mf3 + mf4 + mf5 + mf6 + mf7 + mf8 + mf9 + mf10

Table 34: then-part of the fuzzy rule

then- part – Δ performance parameters MFs
mf1 , mf2 , mf3 , mf4 , mf5 , mf6 , mf7 , mf8 , mf9 , mf10 , mf11 , mf12

4.5.2.4 Fuzzy inference: functional and system parameters

Fuzzy inference is the process used to perform pattern recognition and therefore to compute mapping between input values and output values.

The inference process consists of feeding an input set of % changes of the 10 measurements that are taken along the gas-path (or simulated with the engine model to generate a test case) into the fuzzy logic system that calculates the output performance parameters % changes. The fuzzy inference process includes the following 5 phases: (i) fuzzification of the input variables, (ii) application of the AND fuzzy operator in the if-part of the rule, (iii) implication from the if-part to the then-part of each rule, (iv) aggregation of the then-parts across the rules, and (v) defuzzification, previously described in detail in section 4.2.4.2.

The following parameters are referred to as **functional parameters** and can be combined in several ways in designing a fuzzy system:

- AND operator, implemented as: product, minimum.
- Implication method, implemented as: product, minimum.
- Aggregation method, implemented as: summation, maximum
- Defuzzification method, implemented as: centroid, centre of maximum.

The functional parameters were identified as those parameters that characterise the functionality of the inference process. An initialsensitivity study is performed in section 4.5.5 to identify the combination of parameters most suitable to design a fuzzy-logic based diagnostic system. There is no reason to think that when the type of engine diagnosed changes this optimal combination of functional parameters should vary. Hence, the outcome of this first investigation is the choice of the fuzzy functional parameters for a generic diagnostics system.

On the other hand, we define the following **system parameters**:

- Number, type, width of the input MFs. To take into account sensor noise the value of amplitude (s or σ) for the i-th measurement can be expressed as a multiple of its sensor noise RMS_i ($a \cdot RMS_i$)
- Number, type, width of the output MFs. The number of output MFs is always a result of the search space definition. For each of the 12 performance parameters (involved in this application), for a given range

of variability, this number depends on the increment value (as defined in section 4.5.2.1) once the search space is defined. This number corresponds to the number f of fault levels that the range is divided into.

Summarizing, for the application described in this chapter, with fixed inputs and outputs, the system parameters to be optimized are 6: Number, type and width of the input MFs, type and width of output MFs and increment value in the search space.

A second sensitivity study will be described in section 4.5.5 aimed at identifying their best values to set up a system for the Trent 800 engine. When implementing a new diagnostic system a new sensitivity study maybe required to identify their optimal values. Nevertheless, the logic and the procedure to choose the parameters remains suitable and the parameters chosen in this work can be used as first attempt values.

4.5.3 Automated procedure

The procedure described in section 4.2.4 was automated via the graphical user interface (GUI) shown in Figure 75. This GUI constitutes the first of two windows of the diagnostics module of the HMP 1.1 software developed to provide a heath monitoring and prognostics (HMP) framework for performance analysis – see Appendix C. This first GUI is aimed at setting up fuzzy logic diagnostics models for a given engine. A second interface was implemented to operate the diagnostics models created in order to diagnose the possible faults – see section 4.9.

The first GUI of the diagnostics module, as shown in Figure 75, is able to set-up a diagnostics model given an engine model (Turbomatch), an operating condition and a search space. The main elements that must be specified are:

- *In the engine model set-up frame of the GUI*: engine model used, operating condition, selection of the measurement set (number and type).
- *In the search space definition frame of the GUI*: the ranges of variability of the performance parameters, the number of components simultaneously degraded and the increment value in the ranges.
- *In the system parameters definition frame of the GUI*: number, type and width of the input MFs, type and width of output MFs (Note that the increment value is defined with the search space).
- *In the functional parameters definition frame of the GUI*: AND operator, implication, aggregation, and defuzzification algorithms among the techniques discussed in 4.2.4.

Once these selections are made, a fuzzy logic inference system (FIS) is generated and saved. An ulterior frame of the GUI (in the right bottom corner) was designed to test FISs by simulating test data with implanted faults as well as measurement noise.

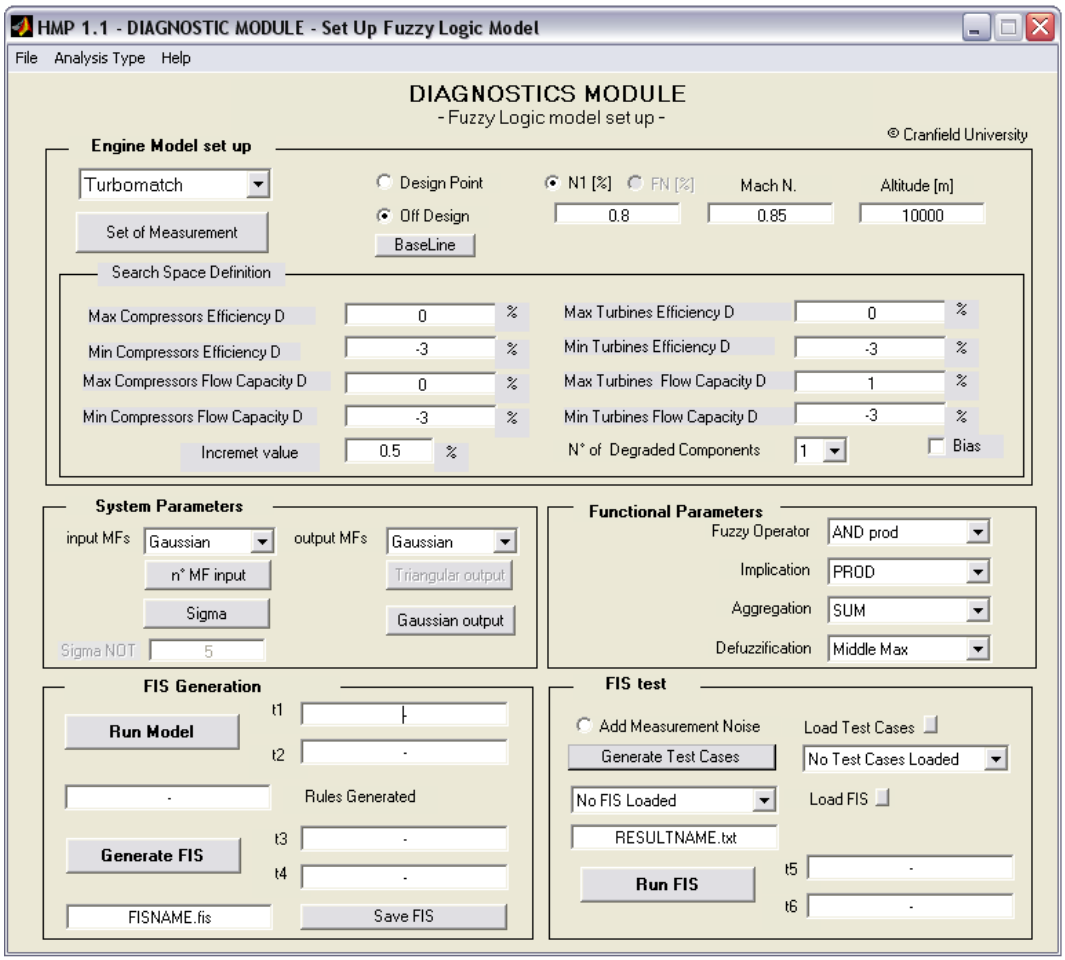


Figure 75: fuzzy diagnostic model set-up GUI of the HMP 1.1 Diagnostics Module

An ulterior feature of this interface is its capability of generating a diagnostics FIS able to diagnose component faults in the presence of systematic errors in the measurements (bias) while identifying the faulty sensor as well. A checkbox in the search space definition frame of the GUI enables the input of an ulterior system parameter called sigma NOT. This feature is discussed in detail in section 4.8.

4.5.4 Sensitivity study: strategy

4.5.4.1 Reasons of the study and anticipation of the results

Section 4.5.5 will present the results of a sensitivity study that was undertaken in order to identify our choice of optimal combination of system and functional parameters for an effective approximation capability of the diagnostics system. The approximation capability is defined as the ability of the method to model and approximate the functional relationship between sets of inputs (fault signature in the measurements) and the right sets of outputs (variations in the performance parameters), without considering, for the moment, the additional complication associated with measurement uncertainties. Subsequently, in section 4.6, noise is added to the test cases and our optimal selection of the

system parameters is modified accordingly, to achieve an enhanced accuracy of the diagnosis.

The sensitivity study, performed in order to assess the method's approximation capability, includes two sets of tests aimed at carrying out: (i) optimization of the functional parameters, and (ii) optimization of the system parameters. For the benefit of the reader, we anticipate here the results that are justified in the next section. Our choice of optimal functional parameters is the following:

- AND operator, implemented as: product.
- Implication method, implemented as: product.
- Aggregation method, implemented as: summation
- Defuzzification method, implemented as: centroid (centre of maximum as second best).

This choice identifies a SAM (standard additive model), as discussed 4.2.5.
On the other hand, the optimal selection of system parameters is:

- Gaussian MFs for input and output.
- Maximum N. of MF fixed to 500. It was found that the more input MFs are defined, the better; in fact this value is bigger than the number of input MFs that correspond to the combinations in the search space identified for a SFI capability. Nevertheless, in case of system with MFI capability, in the opinion of the author, a sensible value (e.g. 500) must be fixed to limit the computational burden.
- Width of MFs equal to 0.15 for the input MFs and equal to 0.5 for the output MFs. Note that the optimal value of the input MFs width to achieve an effective approximation capability is different to the case in which noise is added to the test cases. In the presence of noise, the optimal value is found to be different for each input measurement and it corresponds to the value of the sensor' noise RMS (under the assumption that the noise is normal distributed, as discussed in section 4.6).
- The number of output MFs is identified by the choice of the increment value in the search space. A smaller increment value is associated with a higher number of rules. Even though it is proved that this is advantageous for the accuracy of the system, it considerably reduces the speed of the calculation.

4.5.4.2 Description of the case studies

Test cases were generated, implanting 1771 combinations, deteriorating the 6 components independently (varying 2 parameters simultaneously) in the ranges of variability defined for the examined search space (see section 4.5.2.1) with an increment value of 0.2.
In the sensitivity study described in section 4.5.5, a first series of 16 tests were performed to identify the optimal system parameters. The test cases were used to assess the approximation capability of 16 different systems whose layouts were designed according to the combinations of functional parameters listed in Table 35. For these 16 systems, the system parameters were fixed to the following first guess values: Gaussian MFs in input and output, Maximum N. of

MF fixed to 500, width of input MFs equal to 0.25, width of output MFs equal to 0.5, increment value of the search space equal to 0.5%.

Table 35: combinations of functional parameters.

case	AND	Implication	Aggregation	Defuzzification
1	Product	Product	Summation	Centroid
2	Minimum	Product	Summation	Centroid
3	Product	Minimum	Summation	Centroid
4	Minimum	Minimum	Summation	Centroid
5	Product	Product	Maximum	Centroid
6	Minimum	Product	Maximum	Centroid
7	Product	Minimum	Maximum	Centroid
8	Minimum	Minimum	Maximum	Centroid
9	Product	Product	Summation	C.O.M
10	Minimum	Product	Summation	C.O.M
11	Product	Minimum	Summation	C.O.M
12	Minimum	Minimum	Summation	C.O.M
13	Product	Product	Maximum	C.O.M
14	Minimum	Product	Maximum	C.O.M
15	Product	Minimum	Maximum	C.O.M
16	Minimum	Minimum	Maximum	C.O.M

Once a best choice of functional parameters was found, it was kept unchanged in the subsequent tests: the second group of tests was undertaken using the same 1771 test cases to evaluate the optimal system parameters among the following possible selections.

- Input MFs type= Gaussian, Triangular.
- Input MFs width= 0.1, 0.15, 0.25, 0.5
- Output MFs type= Gaussian, Triangular.
- Output MFs width= 0.25, 0.5, 1%
- Increment Value= 0.25, 0.5, 1 %
- Input MFs number= 50, 100, 500

The strategy used to carry out these tests follows: starting from the first guess values of system parameters used in the first series of tests (Gaussian MFs in input and output, Maximum N. of MF fixed to 500, width of input MFs equal to 0.25, width of output MFs equal to 0.5, increment value of the search space equal to 0.5%), the changes listed in Table 36 were undertaken in sequence. For each change in system parameters, the system so generated was tested. The change was carried forward to the successive test only if it over-performed the results from the previous system.

Table 36: List of system parameters changes for the sensitivity study.

N.	Change in system parameters
1	Input MFs type changed to triangular (from Gaussian)
2	Output MFs changed to triangular
3	Input MFs width increased to 0.5 (from 0.25)
4	Input MFs width reduced to 0.15
5	Input MFs width reduced to 0.1
6	Output MFs width reduced to 0.25 (from 0.5)
7	Output MFs width increased to 1
8	Increment value increased to 1 % (from 0.5%)
9	Increment value reduced to 0.25 %
10	Input MFs number reduced to 100

4.5.4.3 Three methods to estimate the system accuracy

This section introduces three methods that were used to assess the performance parameters' estimation error and therefore the capability of a given diagnostics system to meet the requirements, as discussed below.

For each input set of 10 measurement deviations, 12 deviations in performance parameters are computed by the diagnostics process. The difference between the implanted deviation in each performance parameter and the corresponding calculated one is computed according to the following equation:

$$Delta = Implanted - Calculated \quad (4.71)$$

Method 1 - This method computes, for each test case, the max|Delta| (maximum value of |Delta|) calculated for the 12 parameters estimated. Then it assigns to this value different levels of severity according to its value. Three severity ranges were considered:

- Low severity (LS): max|Delta| < 0.5%
- Medium severity (MS): 0.5%<max|Delta| < 1%
- High severity (HS): max|Delta| > 1%

Therefore, with regards to the 1771 test cases generated, for each system assessed, the following quantities are calculated: number and % of MS cases and number and % of HS cases (the number and % of LS cases can be obviously deduced).
This method is aimed at evaluating local errors of the system in estimating the performance parameters, pointing out when in each test case the maximum error overcome fixed thresholds.

Method 2 - This technique is only used to assess SFI capabilities, when a fault is implanted in only one component at a time (two parameters simultaneously faulty). The 1771 test cases are divided into 6 groups, each of them characterised by having a different faulty component, being 6 the number of components. This method considers, in each group, only the two parameters affected by deterioration and computes the Deltas only for these two paramerers. Therefore, for each parameter in which deterioration is implanted this method computes:

- μ = the mean value of the Deltas across the group of test cases relative to the same deteriorated component.
- σ = the standard deviation of those Deltas.
- $CI_{95\%}$+ = μ + 1.96 σ, the corresponding 95% upper confidence limit.
- $CI_{95\%}$ - = μ - 1.96 σ, the corresponding 95% upper confidence limit.

This approach computes a local error because it considers only the parameters where the deterioration is implanted. It undertakes for these parameters a statistical analysis of the results and therefore it can be used to provide an expected accuracy of the system on them.

Method 3 - This method computes, for each test case, the RMS of the Deltas for the N=12 parameters estimated each calculation, according to the equation

$$RMS = \sqrt{\frac{\sum_{i=i}^{N} (Delta_i)^2}{N}} \qquad (4.\ 72)$$

The average value, mean(RMS)=RMS, of the RMSs calculates for all the test cases (1771 in the sensitivity study) is identified as a global parameter to estimate the accuracy of the diagnosis. This method is particularly useful to highlight a smearing tendency (see section 2.3) or else the propensity of some of the diagnostics methods to distribute the faults over many engine's components even when only a limited number of components are fault affected.

These three methods are employed in this work in the following circumstances:

- Method 1 and Method 3 are used in the sensitivity study reported in the next section (4.5.5) to provide a rapid way of estimating a global accuracy of each system assessed.
- Method 1, Method 2 and Method 3 are then used in section 4.6 to investigate in detail (local and global errors) the approximation capability of the fuzzy diagnostics system and successively its accuracy, in the presence of noisy measurements, for the diagnostics system with the chosen layout.
- Method 1 and Method 3 are used in section 4.7 to assess the partial MFI capability of the system.
- Method 1 and Method 3 are used in section 4.8 to study the capability of the system to diagnose faults in the components and in the sensors simultaneously. In these test cases measurement biasses as well single component faults are implanted.

4.5.5 Sensitivity study: results

4.5.5.1 Choice of the functional parameters

This section is dedicated to report the results of the first part of the sensitivity study to identify the best choice of functional parameters. The 16 different layouts listed in Table 35 (section 4.5.4.2) were investigated and the results are summarized in Table 37, being the case number in the two tables the same.

The table contains the results from two techniques to assess the diagnostics system accuracy: Method 1 and Methods 3 as defined in section 4.5.4.3. In the table, for each system, the results form Method 1 are the number (N.) and the percentage (%) of the cases with medium severity (MS) and high severity (HS) errors. Besides, Method 3 provides the average value of the RMS error, for the 1771 test cases.

Table 37: Results form Method 1 and 3 to assist the best choice of functional parameters.

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	**27 // 0.0152**	**0 // 0**	**0.048**
2	79 // 0.0446	2 // 0.0011	0.065
3	35 // 0.0198	0 // 0	0.084
4	96 // 0.0542	3 // 0.0017	0.097
5	43 // 0.0243	1 // 0.0005	0.058
6	48 // 0.0271	2 // 0.0011	0.060
7	57 // 0.0322	1 // 0.0005	0.068
8	106 // 0.0599	2 // 0.0011	0.079
9	**31 // 0.0175**	**8 // 0.0045**	**0.046**
10	103 // 0.0582	20 // 0.0113	0.055
11	80 // 0.0452	0 // 0	0.065
12	134 // 0.0757	6 // 0.0034	0.095
13	51 // 0.0288	7 // 0.004	0.074
14	49 // 0.027	8 // 0.0045	0.075
15	50 // 0.028	7 // 0.004	0.076
16	51 // 0.0288	10 // 0.0056	0.075

The outcome of this analysis highlighted two optimal combinations of functional parameters that show a minimum number of MS and HS cases and a minimum average value of RMS. These best layouts are to the cases 1 and 9 that correspond respectively to the following layout:

- **Best choice**: AND=Product, Implication=Product, Aggregation=Summation, Defuzzification=Centroid.
- **Second best choice**: AND=Product, Implication=Product, Aggregation=Summation, Defuzzification=Centre of Maximum.

The case 1 was selected as best choice because it showed: minimum number of MS and zero HS cases. As far as the RMS is concerned, case 1 does not over-perform case 9 that is considered to be the second best selection. Nevertheless the difference in RMS for the two systems is minimal. It is worthwhile noticing that the small value of RMS for the case 9 indicates a strong concentration capability on the actual fault.

4.5.5.2 Choice of the system parameters

The procedure to identify the most suitable combination of system parameters was presented in section 4.5.4.2. It consists of a sequence of 10 modifications to the first guess values. After each change in system parameter the layout was tested with the 1771 test cases introduced in section 4.5.4.2 and the change was carried forward to the successive layout only if it over-performed the results from the previous system. This procedure was applied starting from the best choice of layout identified in section 4.5.5.1. The outcome of this sensitivity study is summarized in Table 38.

Table 38: Results form Method 1 and 3 to assist the best choice of system parameters.

case	Method 1		Method 3	Set up time	Keep (K) / Reject (R) the change
	MS cases (N. // %)	HS cases (N. // %)	RMS		
1	339 // 0.1914	310 // 0.175	0.282	1 min, 12 sec	R
2	29 // 0.016	0 // 0	0.049	unchanged	R
3	305 // 0.1722	24 // 0.0136	0.112	unchanged	R
4	26 // 0.0147	0 // 0	0.045	unchanged	**K**
5	41 // 0.0232	4 // 0.0023	0.064	unchanged	R
6	26 // 0.0147	2 // 0.0011	0.048	unchanged	R
7	58 // 0.0327	2 // 0.0011	0.237	unchanged	R
8	334 // 0.1942	44 // 0.0248	0.129	23 sec	R
9	10 // 0.0056	0 // 0	0.117	4 min, 8 sec	R
10	28 // 0.0158	2 // 0.0011	0.055	1 min, 12 sec	R

The table case number corresponds to the layout change numbers of Table 36. Table 38 presents the results of the Method 1 and 3 (see section 4.5.4.3) and the set up time or else the time to generate a new fuzzy logic inference system, with the new layout, for search space under investigation. Eventually, in the last column of the table a sumble is reported that indicates whether the layout with the change over-performed (K) or not (R) the previous one, and consequently whether the change was kept (K) or rejected (R).

The following change was introduced in the system parameters:

- Input MFs width reduced to 0.15 (case 4), because reduces the number of MS cases and the RMS.

It is worthwhile noticing that the change associated with case 9 (Increment value reduced to 0.25 %) were not introduced, althought the corresponding system showed promising results. The reasons are that even though the corresponding number of MS cases appreciably reduces, the RMS increases indicating a higher tendency to smear the fault in the 12 parameters. Moreover,

the set-up time increases significantly. It is an ambition of this work to extend the SFI capability of the system to a MFI capability; therefore concerns about the set-up time are vital to enable this additional feature. In fact, the number of rules that needs to be generated increases dramatically in implementing a system able to identify more that 2 faulty components simultaneously, and so does the set-up time, accordingly.

Similarly, this procedure was applied starting from the layout with the second best combination of functional parameters that was identified in section 4.5.5.1, in order to complete the identification of a second optimal layout. The outcome of this second sensitivity study is summarized in Table 39. The table case number corresponds to the layout change number of Table 36.

Table 39: Results form Method 1 and 2 to assist the best choice of system parameters for the second optimal selection of the functional parameters.

case	Keep (K) / Reject (R) the change
1	R
2	R
3	R
4	**K**
5	R
6	R
7	**K**
8	R
9	R
10	R

The following two changes were introduced in the system parameters:

- Input MFs width reduced to 0.15 (case 4).
- Output MFs width increased to 1 (case 7)

4.6 SFI accuracy and tuning

This section is dedicated to a thorough analysis of the SFI accuracy of the fuzzy logic based diagnostic system in the following cases:

- To approximate and model the functional relationship between the input measurements (deviations) and the output performance parameters (deviations), without the additional complication of measurement errors. The best layout identified in section 4.5.5.2 is studied more in detail in section 4.6.1.
- To diagnose a fault in one component (SFI) in the presence of noise in the measurements. The accuracy of the system is tested and how this accuracy can be enhanced changing the input MFs amplitude according to realistic values of sensor noise RMSs is shown in section 4.6.2.
- To diagnose considerable changes in two health parameters of one component respect to a previously assessed deteriorated condition. A way of tuning the diagnostics system capable of SFI to estimate such changes and the method's accuracy is reported in 4.6.3.

4.6.1 Approximation capability: accuracy

In section 4.5.5.2 an optimal layout for a fuzzy diagnostics system was identified via a sensitivity study. The system has the following features:

- Functional parameters: AND=Product, Implication=Product, Aggregation=Summation, Defuzzification=Centroid.
- System parameters: Gaussian MFs in input and output, Maximum N. of MF fixed to 500, width of input MFs equal to 0.15, width of output MFs equal to 0.5, increment value of the search space equal to 0.5% (this identifies indirectly the output MFs number – see section 4.5.2.4)

This section presents a more in depth study of the accuracy of the devised diagnostics process by means of Method 1 and Method 3, introduced in section 4.5.4.3, in order to assess the system estimation error: This section is entirely dedicated to the analysis of the system's capability of approximating and model the functional relationship between inputs and outputs without considering measurement errors.

4.6.1.1 Accuracy results: Method 2

Figure 76 presents Deltas between implanted and calculated performance parameter deteriorations for the 1771 cases. For each case, efficiency and flow capacity changes were implanted simultaneously in one component at a time: starting from the FAN, on the left of the diagram, to the LPT on the right, the results are shown. Therefore, for each test case reported on the x axis, two values are plotted on the y axis: the corresponding Deltas (errors) in estimating the efficiency and the flow capacity of the component simulated as faulty (the name of the component appears on the top of the diagram for each group of test cases). For each component, a statistical analysis of the result was carried out according to Method 2 and summarized in Table 40.

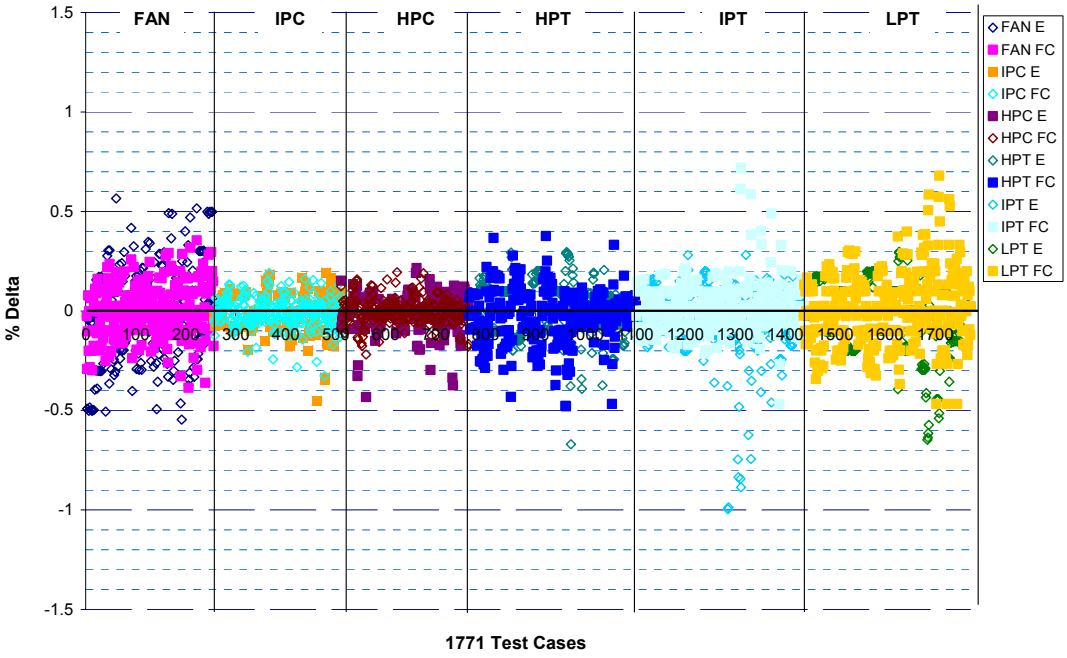


Figure 76: SFI capability of the diagnostics system. Results from 1771 test cases.

Table 40: Statistics of the diagnostics results, Method 2

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}
μ	-0.009	-0.003	-0.007	-0.007	-0.009	0.001	0.006	-0.026	-0.040	0.014	-0.032	0.017
σ	0.231	0.136	0.091	0.075	0.089	0.065	0.127	0.131	0.175	0.123	0.165	0.184
$CI_{95\%}$+	0.444	0.264	0.173	0.141	0.166	0.129	0.256	0.230	0.302	0.255	0.292	0.377
$CI_{95\%}$ -	-0.461	-0.269	-0.186	-0.154	-0.184	-0.127	-0.243	-0.282	-0.382	-0.227	-0.355	-0.344

For each component degraded, the table reports, for each health parameter: the mean value (μ) of the errors between the calculated and the implanted performance parameter changes, over the test cases relative to that specific component, the standard deviation (σ) of such an error, and the derived 95% confidence intervals ($CI_{95\%}$). For each parameter, it can be concluded that, with 95% confidence, the error is contained between $CI_{95\%}$+ and $CI_{95\%}$-.

4.6.1.2 Accuracy results: Method 3

A second performance parameters' estimation error is introduced by computing, for each test case, the RMS of the Deltas for the 12 parameters at each calculation, according to the procedure previously described in the Method 3. From this analysis, it results that the fuzzy logic system has a good accuracy in calculating the parameters not affected by the implanted faults, or else it has a good 'concentration' capability on the actual fault. The average value of the RMS error, for the 1771 test cases, resulted 0.045, which is a considerably low value.

4.6.1.3 Computational time required

One of the most favourable aspects of using fuzzy logic to implement a system capable of SFI, is its speed: once an automated set up procedure is designed (see GUI section 4.5.3) a new fuzzy diagnostics model can be quickly and easily set-up and it results equally fast when operated to diagnose a fault. The

computational time obviously depends on the computer used but sensible figures for a current average computational capability are listed in Table 41. The table reports the set-up time and the diagnostics time, and it relates them respectively to the number of rules to set-up and the number of test cases to diagnose. These represent the elements which the computational time has stronger dependency on. The diagnostics time for a single calculation is on the order of 0.1 seconds, as can be noticed in the table.

Table 41: Computational time with current computational capability

Processing	Time	Dependency
Set-up time	1 min, 12 sec	331 rules
Diagnostics Time	2 min, 50 sec (0.1 sec/case)	1771 test cases

4.6.2 Diagnostics capability in the presence of sensor noise: accuracy

The sensitivity study illustrated in section 4.5.5 provided us with two best choices of layout for a fuzzy diagnostics system designed and tested in order to approximate and model the diagnostics input-output functional relationship, as defined in section 4.5.2. This section studies how these two systems perform when they are demanded to diagnose a fault given in the presence of noise that affects the measurements. Moreover a way to enhance the diagnostics accuracy changing the input MFs amplitude according to sensor noise RMSs is discussed. The systems investigated have the following features:

- System 1 (best choice):
  - Functional parameters: AND=Product, Implication=Product, Aggregation=Summation, Defuzzification=Centroid.
  - System parameters: Gaussian MFs in input and output, Maximum N. of MF fixed to 500, width of input MFs equal to 0.15, width of output MFs equal to 0.5, increment value of the search space equal to 0.5% (this identifies indirectly the output MFs number – see section 4.5.2.4)
- System 2 (second best choice):
  - Functional parameters: AND=Product, Implication=Product, Aggregation=Summation, Defuzzification=Centre of Maximum.
  - System parameters: Gaussian MFs in input and output, Maximum N. of MF fixed to 500, width of input MFs equal to 0.15, width of output MFs equal to 1, increment value of the search space equal to 0.5%.

As far as the functional parameters are concerned, System 1 belongs to the category of SAM systems. It was proven in section 4.2.5.1 that such a system computes a conditional expectation and therefore an optimal non-linear estimation. On the other hand, System 2 is a quasi-SAM system: the main difference lies in the defuzzification algorithm implemented as centre of maximum (COM) function. Even though no formal proof of the optimality of its solution was provided, its suitability to the GPD problem will be demonstrated.

Random noise was added to the 1771 test cases used in this study: a random number that represents a realistic noise level, according to the type of sensor required, was added to the i-th element of the measurement set. The random number is generated as follows. Table 42 lists, for different type of sensors, realistic values of sensor noise standard deviations $SDTV_i$ in % of the measured value, being the noise assumed to follow a Gaussian distribution. For each measurement of the 1771 test cases, a random number is generated from a normal distribution with mean zero, and standard deviation $SDTV_i$, according to the value of the table. This random number represents the % deviation the corresponding measurement must be varied to simulate the noise.

Table 42: Sensor noise standard deviations in % of the measured value

Sensor type	$STDV_i$
Temperature	0.4%
Pressure	0.25%
Fuel Flow	0.5%
Shaft Speed	0.05%

Once the random component is added to the measurements of the 1771 test cases to simulate the presence of noise, they are used to test the System 1 and 2. Figure 77 represents the Deltas between implanted and calculated performance parameter deteriorations for the 1771 cases.

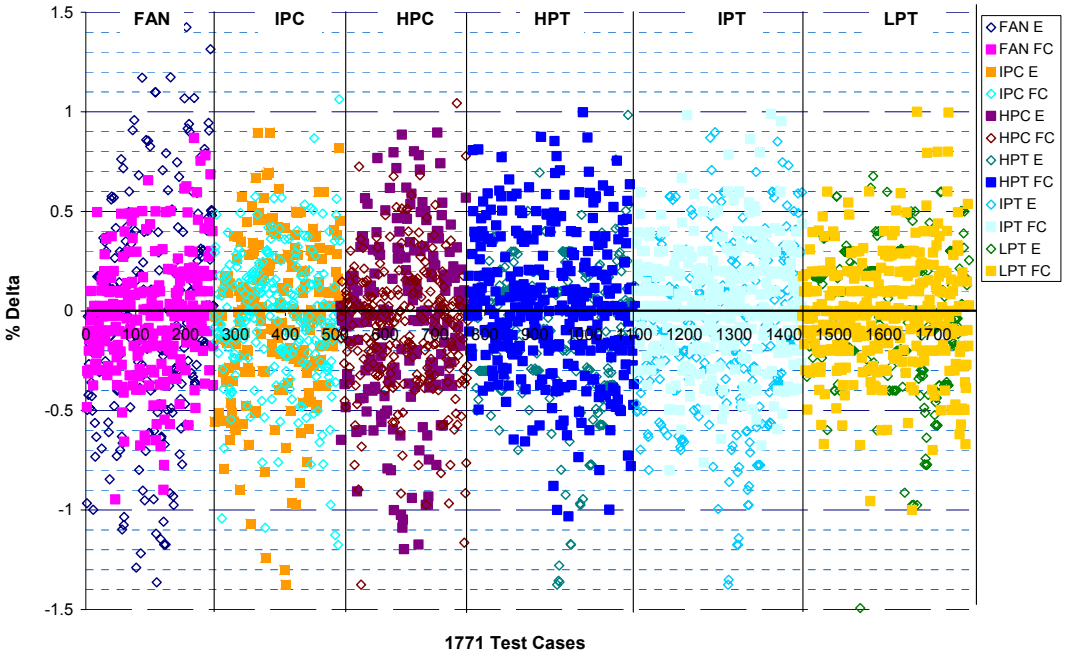


Figure 77: SFI capability of System 1. Results from 1771 test cases.

The test cases are divided into 6 groups characterised by a different faulty component. Figure 77 considers, in each group, only the two parameters affected by deterioration and shows the Deltas only for them. Moreover, for each parameter in which deterioration is implanted, Table 43 reports the statistical results according to Method 2. It can be noticed in Figure 77 how the

values of Deltas are much higher compared to the case without noise. This can be also observed in Table 43 where high values of **σ** are reported. The RMS increased as well up to 0.147 (Method 3) and the results showed 483 cases (27%) with MS errors and 105 cases (5.9%) with HS errors (Method 1) – see Table 44.

Table 43: Statistics of the diagnostics results for System 1, Method 2

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}
μ	-0.08	-0.03	-0.04	-0.05	-0.14	-0.09	-0.12	0.02	-0.08	0.04	-0.04	-0.01
σ	**0.64**	**0.30**	**0.39**	**0.35**	**0.58**	**0.34**	**0.41**	**0.37**	**0.41**	**0.30**	**0.33**	**0.29**
$CI_{95\%}$+	1.16	0.56	0.72	0.64	1.01	0.57	0.68	0.75	0.73	0.62	0.61	0.56
$CI_{95\%}$ -	-1.33	-0.62	-0.81	-0.74	-1.28	-0.75	-0.92	-0.70	-0.89	-0.54	-0.70	-0.58

Table 44: Summary of accuracy results for System 1 via Methods 1 and 3 over 1771 cases

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	483 // 0.27	105 // 0.059	0.147

To improve the system accuracy that resulted dramatically affected when noisy data were introduced, the input MFs amplitude were modified. Different values of amplitude were used for different inputs. The most suitable choice of input MFs amplitude (for the different measurements) was found to be the values of sensor noise standard deviation listed in Table 42.

The improved results obtained with System 1 with enhanced capability of dealing with noisy data are sown in Figure 78.

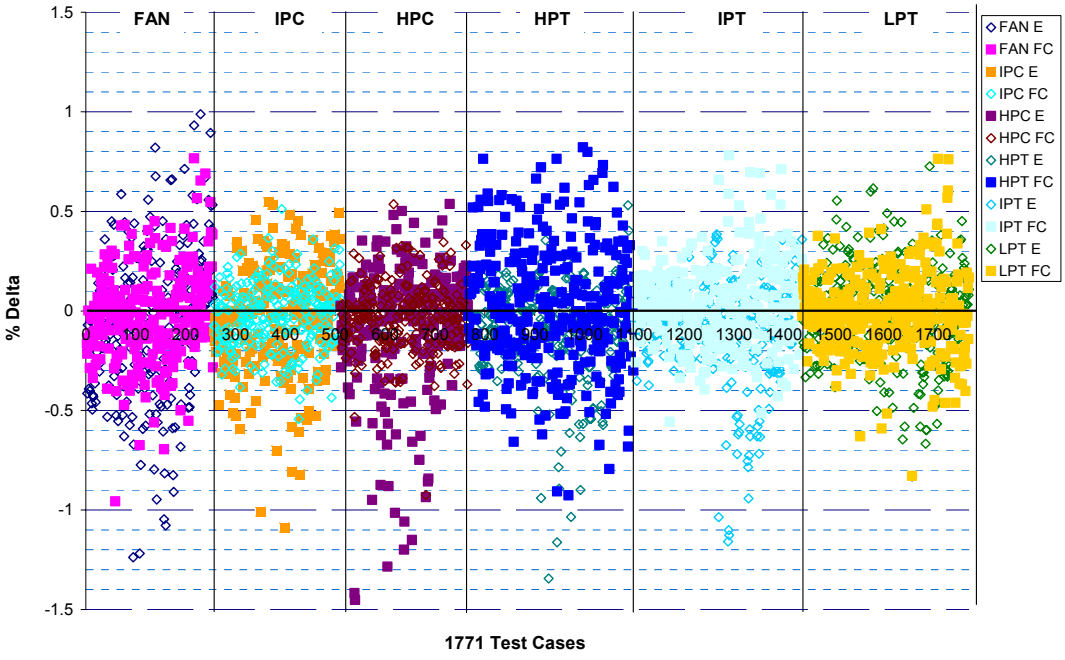


Figure 78: SFI capability of System 1 with enhanced capability of dealing with noisy data. Results from 1771 test cases.

The deltas are considerably more localised within the range between ± 0.5%, and considering that this is also the order of magnitude of the noise introduced in some of the measurements, it is in the opinion of the author, a positive outcome. The improvement can also be appreciated in Table 45 noticing the considerable reduction the values of **σ**. The RMS obtained with the enhanced system was reduced to 0.08 (Method 3) and the results showed 201 cases (11%) with MS errors and 33 cases (1.8%) with HS errors (Method 1) – see Table 46.

Table 45: Statistics of the diagnostics results for System 1 with enhanced capability of dealing with noisy data, Method 2

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}
μ	-0.07	-0.02	-0.03	-0.02	-0.12	-0.03	-0.09	0.02	-0.07	0.04	-0.02	-0.01
σ	**0.42**	**0.24**	**0.26**	**0.17**	**0.40**	**0.17**	**0.25**	**0.31**	**0.24**	**0.20**	**0.26**	**0.20**
$CI_{95\%}$+	0.75	0.46	0.48	0.30	0.67	0.29	0.41	0.64	0.40	0.44	0.49	0.39
$CI_{95\%}$ -	-0.89	-0.49	-0.54	-0.35	-0.90	-0.36	-0.58	-0.59	-0.55	-0.36	-0.53	-0.40

Table 46: Summary of accuracy results for enhanced System 1 via Methods 1 and 3 over 1771 cases

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	201 // 0.11	33 // 0.018	0.08

Due to the fact that System 1 and System 2, as defined at the beginning of this section, provided not dissimilar type of outcomes, it was considered here worthwhile studying the behaviour of System 2 in the presence of noise in the measurements, as well. In the same way that System 1 was adapted to deal with noisy data, also for System 2 it was necessary to change the amplitudes of the input MFs according to the noise level implanted. Figure 79 shows the results obtained with the enhanced System 2.

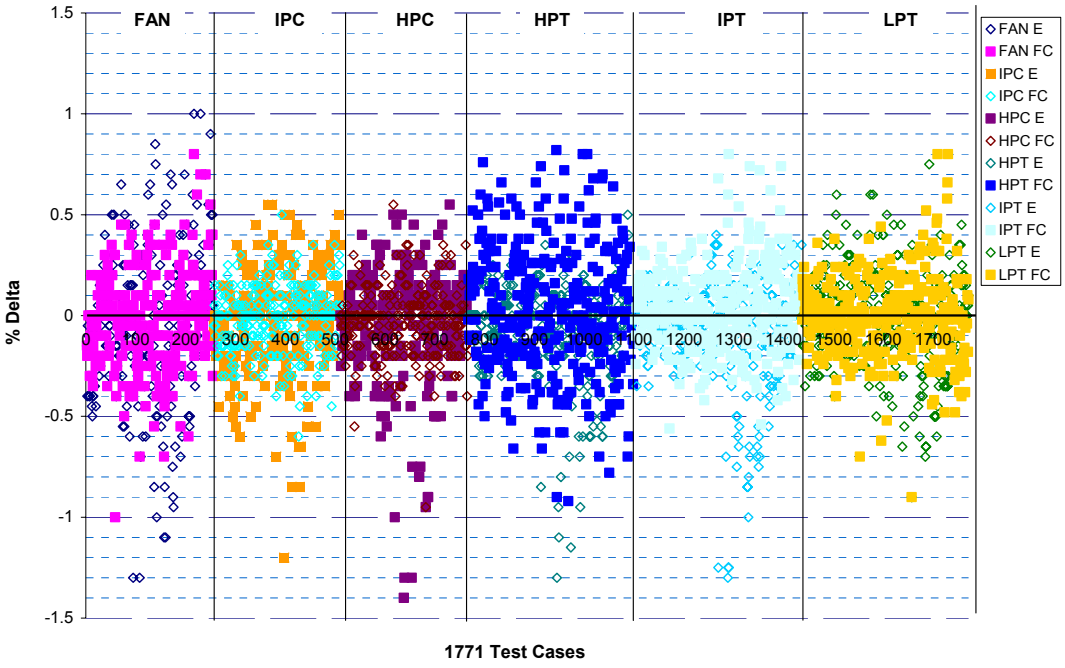


Figure 79: SFI capability of System 2 with enhanced capability of dealing with noisy data. Results from 1771 test cases.

The outcome as expected is similar to the one previously reported for the enhanced System 1. The values of **σ** detailed in Table 47 (Method 2) are comparable in magnitude to the values of Table 45 for the enhanced System 1 even though slightly worse. The RMS obtained with the enhanced System 2 calculated for the 1771 cases was equal to 0.09 (Method 3) but the results showed 183 cases (10%) with MS errors and 30 cases (1.6%) with HS errors over-performing the enhanced System 1 when evaluating the system accuracy with the Method 1 – see Table 48.

Table 47: Statistics of the diagnostics results for System 2 with enhanced capability of dealing with noisy data, Method 2

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}
μ	-0.06	-0.02	-0.03	-0.02	-0.10	-0.04	-0.08	0.02	-0.08	0.04	-0.02	0.00
σ	0.44	0.25	0.28	0.17	0.43	0.17	0.27	0.32	0.26	0.21	0.27	0.21
$CI_{95\%}$+	0.81	0.47	0.51	0.31	0.74	0.31	0.44	0.65	0.43	0.45	0.51	0.40
$CI_{95\%}$ -	-0.92	-0.51	-0.58	-0.35	-0.95	-0.38	-0.61	-0.60	-0.58	-0.37	-0.56	-0.41

Table 48: Summary of accuracy results for enhanced System 2 via Methods 1 and 3 over 1771 cases

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	183 // 0.10	30 // 0.016	0.09

4.6.2.1 Remarks

Concluding in this section an important milestone in this project was proven. Two fuzzy system layouts were identified as capable of performing SFI capability in the presence of noisy measurements and their accuracy was evaluated with the three different Methods introduced in section 4.5.4.3. The enhanced System 1 over-performed the enhanced System 2 in the accuracy tests provided by the Methods 2 and Method 3 but under-performed when the accuracy was estimated with the Method 1.

4.6.3 Tuning capability to enhance the SFI role in GPD

A SFI system is used to evaluate considerable changes in only 2 performance parameters of one component. The application of a SFI approach in a real life case becomes useful under the assumption that only one component is faulty at a time. This assumption becomes more realistic if the changes are estimated in a short period of time, or else the diagnosis is made to assess only changes in the performance parameters from a very recent known condition. In fact if on the contrary the time scale increases, it is more likely that two or more gas-path components are degraded.

These considerations create a new opportunity of using SFI systems coupled with MFI systems (e.g. linear estimation methods). MFI approaches are limited when estimating considerable changes (i.e. > 1%) but are advantageous when calculating small deteriorations that inevitably affect all the parameters simultaneously over the engine operating time. The procedure represented in Figure 80 is an attempt to suggest how this coupling could be implemented. The procedure described relies on the idea that a SFI system and a MFI system

compute a solution in parallel every mission of the engine. The two systems at the flight n calculate deltas in measurements from a baseline that does not referes to the clean engine condition but that is calculated taking into account the global deterioration level accumulated up to the flight n-1. Therefore the two systems do not calculate the absolute changes in performance parameters respect to a clean engine but the relative changes respect to the deteriorated condition evaluated at the previous flight. The relative changes computed at the flight n are then added to the global deterioration level at the flight n-1 to obtain the absolute changes respect to the clean condition.

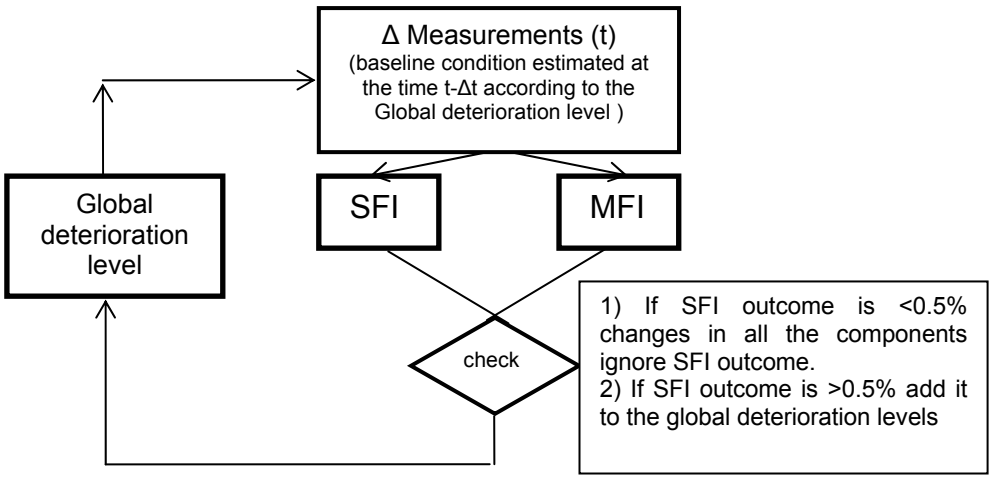


Figure 80: MFI and SFI coupling

Let us assume that at the flight number one the engine is clean and no deterioration is detected. At a given point in time (flight n) the MFI system detects small deteriorations in all the performance parameters, no considerable changes (<0.5) are detected by the SFI and therefore it is ignored. At the flight n+1 instead something happens and 1 component gets severely damaged. The SFI estimates changes > 0.5% (In a real application the value 0.5% should be replaced with a more correct value obtained by validating the suggested procedure). Therefore, the SFI outcome is used to update the global deterioration level instead of the MFI result.

In the light of this proposed framework, in this work an automated procedure (see GUI 4.5.3) was devised to tune the rules of the fuzzy diagnostics system on top of a known baseline characterized by deterioration levels for all the 12 performance parameters. This baseline is assumed to be calculated at the previous flight with a MFI method and represents the global deterioration level in Figure 80. Let us assume, for example, that the values listed in Table 49 represent the baseline of deterioration. The SFI is now required to assess whether there are considerable changes from this already existing level of deterioration.

Table 49: Global deterioration level, baseline

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
-0.5	-0.4	-0.2	-0.5	-0.3	-0.2	-0.3	0.5	-0.4	0.3	-0.6	0.5

The results shown in Figure 81 are obtained by using the enhanced System 1 as defined in the previous section. However, the rules were now tuned to the baseline of Table 49. A new set of 1771 test cases were generated with fault implanted in the ranges defined by the search space identified in section 4.5.2.1 but superimposed on the global deteriorations of Table 49; the measurements calculated running the engine model were disturbed adding a random component according to the same procedure described in the previous section. It is important to observe that these results cannot precisely (i.e. case by case) be compared to the results from the previous set of test cases because having added a random component the two sets could have slightly different severity of noise level. But a comparison can be made looking at the statistical figures. Table 50 presents analogous results to Table 47 (Method 2). The RMS obtained with the tuned diagnostics system calculated for the 1771 cases was equal to 0.089 (Method 3) and the results showed 172 cases (9%) with MS errors and 22 cases (1.2%) with HS errors (Method 1) – see Table 51.

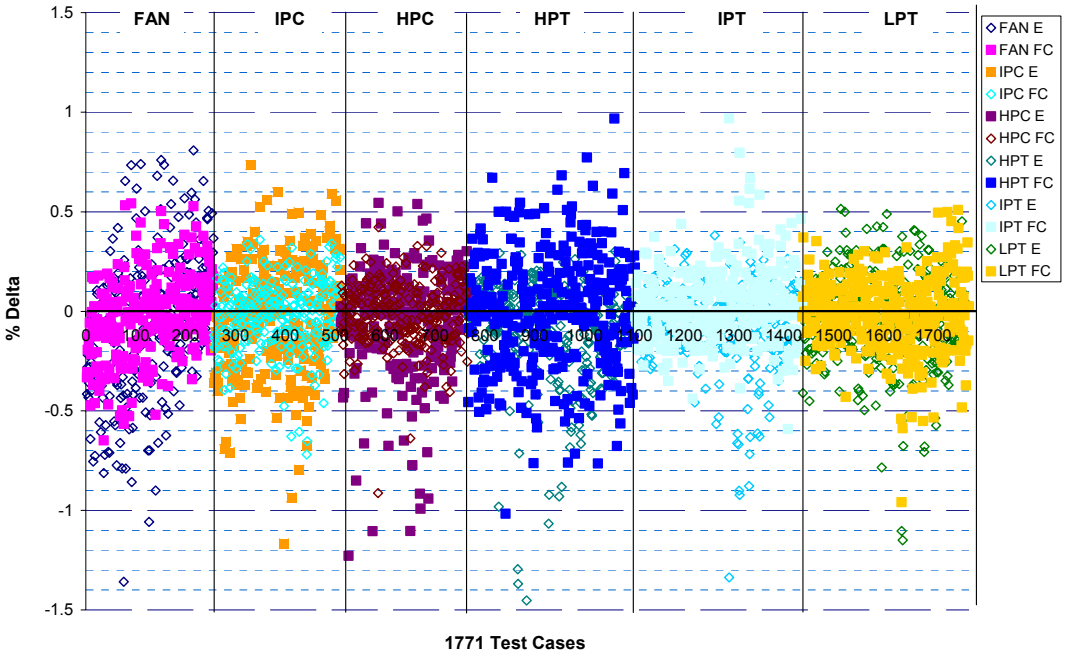


Figure 81: SFI capability of the tuned enhanced System 1. Results from 1771 test cases.

Table 50: Statistics of the diagnostics results for tuned enhanced System 1, Method 2

	η_{FAN}	Γ_{FAN}	η_{IPC}	Γ_{IPC}	η_{HPC}	Γ_{HPC}	η_{HPT}	Γ_{HPT}	η_{IPT}	Γ_{IPT}	η_{LPT}	Γ_{LPT}
μ	-0.07	-0.02	-0.03	-0.02	-0.07	-0.01	-0.09	0.00	-0.06	0.02	-0.04	-0.02
Σ	**0.36**	**0.22**	**0.28**	**0.18**	**0.32**	**0.16**	**0.27**	**0.30**	**0.20**	**0.19**	**0.23**	**0.19**
$CI_{95\%}$+	0.64	0.40	0.52	0.32	0.56	0.30	0.44	0.58	0.34	0.39	0.41	0.35
$CI_{95\%}$ -	-0.78	-0.45	-0.58	-0.37	-0.70	-0.32	-0.62	-0.58	-0.46	-0.34	-0.50	-0.38

Table 51: Summary of accuracy results for tuned enhanced System 1 via Methods 1 and 3 over 1771 cases

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	172 // 0.09	22 // 0.012	0.089

4.7 A fuzzy diagnostics system able of a partial MFI capability

In section 4.6.3 it was discussed how a SFI system can be used in a real life application to evaluate considerable changes in only 2 performance parameters, under the assumption that only one component can become significantly faulty in the considered time interval. It was recognised that this assumption becomes more realistic if the diagnosis is made to assess only changes from a very recent known condition. In fact if on the contrary the time scale increases, it is more likely that two or more gas-path components are degraded. With the intention of making the procedure summarized in Figure 80 more robust, in this section a fuzzy diagnostics system with a partial MFI capability was devised, to substitute the SFI process in the coupling procedure (Figure 80).

The process with a partial MFI capability is in principle similar to the SFI systems described so far. It is able to quantify considerable deviation in performance parameters and it uses the non-linear approach based on fuzzy logic. Moreover, it is able to quantify changes in more that two parameters simultaneously: in this work the system was tested with up to 2 components degraded at a time, which implies 4 parameters simultaneously deteriorated. In the context of section 4.6.3, this allows to relax the previously stated assumption requiring that no more than 2 components can become considerably degraded in 1 mission.

4.7.1 System Layout

A fuzzy diagnostics system with a partial MFI capability was devised in this work for the Trent 800 engine. The inputs and outputs of the diagnostic process are the same shown in Figure 71 (section 4.5.2). The system is designed to assess performance parameters percentage changes from a clean engine condition (12 outputs) given the knowledge of the measurement changes (10 inputs) calculated as percentage deviations with respect to a baseline determined by means of an engine model run at the specific power setting and environmental conditions (defined in section 4.5.2).

This section describes a system able to quantify considerable changes in up to 2 components degraded simultaneously (4 performance parameters) according to the considerations made in section 4.2.6 – see Table 9. The search space was defined as follows:

- Maximum variation in compressors' efficiencies = -1%
- Minimum variation in compressors' efficiencies = -3%
- Maximum variation in compressors' flow capacities = -1%
- Minimum variation in compressors' flow capacities = -3%
- Maximum variation in turbines' efficiencies = -1%
- Minimum variation in turbines' efficiencies = -3%
- Maximum variation in turbines' flow capacities = -1%
- Minimum variation in turbines' flow capacities = -3%

besides, the following additional parameters were fixed:

- Number of components simultaneously deteriorated = 2
- Step of increment= 0.5%
- Number of rules 19440

To limit the number of rules and therefore the complexity of the system no rules were stated to provide the input–output functional relationship corresponding to fault levels between 0% and -1%. Note that even though the ranges in the search space are defined between -1% and -3%, the 0% fault level is always included in the search space. Therefore, the above definition of search space only excludes the -0.5% fault level compared to the search space defined in section 4.5.2. This choice slightly affects the accuracy at low deterioration levels (around 0.5%) but it was recognised that a higher accuracy is required when assessing higher changes in the performance parameters (e.g. 3%). Besides, in this work a strong commitment was devoted to meet the requirement of devising a fast system for on-wing applications, and therefore a reduction in the number of rules (excluding the -0.5% fault level) was driven by time related concerns.

4.7.2 Partial MFI capability: results

4.7.2.1 Test cases

A series of 1201 test cases resulting from the combinations of 3 fault levels (0, -1.2, -2.7) taken 4 at a time (being 4 parameters deteriorated at a time) was generated. A random component was added to the measurements of the test cases to simulate the presence of noise, according to the procedure described in section 4.6.2.

4.7.2.2 Results: accuracy and computational time

Method 1 and 3 introduced in section 4.5.4.3 were used here to assess the system accuracy in performing a partial MFI capability. The RMS obtained considering only the 12 outputs relative to the performance parameters, for the 1201 cases, was equal to 0.1123 (Method 3) and the results showed 201 cases (16.7%) with MS errors and 70 cases (5.8%) with HS errors (Method 1) – see Table 52.

Table 52: Summary of accuracy results for System 1 via Methods 1 and 3 over 1201 cases

case	Method 1		Method 3
	MS cases (N. // %)	HS cases (N. // %)	RMS
1	201 // **0.1674**	70 // **0.0583**	0.1123

A typical result, in addition to the 1201 cases, is presented in Table 53 and Table 54. Table 53 lists the implanted faults in the FAN and HPC. The 12 outputs of the diagnostics system are shown in Table 54: the results are in trend with the implanted faults. A remarkable concentration capability of the fuzzy diagnostics system can be noted.

Table 53: Implanted deterioration (partial MFI)

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
-1.8	**-2.2**	0	0	**-2.3**	**-2.7**	0	0	0	0	0	0

Table 54: Estimated deterioration (partial MFI), typical result

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
-1.51	**-2.43**	-0.01	0.00	**-2.38**	**-2.54**	0.00	0.02	0.00	-0.00	0.01	0.03

As far as the computational time is concerned, Table 55 reports the set-up time and the diagnostics time, and relates them respectively to the number of rules stated and the number of test cases diagnosed, that represent the elements which the computational time has stronger dependency on. The system, designed to provide a partial MFI capability, requires a considerably increased number of rules (19440 in this example) that inevitably affects the computational time. The diagnostics time for a single calculation is approximately 12 seconds, about 100 times the time required by the corresponding system with SFI.

Table 55: Computational time with current computational capability

Processing	**Time**	**Dependency**
Set-up time	18 min, 35 sec	19440 rules
Diagnostics Time	240 min (12 sec/case)	1201 test cases

4.8 A bias-tolerant fuzzy diagnostics system

4.8.1 The methodology: the use of the logical operator NOT to deal with biasses

This section is dedicated to illustrate a novel procedure proposed in this work to enable the fuzzy diagnostics system to perform SFI when one sensor is affected by a systematic error (bias). The bias-tolerant system is able to quantify the fault in the gas-path component while isolating the biased measurement. A bias as discussed in Chapter-2 is defined as a constant systematic error in a measurement, in this work we consider biasses as shifts in a measured parameter greater than 5 times the noise standard deviation of its corresponding sensor – see Figure 13.

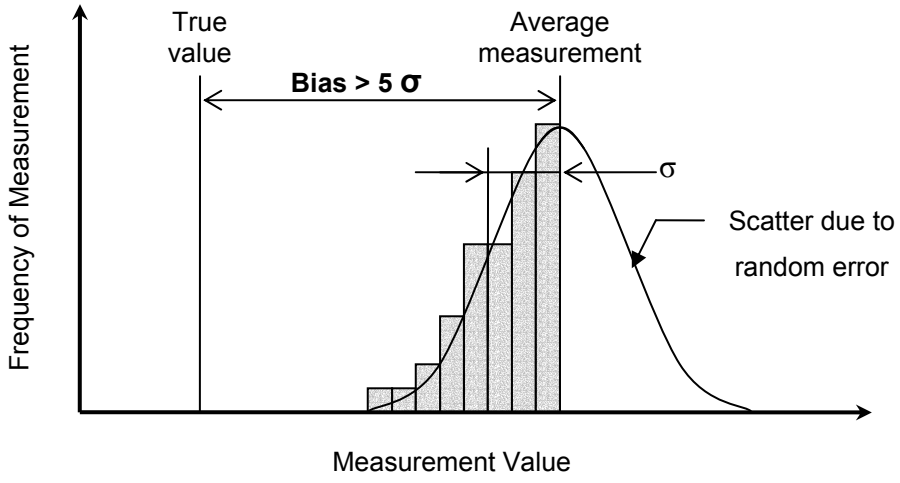


Figure 82: Measurement Bias and Precision error

4.8.1.1 The use of the NOT operator

To describe this procedure we refer again to the simplified example, which was introduced in section 4.2.4. A single shaft turbojet with 3 measurements (shaft speed N, fuel flow FF, and exhaust gas temperature EGT) and 4 gas-path performance parameters (compressor efficiency and flow capacity, and turbine efficiency and flow capacity) is considered – see Figure 45.

Figure 83 provides a fuzzy inference diagram that shows the graphical representation of two particular rules: rule 1 and rule 2. In the diagram we can distinguish the if-part of the rules, with the fault signature in the measurements represented with input MFs, and the then-part of the rules, with MFs of the performance parameters that characterise the fault.

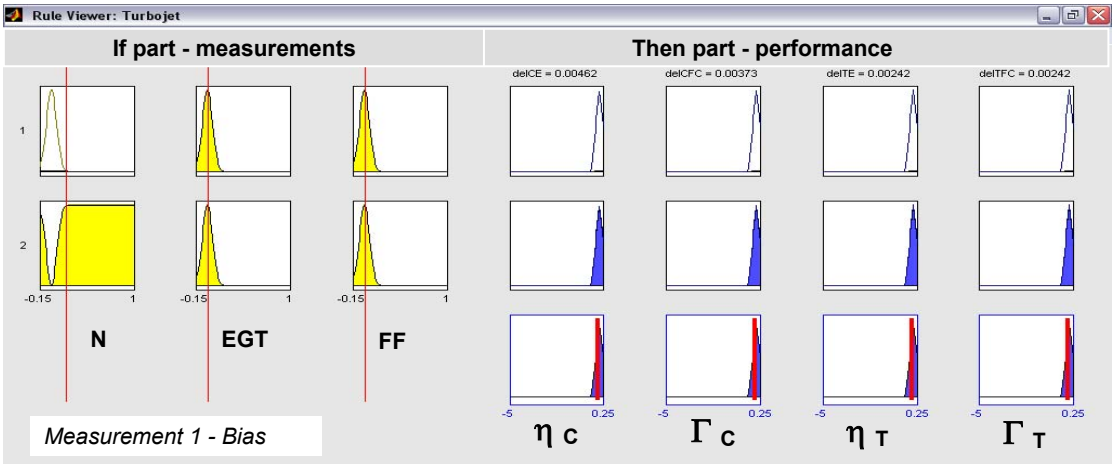


Figure 83: Dealing with Bias in the fuzzy inference diagram

The procedure described below, is aimed at designing a system that, in the case that one measurement (the shaft speed N in the example) is biased, it is able to:

- Recognise the fault signature with the remaining non-biased measurements.
- Isolate the biased measurement.

It will be shown how the information required to do so is taken into account in the rule 2 of Figure 83. The system can perform these two tasks only in the hypothesis that it is possible to diagnose the fault with a reduced number of measurements.

To achieve this goal, a method was designed that uses the NOT logical operator (rule 2 of Figure 83). In fuzzy logic the NOT operator is such that, when applied to a MF, generates its complement. In the second rule in Figure 83 the NOT operator is applied to the MF of the first measurement, N. The rule 2 states that: if N is **NOT** Mf1 AND EGT is Mf2 AND FF is Mf3 then the MFs that characterise the fault are the same as rule 1 (that means same fault), and in addition a bias is present in the first measurement – see Table 56. So, in the case of the example in Figure 83, where the measurement 1 is biased, rule 2 has a higher degree of activation than that of rule one. It is worthwhile reminding that Mf1, Mf2 and Mf3 were defined by running an engine model with the implanted fault associated with rule 1 and by centring for each input set a MF in the values of the computed measurement deltas – see section 4.2.4.

The learning algorithm discussed in section 4.2.4, that provides the formulation of the rules, is modified as it follows – see Table 56: for each rule defined without dealing with biasses, an additional rule is stated for each input measurement. Such a rule contains the complement of the MF that refers to that biased measurement (Mf1 in the rule 2). This is computed by means of the NOT operator, indicated in Table 56 with the minus symbol. Besides, the additional information regarding the location of the bias is provided to complete the rule. In Table 56, the last three columns refer to the three input measurements respectively: 0 is associated with no-bias, while 1 is associated with a bias in the corresponding measurement (being the i-th column relative to

the i-th measurement). In the rules statement of Table 56, 0 and 1 represent respectively a triangular output MF centred in 0 and one centred in 1.

Table 56: New formulation of rules to deal with bias

If part				then part				bias			
Mf1	Mf2	Mf3	,	mf1	mf2	mf3	mf4	0	0	0	*no bias*
- Mf1	Mf2	Mf3	,	mf1	mf2	mf3	mf4	1	0	0	
Mf1	-Mf2	Mf3	,	mf1	mf2	mf3	mf4	0	1	0	
Mf1	Mf2	-Mf3	,	mf1	mf2	mf3	mf4	0	0	1	

It is clear in Table 56 that the new fuzzy diagnostics system has the same 3 inputs (deltas in the measurements) but 7 outputs (instead of 4). Three additional outputs are introduced: they are, for each measurement, a binary code that is equal to 0 if that measurement is recognised as not biased or 1 if a bias is identified.

4.8.1.2 The use of the centre of maximum to devise a bias-tolerant system and the role of the parameter A

A procedure that employs the centre of maximum defuzzification method to design a bias-tolerant diagnostics system is described below. This defuzzification technique proved to be effective in isolating the bias when applied together with the layout described above.

The enhanced System 2, whose design relies on a defuzzification algorithm based on the centre of maximum function, as described in section 4.6.2, was used here. It was proven in section 4.6.2 that the SFI capability of the enhanced System2 in the presence of noisy data was comparable to the best performer, namely enhanced System 1, and therefore it was considered sensible to use the enhanced System2 in this application.

The layout for the 3 inputs and the first 4 outputs (delta in performance parameters) is unchanged. As far as the additional 3 outputs aimed at identifying the bias location are concerned, for each of them only two triangular MFs were designed. One MF is centred in 0 and another is centred in 1, their amplitude is 0.5. Each rule associated with a bias in a measurement has, for that measurement location, the MF centred in 1. When such a rule is partially but sufficiently activated, the centre of maximum defuzzification method will provide in output the centre of the MF centred in one identifying the bias correctly.

Summarizing, according to the above, for each rule that corresponds to an implanted fault level we have to state:

- A rule with no biased measurements associated, which contains:
  - In the if-part of the rule: the fault signature in the measurements in terms of MFs (as in the normal SFI system).
  - In the then-part of the rule: two elements. A first part witht the MFs associated with the fault (as in the normal SFI system). A second part with the additional outputs concerning the location of the bias

with all the MFs centred in zero (typical of a bias-tolerant system – see Table 56).

- A number of additional rules equal to the number of measurements, which contain:
  - In the if-part of the rule: the fault signature in the measurements in terms of MFs (as in the normal SFI system), with the variation of the complement of the MF that refers to that biased measurement (-Mf1 in rule 2 of Figure 83 and Table 56). Note that the amplitude of such a MF is related to the noise according to the procedure described in section 4.6.2.
  - In the then-part of the rule: two elements. A first part witht the MFs associated with the fault (as in the normal SFI system). A second part with the additional outputs concerning the location of the bias, with the output MF that refers to the biased measurement centred in 1 while the others all centred in zero (typical of a bias-tolerant system – see Table 56).

In practice, in the implementation of this idea an additional parameter 'A' was introduced. In each rule stated with a biased measurement, the NOT operator is not directly applied the to the input MF (Mf1 in rule 2 of Table 56) of the biased input, that would have been generated in absence of bias, but on the contrary it is applied to a related MF pertinently modified as shown in Figure 84. The amplitude of such a MF (Mf1 in rule 2 of Table 56 and MF1 in Figure 84) is first multiplied for a parameter A that typically is chosen equal to 5 to create a wider MF with the same centre (MF1A). Then, the complement is computed (-MF1·A). Therefore, Table 56 changes accordingly into Table 57.

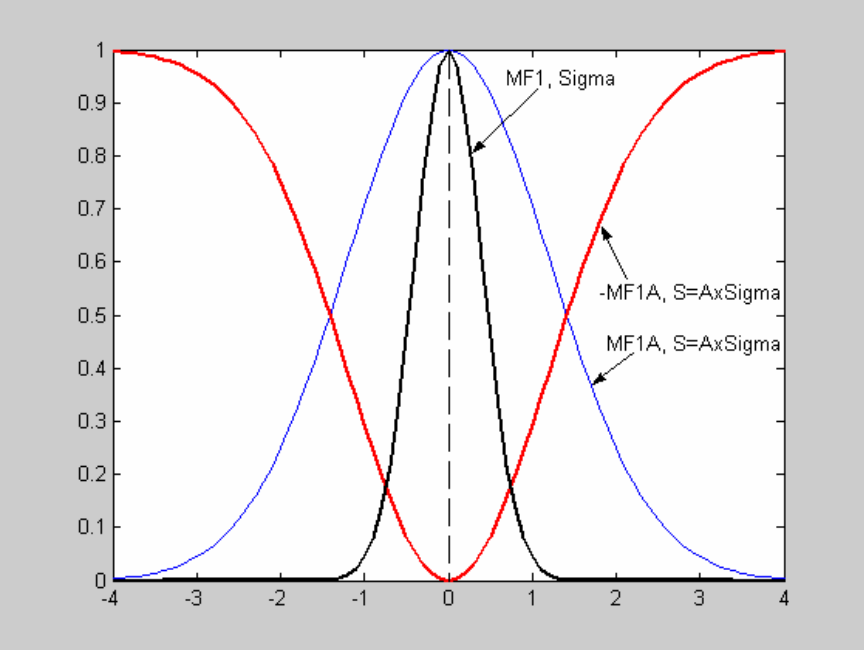


Figure 84: Procedure to calculate the NOT operator used in the learning algorithm

Table 57: New formulation of rules to deal with bias introducing the parameter A

If part				then part				bias			
Mf1	Mf2	Mf3	,	mf1	mf2	mf3	mf4	0	0	0	*no bias*
- Mf1A	Mf2	Mf3	,	mf1	mf2	mf3	mf4	1	0	0	
Mf1	-Mf2A	Mf3	,	mf1	mf2	mf3	mf4	0	1	0	
Mf1	Mf2	-Mf3A	,	mf1	mf2	mf3	mf4	0	0	1	

Note that in the GUI described in section 4.5.3, shown in Figure 75, the value of the parameter A is provided through the fuzzy system parameter called sigma-NOT.

4.8.2 Results dealing with bias: case studies

4.8.2.1 System definition

The procedure illustrated above for the simplified example of a single shaft engine, was then integrated in the fuzzy-logic diagnostic method developed for the Trent 800. This section analyses the capability of dealing with measurement biasses while diagnosing engine degradations, of a system devised for single component fault isolation (SFI). The search space was defined as:

- Maximum variation in compressors' efficiencies = 0%
- Minimum variation in compressors' efficiencies = -3%
- Maximum variation in compressors' flow capacities = 0%
- Minimum variation in compressors' flow capacities = -3%
- Maximum variation in turbines' efficiencies = 0%
- Minimum variation in turbines' efficiencies = -3%
- Maximum variation in turbines' flow capacities = 1%
- Minimum variation in turbines' flow capacities = -3%

besides, the following additional parameters were fixed:

- Number of components simultaneously deteriorated = 1
- Step of increment= 0.5%
- Number of rules 3641

The number of rules increases from 331 of the SFI system developed in section 4.5.2 up to to 3641 for the corresponding bias-tolerant SFI system. For each rule developed without bias 10 additional rules must be stated to take into account the potential presence of a bias in each measurement (331x(10+1)=3641).

The system has 10 inputs (unchanged from the layout of section 4.5.2) but 22 outputs (instead of only 12), among which:

- 12 outputs relative to the 12 performance parameter changes.
- 10 outputs relative to the location of the bias in one of the 10 measurement.

4.8.2.2 Test cases

A series of 275 test cases was generated by implanting all the combinations of two fault levels (i.e. -1.3 and -2.7) in the 6 gas-path components, considered one at a time. The number of test cases (275) can be obtained considering that:

- The tests include 2^2x6= 24 combinations of 2 values of deterioration, i.e. -1.3 and -2.7, for the 2 performance parameters of the 6 components degraded one at a time, plus the additional condition of clean engine.
- These 25 (24+1) combinations are repeated 11 times (1 without bias and 10 respectively with one of the 10 measurements biassed at a time).
- Therefore, (2^2x6+1)x11=275 test cases.

A random component was added to the simulated measurements of the 275 test cases to reproduce the presence of noise, according to the procedure described in section 4.6.2.

4.8.2.3 Results: accuracy and computational time

Method 1 and 3 introduced in section 4.5.4.3 were used here to assess the system accuracy. The RMS obtained considering only the 12 outputs relative to the performance parameters, for the 275 cases, was equal to 0.1435 (Method 3). The results showed 63 cases (22.9%) with MS errors, 22 cases (1.2%) with HS errors (Method 1) and 12 biasses (5.7%) unidentified – see Table 58.

Table 58: Summary of accuracy results for bias-tolerant enhanced System 1 via Methods 1 and 3 over 275 cases

case	Method 1		Method 3	Misidentified biasses (N. // %)
	MS cases (N. // %)	HS cases (N. // %)	RMS	
1	63 // **0.229**	22 // **0.012**	0.1435	9 // **0.032**

Another test case, in addition to the 275 cases, is examined here, as representative of a typical result - see Table 59 and Table 60. The implanted fault, listed in Table 59, characterised by a component and a sensor fault superimposed. The component fault is within the IPC. The sensor fault (bias) affects the FF measurement: a shift equal to 5 times the noise standard deviation assumed for the FF measurement (with reference to Table 42) was added to that measurement as calculated by the engine model. A random component (noise) was added to the 10 measurements, as well.

The 22 outputs of the diagnostics system are shown in Table 60. The biased sensor was identified and the fault in the IPC was quantified with a reasonable accuracy.

Table 59: Implanted Fault (in IPC) and Bias (in FF sensor)

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
0	0	**-1.3**	**-2.7**	0	0	0	0	0	0	0	0

N2	N3	FF	P13	P25	P3	T25	T3	T45	T5
0	0	**1**	0	0	0	0	0	0	0

Table 60: Estimated (bias-tolerant System) Fault and Bias, typical result

$\Delta\eta_{FAN}$	$\Delta\Gamma_{FAN}$	$\Delta\eta_{IPC}$	$\Delta\Gamma_{IPC}$	$\Delta\eta_{HPC}$	$\Delta\Gamma_{HPC}$	$\Delta\eta_{HPT}$	$\Delta\Gamma_{HPT}$	$\Delta\eta_{IPT}$	$\Delta\Gamma_{IPT}$	$\Delta\eta_{LPT}$	$\Delta\Gamma_{LPT}$
0.00	-0.01	**-1.48**	**-2.91**	0.11	0.00	0.00	0.02	0.00	-0.00	0.01	0.03

N2	N3	FF	P13	P25	P3	T25	T3	T45	T5
0	0	**1**	0	0	0	0	0	0	0

As far as the computational time is concerned, Table 61 reports the set-up time and the diagnostics time and relates them respectively to the number of rules stated and the number of test cases diagnosed. The bias-tolerant SFI system requires a considerably increased number of rules that inevitably affects the computational time. The diagnostics time for a single calculation is approximately 1 second, about ten times the time required by the corresponding SFI system without the capability of dealing with biasses.

Table 61: Computational time with current computational capability, bias-tolerant system

Processing	Time	Dependency
Set-up time	3 min, 20 sec	3641 rules
Diagnostics Time	4 min, 50 sec (1.05 sec/case)	275 test cases

4.8.3 Remarks

A procedure was described to implement a bias-tolerant fuzzy diagnostics system capable of quantifying a considerable change in two performance parameters of one faulty component (SFI) while isolating a biased sensor.
The procedure uses the enhanced System 2 developed in section 4.5.2 exploiting its centre of maximum defuzzification method. Besides, the NOT operator and the 'A' parameter are introduced to state rules with superimposed biased measurements. An application of the methodology to the Trent 800 was presented. The diagnostics system has 3641 rules and provided promising results. The ambition of this work is not to state that this is the optimal algorithm to devise a bias-tolerant diagnostics system using fuzzy logic. Nevertheless, in this section it was suggested a possible way of using fuzzy logic technology to cope with the different difficulties that affect the GPD problem in real life applications – namely sensor noise, sensor biasses and small number of measurements available.

4.9 Operating the diagnostics model through the GUI

The HMP 1.1 diagnostics module is constituted by two GUIs, the first presented in Figure 75 of section 4.5.3, was devised to automatically set up a fuzzy diagnostics model. Figure 85 shows the second graphical user interface.

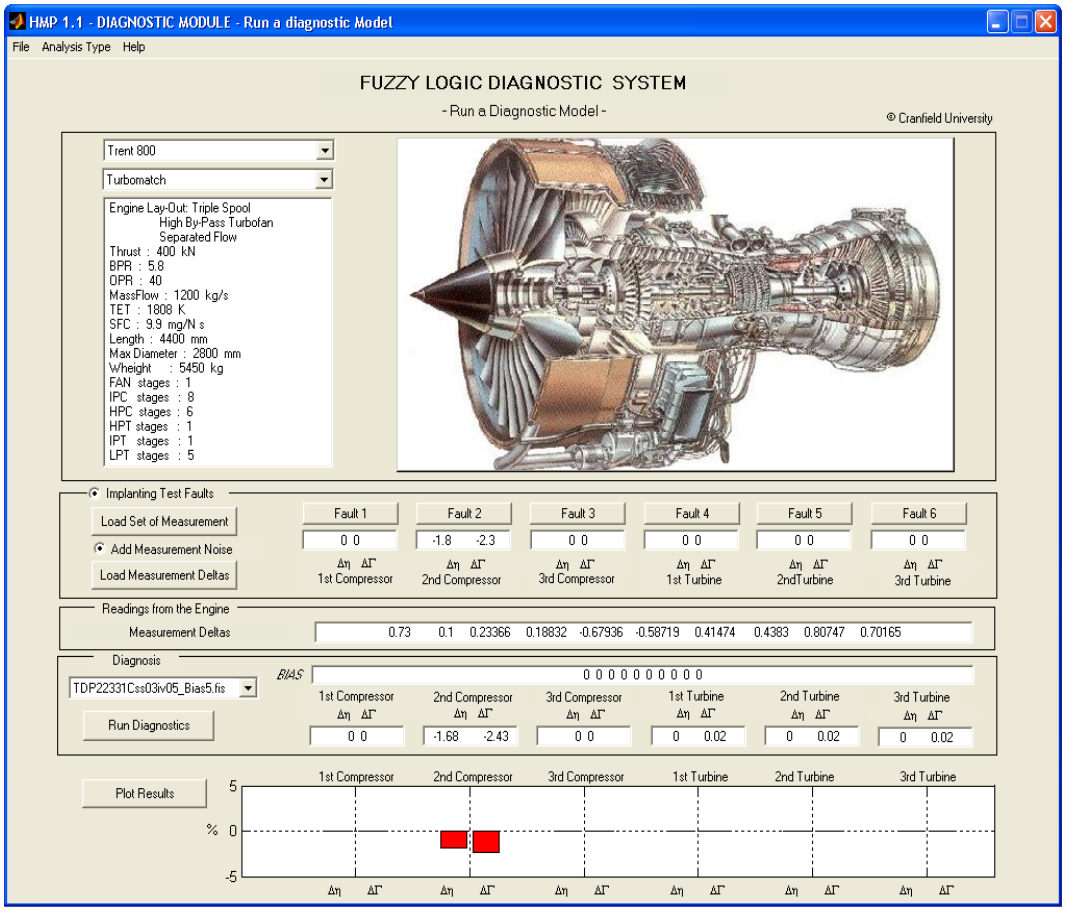


Figure 85: GUI that operates the fuzzy diagnostic models, part of the HMP 1.1 Diagnostics Module, image courtesy of Rolls-Royce

This interface was designed to operate fuzzy diagnostics models previously set-up with the GUI of Figure 75 and to assess the changes in the considered gas-path performance parameters. The GUI can be used in two modes. Once the engine and its simulation model are selected, the readings from the engine can be input and the diagnosis made by means of the diagnostics system previously generated and saved. Alternatively, a fault can be implanted and the corresponding measurements deviations simulated by using the engine model. These are employed to test the fuzzy diagnostics model. This interface can be used to operate diagnostics models with SFI, MFI and bias-tolerant capabilities. The results can be eventually plotted.

4.10 Discussion and summary of the results

4.10.1 Introduction

HMP 1.1 for performance analysis is the software developed within this research to provide an advance health-monitoring-and-prognostics (HMP) framework. It consists of three modules: the observability study module, the diagnostics module, and the prognostics module. The latter and its functionality are discussed in Chapter-5. The procedures implemented within the observability and the diagnostics modules were presented in this chapter, together with the results relative to an application of these procedures to the Rolls-Royce Trent 800 turbofan. The requirements, that were stated in Chapter-3, for these methodologies are now repeated, due to their importance in the discussion that follows. First, it was recognised that the outcome of a reliable diagnosis must not only advise about the quantifiable faults, but must also warn about the faults that are not observable with the given measurement set. Therefore the need for a secure procedure to carry out observability studies was raised. Hence, the first requirement for an advanced diagnostics process was identified. It should be:

1. Capable of providing additional information about the system observability.

Moreover, a thorough analysis of the published literature enabled the statement of the following additional requirements for an advanced diagnostics process that should be:

2. Based on a non-linear model.
3. Designed specifically for SFI and/or MFI.
4. Capable of detecting with reasonable accuracy significant changes in performance.
5. Able to provide a ‘concentration’ capability on the actual fault.
6. Competent to make a worthwhile diagnosis using only a few measurements (N>M).
7. Able to deal with random noise in the measurements and sensor bias.
8. Light in computational requirements and fast in undertaking diagnosis for on-wing applications.
9. Exempt from training and tuning uncertainties, difficulties and dependences for setting-up parameters.
10. Free from a lack of comprehensibility due to black-box behaviour.
11. Capable of data-fusion and able to incorporate expert knowledge.

A total of 11 requirements were identified as the specifications for this first part of the research whose outcomes were the observability module and the diagnostics module of HMP 1.1. The next section discusses the procedures and the results presented in this Chapter in the light of these requirements, in the attempt to assess their fulfilment.

4.10.2 Analysis of the results

This section is aimed at discussing how each requirement is fulfilled by the processes described in this chapter. For some requirements, only a partial fulfilment was obtained: in these cases the underlying principles behind the

choices made are commented and some suggestions for further studies are attempted in section 4.11.

1. **Process capable of providing additional information about the system observability.**

Gelb (1974) and Provost (1995) pointed out the relevance of an observability study in gas-path diagnostics processes, even though the development of a quantitative approach to gas turbine engine faults observability has long been neglected. Hence, section 4.4 described 4 different types of outcomes that can be obtained via the observability module of HMP 1.1:

- The correlations between any two of the available gas-path measurements that respond in a similar way to the component changes. Number and severity of the correlated sensors were identified.
- The correlations between any of the component changes and/or sensor biasses that produce similar change in all (or nearly all) the gas-path measurements. Number and severity of the correlated component changes (and/or sensor bias) were identified.
- The identification and analysis of the smallest eigenvalues and the corresponding eigenvectors of the CORROBS matrix: these eigenvectors provide the combinations of changes in performance parameters and/or sensor biasses that are not observable because they are associated with neglectable changes in the measurements – see Procedure 2 section 4.4.5.2 .
- Identification and study of a parameter able to quantify the quality of the observability of a system. For completely observable systems, the parameter is the product of the eigenvalues of CORROBS, while for not completely observable systems it is the product of the m (= rank of ERT) eigenvalues of CORROBS significantly different from zero. Such a parameter enables the comparison of a large number of similar diagnostics systems (with the same number of inputs and outputs) and the choice of the most suitable measurement set for diagnostics purposes – see Procedure 2 section 4.4.5.2.

The results of these 4 procedures were studied for the Trent 800 in the diagnostics configuration – number and type of inputs and outputs – investigated in this work. These provided the inevitably existing correlations, the combinations of component changes not observable with the available set of 10 measurements and the best set of 10 measurements out of 15 that should be used to gain an improved observability. The limitation of this outcome lies, first of all, in the linearization. Different correlations may appear if the ERT is calculated for changes in the performance parameters different from 1%. Besides, the study is undertaken at a given operating condition: the best measurement set identified at a given flight condition may not result optimal at another. Nevertheless, diagnoses most of the times are required at cruise conditions and therefore the importance of the observability study at that operating condition still holds.

The achievement of a higher level of fulfilment of this first requirement would have demanded more resources to be allocated to the study of the interactions between the observability and the diagnostics modules. Such a choice would have gone beyond the scope of this project, whose core was the development of a novel diagnostics procedure as stated in Chapter-3. Future studies should be focussed on these aspects as suggested in section 4.11.

2. **Process based on a non-linear model.**

Within the dispute on whether a gas-path diagnostics method should be based on a non-linear or a linear model, in this work it was recognised that an advanced diagnostics system must have the capability of detecting considerable changes in at least a limited number of performance parameters, and this necessarily requires a non-linear approach. Gradual and rapid deteriorations can be distinguished and treated separately, as discussed in section 2.1.5. The former implies that all the engine components are deteriorating slowly and by small amounts, whereas the latter may result in significant changes in a few parameters (i.e. single event). Volponi (2003) pointed out the necessity to develop different algorithms to address the problem of estimating gradual and rapid deteriorations, namely MFI (multiple fault isolation) generally based on linear approaches and SFI (single fault isolation) methods necessarily based on non-linear approaches respectively. This study was aimed at the development of a non-linear approach. Sections 4.2.4 and 4.5.2 presented the learning algorithm, relative to the fuzzy rules, developed here. It was discussed how, the use of data obtained from the engine model to generate the rules preserves the linearity of the problem. The method developed is based on a non-linear model but, on the other hand, it carries a limitation typical of this type of methodologies: its layout becomes cumbersome when implemented to isolate more than two deteriorated parameters simultaneously. The system's partial MFI capability based on the non-linear approach was tested in section 4.7: in the context of this study it was limited to the case with up to 4 performance parameters simultaneously deteriorated.

3. **Process designed specifically for SFI and/or MFI.**

The strategy adopted in this investigation was made clear above. The fuzzy diagnostics procedure was initially developed to achieve SFI capability. The results from tests generated implanting different combinations of fault levels in the 6 gas-path components investigated were presented in 4.6. Volponi (2003) highlighted the need for a framework that allows the two processes (SFI and MFI) to operate in concert with one another, automatically, and without corruptive interaction. Section 4.6.3 suggested a possible layout for such a framework.

4. **Process capable of detecting with reasonable accuracy significant changes in performance.**

The evaluation of the fulfilment of this requirement is related to the level of accuracy achieved in the results. The system was devised and tested in the following cases:

- Approximation capability accuracy: this is the accuracy in the assessment of considerable changes in performance parameters (SFI) from the available measurements when these are free from errors – see section 4.6.1.
- SFI capability accuracy in the presence of realistic noise in the measurements – see sections 4.6.2 and 4.6.3.
- A partial MFI capability accuracy with up to 4 performance parameters faulty at a time – see section 4.7.
- SFI capability accuracy of a bias-tolerant system – see section 4.8.

Changes in the performance parameters from zero to 3% were implanted and the accuracy of the system was assessed by using the three methods defined in 4.5.4.3. Deviations between the implanted and the calculated deteriorations were considered to be of very low severity if less than 0.5% errors occurred. This criterion is justified considering that the results are inevitably affected by the following sources of uncertainty:

- The noise in the measurements (e.g. the sensor noise STDV for the FF is typically of the order of 0.5%).
- The fact that we proved with the observability study in section 4.4 that the system is not completely observable (having 10 input measurements and 12 outputs, N>M). It has correlations between the measurements and between the performance parameters. Besides, a more suitable set of 10 measurements was identified that provided a better observability, compared with the available measurement set.

5. **Process able to provide a 'concentration' capability on the actual fault.**

The results achieved with the fuzzy diagnostics system do not show tendency to smear the faults over a large number of the engine's components and sensors (see sections from 4.6 to 4.8). This is a problem that typically affects diagnostics method based on optimal estimation theory (e.g. Kalman filter based). On the contrary, a good concentration capability is achieved in quantifying the implanted fault. These results were achieved by means of a SAM and quasi-SAM diagnostics systems.

In section 4.2.5.1, the role of a fuzzy logic approach compared to a statistic one was discussed. It was proven that a SAM system computes a conditional expectation and therefore a mean-squared optimal non-linear estimation. In addition the suitability of SAM and a quasi-SAM system to solve the GPD problem was demonstrated in section 4.6.2. The type of result achieved with a fuzzy diagnostics system lies both in this optimality of the result and in its model-free structure. Most popular conditional-mean systems – e.g. the Kalman filter – use a mathematical model of the process. Such systems are likely to identify as optimal solution the one with the fault distributed over many engine's components and sensors (Zedda, 1999) – see section 2.3. A fuzzy system computes its optimal solution through its rules. It was shown in section 4.6 that the results of the fuzzy diagnostics systems are not as affected by this problem.

6. **Process competent to make a worthwhile diagnosis using only few measurements (N>M).**

The diagnostics layout described in this document is characterized by 10 input measurements and 12 output performance parameters sought. It was proven that the fuzzy logic based procedure does not require a completely observable system, with same number of inputs and outputs to be effective.

7. **Process able to deal with random noise in the measurements and sensor bias.**

Section 4.2.5.1 provided an interpretation of a fuzzy system not only as a locus of multivalued set values, but also as a locus of two-point conditional probability densities. In the light of these considerations, to gain in diagnostics accuracy in the presence of noisy data, the input MF amplitudes were chosen according to the different values of sensor noise standard deviations available for different types of measurements – see section 4.6. A considerable enhancement of the accuracy of the system was demonstrated.

Section 4.8 presented a bias-tolerant fuzzy diagnostics system capable of quantifying considerable changes in two performance parameters of one faulty component (SFI) while isolating a biased sensor. The basic idea is based on: (i) the use of the centre of maximum defuzzification method, (ii) the use of the NOT logical operator, and (iii) a new formulation of the fuzzy rules that includes the location of the bias. An application of this methodology to the Trent 800 engine was presented. The SFI bias-tolerant system required an increased number of rules (3641 rules in the reported application) compared to the corresponding SFI system not capable of dealing with biasses (331 rules). Therefore, a longer computational time was required.

8. **Process light in computational requirements and fast in undertaking diagnosis for on-wing applications.**

As far as the computational requirements are concerned, the following results were found (with the current average computational power):

- The SFI system, whose accuracy tests are presented in sections 4.6.2 and 4.6.3, required 1 min and 12 sec in the set-up phase (331 rules were generated) and 0.1 sec to provide the result for 1 test case.
- The partial MFI system described in section 4.7, required 18 min and 35 sec in the set-up phase (19440 rules were generated) and approximately 12 sec to provide the result for 1 test case.
- The SFI bias-tolerant system, described in section 4.8, required 3 min and 20 sec in the set-up phase (3641 rules were generated) and approximately 1 sec to provide the result for 1 test case.

Comparing the diagnostics system devised using fuzzy logic with others developed using other artificial intelligence approaches namely neural networks and genetic algorithms, we can state that:

- Fuzzy logic enables the set up of a diagnostics system in a considerably shorter time than neural networks (Ogaji, 2003), that typically require long training periods, and genetic algorithms (Zedda, 1999).

- The diagnostics time required by the fuzzy diagnostics system is comparable to a similar system based on neural networks but significantly shorter than the time demanded by a genetic algorithm based method (Zedda, 1999).

9. **Process exempt from training and tuning uncertainties, difficulties and dependences for setting-up parameters.**

A first sensitivity study was performed in section 4.5.5 to identify the combination of fuzzy functional parameters most suitable for designing a fuzzy-logic based diagnostic system. There is no reason to think that this optimal combination of functional parameters should vary with the type of engine diagnosed. A second sensitivity study was carried out in section 4.5.5 aimed at identifying the best choice of the fuzzy system parameters to set up a diagnostics system for the Trent 800 engine. When implementing a new diagnostic system for a different engine, a new sensitivity study may be required according to the procedure provided. The parameters chosen in this study can be used as first attempt values.

10. **Process free from a lack of comprehensibility due to black-box behaviour.**

One of the advantages in using fuzzy rules to devise a diagnostics system lies in their modular nature that always enables the user to have a clear idea about what is stated and what is not, thereby gaining in comprehensibility. On the other hand, some of the methods described in section 2.3 suffer from black-box behaviour: for example neural networks, instead of requiring the statement of rules, require the training of their synapses. The numerical synaptic values change when the input data make the neurons fire. This makes a network able to learn to recognise patterns and therefore to map inputs into outputs. The major difference is that, in the case of a neural network, a user has no way to know what the network has learnt or forgotten during the learning process. When the network is trained with new information, there is an inevitable tendency to forget the old one.

11. **Process capable of data-fusion and able to incorporate expert knowledge.**

A fuzzy model is beneficial for introducing into the decision-making process different sources of information (i.e. data fusion) and to include expert knowledge or statistical inputs. This feature was not introduced in this project but its potentiality was highlighted. As far as this 11^{th} requirement is concerned, the intention within this project was limited to open up the research subject suggesting some opportunities for future studies (see section 4.11).

4.11 Recommendations for future studies

In section 4.10, a discussion of the results was provided. The discussion was articulated commenting on the level of fulfilment achieved for each research requirement (i.e. specifications for this first part of the study dedicated to observability and gas-path diagnostics), in order to analyse the outcome of the presented processes and the potential research opportunities created. This section provides some suggestions for further studies that should be carried out to gain a higher level of accomplishment of those requirements that have been only partially fulfilled.

The following recommendations for future studies were drawn from the experience achieved through the development of this research:

- A higher level of integration between the observability and the diagnostics processes should be implemented (with reference to the 1st requirement). The reason is twofold. The validity of the observability study can be tested, checking whether its outcome is confirmed by the diagnostics outcome. Diagnostics tests with systems having different measurement sets (in number and type) should be carried out and their accuracies evaluated. It is expected that systems with higher qualities of observability result more accurate during the diagnostics tests. On the other hand, once the observability outcome is pertinently validated, this information could be used to set-up an enhanced fuzzy diagnostics process. Rule weights (introduced in section 4.2) could be exploited for this purpose. When two or more faults are identified as correlated (associated with similar changes in the measurements), different rule weights can be associated with the different rules to differentiate the outcome. Each weight value could be decided in terms of statistical information or expert knowledge.
- Taking advantage of the increasingly available computational power, future studies should be investigating the capability of a fuzzy diagnostics system implemented for a complete multiple fault isolation capability. This would benefit from the non-linear approach extended to all the possible combinations of component deteriorations within a constrained search space (with reference to the 2nd requirement).
- Section 4.6.3 suggested a possible layout for a framework based on the simultaneous operation of SFI and MFI processes (with reference to the 3rd requirement). Future work should look at the opportunity of devising more sophisticated frameworks capable of coupling SFI and MFI, and extensive tests should be undertaken for their validation. A more sophisticated framework should consider the enhancement of the simple SFI capability to a partial multiple fault isolation of considerable variations in performance parameters (based on a non-linear approach). In practice, the assumption that only one component becomes significantly faulty, in a short interval of time, might not be always acceptable. For that reason, a system with partial MFI capability was tested in section 4.7 with up to 4 performance parameters deteriorated simultaneously.

- Future studies should attempt to combine more than one gas-path diagnostics techniques (with reference to the 4th requirement) to offset the limitations of one with the advantages of another, gaining in enhanced diagnostics accuracy within a combined engine-health monitoring (COEHM) scheme (Singh, 2003). This task can be twofold. Firstly, the results from different methods can be used together, so exploiting a potential synergic effect. On the other hand hybrid systems could be designed (Sampath, 2003a).
- An alternative way of combining more techniques to achieve enhanced diagnostics can be devised by effectively performing data fusion (with reference to the 11th requirement). Fuzzy logic data fusion capability should be implemented to introduce, in the decision-making process, different sources of information coming from GPD studies as well as from oil analysis, oil-debris analysis, vibration analysis etc. Besides, expert knowledge and statistical inputs derived from a thorough analysis of in-flight data should be included as well. This capability should be implemented, not only through the statement of fuzzy rules, but also by using the fuzzy rule weights to differentiate the relative weights of different rules.

CHAPTER 5 - GAS-PATH PROGNOSTICS USING TIME-SERIES ANALYSIS

5.1 Introduction

This chapter presents a novel gas-path prognostics (GPP) framework. Prognostics is the capability of providing forecasts regarding the evolution of engine health and decision-making competences based on them. The core of the research presented in this thesis is the diagnostics procedure based on fuzzy logic presented in Chapter-4. Nevertheless, the considerable importance that GPP is going to have in the future, within the gas turbine aftermarket, was recognised, and therefore, in Chapter-3 a research opportunity was highlighted. The work developed regarding gas-path prognostics, presented in this Chapter, was aimed at opening up the advanced GPP research subject identifying benefits and difficulties, rather than at completing it.

First, it was recognised that any attempt to GPP requires an associated business intention: different parameters may be required to be forecasted with different confidence levels according to different business strategies. This study was dedicated to devise a general framework able to make forecasts that enable the decision-making process. Under some assumptions relative to the potential scenarios of application, the framework was used in the study of Rolls-Royce Trent 800 engine simulated and real data and the results are shown in this chapter.

The novelty of the framework, presented in the next sections, lies in the use of two different statistic-based approaches that employ different time-series analysis methods for short and long term investigations (specific to different prognostic problems). These methods can forecast the performance deterioration of civil aero-engines providing significant benefits in mission scheduling and maintenance planning. Differently from previous studies described in the literature (see section 2.4) which are mainly qualitative and lack of statistical rigour, in this investigation prediction intervals are introduced to take into account errors in deterioration modelling as well as forecasting errors in the prognoses concerning the health of the engine. Besides, a strategy is proposed to use measurement forecasts in the input of the diagnostics process in order to predict performance parameters, gaining in statistical information. The procedures presented in this Chapter were implemented in order to constitute the prognostics module of the Heath Monitoring and Prognostics (HMP) framework developed within this research project – see Appendix C.

5.1.1 A guide through the chapter

Section 5.2 provides an introduction to gas-path prognostics and the point of view that has been taken in this research to deal with this problem. Firstly, an ARIMA based forecasting procedure for short term investigations is discussed in section 5.3. Two prognostics scenarios were created that required the forecasts of respectively the take-off values of TGT and N2 (considered as representative of the global health of the engine) for a short term horizon to

undertake dependent parameter-based prognostics. Their forecasted values and the corresponding prediction intervals were compared along side the TGT and shaft speed warning and safety margins: typical prognostics considerations drawn from their study were discussed. Then a long term investigation method based on regression analysis and some considerations regarding the nature of the gas-path component degradation are presented in section 5.4. Two prognostics scenarios were simulated through a list of assumptions regarding the maintenance practice and costs, and the long term forecasts of fuel flow at cruise condition was computed. A procedure that calculates the optimal TBO, considering the maintenance costs and increased fuel costs per flight due to the engine deteriorations was devised and applied in the cases of normal and severe deterioration rates. The practical opportunity of using such a procedure to optimise the time between compressor washings was discussed. In section 5.5, the outcomes of the prognostics framework are discussed and recommendations for future studies are provided.

5.2 Prognostics and forecasting

Prognostics consists of prospects and decision-making competences based on them: so forecasting plays a key role in decision making. In this section a prognostic process framework based on two quantitative forecasting techniques is introduced and its functionality for gas turbine degradation analysis demonstrated. It is recognised that the magnitude of the prediction errors experienced depends upon the forecasting method used. Hence, the development of an appropriate forecasting capability is aimed at improving the prediction accuracy and thereby eliminating some of the losses that are associated with the uncertainties in the decision-making process. As shown in Figure 86, by allocating more resources to forecasting, we reduce the losses due to uncertainty.

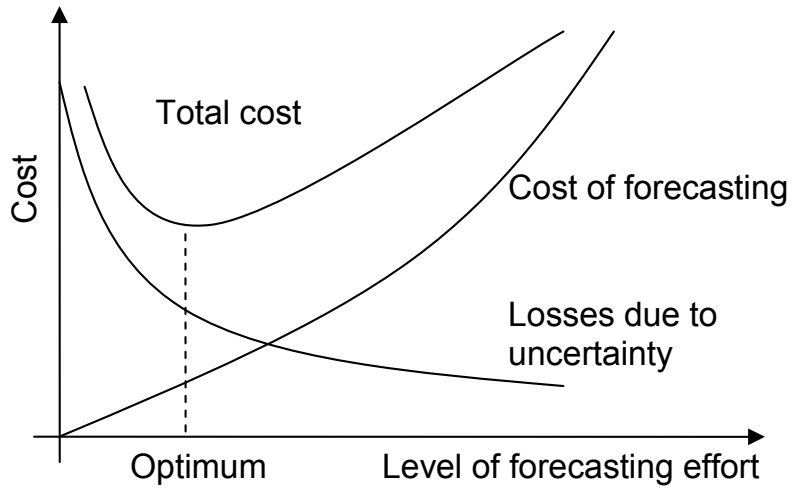


Figure 86: Optimal resources allocated to forecasting

It was an ambition of this study to identify an optimal advisable effort to invest in GPP prediction procedures, without incurring unjustifiable complex methods.

Nevertheless, prediction techniques can never entirely eliminate this uncertainty; therefore in the procedure here described, the prediction error is estimated and explicitly considered in the decision process. So, decision is related to the prediction in the following way: the “actual decision” is equal to the “decision assuming the forecast is correct” plus an “allowance for forecast error” (Montgomery, 1990).

5.2.1 Introduction

A conceptual layout for a gas-path HMP process is proposed in Figure 87.

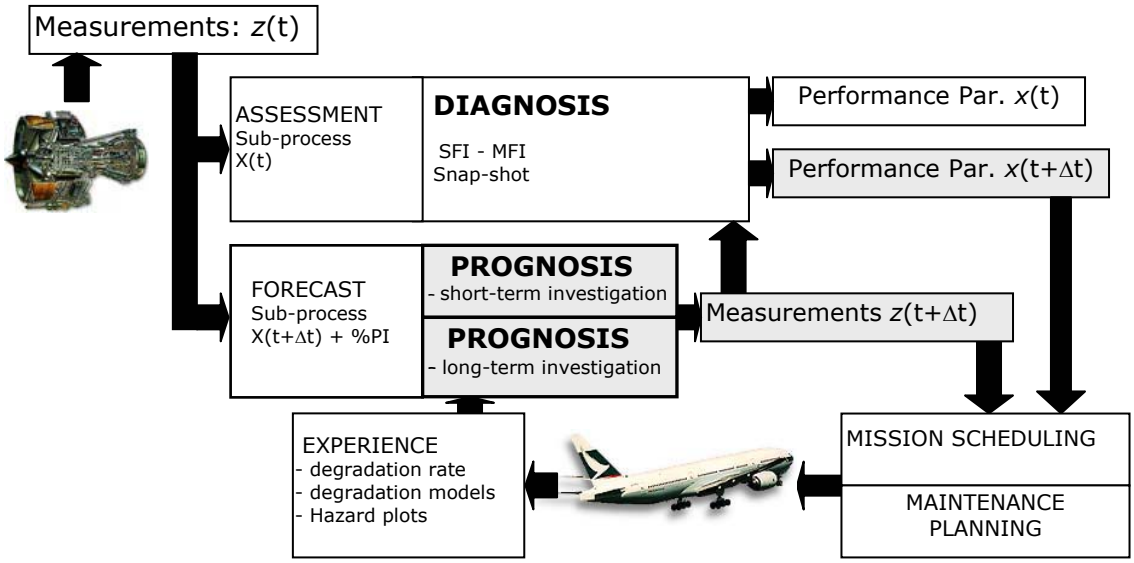


Figure 87: Performance Health-Monitoring and Prognostics (HMP) process, (Marinai et al., 2003), image courtesy of Rolls-Royce

The procedure consists of three articulated sub-processes that communicate with each other: (i) assessment and diagnosis, (ii) forecast and prognosis, and (iii) operation and maintenance management. The engine gas-path measurements are fed into both the diagnostics and prognostics modules.
The diagnostics process performs fault identification, classification and quantification and computes the non-measurable performance parameters.
The prognostics process employs forecasting algorithms to predict the gas-path measurements for future missions based on historical data. The process incorporates two techniques in order to perform short and long term prognosis. Depending on the situation and the reason for making forecast, the time horizon, for which the prediction of a given engine parameter is demanded, may vary from 10 up to more than 1000 cycles. For example, a possible scenario could involve the determination of whether the engine is safe flying the next 10-20 missions by forecasting the TET and checking whether and when the safety margin is overtaken. Conversely, the optimal maintenance interval can be evaluated by predicting the increase in fuel flow until the next scheduled maintenance, which might involve 1000 cycle forecasting horizon.
The forecasted gas-path measurements can be used directly as representative of the global health condition of the engine (e.g. TGT, shaft speed, fuel flow) or can then be fed into the diagnostics module to predict future performance

parameters at a component level. This information can be used to optimise maintenance intervals and mission scheduling. It should be noted that the performance parameters, are not directly forecasted. Two options were considered to predict the gas-path components' performance parameters, which consist on: (i) forecasting the gas-path measurements $z(t+\Delta t)$ and inputting them in the diagnostic algorithm to calculate $x(t+\Delta t)$, or (ii) diagnosing the performance parameters and then use a time series of them to estimate $x(t+\Delta t)$. The former has been here recognised beneficial compared to the latter that is affected by additional uncertainty in the forecasting procedure due to the nature of the result of the diagnostic process; this causes unacceptably high prediction intervals.

Before embarking upon the discussion of the algorithms developed to make prognoses, it is worthwhile remarking and summarizing some useful basic distinctions made in this document. A gas-path prognostics process can be classified as follows:

- Dependent parameter-based prognostics: relies on the prediction of engine global variable, say measurements strictly related with the global health of the engine such as the TGT, the shaft speeds or the SFC.
- Independent parameter-based prognostics: relies on the prediction of health performance parameters (e.g. efficiency, flow capacity of the engine-components).

Another distinction is made between:

- Investigation at T.O., usually involves TGT and shaft speed margins.
- Investigation at cruise conditions, usually involves a study at component level of the independent parameters. Besides, the interest can be extended to a global variable such as FF change, or the contributions to the FF changes due to the various variations at components level (independent parameters).

A third distinction is made between:

- Short term investigation.
- Long term investigation.

Eventually a distinction can be made between:

- Prediction of gradual performance deteriorations. This includes the type of scenarios investigated in this work by means of time series analysis techniques.
- Prediction of considerable faults that manifest as instantaneous deterioration due to events.

The procedures described below are specially made for dealing with a wide range of variables, and some applications with different dependent gas-path parameters will be given. The study is mainly limited to the prediction of gradual deteriorations; nevertheless, the problem of events that cause considerable changes in the variables is highlighted. This can be taken into account through hazard plots that influence the pertinent gas-path variables' safety margins.

5.2.2 Definition of the forecasting problem

Different prognostics problems require different forecasting approaches. Firstly, because forecasting algorithms which are reliable for a short-term predictions tend to be inaccurate for long-term horizons and vice versa. Moreover the nature of the decision to be made will dictate many of the desired characteristics of the forecasting process such as which variables should be investigated, what time elements involved, what form the forecast should take, what accuracy is desired and what is the availability of the data.

The **variables** investigated are component performance parameters and measurements. From one hand, performance parameters forecasts are investigated at cruise condition: the aim is to predict the components' degradation levels and their probability in the next future. From the other hand, measurements trends and forecasts are usually studied at take-off (and climb) to guarantee turbine gas temperature (TGT) and shaft-speed margins. According to the prognostic process's layout suggested in Figure 87, in this work, only measurements are directly inputs of the forecasting algorithm.

As far as the **time elements** are concerned, we distinguish the forecasting period, the horizon, and the interval. (i) The forecasting period is the basic unit of time for which the analysis is made. For example we might wish to forecast by number of missions (i.e. cycles) or by operating hours. (ii) The forecasting horizon or lead time is the number of periods in the future covered by the forecast. (iii) Finally the forecasting interval is the frequency with which new forecasts are performed. The forecasting interval is often the same as the forecasting period.
Although civil engine's running-time is mainly spent under cruise condition, this condition is related to an operation of the engine at much lower rotational speeds, pressures and temperatures and in a less aggressive environment than take-off or climb phases and hence contribute little to deterioration. Thus deterioration is usually correlated against cycles (forecasting period); where 1 cycle is a flight consisting of a take-off, a cruise period, a descent, and a landing. An additional cycle per flight must be considered if thrust reversers are used. The forecasting interval is assumed to be the same as the forecasting period (i.e. one cycle). For each mission a new calculation is undertaken and the forecast is updated.

The **form of the forecast**, as conceived in this work, includes an estimate of the expected value of the variable, and the prediction interval associated with the level of confidence in the predicted future value.

As far as the **prediction accuracy** is concerned, it was recognised that forecasts are usually wrong but the magnitude of the forecasting errors or accuracy depends on the technique used. Typically, a higher accuracy can be achieved by using a more complex forecasting technique. While some of the losses resulting from the uncertainty in the decision-making process can be eliminated, the cost associated with the forecasting (e.g. in terms of computational requirements) increases.

5.2.3 Forecasting with time series analysis

Forecasting techniques can broadly be divided into qualitative and quantitative according to the mathematical and statistical methods used (Montgomery, 1990). Qualitative methods are generally characterised by the fact that little or no quantitative information is available and the forecast is mainly based on sufficient qualitative knowledge (Makridakis et al., 1998). They "involve subjective estimation through the opinions of experts" (Montgomery, 1990). On the other hand, quantitative forecasting procedures rely on historical data. Statistical methods clearly define how the forecast is determined. When dealing with quantitative forecasting methods, technical literature distinguishes: causal models and time series models. Causal models assume that there is a relationship between the variable to be forecast and one or more other independent variables. A time series model instead, is a time-ordered sequence of observations of a variable, which always assumes that some sort of pattern exists in the data. The objective of such time series forecasting methods is to discover the pattern in the historical data series and extrapolate that pattern into the future. It follows that an underlying assumption of time series forecasting is that, over the forecast horizon, the variable will behave in the same way as it did in the past (Wheelwright et al., 1977, Montgomery, 1990, and Makridakis et al., 1998).

In view of the requirements discussed in Chapter-3, time-series-based quantitative techniques able to make predictions that rely on historical measurements data are considered. Generally speaking, time-series analysis uses the history of the variable being forecasted in order to develop a model for predicting future values. To forecast using a time-series, it is necessary to represent the behaviour of a stochastic process by a mathematical model that can be extrapolated into the future, with a given prediction interval (Montgomery, 1990).

The accuracy of a forecasting procedure can be quantitatively described by the variance of the forecast error. As practical matter, it is always desirable to have an estimate of the forecast error variance in order to quantify the uncertainty, or risk, associated with the forecast. By adding and subtracting a multiple of the standard deviation of forecast error, this point estimate can be converted in prediction interval (PI) – see Figure 88.

Three elements constitute a time-series based prediction procedure of gas turbine performance, namely a mathematical model of the deterioration, a forecasting technique and a procedure to derive the prediction intervals. These three elements are covered in the following sections with regards to two techniques specifically design to deal with the different prognostic problems faced in performance analysis.

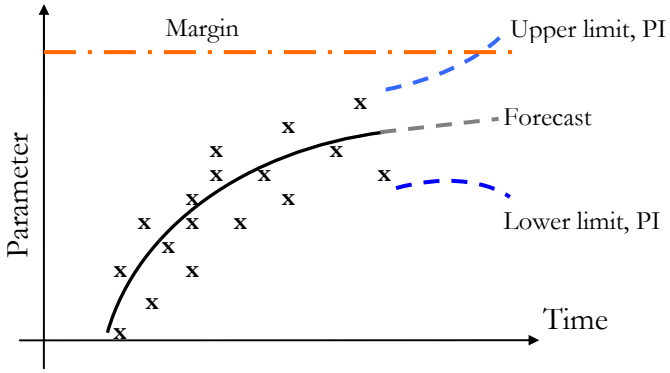


Figure 88: Forecast and prediction intervals.

5.2.3.1 Quantitative forecasting methodologies: choice of the right approach

Many quantitative forecasting methodologies are available to solve a variety of prediction problems and it is important to realise that no single methods is universally applicable. These include exponential smoothing, moving average, regression, Box-Jenkins and Bayesian forecasting techniques. In order to select the most suitable technique a number of criteria can be considered (Wheelwright et al., 1977): (i) lead time, (ii) pattern of data, (iii) type of model, (iv) accuracy, and (v) complexity & applicability. As far as the data patterns are concerned four different types are usually distinguished: horizontal, trend, seasonal, and cyclical. Table 62 compares a number of forecasting techniques.

The requirements stated in Chapter-3 together with a thorough analysis of the different capabilities of the various forecasting methodologies led to the definition of a prognostic framework based on two different algorithms to deal with different GPP problems regarding performance gradual deterioration. Significant instantaneous deteriorations due to events can be taken into account through hazard plots that influence the safety margins.
Box-Jenkins ARIMA method (*technique 1*) was implemented to provide accurate forecasts for immediate and short-term horizons (e.g. prediction based on more than 50 observations, with a forecasting horizon of a few 10s cycles). Its applications in prognostics may include:

- Failure risk for a short term lead time horizon (performance rejection).
- Optimal managing of the fleet considering short term horizon. Short term prediction of fuel consumption, taking into account the contribution due to component degradation.

Regression analysis (*technique 2*) was designed to handle prognoses that require medium and long term predictions (e.g. prediction based on 500 observations, with a forecasting horizon of a few 100s cycles), focusing on a simplified physics-based mathematical model of the degradation. Application may include:

- Failure risk for a medium or long term lead time horizon.
- Identifying the optimal for the maintenance costs to benefit ratio.
- Optimal managing of the fleet considering a long term horizon.

Regression analysis is highly advisable for engineering applications because it has the unique advantage of simplicity that allows tangible justification of the results and a minimum requirement of assumptions on the statistical behaviour of the variables. On the other hand, in short term investigation, a higher level of accuracy can be achieved with more sophisticated techniques such as ARIMA. In this case the assumption that over the forecasting horizon, the variable will behave in the same way as it did in the past, it is better fulfilled. Besides, ARIMA has the capability of discovering trends and predicting turning points in short term analysis; it was previously pointed out how the accuracy of a forecasting technique depends on this capability.

Table 62: Comparison of quantitative forecasting techniques, (Wheelwright et al., 1977)

		Moving average	**Exponential smoothing**	**Simple regression**	**Multiple regression**	**Box-Jenkins**
Time horizon	Immediate	x	x			x
	Short term	x	x			x
	Medium term			x	x	x
	Long term			x	x	
Pattern	Horizontal	x	x			x
	Trend	x	x	x	x	x
	Seasonal					x
	Cyclical					x
Type of model	Time series	x	x	x	x	x
	Causal			x	x	x
Accuracy (scale from 0 to 10; 0 smallest, 10 highest)	Predicting pattern	2	3	5	8	10
	Predicting turning points	0	0	0	4	8
Applicability & complexity (0 smallest, 10 highest)	Time required to obtain forecast	1.5	1	2.5	6	7
	Easiness to under-stand results	9	7	9	7	4

It is worth bearing in mind that forecasts are conditional statements about the future based on specific assumptions. Thus it is important to be prepared to modify the forecasts as necessary in the light of any external additional information. Several different forecasts based on alternative sets of assumptions should be provided so that alternative scenarios can be explored, specifying the pertinent levels of confidence.

5.3 Short-term forecasting procedure: ARIMA

5.3.1 Introduction

The forecasting procedure described in this section is based on autoregressive integrated moving average (ARIMA) models also known as the Box-Jenkins approach (Box et al., 1976).

According to the Box-Jenkins procedure, AR and MA models are fitted to the time series that results from one or more differentiations (transformations). The differentiation procedure induces stationarity, and when such a procedure is employed the ARMA model is called autoregressive integrated moving average (ARIMA).

The main stages of the Box-Jenkins forecasting model set-up procedure follow – see Figure 89.

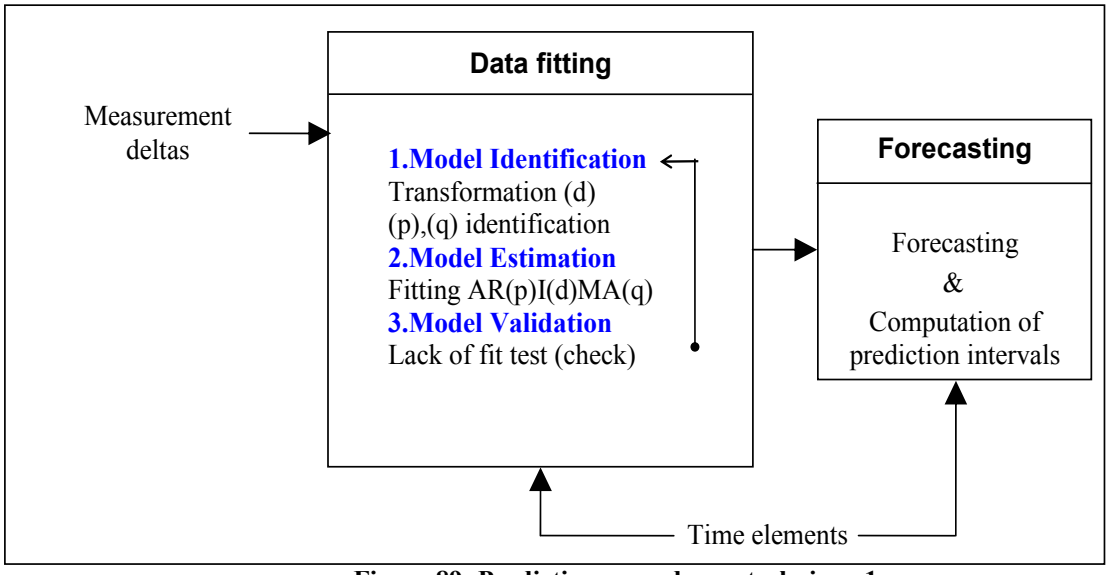


Figure 89: Prediction procedure – technique 1.

- Model identification: the data are examined to see which elements of the class of ARIMA processes appear to be most appropriate.
- Estimation: the parameters of the chosen model are estimated as described in the next section.
- Diagnostic checking: the residuals from the fitted model are examined to see if it is adequate.
- If the first model appears to be inadequate, then other ARIMA models may be tried until a satisfactory model is found.
- Forecasting: the model is then used to predict future observations.

The major contribution of Box and Jenkins (1976) was to provide a general strategy for time-series forecasting, in which the different stages of model building as listed above are all given appropriate importance. In addition they showed how the use of differencing extends the approach to deal with non-stationary series. When a satisfactory ARIMA model is found, it is relatively

straightforward to calculate forecasts as conditional expectations. In this work seasonal effects were not considered, but the method can deal with them, too.

The first step in the Box-Jenkins procedure is to difference the data. This is done examining the correlograms of various differenced series until one is found which comes to zero fairly quickly, the non stationary element of the series have been hence eliminated. For non seasonal data first order differencing is usually sufficient. Then, an ARIMA model can be fitted and the parameters of the model estimated by means of least-squares estimates. The adequacy of the model is checked by examining the residuals from the fitted model to see if there is any evidence of non-randomness. The correlograms of the residuals are calculated: the analyst studies the number of coefficients significantly different form zero and whether any further terms are indicated for the ARIMA model. Once the fitted model appears to be adequate, forecasts may be readily computed.

Numerous test-books (O'Donovan, 1983, Box et al., 1976, Montgomery, 1990, and Makridakis et al., 1998) concentrate on the model estimation process. Besides, many codes are available and capable of providing straightforward estimation of the ARMA models parameters. Section 5.3 focuses rather on the procedure and specifically on the model identification and checking phases as well as on some applications. It will be highlighted how a simple use of a code might often be inadequate, and a thorough analysis of the auto-correlation (AC) plots, partial auto-correlation (PAC) plots and the residuals plots is crucial to achieve an accurate prediction – as discussed in sections 5.3.2.2 , 5.3.2.3 and 5.3.4.

5.3.2 ARIMA: theory

This class of models includes auto-regressive (AR) models, moving average (MA) models and their combinations usually identified as ARMA models. If a transformation process is introduced to deal with non-stationary time series ARMA models are commonly addressed as ARIMA (auto-regressive integrated moving average). They provide a good fit to many different types of time series and should be generally considered when more than 50 observations are available.

A typical time series of gas-path measurements necessarily involves, as discussed earlier, a random element and hence represent a stochastic process. We will denote the generic variable under analysis at the time t corresponding to the i-th observation by $y(t)=Y_i$, (t=0,1,2,…).

The procedure to build an ARIMA model, as summarized in Figure 89, provides for three steps: (i) model identification, (ii) model estimation and, (iii) model checking. A proper use of the ARIMA algorithm requires a critical analysis of the results. A thorough investigation of the AC plots, PAC plots and the residuals plots is crucial to achieve an accurate prediction – see sections 5.3.2.2 and 5.3.2.3.

5.3.2.1 Auto-regressive integrated moving average model

ARIMA models, first introduced in the seventies by Box and Jenkins, are a combination of the autoregressive (AR) and moving-average (MA) models (Box et al., 1976).

Autoregressive model

A process Y_t is said to be auto-regressive with order p if can be expressed as:

$$Y_t = \delta + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + A_t \quad (5.73)$$

where Y_t is the time series, A_t represents white noise, and $\Phi_1..\Phi_p$ are the parameters of the model, and δ expressed as follows:

$$\delta = (1 - \sum_{t=1}^{p} \Phi_t)\mu \quad (5.74)$$

where μ is the mean of the process. The autoregressive model involves a linear regression of the current value of the series against one or more prior values of the series. The value of p is called the order of the AR model.

Moving Average model

Instead the MA model equation can be written as:

$$Y_t = \mu + A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + ... + \theta_q A_{t-q} \quad (5.75)$$

where Y_t is the time series, μ is the mean of the series, A_{t-i} are random shocks to the series, and $\theta_1...\theta_q$ are the parameters of the model. The value of q is called the order of the MA model.

A moving average model is a linear regression of the current value of the series against the random shocks of one or more prior values of the series. The random shocks at each point are assumed to come typically from a normal distribution. In this model, these random shocks are propagated to future values of the time series. Fitting the MA estimates is more complicated than with AR models because the error terms depend on the model fitting. This means that iterative non-linear fitting procedures need to be used in place of a linear least squares fit.

In the standard regression situation, the error terms, or random shocks, are assumed to be independent. That is, the random shocks at the *i*-th observation only affect that *i*-th observation. However, in many time-series, this assumption is not valid because the random shocks are propagated to future values of the time series. Moving-average models accommodate the random shocks in previous values of the time series by estimating the current value of the time series.

However the error terms after the model is fit should be independent and follow the standard assumptions for a univariate process.

ARMA model

Therefore, according to the autoregressive and moving average models, the ARMA model equation can be written as:

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + A_t + \theta_1 A_{t-1} + \theta_2 A_{t-2} + ... + \theta_q A_{t-q} \quad (5.76)$$

where the parameters have the same meanings as for the AR and MA models.

5.3.2.2 Auto-correlation plot

Auto-correlation plots are used in the model identification stage for the Box-Jenkins procedure. The auto-correlation (AC) coefficient for the time series Y_t is defined as:

$$r_k = \frac{c_k}{c_0} \quad (5.77)$$

Where c_k is the auto-covariance coefficient at lag k, given by:

$$c_k = \frac{1}{n}\sum_{t=1}^{n-k}(Y_t - \mu)(Y_{t-k} - \mu) \quad (5.78)$$

and c_0 is the variance:

$$c_0 = \frac{\sum_{t=1}^{n}(Y_t - \mu)^2}{n} \quad (5.79)$$

Suppose that $Y_1....Y_n$ are independent random variables with arbitrary mean, it can be shown (Box et al., 1976) that $E(r_k) \cong -1/n$ and $Var(r_k) \cong 1/n$ and that r_k is normally distributed. Thus, having plotted r_k as said above, the 95% confidence limit can be plotted (approximated) at $-1/n \pm 2/\sqrt{n}$, which can be further approximated to $\pm 2/\sqrt{n}$. Observed values that fall outside these limits are significantly different from zero at the 5% level. However, when interpreting a correlograms, it must be remembered that the overall probability of getting a coefficient outside the limits increases with the number of coefficients plotted (Chatfield C., 1989).

The presence of randomness in time series can be checked by plotting the auto-correlation coefficient versus different time lags k (k=1,2,3..). If the computed r_k are near zero, the time series is constituted of random data. Otherwise one or more of the autocorrelations will be significantly non-zero.

Figure 90 shows an example of auto-correlation plot for 100 observations, simulated on a computer, which are intended to be independent. The confidence limit are approximately ± 0.2.

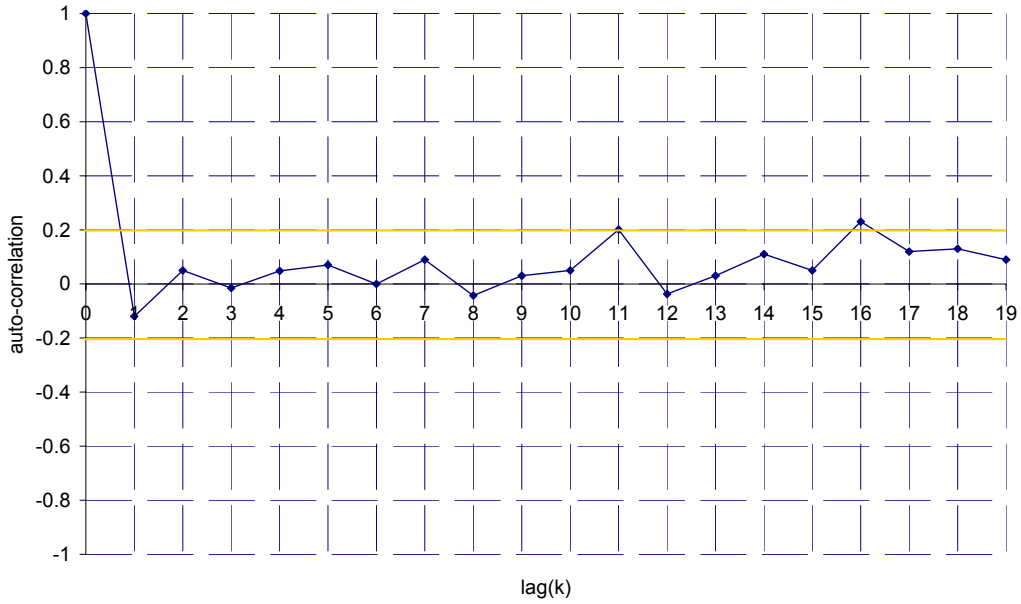


Figure 90: Typical auto-correlation plot for 100 independent normally distributed observations

If the auto-correlation plot shows r_k significantly different from zero, a non-stationarity in the data is pointed out, and the time series needs to be differentiated (Box et al., 1976). Besides, the auto-correlation plots are used in the phase of model identification – see next section.

Interpreting the plot

The correlograms is compared with the theoretical AC plot of different ARMA processes in order to choose one which is most appropriate. The AC plot of an MA(q) process can be recognised because it cuts off at lag q, while for the AR(p) process is a mixture of damped exponentials and sinusoids and dies out slowly (or attenuates). The AC plot of a mixed ARMA model will also generally attenuate rather than cut off. For example, suppose we find that r_1 is significantly different from zero but that subsequent values of r+k+ are all close to zero. Then an MA(1) model is indicated since its theoretical AC plot is of this form. The interpretation of correlograms is one of the hardest aspects of time series analysis and practical experience is a must. Numerous examples can be found in Chatfield, (1989); this section is only aimed at giving an overview of the subject.

5.3.2.3 Partial auto-correlation plot

As far as the order of an AR model is concerned, the analysis of the partial auto-correlation (PAC) coefficients, as defined below, is generally helpful. When fitting an AR(p) model, the last coefficient Φ_p corresponds to the PAC coefficient at lag p and measures the excess of correlation at lag p which is not accounted for by an AR(p-1) model. The PAC plot is generated plotting the p-th PAC coefficient against p, for p=1,…,n. Details for the estimation of the PAC can be found in Visual Numerics (1997).

Interpreting the plot
The PAC plot is estimated by fitting AR processes (Visual Numerics, 1997) of successively higher order. PAC coefficients, that are bigger than $\pm 2/\sqrt{n}$, are, in first approximation, different from zero with a confidence level of 5%. It can be shown (Box et al., 1976) that the PAC of an AR(p) process cuts off at lag p. So, considering the generic time series Y_t, the order of the AR model that better fits the data can be chosen as that value of p beyond which the sample values of PAC coefficients are not significantly different from zero. On the other hand, the PAC coefficients of a MA process will generally attenuate, and in this sense the PAC coefficients have opposite features to the AC coefficients.

5.3.3 Estimating the parameters of an ARIMA model

Typical time series regarding gas-path monitoring are non-stationary. According to the Box-Jenkins procedure, AR and MA models are therefore fitted to the time series that results from one or more differentiations (transformations). Such a procedure is referred to as autoregressive integrated moving average (ARIMA). The estimation of the parameters is performed on the stationary time series.

$$W_t = \nabla_s^d Y_t \qquad (5.\ 80)$$

where ∇_s^d is the backward difference operator with period s (often fixed to 1) and order d, d > 0. Consequently the model AR(p)I(d)MA(q) is said to have orders p,d,q.

The first attempt values of p,d,q is provided by the study of the AC and PAC plots, as discussed earlier. Once an AR(p)I(d)MA(q) models is thought to be appropriate for a given time series, the parameters of the model are estimated. The most common estimation procedure is based on least squares, many computer packages are available for this purpose. In this study IMSL Fortran subroutines were used (Visual Numerics, 1997).

5.3.4 Diagnostic check

When a model has been fitted to a time series, the third step in the procedure provide for the diagnostic check to validate the adequacy of the model to describe the data.
A variety of model validation procedures are suggested in the literature (Montgomery, 1990). In my experience the Q test (Box et al., 1976), a lack of fit test based on the summation of the first m (m is typically chosen between 15 and 30) values of the AC plot of the residuals, all at once, gives poor results. One could also plot the observed values in the time series against the one-step-ahead forecasts of them. Whereas, my preference, that accords with Chatfield (1989), is to consider the residuals singularly. In fact, as with most of the statistical models, the lack of fit can be evaluated by looking at the residuals defined as the difference between the observation and the fitted value. If the model has been properly chosen we expect the residuals to be random and close to zero. In this case model validation usually consists of plotting the residuals in various ways. The time plot will reveal any outliers and any obvious auto-correlations. Besides, the AC plot of the residuals will enable auto-

correlation effects to be examined more closely. It turns out (Box et al., 1976) that residuals AC values that lie outside the range $\pm 2/\sqrt{n}$ (approximated value) are significantly different from zero at the 5% level and give evidence that the wrong model has been fitted. In practice in this work we have considered the first few values (lags 1, 2) and seen whether they were inside the band, if not the model would be modified.

5.3.5 Forecasting with ARIMA models

When a satisfactory model is found, forecasts and their associated probability limits can be obtained. We can observe (Chatfield C., 1989) that for a univariate time series model, the fitted value is the one-step-ahead forecast so that the residual is the one-step-ahead forecast error. For example, with an AR(1) model, the fitted value at time t is $\Phi_1 Y_{t-1}$ so the residual corresponding to Y_t would be $R_t = Y_t - \Phi_1 Y_{t-1}$. So, given data up to the time t=n, these forecasts will involve the observations and the fitted residuals (i.e. one-step-ahead forecast error) up to and including the time t=n. The minimum mean square error forecast of Y_{n+k} at time n is the conditional expectation of Y_{n+k} at time n, namely $Y(n,k) = E(Y_{n+k} | Y_n, Y_{n-1}\ldots)$. Box and Jenkins (1976) describe three general approaches that can be used in order to compute forecasts, which we will only mention: (i) using the difference equation form, (ii) using the Ψ weights, and (iii) using the π weights. In this work, an IMSL Fortran subroutine (Visual Numerics, 1997) was used that is based on the second approach.

5.3.5.1 Psi (Ψ) weights

Using the backward shift operator B defined as:

$$B^j Y_t = Y_{t-j} \text{ for all j} \qquad (5.\ 81)$$

equation (5.76), according to (Box et al., 1976) can be written in the form:

$$\Phi(B) Y_t = \theta(B) A_t \qquad (5.\ 82)$$

Where $\Phi(B)$ and $\theta(B)$ are polynomials of order p,q respectively, such that:

$$\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \ldots . \Phi_p B^p \qquad (5.\ 83)$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots . \theta_q B^q \qquad (5.\ 84)$$

It is sometimes helpful to express an ARMA model as a pure MA process in the form:

$$Y_t = \Psi(B) A_t \qquad (5.\ 85)$$

Where $\Psi(B) = \sum \Psi_i B^i$ is the MA operator. By comparing the equations (5. 82) and (5. 85) we see that $\Psi(B) = \theta(B)/\Phi(B)$. This new formulation with the psi weights Ψ_l is useful in calculating forecasts.

5.3.5.2 Forecasting with the Ψ weights

An IMSL Fortran subroutine was used to compute forecasts with ARIMA models, based on the concepts described below.

The Ψ weights defined earlier are used to compute forecasts and are particularly helpful in calculating forecast error variances and therefore the prediction intervals (PI). Since:

$$Y_{n+k} = A_{n+k} + \Psi_1 A_{n+k-1} + ... \quad (5.\ 86)$$

It is clear that $Y(n,k)$ is equal to $\sum \Psi_{k+j} Y_{n-j}$ (j=1,∞), evidently without including future observations Ys. Thus ($Y_{n+k}+\Psi_1 Y_{n+k-1}+..+\Psi_{k-1} Y_{n+1}$) is the k-steps-ahead forecast error. Hence the variance of the k-step-ahead error is $(1+\Psi_1^2+\ldots+\Psi_{k-1}^2)\sigma_Y^2$, being σ_Y^2 the var(Y_t). The 100(1-α)% prediction interval for the forecast Y_{t+k} are given by:

$$Y_{t+k} \pm t_{n-2,1-\frac{\alpha}{2}} \left\{ 1 + \sum_{j=1}^{k-1} \Psi_j^2 \right\}^{\frac{1}{2}} \sigma_Y \quad (5.\ 87)$$

where t(n-2,1-½α) is the 100(1-½α) percentage point of a t-distribution with (n-2) degrees of freedom. It is well known that the value of t depends on the number of observations. The model is not known exactly, and the model parameters need to be estimated (and hence the Ψs). Although some commercial packages are available to provide automatic ARIMA modelling (with dubious results – see Chatfield C., 1989), the ARIMA method is primarily intended for a non-automatic approach where the analyst uses subjective judgment and experience to select an appropriate model from the large family of ARIMA models according to the proprieties of the individuals series being analysed. Therefore, even if the procedure is more versatile than many competitors, it is also more complicated and considerable experience is required to identify an appropriate ARIMA model. Another drawback is that the method requires at least 50 observations to give a reasonable answer. The method works well for series showing short-term correlations (see section 5.3.7)

5.3.6 ARIMA model suitability

The advantage in using ARMA models is that they are quite flexible due to the inclusion of both autoregressive and moving-average terms. The sample correlation plots are used to identify the presence of non-stationary behaviours in the data, as well as to indicate the type of transformation required to remove them. Power transformations coupled with differencing transformations are included in the prognostics algorithm described in this work; they afford convenient methods of transforming a wide class of non-stationary time-series. The sample autocorrelation functions and the sample partial autocorrelation functions are then used to identify the indexes p, q that are the orders of the AR and MA models.

According to Box et al. (1976) the variance of the l steps ahead forecast for any origin t can be estimated and assuming that A_{t-i} are normally-distributed, it follows that, given information up to time t, the conditional probability distribution $p(X_{t+l}/X_t,X_{t-1},\ldots)$ of a future value of the time series will be normal with mean $\hat{X}t(l)$ and estimated standard deviation. Consequently the prediction intervals (PIs) for a specified significance level α can be calculated, a thorough investigation of the theory of the PIs follows in section 5.4.5.1.

5.3.7 Case study: TGT margin before returning to the maintenance facility

In this section, an application of the ARIMA method to the prediction of the failure risk for a short term lead time horizon, for a Trent 800 engine, is investigated. The case study considers the problem of performance rejection prescribed when the engine TGT exceeds a safety margin.

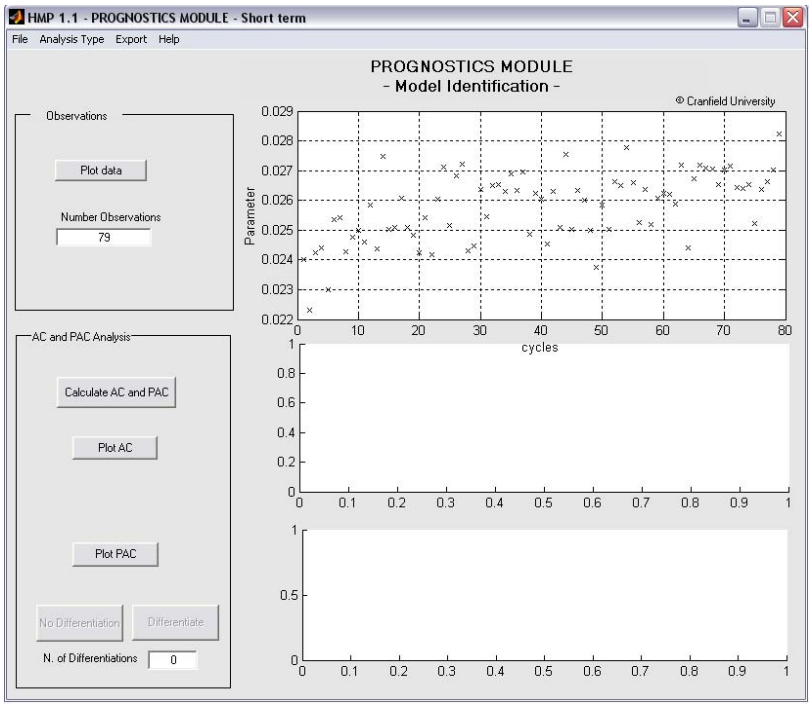


Figure 91: HMP 1.1 prognostics module, ARIMA model identification GUI

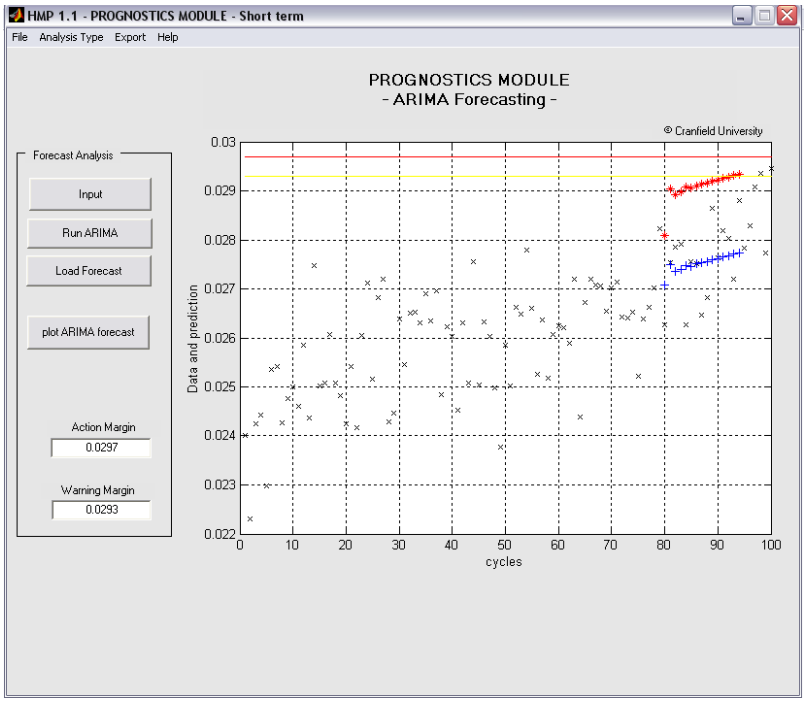


Figure 92: HMP 1.1 prognostics module, ARIMA forecasting GUI

The short term investigation framework part of the prognostics module of HMP 1.1 is constituted by two GUIs shown in Figure 91 and Figure 92. The first GUI is used in order to perform the ARIMA model identification, while the second computes the ARIMA forecast.

5.3.7.1 Performance rejections

These happen when the engine, because of the degradation process, is not able to guarantee the required thrust under the hot day limit condition. Civil aero-engines require for certification, a 'maximum take off rating' a 'maximum continuous rating', and a defined 'idle'. Engines ratings are prescribed maximum levels of thrust appropriate to different phases of flight. As far as the take off is concerned, in order to achieve the same payload and range requirements for both sea level and higher altitude runways the engine has to be run hotter for the airfields at higher altitude, in order to keep approximately the same thrust. The impact of ambient condition on engine performance can be offset by "flat rating" the T.O. thrust – see Figure 93. This results in an increase in turbine entry temperature (TET) and high pressure shaft speed with ambient temperature, at least up to the hot-day limit of flat rating. For day temperatures exceeding the limit of the flat rating, known as the "Kink Point", TET is held constant and therefore the guaranteed thrust/p_o reduces.

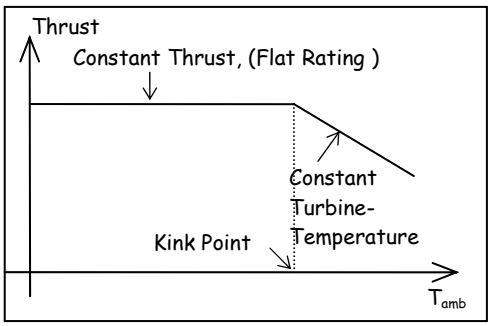


Figure 93: Typical rating curve.

When the engine degrades, the airlines have still to guarantee that it is able to deliver the prescribed thrust under the prescribed conditions (i.e. hot-day thrust). The engine is serviced when this capability cannot be assured anymore, unless the airline reduces the operating-temperature range.

5.3.7.2 Forecasting TGT changes

Table 63 shows 100 simulated data relative to % changes in TGT from the engine model at the actual power level.

The simulated data are realistic because based on the study of real on-wing data gathered in-flight from a Trent 800. So, a case study is created, considering the observations up to flight 79 as representative of the 'past', or else the time-series (historical data) that the prediction is based on. The observations from flight 80 up to flight 100 are then used to evaluate the accuracy of the prediction. The data are plotted in Figure 94. In-flight changes in TGT are kept under control since are measurable quantities strictly related to changes in TET and hence to the risk of overheating the engine.

Table 63: Simulated data. TGT % changes data from engine performance model at actual power level.

flight n.	delta TGT	flight n.	delta TGT	flight n.	delta TGT	flight n.	delta TGT
1	0.0240	26	0.0268	51	0.0250	76	0.0264
2	0.0223	27	0.0272	52	0.0266	77	0.0266
3	0.0242	28	0.0243	53	0.0265	78	0.0270
4	0.0244	29	0.0245	54	0.0278	79	0.0282
5	0.0230	30	0.0264	55	0.0266	***80***	*0.0263*
6	0.0254	31	0.0255	56	0.0253	*81*	*0.0275*
7	0.0254	32	0.0265	57	0.0264	*82*	*0.0279*
8	0.0243	33	0.0265	58	0.0252	*83*	*0.0279*
9	0.0248	34	0.0263	59	0.0261	*84*	*0.0263*
10	0.0250	35	0.0269	60	0.0263	*85*	*0.0276*
11	0.0246	36	0.0263	61	0.0262	*86*	*0.0275*
12	0.0259	37	0.0270	62	0.0259	*87*	*0.0265*
13	0.0244	38	0.0248	63	0.0272	*88*	*0.0268*
14	0.0275	39	0.0262	64	0.0244	*89*	*0.0286*
15	0.0250	40	0.0260	65	0.0267	*90*	*0.0276*
16	0.0251	41	0.0245	66	0.0272	*91*	*0.0282*
17	0.0261	42	0.0263	67	0.0271	*92*	*0.0280*
18	0.0251	43	0.0251	68	0.0271	*93*	*0.0274*
19	0.0248	44	0.0276	69	0.0265	*94*	*0.0273*
20	0.0242	45	0.0250	70	0.0270	*95*	*0.0284*
21	0.0254	46	0.0263	71	0.0271	*96*	*0.0271*
22	0.0242	47	0.0260	72	0.0264	*97*	*0.0288*
23	0.0260	48	0.0250	73	0.0264	*98*	*0.0288*
24	0.0271	49	0.0238	74	0.0265	*99*	*0.0272*
25	0.0252	50	0.0258	75	0.0252	*100*	*0.0276*

The problem addressed in this example regards the probability that the aircraft will complete the next 15 missions, which, in the scenario simulated, is the number of flights before the aircraft returns to the location of the operator's maintenance facility. This investigation is limited to the TGT margin. For the required short-time prediction in this work we use the ARIMA algorithm. Hence, the problem is to find a suitable time-series model for the data and compute forecasts for up to 15 flights ahead.

ARIMA model identification

The first step is to plot the data. First we note that there is no discernable seasonal variation (i.e. regular pattern which repeats itself after a certain number of basic time units) and this is typical of the deterioration phenomena we are studying. Besides we can observe a trend that has by no means a constant rate, so it would be a gross simplification to simply fit a straight line to

the data. With a single important time series like this it is worth investing some effort on the analysis, and the use of the Box-Jenkins procedure is indicated.
The time-series is clearly non-stationary and some sort of differencing is required. To confirm this, the auto-correlation coefficients up to the lag 20 were calculated and, as expected, the values at low lags were all large and positive – see Figure 95. Here we used the rule that values outside the range $\pm 2/\sqrt{79} = \pm 0.22$ are significantly different from zero, where 79 is the number of observations.

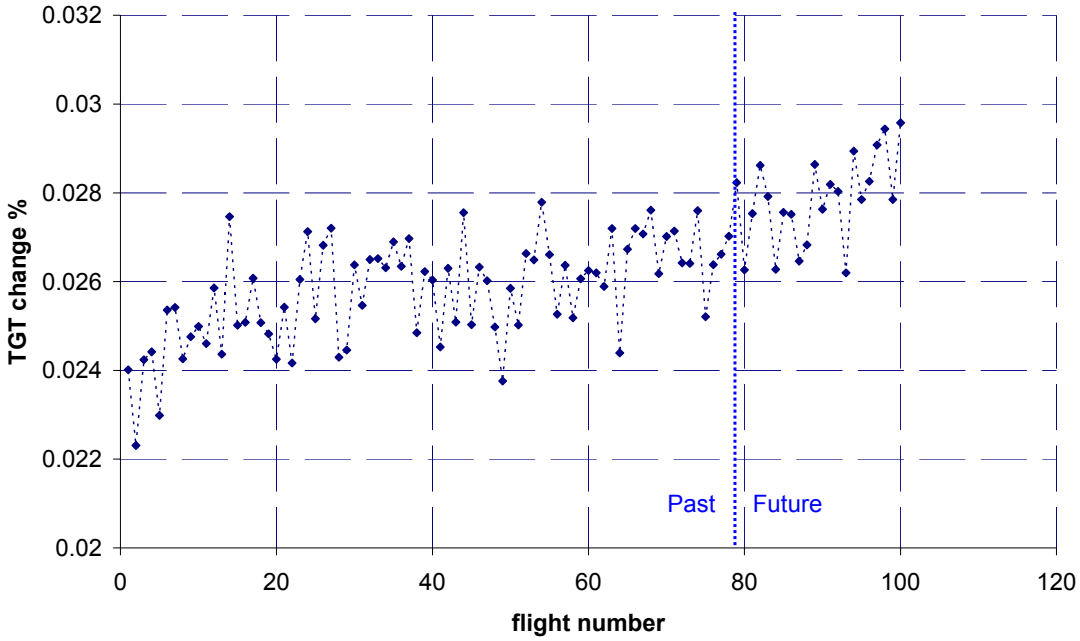


Figure 94: Simulated data. Plot of TGT % changes from engine performance model at actual power level against flight number. Future is assumed to start at flight 80.

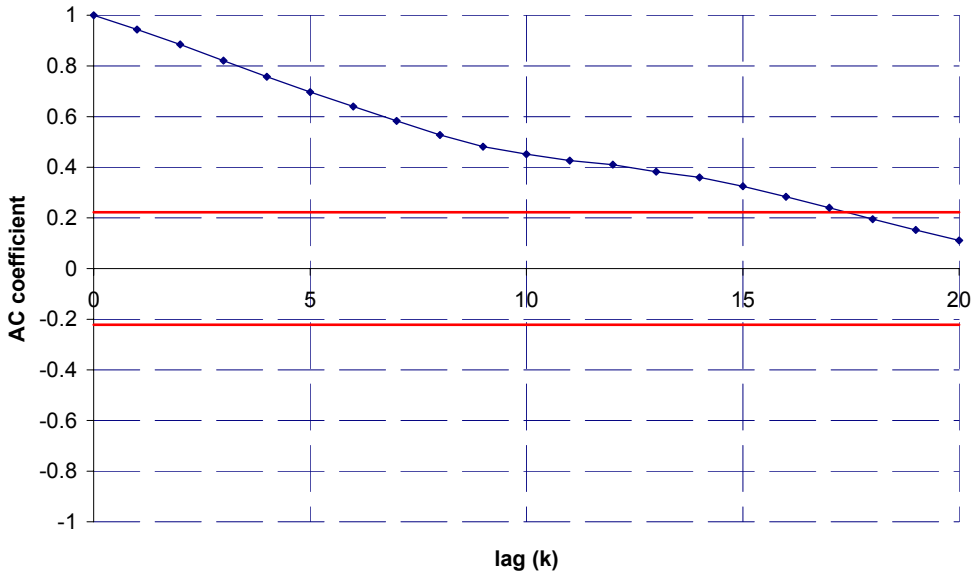


Figure 95: AC plot for the data, TGT

The simplest type of differencing procedure is to take first differences of the data, and this was done next – see Figure 96.

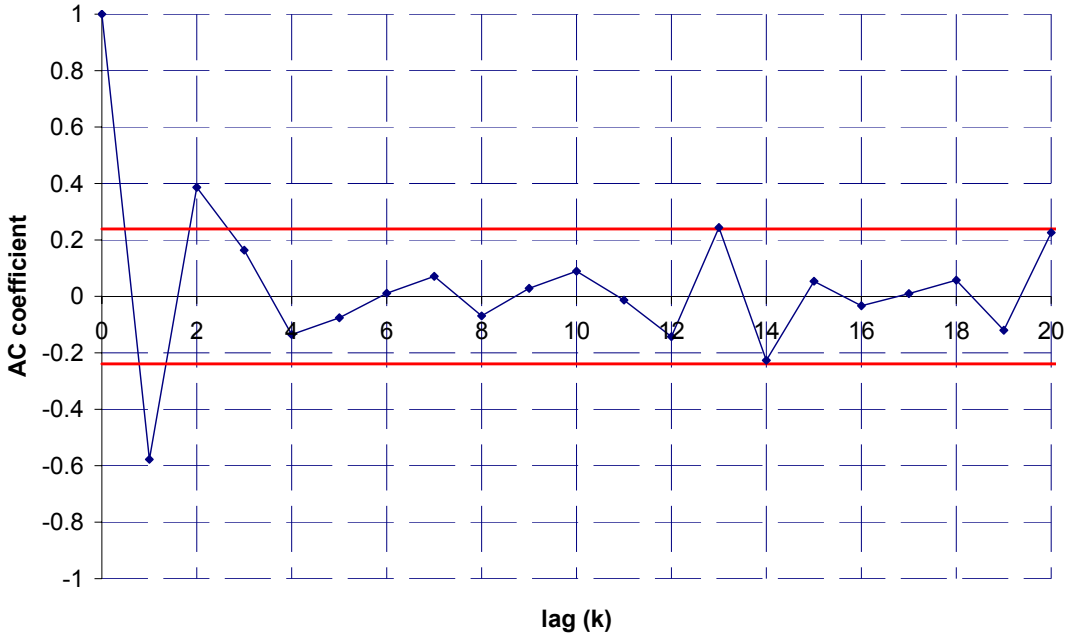


Figure 96: AC plot for the transformed data, TGT

The auto-correlation plot of the first differences turned out to be of a particularly simple form. The coefficient at lag 1 and 2 were -0.57 and 0.39 respectively, which are significantly different from zero, whereas all subsequent coefficients up to lag 20 were not significant different form zero. This tells us two things. First, the differenced series is now stationary and so no further differencing is required. Secondly, the order of the autoregressive part of the ARIMA model can be chosen to be 2.

In order to identify the order of a suitable MA model as part of the ARIMA model, we need to calculate the partial auto-correlation coefficients. The PAC plot is given in Figure 97. As for the AC coefficients, values whose modules exceed 0.22 may, as a first approximation, be taken to be significantly different from zero. In the PAC plot we note significant values at lags 1 and 2 (respectively 0.39 and 0.33). Hence an AR(2) model was identified as suitable.

An appropriate ARIMA model can now be identified and fit to the data. According to the previous analysis of the AC and PAC plots, the orders p, d, q are set to be 2,1,2. The model is fitted and an estimate of the residuals variations is provided. According to the diagnostics check method proposed in section 5.3.4, in this example there is no evidence that the fitted model is inadequate and so no alternative models was tried.

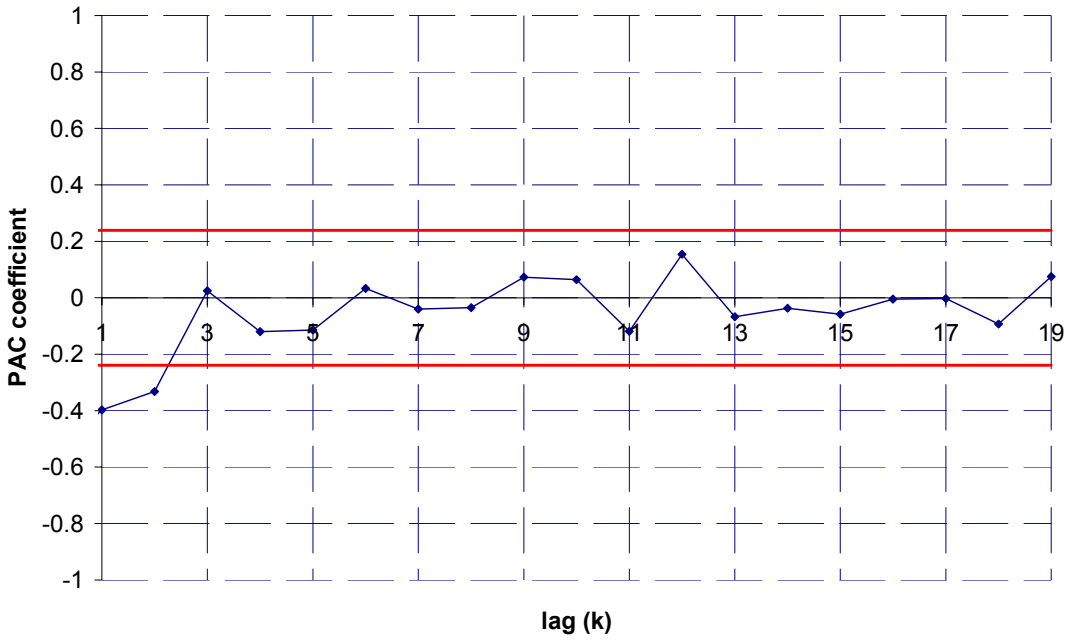


Figure 97: PAC plot for the transformed data, TGT

Forecasting and prognosis

Forecasts and prediction intervals can now be computed. The diagram in Figure 98 shows TGT percent changes from engine performance model at actual power level (simulated data) plotted against the flight number.

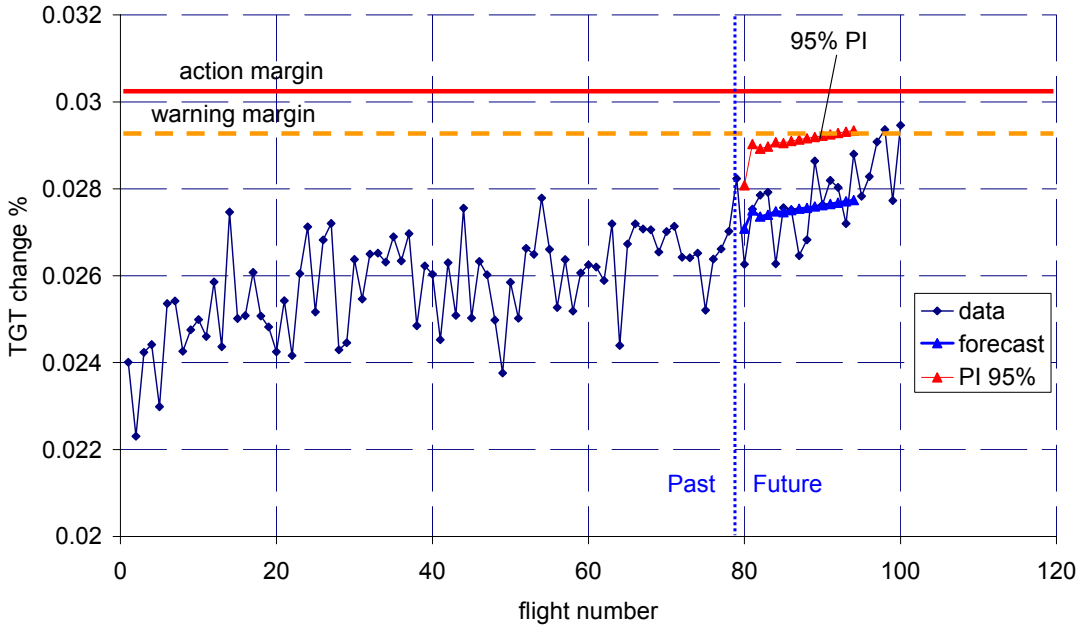


Figure 98: Simulated data – TGT plot. Changes from engine performance model at actual power level. 15 flights forecast ahead and 95%PI

The ARIMA outcomes are 15 flight-ahead forecasts and 95% (upper) prediction interval. In the diagram, a safety margin indicating when a corrective action should take place and a warning margin are plotted. The above model was

fitted to the first 79 flights data, with the intention of comparing the forecasts for 15 flights ahead with the observed values. It is interesting to notice how the projection predicts fairly well the trend in the observed data. The PIs get wider as the distance ahead increases to take into account an additional allowance for the forecast error. The conclusion that can be drawn analysing the time-series is that no warning is expected for the next 11 missions with 95 % confidence. This information enables a more cost-effective maintenance plan and mission schedule to be devised. This situation may result, for example, in the prognosis that the aircraft can continue to operate between city pairs in cooler climates, in which the engine operates at cooler internal temperatures.

5.3.8 Case study: shaft speed in-flight data

The variations of speed of a gas turbine shaft are regularly monitored and their values at TO condition must not overtake a prescribed margin for the safety of the engine. Hence, a second case study is here dedicated to the investigation of a time series constituted by 86 Trent 800 in-flight real data regarding percentage changes of intermediate pressure shaft speed from the engine model at the actual power level, courtesy of Rolls Royce Plc. The safety and warning margins used are arbitrarily fixed for the scope of this example. The observations up to flight 70 are assumed to be the time-series, which the forecasts are based on. The data from flight 71 up to flight 86 are then used to compare forecasts with the observed values. As previously, we begin by plotting the data - see Figure 101. This shows that there is no seasonal variation but a trend is present in the data. The first differences of the data are computed to get rid of the non-stationarity present in the observations. The auto-correlation and partial auto-correlation plots of the first differences are computed and shown in Figure 99 and Figure 100, in order to find a suitable ARIMA model.

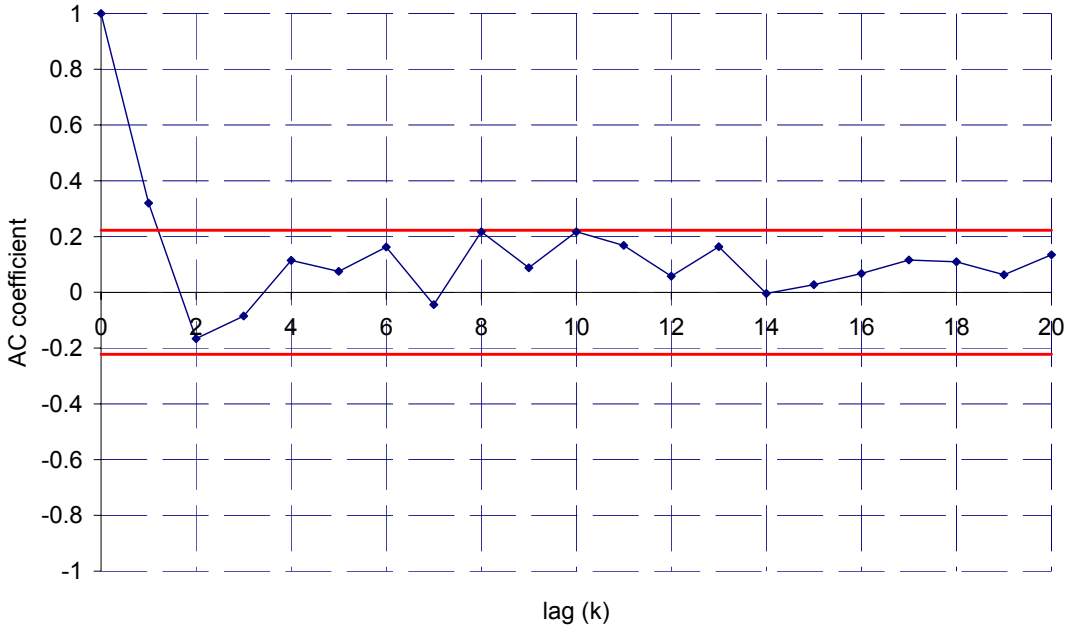


Figure 99: AC plot for the transformed data, N_2

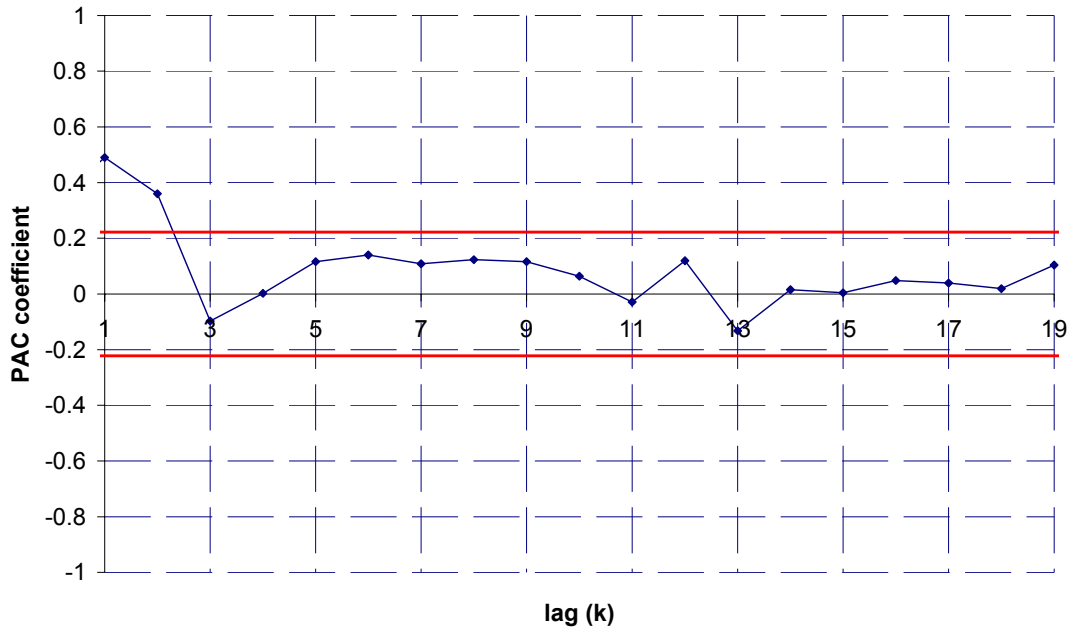


Figure 100: PAC plot for the transformed data, N_2

The analysis of the AC and PAC plots leads to the identification of an ARIMA(2,1,1) model: one AC and two PAC coefficients result significantly different from zero, exceeding $\pm 2/\sqrt{70} = \pm 0.24$. The model is fitted and from the analysis of the residuals no evidence of lack of fit is shown. So, forecasts and PIs are computed for the next 16 flights. A warning is expected at flight 81 with 95% confidence.

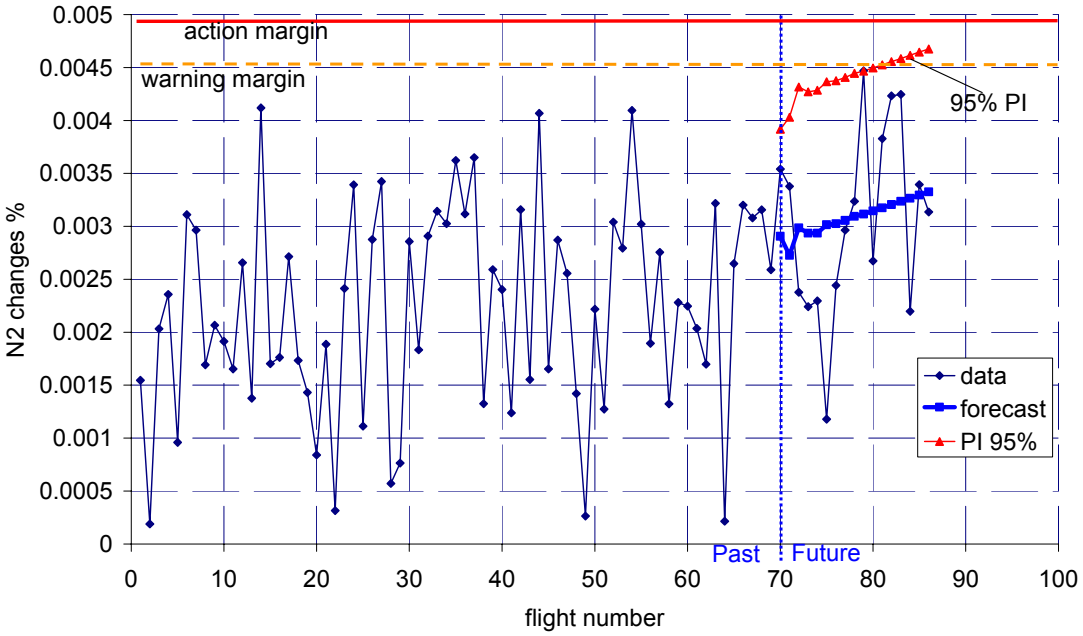


Figure 101: In- flight data, Curtsey of Rolls-Royce Plc. – N_2 plot. Changes from engine performance model at actual power level. 15 flights forecast ahead and 95%PI, the margins are arbitrarily fixed.

5.3.9 Summary

This section is aimed at summarizing the concepts discussed while providing some comments and recommendations. If short term projections are required with regards to engine performance analysis, ARIMA models are indicated as suitable to achieve a significant level of accuracy. The method has a remarkable capability of discovering trends and predicting turning points in short term analysis. The following observations can be made:

- The best fit model in the data does not guaranty the minimum error in the forecast. The further the forecasting horizon is (long-term investigation) the more it might lead to wrong predictions. Besides while for short term forecasting complex models can be justified, in a long term they give forecast that might not be better than simple models.
- Combinations of forecasts from different techniques or from the same method but with different assumptions can allow better results. Besides adjustments in view of any additional information is advisable.
- The confidence level of the prediction interval is strictly related to the business intention and on the experience.

A strategy for non-automatic univariate forecasting was proposed above based on Box-Jenkins approach.

Two examples have been considered to provide some indication of the type of problem which may be addressed with Box-Jenkins ARIMA analysis with regards to gas-path prognostics, and the type of judgement which has to be made. The applications concern the prediction of the failure risk for a short term lead time horizon concerning performance rejection. The first case study regards simulated data of TGT, while in the second case real data of intermediate pressure shaft speed are considered.

A critical assessment of the method was made in this chapter. The accuracy that this method can provide in short-term forecasting is nevertheless affected by the ability of the analyst. As matter of fact, some of the practical difficulties which arise in using ARIMA models were highlighted, particularly in regard to the differentiation and the subjective interpretation of the correlograms.

5.4 Long-term forecasting procedure: regression

5.4.1 Introduction

Figure 102 illustrates a forecasting procedure devised for long term engine performance prognoses based on the use of regression analysis.

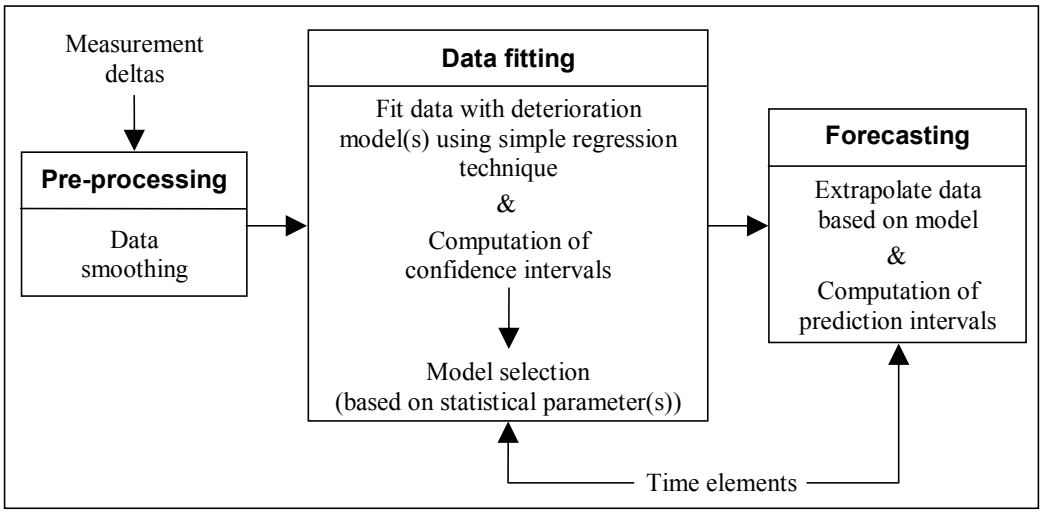


Figure 102: Forecasting procedure – with Regression

The procedure is constituted of three elements: (i) pre-processing, (ii) data fitting and (iii) forecasting modules. Measurements are pre-processed in order to treat the sensor-noise that affects them. This significantly improves the performance of the forecasting operation (Mathioudakis et al., 2002). Note that the smoothing operation is applied on the measurement deltas and not on the measurements themselves. A simple regression technique is then used to fit the data with one of the three deterioration models (i.e. soft, linear or severe model) considered in the next section. The forecasting period is required in order to perform the regression operation during which the coefficients of the equation are also evaluated. The procedure also evaluates the coefficient of determination which is used for model selection. After the most suitable model has been chosen, the data can be extrapolated into the future to generate a forecast. The forecasting horizon must obviously be specified in order to perform this operation during which the prediction intervals are also computed.

5.4.2 Deterioration modelling

The failure rate versus time relationship for most mechanical equipment can be modelled with the typical ‘bath-tub’ curve – see Figure 103. Three phases are distinguished: run-in, design-life and wear-out phases. The failure rate is high at the beginning, then stabilises to a constant value and increases again at the end of the life (Abernethy, 2000, and Li, 2002).

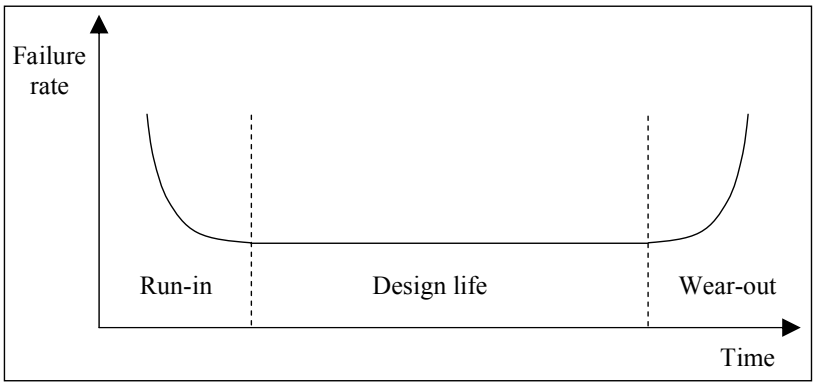


Figure 103: Typical bath-tub curve showing operational phases (Li, 2002)

As far as the gas-turbine performance deterioration is concerned, we can observe that in average it starts off at high rate and than it tends to settle down to a lower rate. It has been recognised that service use imposes external distortion loads and differential thermal growths, slightly in excess of those experienced during the post build engine running-in testing, which further wear seal clearances. There comes a point, in service use, at which no further wear due to the above occurs except under the rare instances of a violent manoeuvre. Therefore, the deterioration rate diminishes (Crosby, 1986). What of course will still continue, but generally at low rate, will be (i) the corrosion of blades, (ii) the erosion of seals and blades due to contaminants and particles in the air, and (iii) the fouling or deterioration caused by the adherence of particulate contaminants to the gas turbine airfoil and annulus surfaces.

To study and to model the evolution of gas turbine performance deterioration, we introduce two curves, one representative of the health level trough the generic parameter, Figure 104, and one for the deterioration rate, i.e. its derivative. For simplicity the diagrams that follow refer only to the case of a parameter that deteriorating decreases its value (e.g. compressor delivery pressure, P3). The generalization to the case of a parameter (e.g. TGT) that increases its values can be simply done inverting the slopes of all the curves respect to the x axis. The curve in Figure 104 shows the trajectory of a generic deteriorating parameter: its derivative, that represents the deterioration rate, follows a typical bath-tub curve Figure 103.

The curves considered so far represent the statistical behaviour of a given sample of engines (Sasahara, 1986) and not the behaviour of any single engine. The technique described in this work is based on the assumption that, for the single engine under analysis the mathematical model of the deterioration does not vary during the short-term that includes the data used for the forecasts and the forecasting horizon, while the parameters that make that model specific for the particular engine provide the adaptation of the model. Nevertheless, any error related to this assumption can be accounted for in the calculation of the prediction intervals.

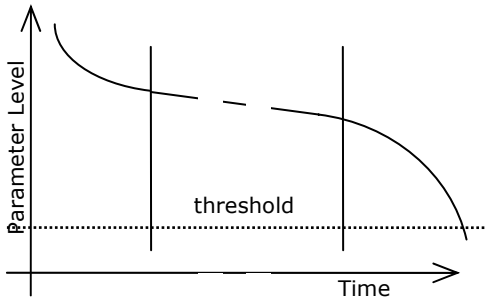


Figure 104: Typical evolution of deterioration, generic parameter. A typical threshold of acceptance for a parameter is shown.

So, three rate trends are taken into account in the method proposed, as shown in Figure 105**:**

a) decreasing trend (linear or non-linear)
b) constant trend
c) increasing trend (linear)

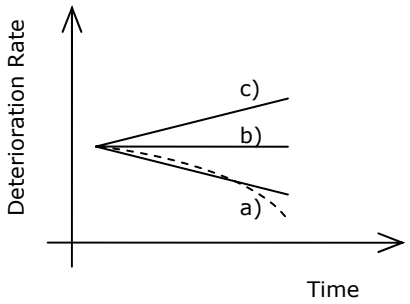


Figure 105: Generic gas-path parameter deterioration rate: three models

The corresponding level curves are:

a) trajectory that curves upwards (positive convexity)
b) straight line trajectory (linear)
c) trajectory that curves downwards (negative convexity)

In the prediction procedure described subsequently, in accordance to other authors (Sasahara, 1986), the decreasing rate trend (case a) has been chosen to be non-linear. Typical model trajectories over time of the gas-path parameters are assumed to be the following (Figure 106):

a) $Y = a\sqrt{t} + b$ (5.88)
soft deterioration model (positive convexity)

b) $Y = at + b$ (5.89)
linear deterioration model

c) $Y = at^2 + b$ (5. 90)
severe deterioration model (negative convexity).

Where *Y* identifies a time-series of a generic parameter; *t* is the time unit; and *a* and *b* are two coefficients, with *a*<0 for decreasing curves.

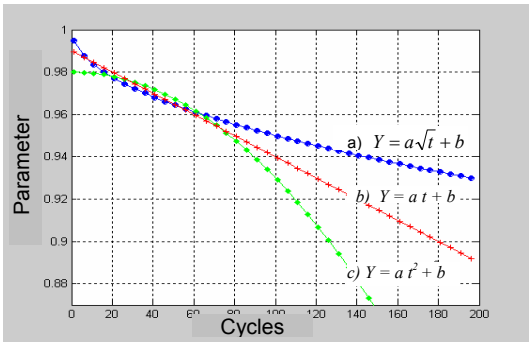


Figure 106: Three models for deterioration level curve: generic parameter.

Most commonly gradual deteriorations of a generic parameter during engine's life follow a 'soft deterioration model' with positive convexity. At a certain point in time the trajectory might change according to a 'severe deterioration model' with negative convexity (Figure 107). Life in the context of performance analysis is meant as that period of gas turbine usage that terminates with the performance being inadequate for the prescribed requirements, i.e. the engine is rejected from service until a maintenance action is undertaken.

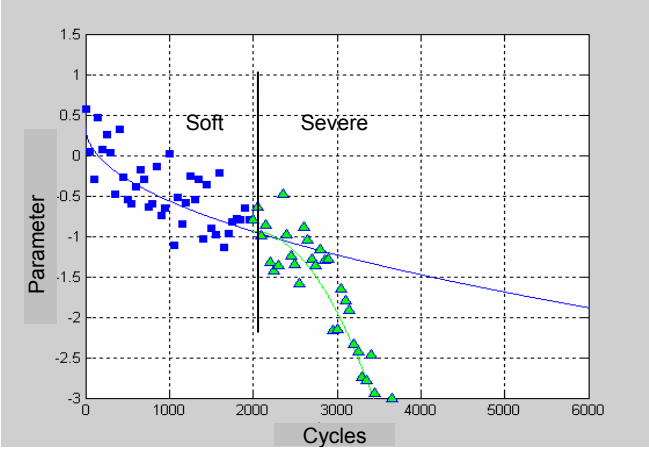


Figure 107: Transition from soft to severe rate of deterioration: generic parameter

5.4.3 Smoothing pre-processor

An algorithm that functions as smoothing pre-processor is required before any attempt of extracting information from historical on-wing data is pursued in order to perform long-term forecasts. Sensors are sources of considerable noise, and this can be treated smoothing the data, which makes the underlying pattern easier to detect. Three different smoothing techniques have been introduced. It has been recognised that if a versatile prognostic process is required, the opportunity of selecting the most appropriate smoothing technique for different situations increases the functionality of the process.

The three techniques are simple moving average (MA), weighted MA and exponential smoothing. It is worthwhile noticing that MAs and exponential smoothing techniques, as mentioned before, can be used as forecasting in their own right (Wheelwright et al., 1977). In this work those methods are used for smoothing only, so the second step is disregarded.

Simple moving average

This technique smoothes the data by taking a set of observed values, finding their average and then replaces the noisy measurement data with this average value. It follows that the idea behind moving averages is that observations that are nearby in time are also likely to be close in value (Makridakis et al., 1998).

A decision has to be taken regarding the number of data points included in the calculation of the MA. This is done by specifying values for the lead and lag variables. They indicate the number of previous (lead) and following (lag) data points that are used together with the current data point in calculating the MA. The number of data points included in a MA affects the smoothness of the resulting curve. The more observations are used in calculating the MA, the smoother will be the computed curve and the slower will the MA react to changes in the parameter.

Weighted moving average

Simple MA, as considered above, assigns equal importance to each of the observations. However, since the values of the nearest observations are closest to the considered data point, an argument can be made that they should be given relatively more weight. Additionally, instead of observations entering and leaving the average abruptly, they can be slowly down-weighted, resulting in a smoother curve (Makridakis et al., 1998).

For the weighted MA, weights have to be specified for all of the observations that are to be included in the calculation of the average. Hence, the number of specified weights automatically defines the size of the "moving window". Additionally, the pivot has to be specified. This parameter defines what weight is to be applied to the current data point. The best weights might be found by trial and error or alternatively, using widely used sets proposed in the literature (Makridakis et al., 1998).

Exponential smoothing

If the weights are chosen decreasing exponentially going further into the past, the approach is commonly known as exponential smoothing. For this technique only one variable, namely alpha, has to be specified. The smoothed (sm) value for the data point under consideration (at the origin 0) is indicated as $Y_{0,\mathrm{sm}}$ and is computed as follows:

$$\mathrm{Y}_{0,\,\mathrm{sm}} = \alpha \mathrm{Y}_{-1} + \alpha(\alpha-1)\mathrm{Y}_{-2} + \alpha(\alpha-1)^2 \mathrm{Y}_{-3} + \alpha(\alpha-1)^3 \mathrm{Y}_{-4} + \ldots \qquad (5.91)$$

Since alpha is a number between 0 and 1, the weights of the observations decrease as you go back in time. The main difference with weighted MA methods is that exponential smoothing relies solely on past data (Makridakis et al., 1998).

5.4.3.1 Functionality of the smoothing pre-processor

The functionality of the techniques discussed above is shown via an example with simulated data based on real Trent 800 data. For the scope of this section the compressor delivery pressure (P3) is chosen as the dependent parameter to study. The sensor noise level is equivalent to an RMS value of 0.25.

The different smoothing methods are illustrated using 500 cycles data. The more the data are smoothed, the smaller will generally be the error in fitting the

deterioration model. However, care must be taken not to over smooth the data and with them the information carried in the observations. Figure 108,Figure 109 and Figure 110 show the results of smoothing the P13 data with each of the three methods. It can be seen that the three techniques have similarities but provides different smoothing capabilities.

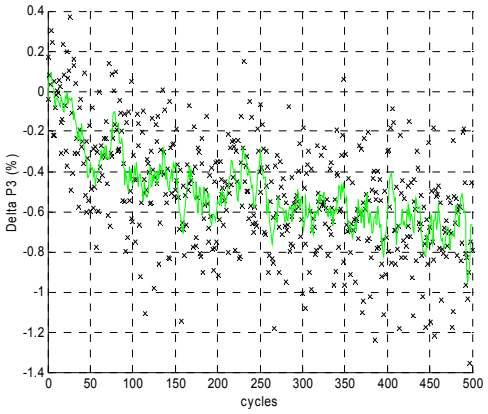


Figure 108: 7 point simple MA

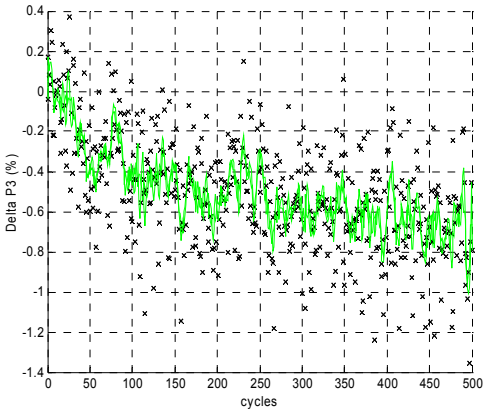


Figure 109: Weighted (weigths:0.555, 1.11, 1.665, 1.11, 0.555)

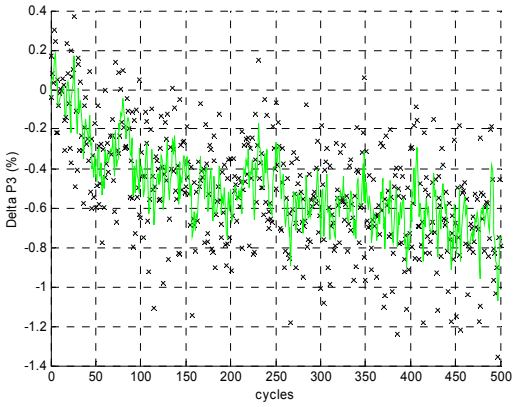


Figure 110: Exponential smoothing (alpha=0.4)

5.4.4 Trending algorithm

This section considers a trending algorithm based on regression analysis, employed to fit the different deterioration models (discussed in section 5.4.2) to the engine data. The selection of the most appropriate model, based on two statistical parameters, is examined here.

In regression analysis of time-series, historical data are represented by mathematical models that are analytical functions of time. A general form of model can be expressed as:

$$Y_i = \beta_0 + \beta_1 f_1(X_i) + ...\beta_k f_k(X_i) + \varepsilon \qquad (5.\ 92)$$

where:

Y_i is the dependent variable (engine measurement) for the i-th observation of the independent variable (cycle) X_i.

f_k is the k-th function related to the k-th coefficient.
β_k are the k coefficients, to determine.
ε_i denotes the error or random component (deviation of the observation from the assumed relationship) The random component is assumed to have an expected value $E(\varepsilon) = 0$, and variance $V(\varepsilon) = \sigma^2$.

In particular, the three mathematical models of the deterioration, described have the following simplified form:

$$Y_i = \beta_0 + \beta_1 X_i^r + \varepsilon_i \qquad (5.\ 93)$$

where

r = 1, 0.5 or 2 depending on the selected deterioration model
β_0 is the intercept
β_1 is the slope coefficient

Forecasting consists of estimating the unknown parameters in the appropriate model and using these estimates, to project the model into the future and obtain a forecast and its prediction interval. Therefore, the prediction procedure is divided into two steps: the fitting procedure and the forecasting procedure. A wide range of techniques can be used to estimate the unknown parameters in the way that the model best fits the observations within the interval observed. In this work, the least-square method has been used. The estimates are chosen to minimise the error or residual. The procedure proposed is based on the assumption that the mathematical model fitted in the data is correct. The analysis includes a method to consider the validity of this assumption, which enables the identification of the best fitted model among the three considered by calculating and comparing the Coefficient of Determination (Montgomery, 1990). The fitted model is given by:

$$\hat{Y}_i = b_0 + b_1 X_i^r \qquad (5.\ 94)$$

where: $\hat{Y}_i$ denotes the predicted value of Y for a given X
b_0 is the estimate of the intercept
b_1 is the estimate of the slope coefficient.

The difference between the fitted model and the data is the estimated error or residual and denoted by e_i.

A least squares method is employed to estimate the values of b_0 and b_1 so that the resulting curve represents the best fit to the data. This is done by minimising the sum of the squared errors, SSE.

$$SSE = \sum_{i=1}^{n} \varepsilon_i^{\,2} = \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i^r\right)^2 \qquad (5.\ 95)$$

where n is the number of observations.

5.4.4.1 Linear model

In the case of linear model the exponent r is fixed to 1. The derivation of the estimates for the linear model is shown below.
In order to minimise SSE it is necessary that b_0 and b_1 satisfy

$$\left.\frac{\partial SSE}{\partial \beta_0}\right|_{b0,b1} = 0 \qquad \left.\frac{\partial SSE}{\partial \beta_1}\right|_{b0,b1} = 0 \tag{5. 96}$$

This results in the following equations:

$$-2\sum\left(Y_i - b_0 - b_1 X_i\right) = 0 \qquad -2\sum X_i\left(Y_i - b_0 - b_1 X_i\right) = 0 \tag{5. 97}$$

From equation (5. 97)

$$-\sum Y_i + nb_0 + b_1\sum X_i = 0 \tag{5. 98}$$

And so

$$b_0 = \frac{1}{n}\left(\sum Y_i - b_1\sum X_i\right) \tag{5. 99}$$

The expression for the intercept can be rewritten in terms of the mean X and Y values.

$$b_0 = \overline{Y} - b_1\overline{X} \tag{5. 100}$$

where

$$\overline{Y} = \frac{\sum Y_i}{n} \qquad \overline{X} = \frac{\sum X_i}{n} \tag{5. 101}$$

From equation (5. 97)

$$-\sum X_i Y_i + b_0\sum X_i + b_1\sum X_i^2 = 0 \tag{5. 102}$$

Substituting the value of b_0 from equation (5. 99) into equation (5. 102)and solving for b_1 yields

$$b_1 = \frac{n\sum X_i Y_i - \sum X_i\sum Y_i}{n\sum X_i^2 - \left(\sum X_i\right)^2} \tag{5. 103}$$

Similarly to b_o, b_1 can also be expressed in terms of the mean Y and X values. After some algebraic manipulation the following equation is obtained

$$b_1 = \frac{\sum\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum\left(X_i - \overline{X}\right)^2} \tag{5. 104}$$

Substituting equation (5. 100) into equation (5. 94) gives:

$$\hat{Y} = \overline{Y} + b_1(X - \overline{X}) \tag{5. 105}$$

This form of the regression equation will be used later when computing the prediction intervals. Once the coefficients have been estimated, the fitted curve can be computed by inserting the coefficients into the linear equation.

5.4.4.2 Non-linear models

The non-linear models used in this work correspond to equation (5. 94) with the exponent r = 0.5 and 2 as discussed earlier. In those circumstances it has to be recognised that the origin of the curve affects the shape of the models. For the severe model, the origin will typically be the start of the severe deterioration phase, while for the soft model it is usually the point of the last overhaul. (Figure

111) The capability of fixing an origin outside the time window that contains the data used to fit the model has been included – see Figure 111. Therefore, a particular approach had to be developed for estimating the coefficients and computing the fitted curve. The method involves four basic steps.

1. The origin is placed at the data point of the cycle of the last maintenance or at the start of the severe deterioration phase (O_{soft} and O_{severe} in Figure 111). The X values (on the time axes) are adjusted by subtracting the cycle of last maintenance, or the cycle at which the severe deterioration starts, from the X values in the time window. The modified Y values are obtained by deducting the Y value at the point of last maintenance (or start of severe deterioration) from the Y values of the time series.
2. The slope coefficient of the non-linear model is then evaluated using the least squares method using the modified X and Y values. The intercept is assumed to be 0, so that the curve is forced through the local origin.
3. The fitted Y values are computed based on the calculated slope coefficient.
4. Finally, the Y value at the point of last maintenance (or the start of severe deterioration) is added to the Y values of the fitted model.

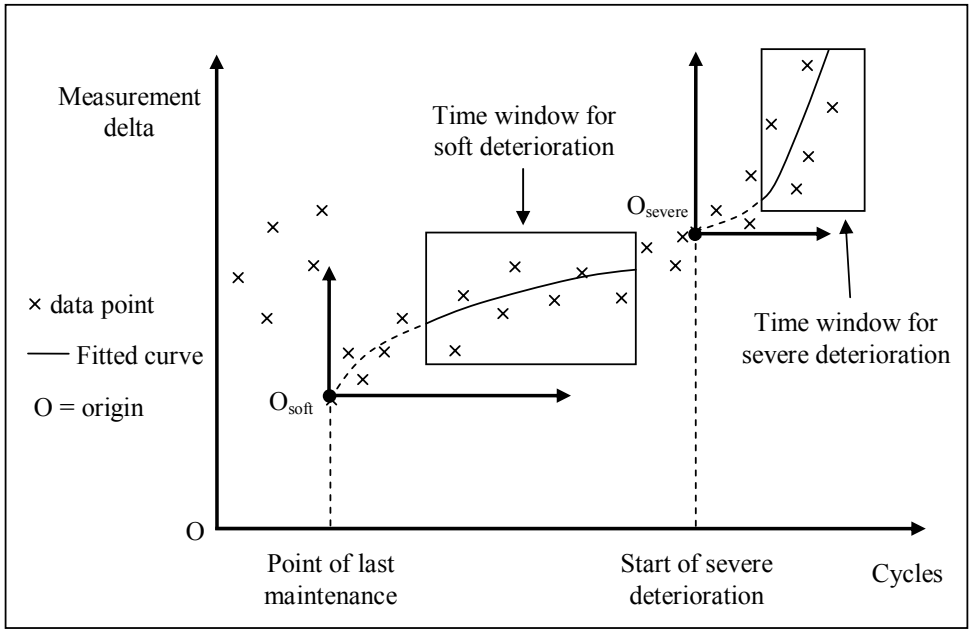


Figure 111: Data fitting for non-linear models, (Moes T., 2003)

The derivations of the non-linear models, follows the same logic of the linear model with the only difference that for the non-linear models the curves are forced through a local origin (see Figure 111), by modifying the Y and X data.

$$X_{m,i} = X_i - X_{ref} \qquad Y_{m,i} = Y_i - Y_{ref} \tag{5. 106}$$

where

Y_i is the dependent variable (engine measurement) for the *i*-th observation

X_i is the independent variable (cycle) for the *i*-th observation

$Y_{m,i}$ is the modified dependent variable for the *i*-th observation
$X_{m,i}$ is the modified independent variable for the *i*-th observation
X_{ref} is the cycle for the point of last overhaul or the start of severe deterioration.
Y_{ref} is the Y value at the point of last overhaul or the start of severe deterioration.

These modified values are used for deriving the slope coefficient, the intercept being equal to zero. The time series can then be described by the following equation.

$$Y_{m,i} = \beta_1 X_{m,i}^{r} + \varepsilon_i \qquad (5.\ 107)$$

where r= 0.5 or 2 depending on the selected non-linear model
β_1 is the slope coefficient
ε_i denotes the error (deviation of the observation from the assumed relationship)

The fitted model is then given by

$$\hat{Y}_{m,i} = b_1 X_{m,i}^{r}$$

where $\hat{Y}_i$ denotes the predicted relative value of Y for a given modified X
b_1 is the estimate of the slope coefficient.

The slope coefficient is estimated using the least squares method, which aims to minimise the sum of the squared errors, SSE.

$$SSE = \sum_{i=1}^{n} \varepsilon_i^{\,2} = \sum_{i=1}^{n} \left(Y_{m,i} - \beta_1 X_{m,i}^{r} \right)^2 \qquad (5.\ 108)$$

where n is the number of observations.
In order to minimise SSE it is necessary that b_1 satisfies

$$\left. \frac{\partial SSE}{\partial \beta_1} \right|_{b1} = 0 \qquad (5.\ 109)$$

This results in the following equation:

$$-2 \sum X_{m,i}^{r} \left(Y_{m,i} - b_1 X_{m,i}^{r} \right) = 0 \qquad (5.\ 110)$$

The slope coefficient can then be found from the following expression,

$$b_1 = \frac{\sum X_{m,i}^{r} Y_{m,i}}{\sum X_{m,i}^{2r}} \qquad (5.\ 111)$$

The fitted curve can now be computed from,

$$\hat{Y}_i = \hat{Y}_{m,i} + Y_{ref} = b_1 X_{m,i}^{r} + Y_{ref} \qquad (5.\ 112)$$

5.4.4.3 Model selection

The analysis includes a method, based on two parameters, to quantify the validity of the assumption that the mathematical model fitted to the data, is correct and therefore provides a good representation of the deterioration behaviour. This practice gives in output the best fitted model among the three

considered by calculating and comparing the Coefficient of Determination and the standard deviation of the residuals (Montgomery, 1990).

Coefficient of determination

The coefficient of determination (R^2) is the most common method for assessing the goodness of the fit. It separates the total amount of variation into two parts: explained and unexplained variation. R^2 is then defined as the ratio of the explained variation over the total variation.

$$R^2 = \frac{\text{explained variation in Y}}{\text{total variation in Y}} = \frac{\sum\left(\hat{Y}_i - \overline{Y}_i\right)^2}{\sum\left(Y_i - \hat{Y}_i\right)^2 + \sum\left(\hat{Y}_i - \overline{Y}_i\right)^2} \qquad (5.\ 113)$$

R^2 varies between 0 and 1, where 0 indicates complete lack of fit and 1 a perfect fit. If, for example, R^2 is equal to 0.6, this means that the model accounts for 60% of the variation in the value of Y. Sometimes similar values of R^2 might be obtained for two models. In this case, it is desirable to have a second criterion for describing the goodness of fit.

Standard deviation of the residuals

The variance of the estimated error or residuals is given by

$$s^2 = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2} = \frac{\sum e_i^2}{n-2} \qquad (5.\ 114)$$

The denominator of the above expression denotes the degrees of freedom, which can be defined as the number of data points minus the number of estimated parameters (β_0 and β_1). The standard deviation, s, is simply the square root of the variance.

$$s = \sqrt{\frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}} \qquad (5.\ 115)$$

This measure is useful when two models give similar high values for R^2. One might think that there is not a great deal to choose between the two models. However, there could be a large difference in the standard deviation of the residuals for the two models. In this case a comparison of the two methods suggests that the standard deviation of the residuals is a more appropriate and useful measure of fit (Wittink, 1988).

In order to make a suitable choice, both measures are computed by the forecasting program.

5.4.5 Forecasts and prediction intervals

Once the coefficients of the model have been evaluated the computation of the forecast is straightforward. The equation of the model is simply extrapolated for future X values. A measure of the uncertainty in the predictions is desirable. This can be provided by means of the prediction intervals, which are defined as the interval in which a given percentage of the predictions are expected to lie.

5.4.5.1 Prediction intervals: linear model

In order to compute the prediction intervals the variance of the predicted Y values has to be known. The following section describes the derivation for the linear model. It has been shown before that:

$$\hat{Y} = \overline{Y} + b_1(X - \overline{X}) \tag{5. 116}$$

The variance of the computed values can then be found from

$$V(\hat{Y}_0) = V(\overline{Y}) + (X_0 - \overline{X})^2 V(b_1) \tag{5. 117}$$

where V(A) is the variance of parameter A

$\hat{Y}_0$ is the predicted value at a specific value X_0

The variance of the mean Y value is equal to the variance of the Y value divided by the number of observations.

$$V(\overline{Y}) = \frac{V(Y_i)}{n} = \frac{\sigma^2}{n} \tag{5. 118}$$

Variance of b_1

An expression for b_1 in terms of the mean values was derived before

$$b_1 = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} \tag{5. 119}$$

Further it can be shown that

$$\sum (X_i - \overline{X})\overline{Y} = 0 \tag{5. 120}$$

Thus

$$b_1 = \frac{\sum (X_i - \overline{X})Y_i}{\sum (X_i - \overline{X})^2} = \{(X_1 - \overline{X})Y_1 + ... + (X_n - \overline{X})Y_n\} / \sum (X_i - \overline{X})^2 \tag{5. 121}$$

Now the variance of a function

$$a = a_1 Y_1 + ... + a_n Y_n \tag{5. 122}$$

is

$$V(a) = a_1^2 V(Y_1) + ... + a_n^2 V(Y_n) \tag{5. 123}$$

If the Y_i are pair wise uncorrelated and a_i are constants

$$V(a) = (a_1^2 + ... + a_n^2) V(Y_i) = \left(\sum a_i^2\right)\sigma^2 \tag{5. 124}$$

In the expression for b_1

$$a_i = \frac{(X_i - \overline{X}_i)}{\sum (X_i - \overline{X})^2} \tag{5. 125}$$

Since the X_i can be regarded as constants, the variance of b_1 can be found after reduction

$$V(b_1) = \frac{\sigma^2}{\sum\left(X_i - \overline{X}\right)^2} \qquad (5.\ 126)$$

Variance of predicted Y values

Substituting equations (5. 118) and (5. 126) into equation (5. 117),

$$V(\hat{Y}_0) = \frac{\sigma^2}{n} + \frac{(X_0 - \overline{X})^2 \sigma^2}{\sum\left(X_i - \overline{X}\right)^2} \qquad (5.\ 127)$$

Note that the variance of the random error, σ^2 is unknown. However, if the selected model adequately describes the data, then it may be estimated from the standard deviation of the residuals, derived before. Hence, the estimated standard deviation of the predicted value of Y is

$$s.d.(\hat{Y}_0) = s\sqrt{\frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum\left(X_i - \overline{X}\right)^2}} \qquad (5.\ 128)$$

The above equation gives the standard deviation for the conditional mean of Y. However, with regards to engine deterioration we are more interested in quantifying the statistical uncertainty of the estimation for one individual observation rather than the conditional mean. The uncertainty of individual predictions is always larger because for an individual observation the uncertainty results not only from the estimates of the coefficients but also from the error term in the equation. For the conditional mean the error can be neglected because on the average it is zero.

The standard deviation for an individual observation is then given by,

$$s.d.(\hat{Y}_0) = s\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum\left(X_i - \overline{X}\right)^2}} \qquad (5.\ 129)$$

The above expression is a minimum when X_0 is equal to the mean value and it increases as X_0 moves away from the mean value. This is what you would expect because the further you look into the future the more uncertain becomes the prediction. The upper and lower limits of the prediction interval are calculated by adding and subtracting a multiple of the standard deviation to/from the predicted value of Y for a given X_0.

$$\hat{Y}_0 \pm t(n-2, 1-\tfrac{1}{2}\alpha)s\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{\sum\left(X_i - \overline{X}\right)^2}} \qquad (5.\ 130)$$

where t(n-2,1-½α) is the 100(1-½α) percentage point of a t-distribution with (n-2) degrees of freedom. It is well known that the value of t depends on the number of observations.

5.4.5.2 Prediction intervals: non-linear models

Similarly for the non-linear models, the variance of the predicted Y values can be expressed as:

$$V(\hat{Y}_0) = V(b_1)X_{m,0}^{2r} + V(Y_{ref}) \quad (5.\ 131)$$

where V(A) is the variance of parameter A. $\hat{Y}_0$ is the predicted value at a specific value $X_{m,0}$ ($X_{m,0}=X_0-X_{ref}$). The variance of Y_{ref} is zero as it is just a constant.

Variance of b_1

A similar method, to the one considered for the linear model, is used to evaluate the variance of the slope coefficient for the non-linear models.

$$b_1 = \frac{\sum X_{m,i}^{r} Y_{m,i}}{\sum X_{m,i}^{2r}} = \left\{X_{m,1}^{r} Y_{m,1} + ... + X_{m,n}^{r} Y_{m,n}\right\} \Big/ \sum X_{m,i}^{2r} \quad (5.\ 132)$$

The variance of a function

$$a = a_1 Y_1 + ... + a_n Y_n \quad (5.\ 133)$$

is

$$V(a) = a_1^2 V(Y_1) + ... + a_n^2 V(Y_n) \quad (5.\ 134)$$

If the Y_i are pair wise uncorrelated and a_i are constants

$$V(a) = (a_1^2 + ... + a_n^2)V(Y_i) = \left(\sum a_i^2\right)\sigma^2 \quad (5.\ 135)$$

In the expression for b_1

$$a_i = \frac{X_{m,i}^{r}}{\sum X_{m,i}^{2r}} \quad (5.\ 136)$$

Since the $X_{m,i}$ can be regarded as constants, the variance of b_1 can be found after reduction

$$V(b_1) = \frac{\sigma^2}{\sum X_{m,i}^{2r}} \quad (5.\ 137)$$

Variance of predicted Y values

Substituting equation (5. 131) into equation (5. 137),

$$V(\hat{Y}_0) = \frac{X_{m,0}^{2r}\sigma^2}{\sum X_{m,i}^{2r}} \quad (5.\ 138)$$

The variance of the random error, σ^2 is unknown. As with the linear model, σ^2 can be estimated from the standard deviation of the residuals, if the selected model adequately describes the data.

$$s = \sqrt{\frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-1}} \quad (5.\ 139)$$

Note that for the non-linear models, only one parameter is estimated and so the number of degrees of freedom is just n-1, rather than n-2. Hence, the estimated standard deviation of for the conditional mean of Y is

$$s.d.(\hat{Y}_0) = s\sqrt{\frac{X_{m,0}^{2r}\sigma^2}{\sum X_{m,i}^{2r}}} \quad (5.\ 140)$$

The standard deviation for an individual observation is then given by,

$$s.d.(\hat{Y}_0) = s\sqrt{1+\frac{X_{m,0}^{2r}\sigma^2}{\sum X_{m,i}^{2r}}} \quad (5.\ 141)$$

The upper and lower limits of the prediction interval are calculated by adding and subtracting a multiple of the standard deviation to/from the predicted value of Y for a given X_0.

$$\hat{Y}_0 \pm t(n-1,1-\tfrac{1}{2}\alpha)s\sqrt{1+\frac{X_{m,0}^{2r}\sigma^2}{\sum X_{m,i}^{2r}}} \quad (5.\ 142)$$

$$\hat{Y}_0 = \hat{Y}_{m,0} + Y_{ref} = b_1 X_{m,0}^{r} + Y_{ref} \quad (5.\ 143)$$

Where t(n-1,1-½α) is the 100(1-½α) percentage point of a t-distribution with (n-1) degrees of freedom.

5.4.6 Prediction procedure applications: functionality demonstration

The long term investigation framework part of the prognostics module of HMP 1.1 is shown in Figure 112.

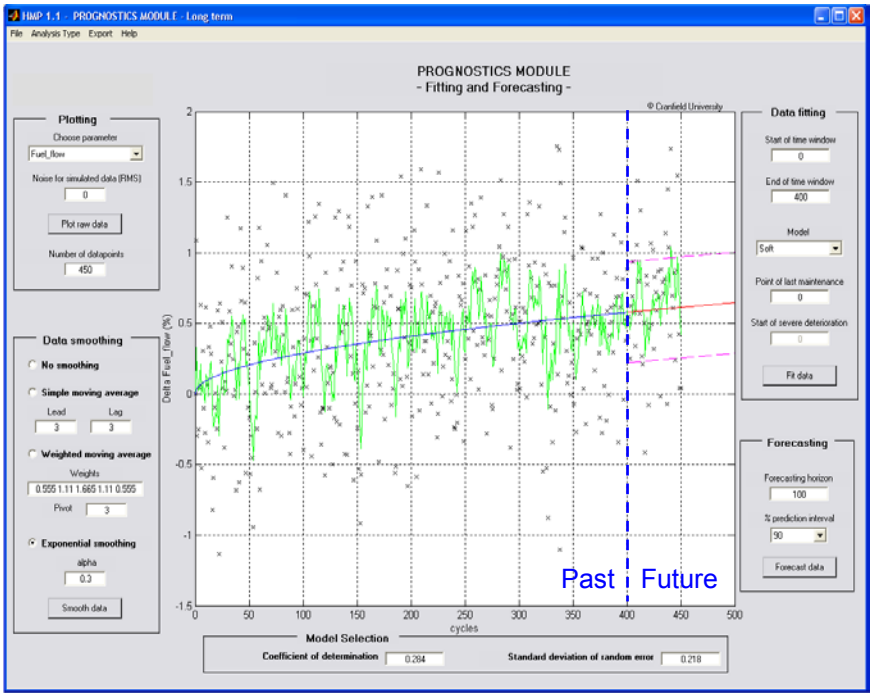


Figure 112: HMP 1.1 prognostics module, long-term Regression fitting and forecasting GUI

This section is dedicated to demonstrate the functionality of the above underlying theory. As stated earlier, the procedure here proposed is made of four elements: (i) deterioration models: three models were introduced, (ii) smoothing pre-processing algorithm, (iii) fitting procedure based on regression analysis, (iv) forecasting with prediction intervals. Simulated data were used in order to show the functionality of the process. Initially, data representing a slow deterioration phase are simulated. Then, a typical severe deterioration phase is considered, which changes the trend of the data. The capability of the method is shown via two examples. The simulated data are based on the analysis of real on-wing Trent 800 data but properly modified for reasons of confidentiality. According to the procedure summarized in Figure 87, the analysis could be subsequently extended by inputting the measurement forecasts into the diagnostics process in order to predict the performance parameters.

5.4.6.1 Functionality of the prediction procedure

The functionality of the prediction procedure is illustrated through two examples. The exponential smoothing method is used as pre-processing technique.

HPC exit pressure forecast

Delta HPC exit pressure data are plotted against flight number in Figure 113. Observations in a time window of 1000 flights are simulated: data up to fight 400 are considered to be the past or else the time series that contains the information to create a projection. As typically happens in real data a convexity is present in the data. Initially a linear model is fitted in the data. Figure 113 shows how when a convexity is present in the data the use of a linear model of the deterioration can be misleading compared to a soft deterioration model that for example is used in Figure 114. The ability of finding out the sign of the convexity of the data (i.e. + or -) and the use of this information to compute forecasts has a significant impact in its accuracy. A number of different statistical measures exist for quantifying the forecast accuracy. One relatively simple parameter is the mean absolute error (MAE), defined below (Makridakis et al., 1998).

$$\mathrm{MAE} = \frac{1}{n_f}\sum_{i=1}^{n_f}\left|e_f\right| \qquad (5.\ 144)$$

where n_f is the number of forecasts and e_f is the error between the forecast and the actual observation. The MAE of 0.45 achieved with a linear model is not satisfactory compared to 0.190 that a soft model can provide.

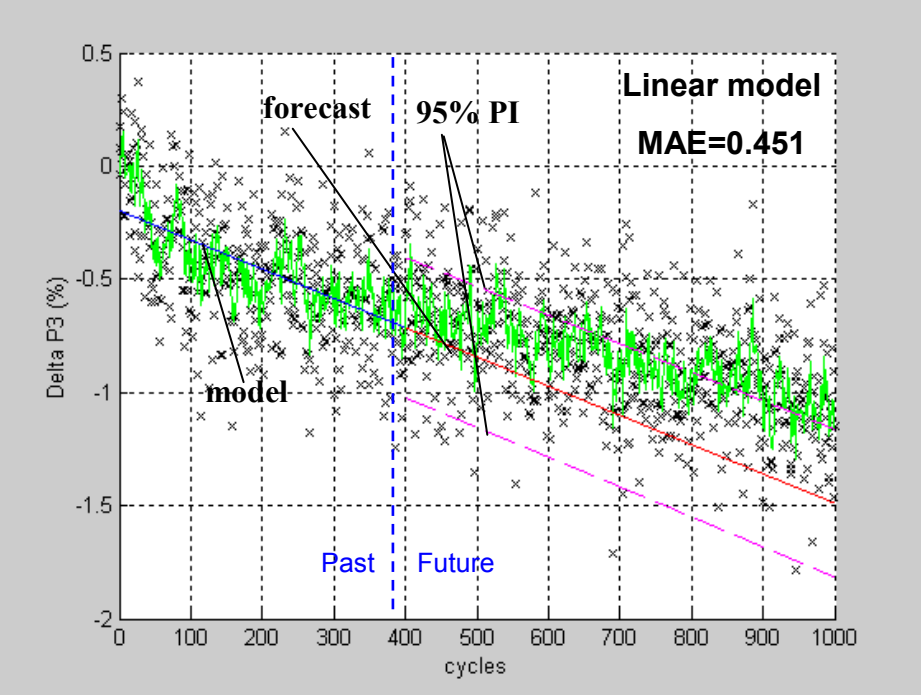


Figure 113: Delta HPC exit pressure forecast with linear model (exponential smoothing, alpha=0.3)

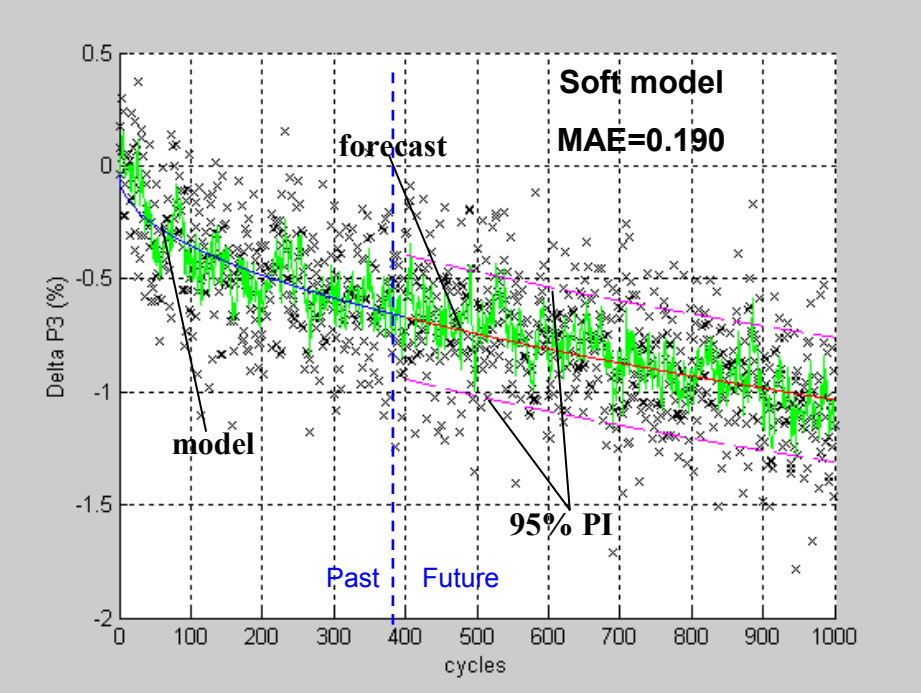


Figure 114: Delta HPC exit pressure forecast with soft model (exponential smoothing, alpha=0.3)

Nevertheless, when the analyst makes a prediction, the choice of the deterioration model is influenced by a thorough analysis of the objectives of the study. A pessimistic but safer prediction (in this case provided by the linear model) could be chosen anyway when the decision-making process has high impact on the safety of the engine's mission. So, in long term forecasting, it is highly recommendable to produce different forecasts associated to different assumptions.

Delta fuel flow forecast

A second case study considers a time series of 450 fuel flow data – see Figure 115 (note that the parameter increases while deteriorating). This is the case of an engine that after undergoing normal deterioration suddenly starts to degrade faster manifesting a change in convexity in the data. This severe deterioration was implanted after 400 cycles and that data were generated for a total of 700 cycles. According to the prediction procedure previously described (Figure 102) the three deterioration models are used, and the problem lies in the identification of the right sign of the convexity in the data. The prognostics procedure has the capability of detecting when a severe deterioration starts. The origin of the severe deterioration model is fixed "*s*" observations before the last. The value of "*s*" depends on the parameter studied, on the objectives of the study, and it is based on experience. In this example it was fixed to 70 cycles. The best-fit is identified by using the coefficient of determination. If a severe deterioration manifests the convexity of the data changes and with a delay that is normally shorter than "*s*" observations is detected: the severe model that is used for the forecast is associated with the highest coefficient of determination. Once a severe deterioration is detected, the engine usually requires an early overhaul because of the very detrimental effect on engine performance.

In Figure 115 three models are fitted in the data and the values of R^2 are 0.49 for the linear model, 0.57 for the soft model and 0.65 for the severe model. Three projections are made with the respective PIs. The third model is therefore chosen in Figure 116: the forecast and the MAE are computed.

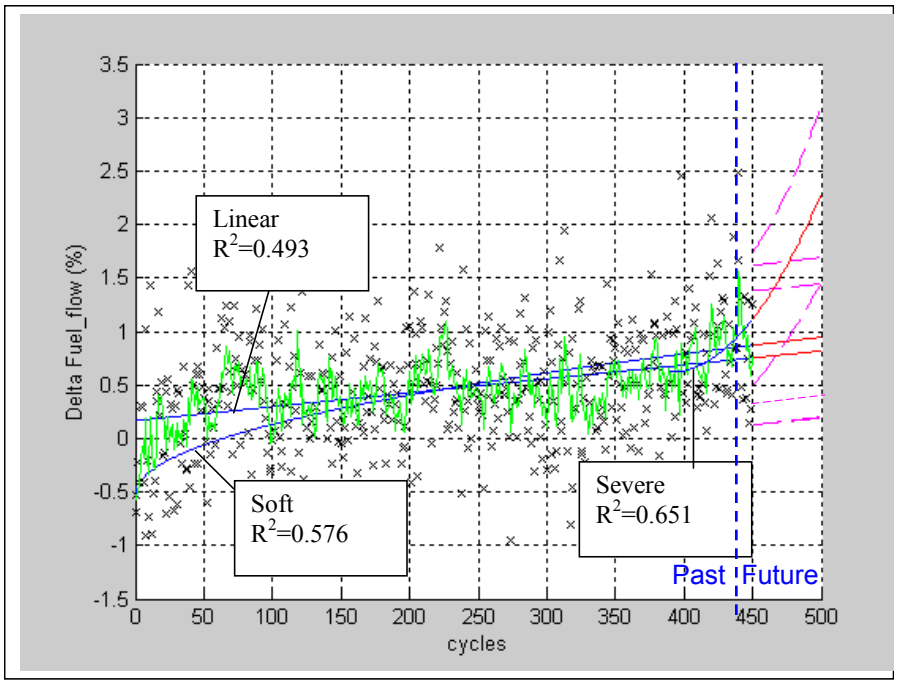


Figure 115: Delta Fuel Flow observations, three deterioration models and forecasts (exponential smoothing, alpha=0.3)

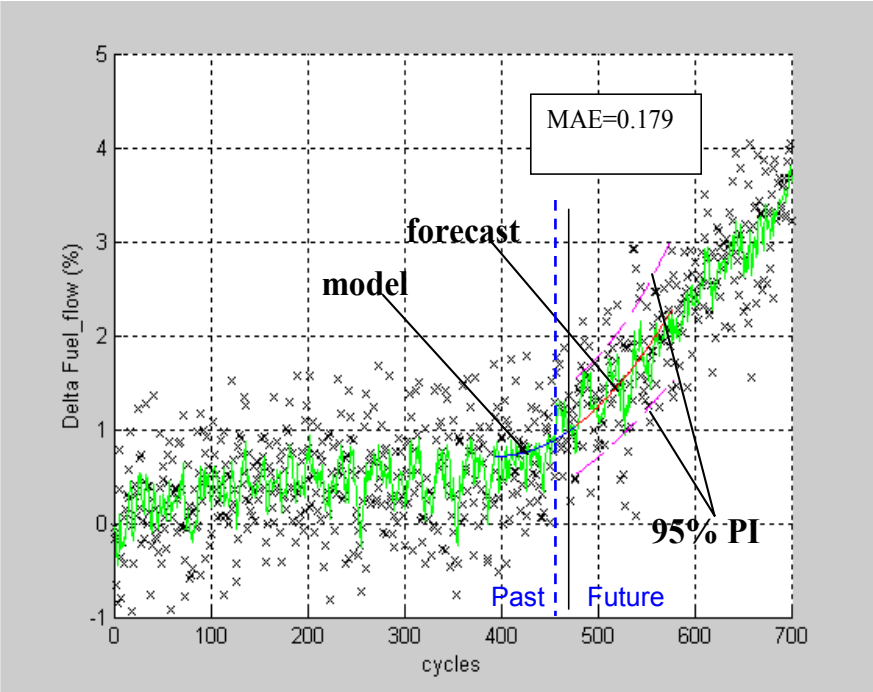


Figure 116: Delta Fuel Flow observations, projection with severe deterioration model and prediction intervals (exponential smoothing, alpha=0.3)

Novelty of the Prediction Intervals

The key advantage of the procedure described in this document lies in the use of the prediction intervals. The concept is similar to confidence intervals commonly used in statistic to define the range of variability, with a fixed confidence level, for an observed quantity (Spiegel, 1972). The prediction interval defines the band in which the observations are predicted to fall in, according to their previous behaviour. In the diagrams previously shown it is evident how a simple extrapolation of the data, no matter how accurate it is, gives an excessively optimistic forecast. The upper and lower PI can then be used to provide a more or less precautionary or pessimistic prediction. Besides the choice of the confidence level is an ulterior parameter to play with. If a cost analysis is pursued low confidence in the prediction and therefore an optimistic prediction is normally required (e.g. 70%). While if a safety related investigation is carried out higher values of confidence in the prediction are demanded (e.g. 95%).

According to equation (5. 130), the prediction intervals diverge as you go further into the future. The reason for this is the relatively large variability in X, which increases with the number of fitted data points, and results in $\frac{\left(X_0-\overline{X}\right)^2}{\sum\left(X_i-\overline{X}\right)^2}$ being much smaller than 1. As 1/n is also much less than 1, the first term under the square root in equation (5. 130), which is 1, becomes the dominant parameter. Therefore, the prediction interval curves diverge even thought slowly.

5.4.7 Case studies: maintenance versus operating costs

A prediction procedure for long tem investigations has been devised and discussed above. Forecasting performance deterioration has many potential applications in civil aircraft maintenance operations. In this section prognoses of optimal time between overhauls are considered. It was recognised in this research that any attempt to GPP requires an associated business intention: different parameters may be required to be forecasted with different confidences. Under some assumptions relative to the potential scenarios of application, the framework was used in this section to study Rolls-Royce Trent 800 engine data and the results are shown below.

In service deterioration causes increase in specific fuel consumption and therefore additional fuel costs associated with it. Maintenance actions can therefore be undertaken to restore the engine performance. Maintenance costs per cycle decrease as the overhaul intervals grow, the less the engine is maintained. So there is a need of calculating the optimal time between overhauls (TBO) that minimizes the total costs that include maintenance costs and additional fuel costs due to the deterioration. Different engines deteriorate differently thus it is not desirable to define an average optimal maintenance interval. A prediction procedure can be used to forecast variations in fuel flow for a given period, and hence to determine the best TBO for the specific engine. Considering the highly increasing fuel price, this is a typical application of gas path prognostics that can potentially result in considerable cost savings for the operators. Detailed demonstration of the GUI developed for this purpose can be found in Appendix C.

The problem has been simplified for the scope of this work with the **assumptions** that follow. Fixed maintenance intervals of the life limited parts are requested by the airworthiness authorities. So we assume that a maintenance action is scheduled and we explore the effect that additional intermediate actions would have on the costs. The investigation is limited to the additional fuel costs due to the deterioration and the extra maintenance costs, with the purpose of finding an optimal TBO. For simplicity all other factors that might be taken into account in a real case when establishing the TBO will be ignored. Besides the following assumptions have been made:

- Each intermediate maintenance interval has the same length, except for the last interval which can be shorter.
- From the previous assumption it follows that the optimal TBO must be larger than the number of engine cycles since the last maintenance.
- The maximum possible TBO is equal to the engine cycle for which the next scheduled maintenance is planned.
- The costs of the intermediate maintenance activities are assumed to be the same for each overhaul.

Two different examples are considered examining two different scenarios. (i) Scenario 1 – normal deterioration: this is the simple case where the engine does not experience severe deterioration phases over the considered time period. (ii) Scenario 2 – normal, followed by severe deterioration: in this second

possible scenario, the engine degrades normally at first but after a certain number of cycles a change in convexity happens resulting in a fast deterioration rate.

Simulated delta fuel flow data are plotted in Figure 117 against cycles. A normal deterioration phase happens during the first 400 cycles (the prediction procedure will find out if it is best fitted by a soft or a linear model) and this is the time series considered in the first scenario. At flight number 400, deterioration continues with a faster rate; the full length time series that includes all the 450 observations is investigated in the second scenario. Each cycle, for the simulated data, was associated with a 7 hour flight, constant for all the data points. This corresponds to the flight time of a typical transatlantic route flown by a Boeing 777.

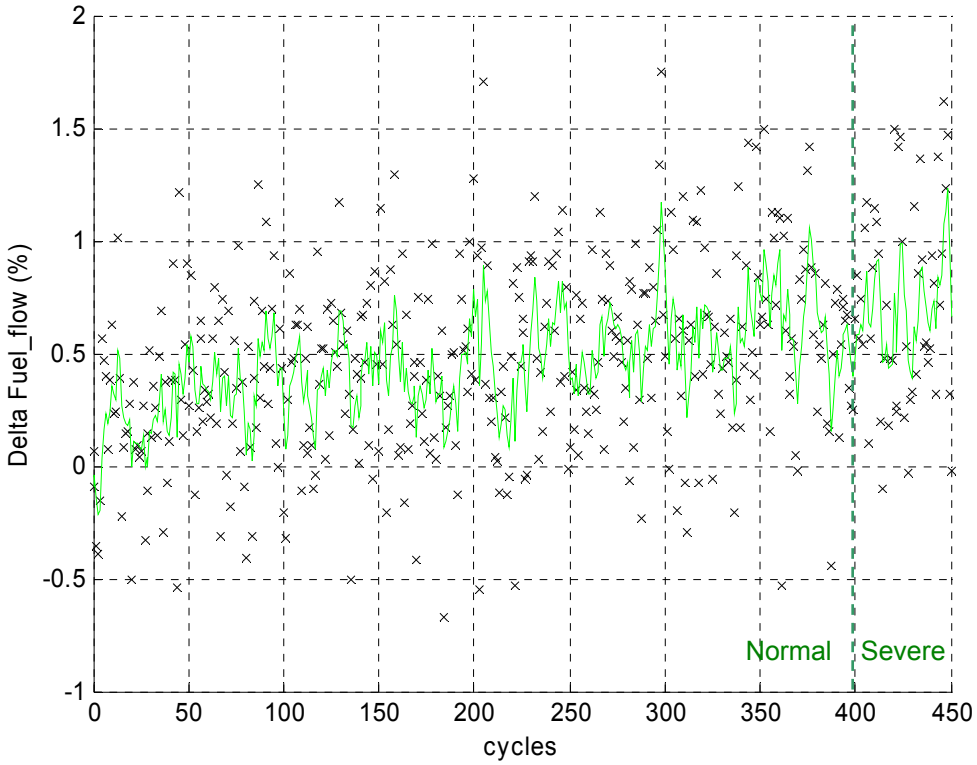


Figure 117: Delta Fuel flow data (exponential smoothing, alpha=0.3)

5.4.7.1 Scenario 1: optimal TBO prognosis

A GUI window was devised for this type of study (see Figure 118 and Appendix C).

In this study the following variables must be defined as input of the calculation:

- Additional maintenance costs: in U.S. dollars, include the costs involved in intermediate maintenance actions in addition to the scheduled overhaul. Consequently they must include the cost of the actual maintenance activities as well as the cost associated with grounding the aircraft and removing the engine. The cost of maintenance is directly related to the deterioration recovered; i.e. a large cost would generally

translate in the restoration of the majority of the performance. In this sense the program can be used to evaluate the optimal TBO for the whole engine as well as for individual modules. In this application a value of 50000 $ has been chosen.

- Additional fuel costs: in U.S. $ per U.S. gallon, are related to the delta in fuel flow from the engine model at actual power level, they include the additional costs due to running the same mission with a dirty engine compared to the clean one. This variable has a significant impact on the results due to the considerable fluctuations in the price of petrol. In the example it was fixed to 1.5 $/gallon.
- Last maintenance time: represents the flight number at which the last major overhaul was performed. It is the first data point used in the calculation of the fuel cost. As previously mentioned, the point of the last maintenance is also the starting point for fitting the soft deterioration model. In the current example the last maintenance time is at the cycle 0.
- Next scheduled maintenance action time: the optimal TBO is computed in the time window included between the previous and the next scheduled maintenance action. In the example, we assume that it is scheduled after 2000 cycles. This is equivalent to 14000 hours, which might be a typical interval specified to inspect and overhaul the hot parts of an engine.
- Recovered deterioration (% of total deterioration): This is the percentage of the additional specific fuel consumption (as a result of deterioration) that is recovered with the action. If 100% of the deterioration is recovered the engine will have the same fuel consumption as the nominal/non-deteriorated engine. As explained above, the degree of performance restoration is closely linked to the maintenance costs. In the example we assume to recover 90% of the deterioration with the intermediate maintenance action.

Besides the following choices have to be made:

- Model for normal deterioration: The soft and linear deterioration models are available for fitting normal engine deterioration. The prediction procedure identifies the best regression model with higher coefficient determination, but the analyst might decide a more optimistic or pessimistic model.
- Prediction interval: that takes into account for the uncertainty in the forecasted fuel flow data. For example, if 99% upper PI is selected, the program will use the fuel flow values of the 99% upper prediction interval boundary. These are obviously significantly higher than the 50% values and therefore the computed TBO would result equivalent to a worst case scenario. This capability of the program of selecting different confidence levels provides the analyst with the competence of investigating different scenarios. In this example 70% was chosen.

The problem studied regards the determination of the most cost-effective maintenance interval by determining the TBO that gives the minimal overhaul additional cost. The code calculates the costs associated with all the discrete TBOs in a considered range and determines the optimal TBO. The total costs

include the additional fuel costs due to deterioration and the extra maintenance expenses incurred by the performance related overhauls.

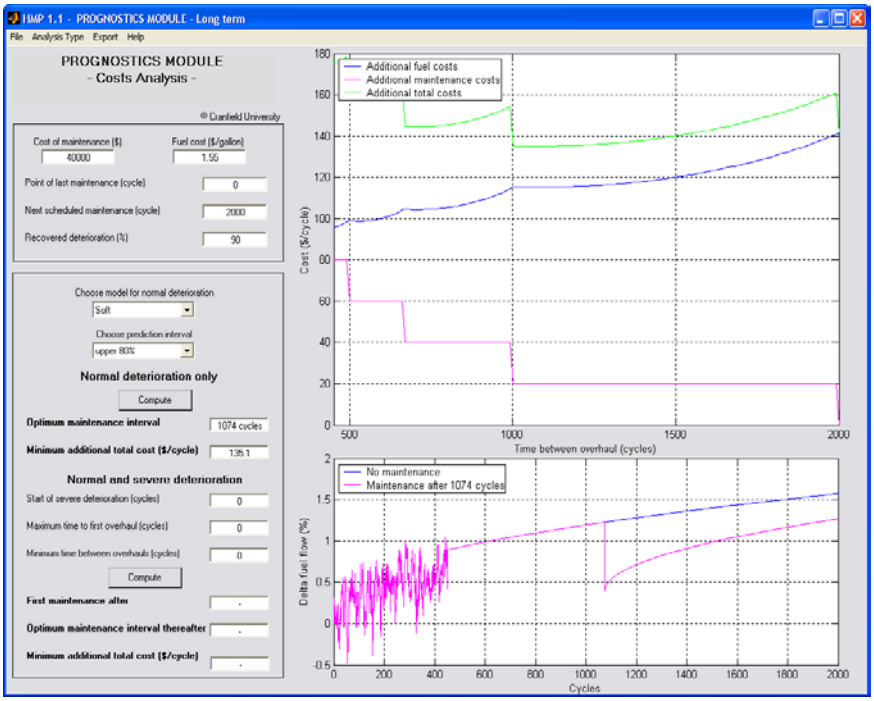


Figure 118: HMP 1.1 prognostics module, long-term Regression cost analysis GUI

In order to calculate the **additional fuel costs** the algorithm evaluates the extra costs for each TBO. The number of maintenance intervals is evaluated as follows. If the point of last maintenance is, for example, at zero engine cycles and the next scheduled maintenance at 2000 cycles, a TBO of 600 will result in four maintenance intervals (0-600, 600-1200, 1200-1800 and 1800-2000). Note that all the intervals are the same length apart from the last one, which can only be 200 cycles because the next scheduled maintenance is at 2000 cycles.

For the first interval the fuel costs are calculated from both engine measurements and predicted fuel flow data. If the fuel flow data are available for the first 400 cycles and the TBO is 600 cycles, then the total additional fuel costs are computed considering the fuel flow measurements of the first 400 cycles plus the fuel flow forecasts for the cycles between 400 and 600. Note that the data used for the predicted measurements depend on the selected prediction interval. For the remaining intervals, the fuel flow data are predicted by using the model that best fits the first 400 data, provided with a new origin that is determined in the following way. The origin of the model corresponds to the prediction of the fuel flow datum after each overhaul and depends on the deterioration removed during the maintenance action. If 100% of the deterioration is removed, all the deterioration accumulated during the previous interval will be recovered. On the other hand, if 80% is specified for recovered deterioration, the deterioration at the beginning of the interval will be 20% of the increase in delta percentage fuel flow during the last interval plus any other non-recoverable deterioration from previous overhauls.

The **additional maintenance costs** are calculated as follows. If for example four maintenance intervals are undertaken, one after 600 cycles, a second after 1200 cycles, a third after 1800 cycles and then the scheduled after 2000 cycles, then three additional intermediate actions are accounted for. The cost of the maintenance activities carried out to restore part of the engine performance are assumed to be the same for each overhaul. Hence, the additional maintenance costs are equal to the number of extra overhauls times the specified cost of maintenance.

Below, a description of how the optimal maintenance interval is computed when the engine deteriorates normally is provided via an example. The fuel flow data in Figure 117 represent the time series studied. Firstly, the soft deterioration model is selected as the best regression model, because it provides the highest value of R^2. Being a cost analysis we assume that a prediction interval of 70% is suitable. This value should be adjusted according to the experience. Hence, fixed the next scheduled maintenance action after 2000 flights, the cost of any intermediate action (50000$) the recovered deterioration expected to be achieved with that action (90%), and successively given the soft deterioration model and the 70% upper prediction interval as shown in Figure 119, the following quantities are calculated and plotted in Figure 120 for each TBO in the time window between 0 and 2000 cycles:

- Additional fuel costs.
- Additional maintenance costs.
- Total costs calculated as addition of the above two quantities.

Therefore, minimising the total costs the optimal TBO is computed with the confidence required. In this specific case an optimal TBO of 1080 cycles was found and the resulting additional total cost per cycle amounted to 106.6 $.

Figure 119 shows the evolution of the fuel flow for the calculated optimal TBO. It can be seen that for the first 400 cycles the engine measurements are used to compute the fuel cost. One intermediate overhaul is performed after 1080 to partially restore the performance of the engine. A non-recoverable delta fuel flow was taken into account after the maintenance action and this is also shown. The blue line represents the increase in fuel flow if no maintenance is performed according to the determined deterioration model.

It is worth noting that such a procedure can be used for example in a real case in which the compressors performances decrease due to fouling, to calculate the optimal time between compressor washing, leading to substantial cost savings for the operators.

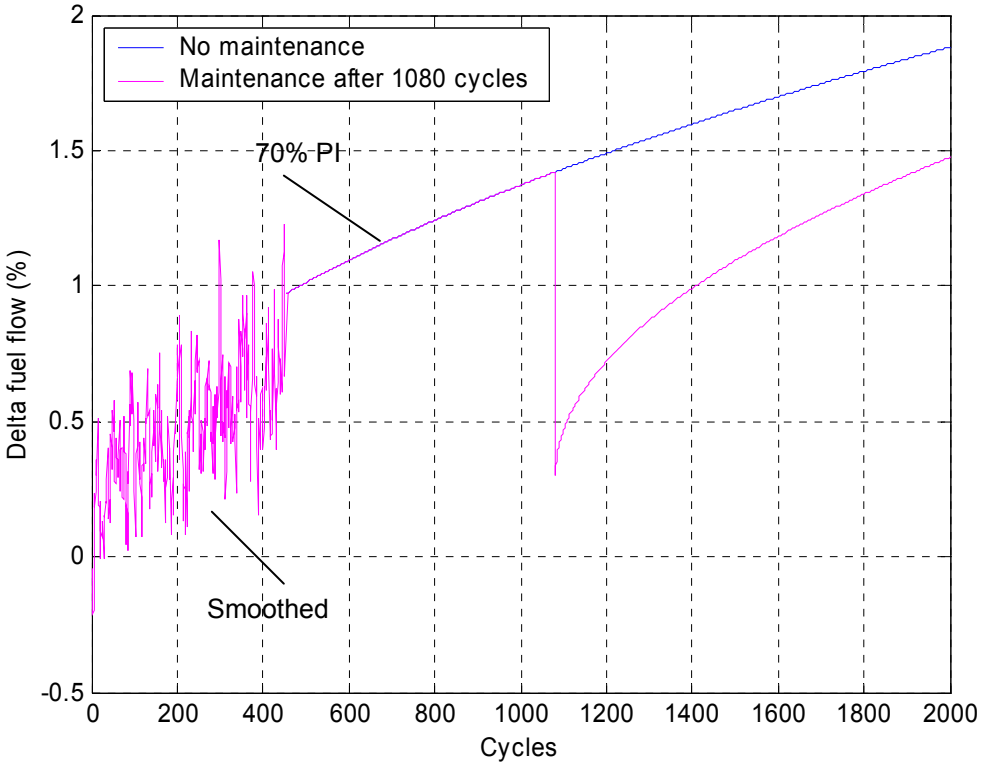


Figure 119: Smoothed Delta Fuel Flow time-series, projection with soft deterioration model upper 70% prediction intervals

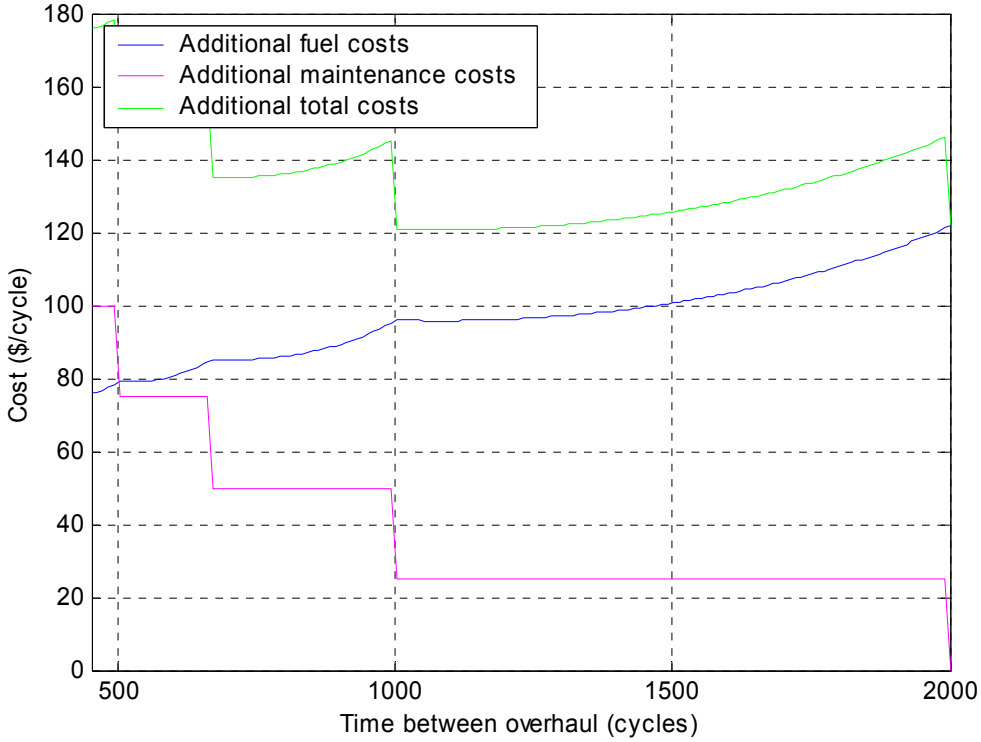


Figure 120: Optimal time between overhaul: additional fuel, maintenance and total costs

5.4.7.2 Scenario 2: optimal TTFO and TBO prognosis

A GUI was devised also for the study of a second type of scenario that is characterized by a severe deterioration phase that manifests during a normal deterioration (see Appendix C): 450 data are plotted in Figure 117. As discussed above, the data show a change in convexity around the flight 400 that leads to a faster deterioration rate. Assumptions and inputs to the calculation for this second scenario are the ones previously specified apart from the differences discussed below. The following additional variables have to be chosen:

- Starting point of the severe deterioration: it can be identified by means of the prediction procedure, as suggested above. This parameter has to be input in the second GUI. In this case this happens at flight 400.
- Maximum cycle for first overhaul and minimum cycle between overhauls: These two inputs are required in order to restrict the optimisation problem and hence improve the computational speed.

The major difference this new scenario and the one previously investigated lies in the fact that now the total costs are function of two variables (whose optimum we are looking for). The first variable is the cycle at which the first overhaul should be performed, which will be referred to as the time to first overhaul (TTFO). In this first interval the fuel flow data follow a severe deterioration path, while normal deterioration paths are assumed to happen during all subsequent intervals. The second variable is the TBO for the cycles remaining after the first overhaul. Note that these two variables are inter-dependent as the TTFO affects the optimal TBO. The optimal value for the TBO and TTFO are found by calculating and minimising the overall costs associated with each value of TTFO and TBO in the range under investigation:

- There is no longer a minimum value for the TBO after the first interval. However, relatively low values of TBO generally result in high overall costs. Therefore, a value is also specified for the minimum TBO. The reason for setting these limits for TTFO and TBO is that it significantly reduces the computational time.
- The maximum value for the TBO is the number of cycles from the end of the first interval to the point of the next scheduled maintenance.

For each value of TTFO, the overall cost for the whole range of TBOs is calculated. The optimal values of TTFO and TBO are those that result in minimum overall costs, which, as before, are made up of the additional fuel and maintenance costs.

Note that, the additional maintenance costs are calculated in the same way as before. The program still assumes that each overhaul costs the same. This might not be a completely realistic assumption because the first maintenance action, that is associated with the end of the severe deterioration phase, is likely to be more comprehensive and costly.

Two models are therefore identified by applying the prediction procedure to the 450 data time series under study in this example: the best model that fits the data during the normal deterioration phase is a soft model while a severe model best fits the data after the change to a faster rate that manifests at the flight 450. In this second scenario the prediction interval is fixed to 50%.

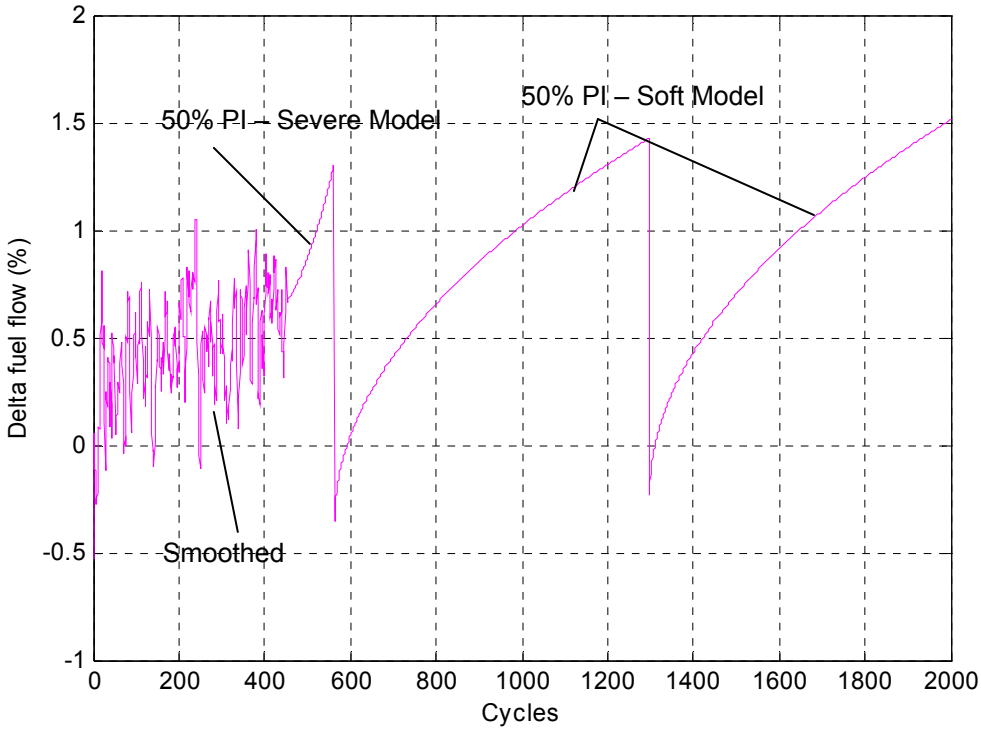


Figure 121: Smoothed Delta Fuel Flow time-series, projection with severe deterioration model upper 50% prediction interval up to the first overhaul and projection with soft deterioration model upper 50% prediction interval for the remaining missions

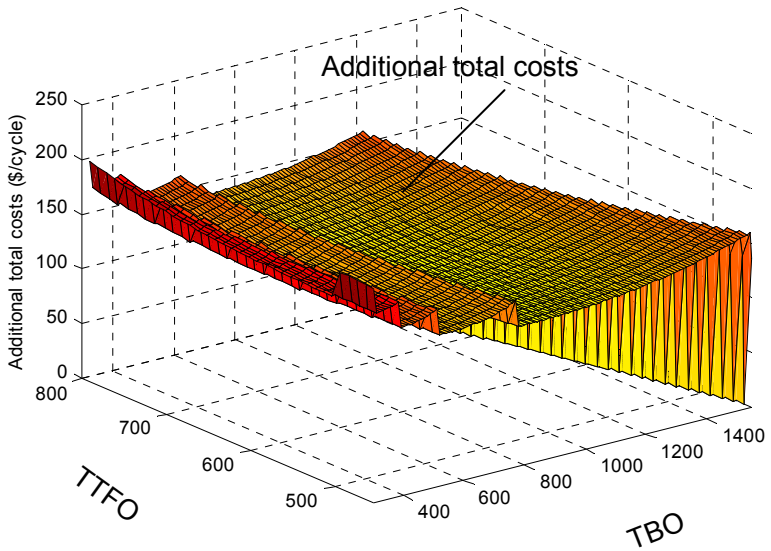


Figure 122: Optimal time between overhaul (TBO) and time to fists overhaul (TTFO): additional total costs that include fuel and maintenance costs

Similarly to the scenario 1, the following quantities are fixed: (i) the next scheduled maintenance action after 2000 flights, (ii) the cost of any intermediate action (50000$), (iii) the (assumed) recovered deterioration after that action (90%). Provided the soft and severe deterioration model as shown in Figure 121, the additional total costs are calculated and plotted as function of the two variables TBO and TTFO in Figure 122. The computation of the minimum total costs provides the optimal TBO and TTFO with the confidence required.
In this specific case a first maintenance is prognosed at flight 562 followed by an optimal maintenance interval of 735 flights. A minimum additional cost of 136.1 dollars is forecasted.

5.4.8 Summary

If long term projections in engine performance analysis are required, section 5.4 presented a procedure based on regression analysis suitable to provide forecasts with a significant level of accuracy and a statistical interpretation of the results. A remarkable advantage of the procedure is in fact the use of the prediction intervals. The functionality of the procedure was demonstrated by studying HPC exit pressure and fuel flow simulated data.
An application of the long-term investigation procedure was described in section 5.4.7 that regards the calculation of the optimal TBO considering maintenance costs and increased fuel costs per flight due to the deterioration.

5.5 Discussion and summary of the results

5.5.1 Introduction

Chapter-5 presented the procedures that are implemented within the prognostics modules of HMP 1.1, and the results relative to an application of these procedures to the Rolls-Royce Trent 800 turbofan.

Contrary to diagnostics, prognostics cannot be considered a mature technology. This research was more aimed at opening up the subject, highlighting its importance in the future rather than to complete it. A general framework was developed to make forecasts that enable the prognostics decision-making process. The requirements, that emerged (in Chapter-3) for such an advanced gas-path prognostics framework, are now repeated due to their importance in the discussion that follows. The framework should be:

1. Able to include the confidence level in the projections.
2. Designed to cope with short and long term investigations.
3. Able to cope with slow and/or rapid deterioration scenarios.
4. Designed with different strategies for dealing with performance parameters and/or measurements.
5. Easy to be understood/explained in management meetings where the cash-flow projections are discussed.
6. Light in computational requirements.

These 6 requirements were identified as the specifications for this second part of the research. The next section discusses the procedures and the results presented in this Chapter in the light of these requirements, in an attempt to assess their fulfilment.

5.5.2 Analysis of the results

A discussion follows aimed at evaluating how each requirement, considered individually, is fulfilled by the processes developed within this project. For some requirements only a partial fulfilment was obtained: the rationale behind the choices made is discussed and some suggestions for further studies are given in section 5.6.

1. Process able to include the confidence level in the projections.

It was recognised in section 3.2 that most of the publications describe approaches to GPP that have a qualitative nature and lack statistical rigour. Some of these studies (Provost, 1995, and Bakken et al., 1996) were able to identify the trends in the data, but no statistical interpretations of the results were provided. One of the advantages and novel aspects of the procedure presented in this chapter lies in the calculation of the prediction intervals. The level of confidence required in the prediction plays a key role in prognostics and its specification is strictly related to the business intention that drives the prognostics activity. In this study, there was no ambition to investigate how prognostics requirements vary according to the different business intentions, but this is considered to be the next step towards achieving an enhanced GPP and so future studies should be devoted to these aspects.

2. **Process designed to cope with short and long term investigations.**

Due to the nature of the GPP problem, that requires both immediate and long term predictions and due to the diverse time-series methodologies available for such different problems, a GPP module needs to differentiate the two typologies of investigations. Therefore, the outcome of this research was two procedures: *technique 1* based on ARIMA models was implemented to provide accurate forecasts for immediate and short-term horizons and *technique 2* based on regression analysis and some considerations on the nature of the gas-path component degradation was designed to handle prognoses that require medium and long-term predictions.

Three typical prognostics approaches employed in engineering were reported in section 2.4.3. It is our opinion that, the time-series based approach, at the present time, is the most suitable for GPP problems due to its simplicity and wide applicability. Future studies, if more sophisticated models of the deterioration become available, should consider other approaches.

3. **Process capable of coping with slow and/or rapid deterioration scenarios.**

The framework presented in this chapter is mainly dedicated to forecast gradual deterioration. For long term predictions, the capability of identifying and forecasting severe deteriorations is provided, but the data must still present a trend that is identifiable (even though a rapid one). The forecast is based on the time series and therefore on such a trend, a delay in the detection is inevitable. On the other hand, if a sudden event happens this cannot be predicted with the procedure presented in this work. An idea is suggested, but further studies are highly advisable on the subject: instantaneous deteriorations due to events can be taken into account through hazard plots that influence the warning and safety margins. Hazard plots represent the probability that an event occurs over time.

4. **Process designed with different strategies for dealing with performance parameters and/or measurements.**

In this study, forecasting techniques were used to predict the gas-path measurements, undertaking dependent-parameter-based prognostics that relies on the prediction of engine global variables, say measurements strictly related to the global health of the engine, such as the TGT, shaft speed or the SFC. On the other hand, independent-parameter-based prognostics relies on the prediction of health performance parameters (e.g. efficiency and flow capacity of the engine-components).

In section 5.2.1, two options were considered for predicting the gas-path components' performance parameters: (i) first forecast measurements and input them into a diagnostic algorithm to predict the performance parameters deviations, or (ii) first make diagnoses and then use a time series of performance parameters to predict their future values. The former was here considered beneficial compared to the latter that is affected by additional

uncertainty in the forecasting procedure due to the nature of the diagnostic process; this causes unacceptably high prediction intervals. Nevertheless, no systematic tests were undertaken to validate such a layout and future studies should be dedicated to prove the optimal configuration.

5. **Process should be easy to be understood/explained in management meetings where the cash-flow projections are discussed.**

A procedure based on regression analysis and some considerations on the nature of the gas-path component degradation was designed to compute long term predictions. This choice was considered highly advantageous for such an engineering application due to its simplicity that allows tangible justification of the results and a minimum requirement of assumptions on the statistical behaviour of the variables. On the other hand, for short term investigations, a higher level of accuracy can be achieved with more sophisticated techniques such as ARIMA. However, it was pointed out that the accuracy that can be achieved with ARIMA models is affected by the ability of the analyst. Some of the practical difficulties which arise are related to the subjective interpretation of the correlograms. Future studies should look at the opportunity of combining together short term forecasting methods to gain in reliability of the predictions.

6. **Process light in computational requirements.**

The GPP framework presented is remarkably low with respect to computational power requirements relative to the average computational power currently available. Difficulties may arise related to the data storage and analysis necessary to achieve an enhanced GPP. Sophisticated methods based on accurate models of the deterioration should be developed and this competence can only be achieved with the systematic statistical analysis of in-flight recorded data. Future studies should investigate which type of in-flight data should be analysed, how this analysis should be carried out and what sort of information should be sought.

5.6 Recommendations for future studies

In section 5.5.2, a discussion of the results was provided. The level of fulfilment of each research requirement (i.e. specifications for this second part of the study dedicated to gas-path prognostics) was discussed in order to analyse the outcome of the presented processes and the risen research opportunities. This section provides some suggestions for further studies that should be carried out to gain a higher level of accomplishment of those requirements that have been only partially fulfilled.

The following recommendations were drawn from the experience achieved during this research:

- The next step towards obtaining an enhanced gas-path prognostics is the study of how the prognostics requirements vary according to the different business intentions. Only such a thorough investigation can define the confidence level requirements according to the different prognostics needs (with reference to the 1st requirement).
- Future studies should be devoted to the development of sophisticated models of the deterioration to base the predictions on (with reference to the 2nd requirement).
- Future prognostics frameworks should include the capability of dealing with sudden faults that inevitably affect the gas-path. Such a level of competency requires the statistical knowledge of the fault occurrences for different type of engines operating in different types of environment on different routes. Besides, the need for an advanced procedure able to provide an early detection of fault event occurrences in an in-flight data time series has been identified (with reference to the 3rd requirement).
- Systematic tests should be carried out to identify the best layout in order to undertake independent parameter-based prognostics (that relies on the prediction of health performance parameters). A procedure that inputs forecasted measurements into the diagnostics process in order to estimate future deviations in performance parameters was proposed (with reference to the 4th requirement).
- Future studies should look at the opportunity of combining together more than one short term forecasting methods to gain in results reliability. Besides, a strong interest should be oriented to the development of a physics-based model also for short term predictions, gaining in reliability and comprehensibility (with reference to the 5th requirement).
- Considerable effort should be dedicated to find out how to manage the enormous amount of in-flight data that can be made available nowadays. The key elements to identify are: (i) the type of in-flight data that should be stored and analysed, (ii) the way how this analysis should be carried out, and (iii) the sort of information that should be sought, in the attempt to minimise the computational storage capability needed (with reference to the 5th requirement).

The HMP 1.1 prognostics module revealed some of the benefits that can be expected from the use of an advanced gas-path prognostics framework. For

example, the optimisation of the TBO and mission costs, in the assumptions made above, can lead to substantial cost savings for the operators. However, there are many other potential applications that should be investigated according to the alternative assumptions and business intentions. The use of an engine-health monitoring system, coupled with an advanced prognostics tool can greatly improve the process of scheduling maintenance, ordering parts and using resources. The ultimate goal is to obtain a complete on-condition maintenance competence maximising the aircraft's operational capability, as well as minimising its support cost.

CHAPTER 6 - CONCLUSIONS & RECOMMENDATIONS

6.1 Introduction

The goal of this research was the development of significant components of a performance health-monitoring-and-prognostics (HMP) framework. Hence, HMP 1.1 for performance analysis was designed. This software is constituted by three modules: observability study module, diagnostics module, and prognostics module (their GUIs are showed in Appendix C). The core of the research presented in this thesis is the fuzzy logic based gas-path diagnostics (GPD) procedure presented in Chapter-4. However, before any diagnostics is computed, an observability study must be undertaken. Hence substantial attention was dedicated to this subject as described in section 4.4. Moreover, the considerable importance that gas-path prognostics (GPP) is going to have in the future, within the gas turbine aftermarket, was recognised, and therefore a research opportunity was highlighted. The work developed on GPP, presented in Chapter-5, was aimed at opening up the advanced GPP research subject identifying benefits and difficulties, rather than at completing it.

The present study provided the following five main conclusions summarized below and expanded in the next five sections:

1. **New role for the gas turbine aftermarket**

A key competitive advantage for manufacturers will be their advanced gas-path diagnostics and prognostics capabilities.

2. **Technical requirements for a health monitoring and prognostics framework for gas turbine performance analysis drawn form a thorough literature study**

A detailed investigation of publications concerning gas-path monitoring methodologies (described in Chapter-2) enabled the identification of three main elements of a health monitoring and prognostics framework for performance analysis. The technical requirements for such an advanced tool were drawn from considerations regarding the gaps between the contemporary technologies, with their advantages and limitations, and the market needs.

3. **Importance of the observability study**

It was recognised that the outcome of a reliable gas-path diagnosis must not only advise about the quantifiable faults, but must also warn about the faults that are not observable with the given measurement set. Therefore the need for a secure procedure to carry out observability studies was raised.

4. **Appropriateness of fuzzy-logic in gas-path diagnostics**

Fuzzy logic is suitable for GPD because of its inherent capability of dealing with GPD problems due to its rule-based nature and its fuzzy approach. A novel fuzzy diagnostics procedure was devised for a Rolls-Royce Trent 800 engine and its promising results were discussed in Chapter-4.

5. **Suitability of time-series analysis in gas-path prognostics**

A general framework able to make forecasts that enables the decision-making process in gas-path prognostics was devised and its functionality was demonstrated in Chapter-5 with Rolls Royce Trent 800 in-flight data. It was recognised that any attempt to GPP requires an associated business intention. Only very limited resources were allocated to study how the prognostics requirements vary according to the business model; but this was identified as the main element that can enable the next step towards an enhanced GPP. Nevertheless, the framework was able to make prognostics in different hypothetical scenarios that were investigated considering safety and cost studies.

6.2 New role for the gas turbine aftermarket

The gas turbine business model has been based extensively on revenue strictly related to the aftercare. Nowadays the airlines demand is for high quality fleet-management and comprehensive engine-aftercare service, based on an agreed, prior to the engine delivery, rate per engine flying hour. The engine is paid for during the period when the aircraft is in the air and so producing revenue. This transfers much of the technical risk from the airline to the gas-turbine manufacturer. Consequently, improvements are required concerning the in-service engine operations, and this has already had significant impacts on the business case. In these circumstances, a key competitive advantage for manufacturers will be their understanding of the market and one consideration, within this, will be related to engine-condition monitoring methodologies. Among them, in consideration of the primary role that advanced gas-path diagnostics and prognostics play in an industry in which the core of the businesses is based on aero-thermal performance, the need for an effective approach cannot be ignored.

6.3 Technical requirements drawn from the literature study

One of the relevant conclusions of this thesis is the definition of the technical requirements for an advanced gas-path diagnostics and prognostics tool. These requirements were identified as the gaps between the capabilities of the contemporary technologies and the market needs. These became the specifications for the present research project: some of them were well fulfilled by the developed technology, while others were only partially fulfilled, so indicating the further steps to peruse in future research.

Chapter-2 presents, through a literature survey, an overview from the early investigations up to the state of the art of gas-path diagnostics and prognostics methodologies. Methods developed in order to undertake observability studies were discussed in section 2.2. Table 7 in section 2.3 provided a critical summary of the current gas-path diagnostics capabilities. Different prognostics approaches were reported in section 2.4.3. Authoritative publications were reviewed with the twofold aim of identifying the pros and cons of the present methodologies as well as specifying the desirable requirements for novel

diagnostics and prognostics systems that are reported below. It was concluded that a modern gas-path diagnostics tool should be:

1. Capable of providing additional information about the observability of the system.
2. Based on a non-linear model.
3. Designed specifically for SFI and/or MFI. An effective integration between SFI and MFI is necessary for the practical use of these techniques.
4. Capable of detecting, with reasonable accuracy, significant as well as small changes in performance.
5. Able to provide a 'concentration' capability upon the actual fault, limiting the tendency to smear the result on the various parameters.
6. Competent to make a worthwhile diagnosis using only a few measurements (i.e. N>M), compatibly with the system observability.
7. Able to deal with random noise in the measurements and sensor bias.
8. Light in computational requirements and fast in undertaking diagnoses for on-wing applications.
9. Exempt from training and tuning uncertainties, difficulties and dependences for setting-up parameters.
10. Free from a lack of comprehensibility due to black-box behaviour.
11. Capable of data-fusion and able to incorporate expert knowledge.

On the other hand, it was concluded that any attempt to GPP requires an associated business intention, and that the development of a reliable forecasting framework plays a key role in providing the pertinent decision-making competences. Considering the current state of the art in GPP, and the absence of physics-based deterioration models in the public domain, the need for a time-series-analysis based framework that provides the necessary statistical approach to the prognoses was pointed out. Its technical requirements identified a process that should be:

1. Able to include the confidence level in the projections.
2. Designed to cope with short and long term investigations.
3. Able to cope with slow and/or rapid deterioration scenarios.
4. Designed with different strategies for dealing with performance parameters and/or measurements.
5. Easy to be understood and explained in management meetings, where the cash-flow projections are discussed.
6. Light in computational requirements.

6.4 Importance of the observability study

During the development of the present study it was concluded that, before the set up of a diagnostics procedure, a preliminary observability investigation must be undertaken if the objective is to have a reliable and effective diagnosis.
The benefits that can be achieved with the methodologies presented in this work are:

- The identification of the correlations (specifying their severities) between any two of the available gas-path measurements that respond in a similar way to the component changes.

- The identification of the correlations (specifying their severities) between any of the component changes and/or sensor biasses that produce similar changes in all (or nearly all) of the gas-path measurements.
- The identification of more complex correlations among combinations of more than two changes in performance parameters and/or sensor biasses that are not observable because they are associated with no changes in the measurements.

The awareness of such correlations can be used for various purposes: (i) to choose among a very limited number of different diagnostics layouts (in term of measurement set selection), (ii) to provide additional information to the diagnosis regarding what faults cannot be distinguished, and therefore (iii) in future developments, to enhance diagnostics, for example, through a procedure that distinguishes two faults that are not discernible by means of a simple gas-path measurement study, providing other sources of information – e.g. statistical information, expert knowledge, oil analysis data, vibration analysis data, etc. Besides, another benefit that can be gained through the methods developed in the present work is:

- The capability of comparing, for a given engine, a large number of similar diagnostics systems, with the same number of inputs and outputs, but different type of measurement choices, aimed at the identification of the most observable and therefore suitable measurement set.

6.5 Appropriateness of fuzzy-logic in gas-path diagnostics

Because of the technical requirements identified for advanced gas-path diagnostics, it can be concluded that fuzzy logic showed significant advantages and inherent features well suited to GPD problems, as discussed below.

- Volponi (2003) pointed out the necessity to develop different algorithms to address the problem of estimating gradual and rapid deteriorations, namely MFI (multiple fault isolation) techniques generally based on linear approaches and SFI (single fault isolation) methods necessarily based on non-linear approaches respectively. The developed fuzzy diagnostics system preserved the non-linearity that is present in the aero-thermal relationships between the performance parameters and the gas-path measurements.
- For fuzzy diagnostics to be effective, an exhaustive number of rules, defined within a performance parameter search space, need to be stated. This becomes cumbersome when the number of parameters that are considered simultaneously changed increases (tests were performed with one gas-path component degraded at a time – SFI, and with up to 2 components and so 4 performance parameters deteriorated at a time – partial MFI)

- A fuzzy diagnostics system, with SFI or partial MFI capability, can operate coupled with a linear MFI algorithm as long as a global deterioration level is updated every flight. The rules must be tuned (see section 4.6.3) over the calculated global deterioration level, which is estimated for the previous flight: this is enabled by the remarkably rapid set-up phase that characterizes the fuzzy diagnostics models implemented in this research.
- Fuzzy diagnostics systems do not show a tendency to smear the results over all the performance parameters (that for example affects the Kalman filter based diagnostics methods), demonstrating on the contrary good concentration capability.
- Fuzzy diagnostics systems do not require completely observable systems with the same number of inputs and outputs.
- A considerable enhancement of the diagnostics accuracy in the presence of noisy data can be obtained by choosing the input measurement MFs amplitudes according to the different values of sensor noise standard deviations available for the different sensors. A statistical interpretation of the fuzzy systems can be formulated.
- A SFI bias-tolerant fuzzy system can be effectively implemented by means of the NOT logical operator and a new formulation of the fuzzy rules that includes the location of the bias. The number of rules and therefore the computational time required is greater than the SFI without the capability of dealing with biasses.
- As far as the computational time required, fuzzy diagnostics systems exhibit:
  - A considerably fast set-up phase (e.g. approximately 1 minute for a SFI system), especially when compared with the very long training period required by a neural network for equivalent diagnostics features. This enables the set up of a new system for a new operating condition or over a calculated deterioration baseline in a short time.
  - Fast diagnostics time suitable for on line applications.

  The computational time depends on the number of rules stated and therefore on the number of parameters simultaneously deteriorated.
- Fuzzy logic diagnostics models are advantageous when different sources of information (e.g. oil-analysis, oil debris analysis, vibration analysis, expert knowledge, statistical inputs, etc.) need to be combined in the decision-making process (i.e. data fusion). Such a feature can also be used to combine results computed with different gas-path diagnostics techniques gaining in accuracy and reliability of the results.
- The modular nature of the fuzzy rules (see section 4.2.6.1) that are stated to devise a diagnostics system enables the user with a high level of comprehensibility.
- The adaptation of a fuzzy diagnostics system to different gas turbines is expected to be simple according to the procedures

described above. However, a sensitivity study to optimise the fuzzy system parameters is strongly advisable.

6.6 Suitability of time-series analysis in gas-path prognostics

The first conclusion that can be drawn in an attempt to devise a GPP framework is that the business intention that drives the prognoses determines the various system's requirements. A detailed study of the different requirements associated with the different business intentions should dictate what future research should be undertaken. In any case, the development of a reliable forecasting framework plays a key role in providing the pertinent decision-making competences. At present, due to the absence of physics-based deterioration models in the public domain, whose future development could enable substantial improvements in the prognostics capability, a time-series-analysis based framework can provide the necessary statistical approach to achieve effective gas-path prognoses. In the light of the technical requirements for such a framework, previously summarized, the following conclusions can be drawn.

- The use of time series forecasting methods, able to compute forecasts and associated prediction intervals, enables gas-path prognoses with additional statistical information to be obtained. The choice of the confidence required in the prediction intervals is strictly related to the business intention. If for example an airline wants to guarantee competitive departure statistics for executive customers, a higher level of confidence might be required compared with a low fare airline that aims at competitive ticket-prices instead.
- Gas-path prognostics problems require both short and long term predictions. When using time-series analysis to assist the prognosis with forecasts, different methods should be used for the 2 diverse problems. The use of regression analysis introducing regression models justified on physics bases related to the nature of the gas-path component degradations, are best suited for long term predictions. A demonstration was given in section 5.4. This choice is advantageous due to its simplicity: it allows tangible justification of the results and minimum assumptions on the statistics of the variables. Conversely, more effective forecasts can be computed for short term analyses with more suitable techniques, such as ARIMA models. Nevertheless, the reliability of the ARIMA forecasts is affected by the analyst's subjective interpretation of the correlograms.
- The decision-making process for dependent-parameter based prognostics relies on the prediction of the engine's global variables, say measurements strictly related with the global health of the engine, such as the TGT, shaft speed or the SFC.
- Independent parameter based prognostics is aimed at the prediction of the performance parameter deviations in short and long term horizons. It is the author's belief that the most suitable strategy to achieve such a level of competency requires an algorithm that feeds

the forecasted measurements into the diagnostics procedure in order to estimate future performance parameter deviations and the associated prediction intervals.

- A time-series analysis based framework can only forecast gradual deteriorations, which nonetheless can be characterized by either slow (soft) or fast (severe) rates – see Chapter-5 . Sudden events, that do not manifest themselves with trends in the time-series, cannot be forecasted by definition. However, such instantaneous fault events can be taken into account within the characterization of the warning and safety margins by using hazard plots that establish the probability of occurrence of an event over time.
- A key issue of prognostics was identified in an intelligent management, storage and analysis of the in-flight data that can be made available nowadays and can be used to enhance the prognostics capability.

6.7 Summary of contributions from this study

A detailed discussion of the novel contributions that the present research makes was given in section 3.2.1. For completeness, a summary of the most relevant ones is reported here:

- The novel development of a SAM fuzzy diagnostics system, by means of extensive rule statements, generated using an engine model was developed to achieve a quantitative solution through a non-linear approach, that is able to:
  - Perform SFI in the presence of noisy data.
  - Be tuned over a known deterioration level for all the performance parameters (baseline). This baseline is assumed to be calculated at the previous flight with a MFI method and represents the global deterioration level. This feature enables SFI and MFI processes to operate in concert with one another – see section 4.6.3.
  - Perform a partial MFI with up to 2 degraded components (4 performance parameters) being significantly faulty simultaneously.
  - Perform a SFI while isolating systematic errors in the measurements (biasses). A bias-tolerant system was devised by means of the NOT logical operator and a new formulation of the fuzzy rules that includes the location of the bias.
- A novel idea was proposed to quantify the quality of the system observability through a pertinent parameter. This enables comparisons among a significant number of systems characterised by different measurement set selections, to identify the most suitable.
- An innovative prognostics framework was devised, that uses two different statistic-based forecasting techniques for short and long term investigations characterised by:
  - The capability of computing the prediction intervals, which take into account errors in the forecast.
  - A procedure that feeds forecasted measurements into the diagnostics process, and gains statistical information on the

predicted output performance parameters, is proposed as a novel logic layout.

6.8 Summary of recommendations for future studies

A thorough analysis of the recommendations for future studies, which matured during the development of the present research, was presented in section 4.11 for advanced gas-path diagnostics (GPD) and in section 5.6 for advanced gas-path diagnostics (GPP). For completeness, a summary of the most relevant recommendations is reported below. As far as GPD is concerned:

- A pertinent integration between the observability and the diagnostics processes should be put in place. If the diagnostics procedure is based on fuzzy logic, fuzzy rule weights could be used for this purpose.
- The complete multiple fault isolation capability achievable within a given search space by means of a fuzzy diagnostics system should be tested.
- Frameworks capable of coupling SFI and MFI and able to detect events when they happen in order to enable the two methods to work in concert with one another should be studied.
- Future studies should attempt to combine more than one diagnostics technique to offset the limitations of one with the advantages of the other: suggestions were given in section 4.11.
- Data fusion techniques, that combine different sources of information (e.g. oil analysis, oil-debris analysis, vibration analysis, expert knowledge, statistical inputs, etc.) in order to enable a higher level of competency in the decision-making process, should be investigated. Fuzzy logic should be considered as a potential successful candidate.

As far as GPP is concerned, the following recommendations were drawn:

- A thorough investigation of how prognostics requirements vary according to the different business intentions should drive future studies.
- Resources should be devoted to the development of reliable physics-based models of the deterioration to enable effective predictions.
- Future prognostics frameworks should include the capability of dealing with fault events.
- The optimal layout to undertake independent parameter-based prognostics able to predict performance parameters deviations in a given time horizon should be identified. Suggestions are provided Chapter-5.
- One of the most difficult and important issues related to engine monitoring remains the management the enormous amount of in-flight data that can be made available nowadays. The key elements to identify are the type of in-flight data that should be stored and analysed, how this analysis should be carried out and what sort of information should be sought, in an attempt to minimise the computational storage requirements.

REFERENCES

Abernethy, R. (2000), *The new Weibull handbook,* by Abernethy, Florida, USA.

Abernethy, R.B. and Thompson, J.W. (1973a), *Handbook, uncertainty in gas turbine measurements,* Report, Arnold air force station, Tennessee.

Abernethy, R.B. and Thompson, J.W. (1973b), 'Uncertainty in gas turbine measurements.', AIAA/SAE 9th Propulsion Conference, 5-7 November 1973,Las Vegas, Nevada, U.S.A., AIAA paper N.73-1230.

Airbus S.A.S. (2002), *Global market forecast 2002-2000,* © AIRBUS S.A.S., Printed in France.

Alexander P. (2003), 'Engine sensor fault diagnostics of a large turbofan gas turbine using fuzzy logic', MSc Thesis, School of Engineering, Cranfield University.

Ali, M. and Gupta, U. (1990), 'An expert system for fault diagnosis in a space shuttle main engine', 26th AIAA/SAE/ASME/ASEE Joint Propulsion Conference, Orlando, FL ,July 16 - 18,AIAA 90-1890.

Bakken, L.E. and Skorping, R. (1996), 'Optimum operation and maintenance of gas turbines offshore', International Gas Turbine and Aeroengine Congress & exhibition, Birmingham, U. K., June 10-13 1996.

Barwell, M.J. (1987), 'Compass-ground based engine monitoring program for general application', Aerospace Technology Conference and Exposition, Long Beach, California, October 5-8 1987, SAE paper No. 87/1734.

Beschorner J. (2003), 'Gas-path prognostics applied to the F404 military turbofan', MSc Thesis, School of Engineering, Cranfield University.

Boeing (2003), *Current market outlook report,* Marketing, Boeing commercial airplanes .

Box, G.P. and Jenkins, G.M. (1976), *Time-series analysis forecasting and control,* Holden-Day, Oakland, California.

Brotherton, T. (2000), ' Prognosis of faults in gas-turbine engines', Aerospace Conference Proceedings, 2000 IEEE Vol. 6, 163-171.

Brown G. (1966), 'Not just observable, but how observable?', Iowa proceedings of the national electronics conference, vol 22, Iowa state university, Ames,

Bryson and Ho (1975), *Applied optimal control,* Hemisphere Publishing Corporation, USA.

Burnell, D. (1995), 'Engine condition monitoring-engine system and aircraft operational effects', 18th Symposium on aircraft integrated monitoring systems, Stuttgart, Germany, September 19-21 1995, 353-363.

Byington, C., Roemer M. , and Galie T. (2002), 'Prognostic enhancements to diagnostic systems for improved condition-based maintenance', Aerospace Conference, Big Sky, MT, March 9-16, 2002 IEEE.

Chatfield C. (1989), *The analysis of time series: an introduction,* Chapman and hall, London, UK.

Cribbes, T.D. (1997), *Changes in engine maintenance management,* Aerospace Engineering, 17(12), 7-9.

Crosby, J.K. (1986), 'Factors relating to deterioration based on Rolls-Royce RB211 in-service performance', Vol. 37, p.41-47, Turbomachinery performance deterioration.

Curnock, B.S. (2000), 'Obidicote programme work package 4: Analysis of transient engine data', Rolls-Royce Defence, Technical Report DNS67190.

DePold, H.R. and Gass F. D. (1999a), 'The application of expert system and neural networks to gas-turbine prognostics and diagnostics', Vol. 121, N.4, p.607-612, Journal of gas-turbine and power.

DePold, H.R. and Gass, F.D. (1999b), 'The Application of Expert Systems and Neural Networks to Gas Turbine Prognostics and Diagnostics ', Volume 121, pp 607-612., Journal of Engineering for Gas Turbines and Power.

Diakunchak, I.S. (1992), 'Performance deterioration in industrial gas turbines', vol 114, p161-168, Journal of Engineering for Gas Turbines and Power.

Doel D. L. (1990), 'The Role for Expert Systems in Commercial Gas Turbine Engine Monitoring', The Gas Turbine and Aero-engine Congress and Exposition, Brussels, Belgium, June 1990, ASME Paper 90-GT-374.

Doel D. L. (1994), 'TEMPER - A gas-path analysis tool for commercial jet engines', vol. 116, pp. 82-89, Journal of engineering for gas turbines and power.

Dunn, J. (1997), 'Faultless plan for planes ', 12th of February, 17, Professional engineering.

English, L. (1995), 'Application of gas-path analysis, gas-path debris monitoring and expert system technology to Allison T56 turboprop engine', MSc Thesis, School of Mechanical Engineering, Cranfield University.

Escher P. C. (1995), *Pythia: an object-oriented gas-path analysis computer program for general applications,* PhD thesis, School of Mechanical Engineering, Cranfield University.

Eustace R. and Merrington G. (1993), 'Fault diagnosis of fleet engines using neural networks ', XII ISABE 95- 7085.

Fuster P., Ligeza A. , and Aguilar Martin J. (1997), 'Adductive diagnostic procedure based on an AND/OR/NOT graph for expected behaviour: application to a gas turbine ', COMADEM'97, Vol.171, pp.511-520. 10th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management, Finland.

Ganguli, R. (2001a), 'Application of fuzzy logic for fault isolation of jet engines', ASME 2001-GT-0013.

Ganguli, R. (2001b), 'A fuzzy logic intelligent system for gas turbine module and system fault isolation', ISABE-2001-1112.

Ganguli, R. (2002), 'Health monitoring of a helicopter rotor in forward flight using fuzzy logic', AIAA Journal, Vol. 40, N.12, pp.2373-2381.

Gelb, A. (1974), *Applied optimal estimation,* The analytic science corporation, U.S.A.

Ghiocel D. and Roemer M. (1999), 'Probabilistic interaction of relevant technologies for risk-based life prediction of GTE components', AIAA-99-1591.

Ghiocel, D.M. (2001), 'Critical modelling issues for the prediction of turbine-performance degradation: use of a stocastic-neuro-fuzzy inference system', AIAA-2001-1452, Seattle, WA, 16-19 April.

Goldberg, D. (1989), *Genetic algorithms in search, optimisations and machine learning,* Addison Wesley, USA.

Green, A. (1997), ' Artificial-intelligence for real-time diagnostics and prognostics of gas-turbine engines', AIAA 97-2899, Seattle, WA, 6-9 July 1997,

Greitzer et al. (1999), 'Gas-turbine engine health-monitoring and prognostics', Las Vegas, Nevada, August 30 – September 2, International Society of Logistics (SOLE) 1999 Symposium.

Greitzer, F.L. and Ferryman, T.A. (2003), 'Predicting remaining life of mechanical systems', Intelligent Ship Symposium IV, April 2-3

Gulati, A., Taylor, D., and Singh, R. (2001), 'Multiple operating point analysis using genetic algorithm optimisation for gas turbine diagnostics', ISABE-2001-1139.

Gulati, A., Zedda, M., and Singh, R. (2000), 'Gas turbine engine and sensor multiple operating point analysis using optimisation techniques', 36th AIAA/ASME/SAE/ASEE Joint propulsion conference and exhibit,17-19 July 2000,Huntsville, Alabama, AIAA 2000-3716.

Healy A., Kerr L., and Larkin L. (1997), 'Model based fuzzy logic sensor fault accommodation', 97-GT-222, ASME.

Hesketh G. and Swann A. (1999), 'Data fusion - reasoning with uncertainty', Rolls-Royce Aerospace, Technical Report DNS63633.

Jaw, L.C. and Friend R. (2001), 'ICEMS: a platform for advanced condition-based health-management', Aerospace Conference, Big Sky, Montana, USA, March 13, 2001 IEEE.

Jazwinski, A.H. (1970), 'Stochastic process and filtering theory', Academic Press, New York.

Kacprzynski, G., Gumina M. , Roemer M. , Caguiat D. , Galie T. , and McGroarty J. (2001), 'A prognostics modelling approach for predicting recurring maintenance for shipboard propulsion systems', ASME TURBOEXPO 2001,4-7 June 2001,New Orleans, USA, 2001-GT-0218.

Kacprzynski, G. and Roemer M. (2002), 'Enhancement of physics-of failure prognostics models with system level features', 2002 IEEE.

Kadamb A. (2003), 'Bayesian belief network for aero gas turbine module and system fault isolation', MSc Thesis, School of Engineering, Cranfield University.

Kandel, A. (1986), *Fuzzy mathematical techniques with applications,* Addison-Welsey publishing company, USA.

Kanelopoulos K. , Stamatis A. , and Mathioudakis K. (1997), 'Incorporating Neural Networks into Gas Turbine Performance Diagnostics ', ASME 97-GT-35, International Gas Turbine & Aero-engine Congress & Exhibition, Orlando, Florida, USA, June 8-13 1997.

Karvounis G. and Frith P. (2003), 'Automated detection of engine heath using hibrid model-based and fuzzy-logic approach', XVI International Symposium on air breathing engines; Cleveland, Ohio; ISABE-2003-1232, 31 Aug-5 Sept 2003.

Kosko, B. (1994), 'The probability monopoly', vol. 2, n.1, 32-33, February 1994, IEEE Transactions on fuzzy systems.

Kosko, B. (1997), *Fuzzy engineering,* Prentice Hall, New Jersey.

Kurz, R. and Brun, K. (2001), 'Degradation in gas turbine systems', vol. 123,pp 70-77, Journal of Engineering for Gas Turbines and Power.

La Grandeur, R.D. (1986), *Instrumentation for aero gas turbine engine condition monitoring systems,* MSc Thesis: School of mechanical engineering, Cranfield University.

Lakshminarasimha, A.N., Boyce, M.P., and Meher-Homji, C.B. (1994), 'Modelling and analysis of gas turbine performance deterioration', vol. 116, pp 46-52, Journal of Engineering for Gas Turbines and Power.

Laviolette, M. and Seaman, J.W. (1994), 'The efficiency of fuzzy representations of uncertainty', vol. 2, n.1, 4-15, February 1994, IEEE Transactions on fuzzy systems.

Law, M. and Kelton, W. (1991), *Simulation modelling and analysis,* McGraw-Hill, U.S.A.

Li, I.Y. (2002a), *Gas Turbine Diagnostics,* School of Engineering, Cranfield University.

Li, I.Y. (2002b), 'Performance analysis based gas-turbine diagnostics: a review', IMechE Journal - Power and Energy, vol 216 n. A5 p. 363-377.

MacLeod, J.C., Taylor, V., and Laflamme, J.C.G. (1992), 'Implanted component faults and their effects on gas turbine engine performance', Vol- 114, pp 174-179, Transactions of the ASME journal of Engineering for Gas Turbines and Power.

Makridakis, S., Wheelwright, S.C., and Hyndman, R.J. (1998), *Forecasting: methods and applications,* 3rd ed. John Wiley & Sons, New York.

Marinai, L., Ogaji, S., Sampath, S., and Singh, R. (2003a), 'Engine diagnostics - fuzzy logic apprach', Seventh International Conference on Knowledge-Based Intelligent Information & Engineering Systems - KES-03169, Oxford, 3, 4 & 5 September 2003.

Marinai L., Probert D., and Singh R. (2004), 'Prospects for aero gas-turbine diagnostics: a review', Applied Energy, Volume 79, Issue 1, September 2004, Pages 109-126.

Marinai, L., Singh, R., and Curnock, B. (2003b), 'Fuzzy-logic-based diagnostic process for turbofan engines', ISABE-2003-1149, XVI International Symposium on air breathing engines; Cleveland, Ohio; 31 Aug-5 Sept 2003.

Marinai, L., Singh, R., Curnock, B., and Probert D. (2003c), 'Detection and prediction of the performance deterioration of a turbofan engine', TS-005, International Gas Turbine Congress 2003 Tokyo; November 2-7 2003.

Mathioudakis, K., Kamboukos, P., and Stamatis, A. (2002), 'Turbofan performance deterioration tracking using non-linear models and optimization techniques ', ASME TURBOEXPO 2002, 3-6 June, 2002, Amsterdam, The Netherlands, GT-2002-30026.

Mathioudakis K. , Romessis C. , and Stamatis A (2003), *Probabilistic methods for gas turbine fault diagnostics ,* Von Karman Institute Lecture Series 2003-01, Gas Turbine Condition Monitoring & Fault Diagnosis, Brussels, Belgium, 13-17 January 2003.

Merrington, G.L. (1989), 'Fault diagnostics of gas turbine engines from transient data', Journal of engineering for gas turbines and power, April 1989, vol. 111 p.237-243.

Michalewicz, Z. (1996), *Genetic algorithms and data structures = evolution programs,* Springer-Verlag, USA.

Moes T. (2003), 'Development of a prognostics tool as a contribution to optimising maintenance of a civil turbofan engine', MSc Thesis, School of Engineering, Cranfield University.
Montgomery, C.e.al. (1990), *Forecasting and time-series analysis,* McGraw-Hill, USA.
Naeem, M., Singh, R., and Probert, D. (2001), 'Consequences of aero-engine deteriorations for military aircraft', 70, 103-133, Applied Energy.
Nguyen, H.T. (1978), 'On random sets and belief functions', vol.65, 531-542, 1978, Journal of mathematical analysis and applications.
O'Donovan, T. (1983), *Short term forecasting,* John Wiley & Son Ltd., Bath, UK.
Ogaji S. (2003), 'Advanced gas-path fault diagnostics for stationary gas turbines', PhD Thesis, School of Engineering, Cranfield University.
Ogaji, S., Sampath, S., Singh, R., and Probert, D. (2002), 'Novel approach for improving power-plant availability using advanced engine diagnostics', Applied Energy 72, 389-407.
Palmer C. (1998), 'Combining bayesian belief network with gas-path analysis for test cell diagnostics and overhaul', ASME-98-GT-168,
Papoulis, A. (1994), *Probability, random variables, and stochastic processes,* McGraw-Hill, 2nd ed.
Passalacque J. (1974), 'Description of automatic gas turbine engine trends diagnostic system', National Research Council of Canada, first Symposium on Gas Turbine Operation and Maintenance.
Provost, M.J. (1987), *Kalman filtering for performance engineering,* Performance technical report, Rolls-Royce, Derby.
Provost M.J. (1988), 'COMPAS: A generalized ground based monitoring system', Engine Condition Monitoring, AGARD-CP-449.
Provost, M.J. (1995), *The use of optimal estimation techniques in the analysis of gas turbines,* PhD Thesis, School of mechanical engineering, Cranfield University.
Roemer, M. and Kacprzynski G. (2000), 'Advanced diagnostics and prognostics for gas-turbine engine risk-assessment', IGTI/ASME Turbo Expo, May 2000, Munich, Germany.
Roemer M. and Kacprzynski G (2001), 'Assessment of data and knowledge fusion strategies for prognostics and health management', Aerospace Conference , Big Sky, Montana, USA, March 13, 2001 IEEE.
Romessis C., Stamatis A, and Mathioudakis K. (2001), 'Setting up a belief network for turbofan diagnosis with the aid of an engine performance model', ISABE-2001-1032.
Rupp, O. (2002), Instandhaltungskosten bei zivilen strahltriebwerken, Hannover, Langenhagen, Germany, MTU Maintenance.
Sallee, G.P. (1978), 'Performance deterioration based on existing (historical) data; JT9D jet engine diagnostic program', CR-135448, NASA, Washington, D. C., NASA.
Sampath, S. (2003), 'Fault diagnostics for advanced cycle marine gas turbine using genetic algorithms', PhD Thesis, School of Engineering, Cranfield University.

Sampath, S., Li, Y.G., Ogaji, S.O.T., and Singh, R. (2003a), 'Fault diagnosis of a two spool turbofan engine using transient data: a genetic algorithm approach', ASME-TE-2003, 16-19 June 2003, Atlanta, Georgia, USA.

Sampath, S., Marinai, L., Ogaji, S.O.T., and Singh, R. (2003b), 'Evolution strategy applied to aero engine fault diagnosis ', Seventh International Conference on Knowledge-Based Intelligent Information & Engineering Systems - KES, 3-5 Sep 2003, Oxford, United Kingdom.

Saravanamuttoo, H. (1985), 'A preliminary assessment of compressor fouling', AGARD-LS-183.

Sasahara, O. (1986), 'JT9D engine/module performance deterioration results from back-to-back testing', Vol. 37, p25-32, Turbomachinery performance deterioration.

Sheuren, W.J. (1998), 'Joint strike fighter prognostics and health management', Cleveland, OH, 13-15 July , AIAA 98-3710.

Singh, R. (2001), *Civil aero gas turbines: strategy and technology,* Chairman's Address, Aerospace Division, Institution of Mechanical Engineers, London, April 2001.

Singh R. (2003), *Advances and opportunities in gas-path diagnostics,* ISABE 2003-1008.

Siu C., Shen Q., and Milne R. (1997), 'TMDOCTOR: a fuzzy rule- and case-based expert system for turbomachinery diagnosis', SAFEPROCESS '97, vol.1, pp. 556-563, proceedings of IFAC symposium.

Spiegel, M.R. (1972), *Theory and problems of statistics ,* McGraw-Hill, New York.

Spina, P.R., Torella, G., and Venturini, M. (2002), 'The use of expert systems for gas turbine diagnostics and maintenance', Proceedings of the ASME Turbo Expo, Amsterdam, The Netherlands. June 3-6 2002.

Stamatis A. , Mathioudakis K , Smith M., and Papailiou K. (1990a), 'Adaptive simulation of gas turbine performance', 90-GT-205, ASME.

Stamatis A. , Mathioudakis K., and Papailiou K.. (1991), 'Jet engine fault detection with discrete operating points gas-path analysis', or Journal of Propulsion, Vol.7, No.6, Nov./Dec. 1991, pp. 1043-1048, ISABE 89-7133.

Stamatis A., Mathioudakis K., S.M., and Papailiou K. (1990b), 'Gas turbine component fault identification of adaptive performance modelling', Paper 90-GT-376, , ASME.

Tang G. (1999), 'A practical intelligent system for condition monitoring and fault diagnosis of jet engines', AIAA 99-2533.

Torella, G. (1993), 'Expert systems for the simulation of turbofan engines. ', ISABE 93-7133., International Society for Air Breathing Engines (ISABE), Tokyo, Japan, September 20 – 24 1993.

Torella, G. (1997), 'Expert systems and neural networks for fault isolation in gas turbines', ISABE 97-7148, International Society of Air Breathing Engines.

Torella, G. and and Torella, R. (1999), 'Probabilistic expert systems for the diagnostics and trouble-shooting of gas turbine apparatuses', AIAA 99-2942, 35th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, LA. California, June 20 -24 1999.

Torella G. and Lombardo G. (1995), 'Utilization of neural networks for gas turbine engines', ISABE 95-7032, XII ISABE.

Urban, L.A. (1969), *Gas turbine engine parameter interrelationships,* Hamilton Standard Division of United Aircraft Windsor Locks, USA.

Urban L. A. and Volponi A. J. (1992), 'Mathematical methods of relative engine performance diagnostics', Transaction Journal of Aerospace, section 1, Vol.101, technical paper 922048, SAE 1992.

Visual Numerics (1997), *IMSL, Tools for professional software developer,* Visual Numerics, Huston, Texas, USA.

Volponi, A. (2003), 'Extending gas-path analysis coverage for other fault conditions', Von Karman Institute Lecture Series 2003-01 , Gas Turbine Condition Monitoring & Fault Diagnosis, Brussels, Belgium, 13-17 January 2003.

Volponi A. , DePold H. , Ganguli R. , and Daguang C. (2000), 'The use of kalman filter and neural network methodologies in gas turbine performance diagnostics: a comparative study', ASME 2000-GT-547.

Wheelwright, S.C. and Makridakis, S. (1977), *Forecasting methods for management,* 2nd ed. John Wiley & Sons, New York.

Wittink, D.R. (1988), *The application of regression analysis,* Allyn and Bacon, Boston.

Wulf, R.H. (1980), ' CF6-6D engine performance deterioration', CR-159786, NASA, Washington, D.C., NASA.

Zadeh, L.A. (1969), *Toward a theory of fuzzy systems,* NASA CR-1432, Washington, D.C.

Zadeh, L.A. (1995), *Fuzzy Logic Toolbox User's Guide,* The MathWorks, Inc., Natick, MA.

Zaita, A.V.B.G.K.G. (1998), 'Performance deterioration modelling in aircraft gas turbine engines', vol. 120,pp 344-349, Journal of Engineering for Gas Turbines and Power.

Zedda, M. (1999a), *Gas turbine engine and sensor fault diagnosis,* PhD Thesis, School of mechanical engineering, Cranfield University.

Zedda, M. and Singh, R. (1998), 'Fault diagnosis of a turbofan engine using neural networks: a quantitative approach', 34th IAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, AIAA 98-3602, Cleveland, OH, 13-15 July 1998.

Zedda, M. and Singh R. (1999b), 'Gas turbine engine and sensor fault diagnosis', ISABE 99-7238, 13th ISABE, Florence, Italy.

Zedda, M. and Singh, R. (1999c), 'Gas turbine engine and sensor fault diagnosis using optimization techniques', 35th AIAA/ASME/SAE/ASEE Joint propulsion conference and exhibit, Los Angeles, California, AIAA 99-2530, 20-24 June 1999.

APPENDIX A – THE TRENT 800

The engine involved with this research project is the Rolls-Royce Trent 800. It powers the Boeing 777-200, -200ER and -300 twinjet. The Trent 800 is a high BPR (bypass ratio) turbofan and entered service in April 1996 with Thai Airways International.
The Boeing 777 is a medium to long haul twin-aisle wide body aircraft. The Trent 800 powered versions have a maximum range of between 9649km (5210nm, -200) and 14316km (7730nm, -200ER) with a seating capacity varying between 301 (-200ER) and 368 (-300) passengers in a standard three class configuration. The typical cruising speed of the 777 is Mach 0.83 at 10668m (35000ft).

In line with the Rolls-Royce design philosophy the Trent 800 has a three shaft layout (see Figure 68 in section 4.3). Currently it is both the largest Rolls-Royce production engine with a fan diameter of 2.794m and also the most powerful one. The engine is currently available in five different performance ratings ranging from the 875, which delivers a take off thrust of 331.9kN (74600lb) to the latest version, the 895, which produces a maximum thrust of 422.6kN (95000lb). The Trent 892 that is considered in the Turbomatch simulations powers the 777-200ER and -300 and provides a take off thrust of 407.5kN (91600lb).

Take off and cruise performance parameters, as well as key dimensions and the weight information are summarized below.

Table 64: Take off performance (SLS, ISA, flat rated to 30°C, 895 to 25°C)

Thrust	331.9-422.6kN (74600-95000lb)
Fan mass flow	1119-1208kg/s (2467-2664lb/s)
OPR (Operating pressure ratio)	34.5-41.9
BPR	6.2-5.8

Table 65: Cruise performance (10668m, Mach 0.83, to ISA+10°C)

Thrust	60.05kN (13500lb)
SFC	16.283mg/Ns (0.575 lb/h/lb st)

Table 66: Engine dimensions and weight

Fan diameter	2.794m (110in)
Length	4.369m (172in)
Weight (dry)	5942kg (13100lb)

APPENDIX B – ENGINE MODELLING

Performance modelling of a gas turbine engine takes place from its first conception to well into the engine's service life. Essentially, performance modelling refers to the analysis of the behaviour of the engine primarily investigating the aerothermodynamics of the design. In the early stages of a new design, the preliminary sizing of components is performed by undertaking a design point analysis, in which sensible values of component efficiencies are used to produce a thermodynamic cycle which will fulfil the basic design criteria, usually relating to thrust and fuel consumption. Subsequent to the preliminary cycle design, further analysis takes place away from the design point conditions, in order to ensure that the operability of the engine, its emissions and noise profiles and various aspects of engine safety are satisfactory. Modelling may then continue beyond the design phase, in order to analyse in-service operation and to schedule preventative maintenance.

Performance modelling is used in order to ensure the reliability of an engine design to verify that it meets the specifications expecially in terms of fuel consumption, emissions, operability and noise.

The performance model primarily employed in this research is Turbomach (developed at Cranfield University) that is a whole-engine steady-state simulation code. As detailed computational models of the performance of the individual components of the engine would be very intensive on computing power, the components are treated as black boxes, otherwise known as bricks. These bricks generally contain experimental or computational maps of the component performances. These bricks are then joined together to build an overall model. The mass flow rate, pressure and temperature are conserved between the blocks and the pertinent relationships between the rotational speeds of the components are set where appropriate. If the simulation is steady-state, power or torque balances are made on the rotating components. The whole system is then iteratively solved.

Many performance simulation packages are available for modelling gas turbines, an example is the Rolls-Royce Aerothermal Performance (RRAP) code.

Both RRAP models and Turbomatch share many similar features. Particularly, they include a relatively robust multi-dimensional mathematical non-linear solver and a set of thermodynamic subroutines. These all form part of the overall gas turbine performance simulation tool, comprising the usual types of bricks to define the performance of individual components, linked together to form an overall system.

APPENDIX C – THE SOFTWARE: HEALTH MONITORING AND PROGNOSTICS (HMP)

The outcome of this research is a health-monitoring-and-prognostics (HMP) framework for performance analysis that consists of three main modules able to carry out:

- Observability study.
- Fault assessment or diagnostics.
- Forecasts and prognostics.

The modules compute the actual calculations and communicate with each other, they constitute the software named HMP 1.1 for gas turbine performance analysis – see its logo in Figure 123.

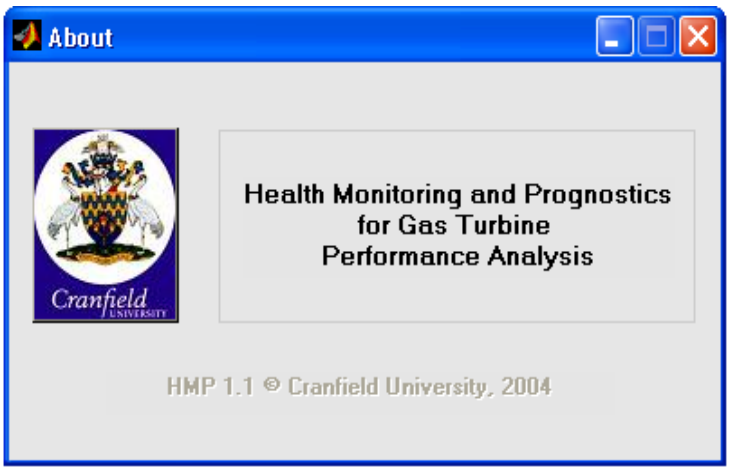


Figure 123: HMP 1.1 logo

The software's main window is shown in Figure 124. Firstly, the engine to be analysed and the relative engine model are selected. In Figure 125 the Rolls-Royce Trent 800 is chosen. This thesis was totally dedicated to the study of the Trent 800 turbofan engine. Nevertheless, the following engines were investigated outside this research program and their analysis capability is currently available within the HMP 1.1 framework:

- Rolls-Royce two-shaft high-bypass-ratio V2500 engine for civil applications.
- Rolls-Royce two-shaft reheated turbofan EJ200 engine for military applications.

Once the engine is chosen a type of analysis and its corresponding module can be selected among the three available: (i) observability module, (ii) diagnostics module, and (iii) prognostics module. Figure 126 and Figure 127 show the three options and the further brake-down of the diagnostics and prognostics modules:

- The observability module is shown in Figure 128. The theory and the procedures implemented within this module are described in section 4.4.
- The diagnostics module is constituted by two GUIs: one capable of setting-up a fuzzy logic diagnostics model for a given engine and operating condition represented in Figure 129, and another capable of operating that diagnostics model Figure 130 – see Chapter-4 for the the underpinning theory.

- The prognostics model includes two main GUIs: the first is designed to undertake short-term investigations using ARIMA models – see Figure 131, Figure 132 and Figure 133, and the second dedicated to the study of long term investigations using regression analysis – see Figure 134 and Figure 135. Theory and procedures are described in Chapter-5.

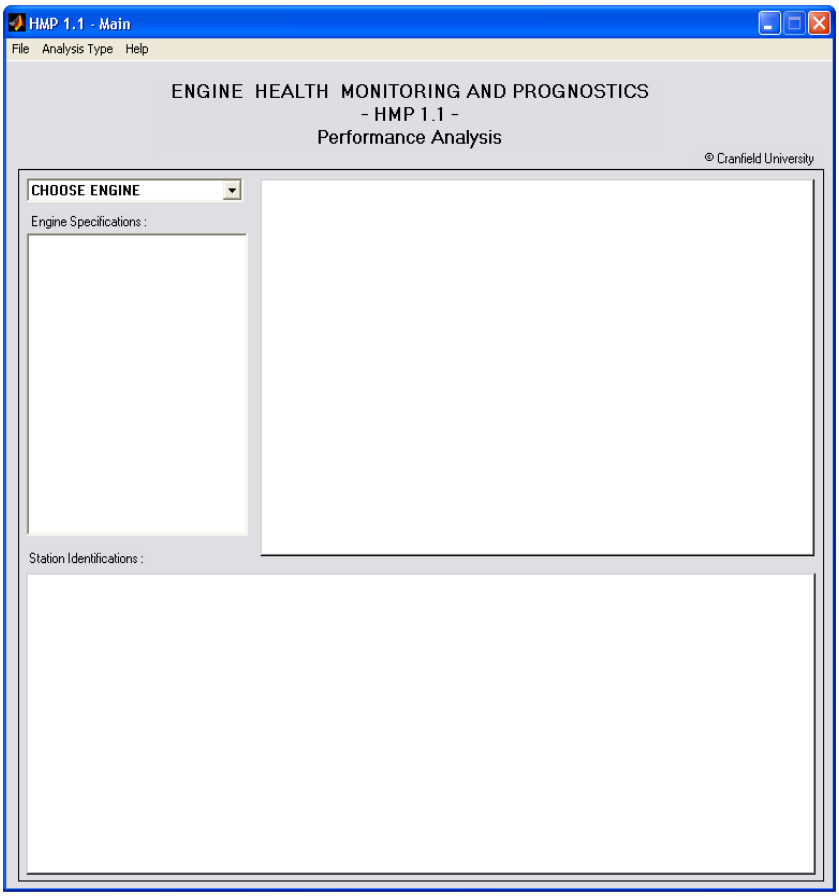


Figure 124: HMP 1.1 Main window

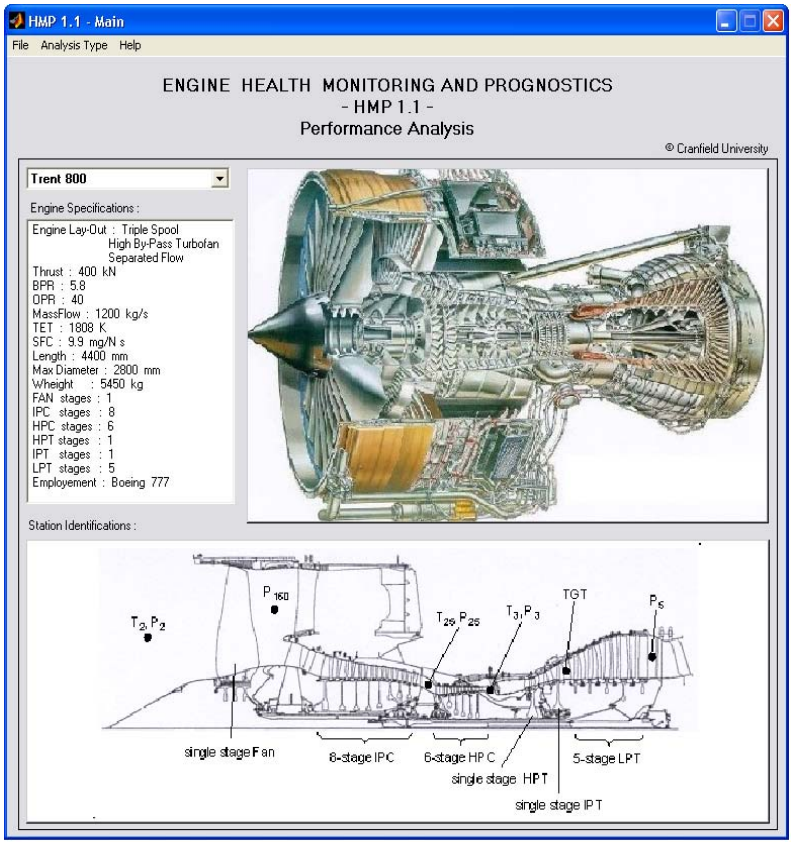


Figure 125: HMP 1.1 Main window, engine selection, image courtesy of Rolls-Royce

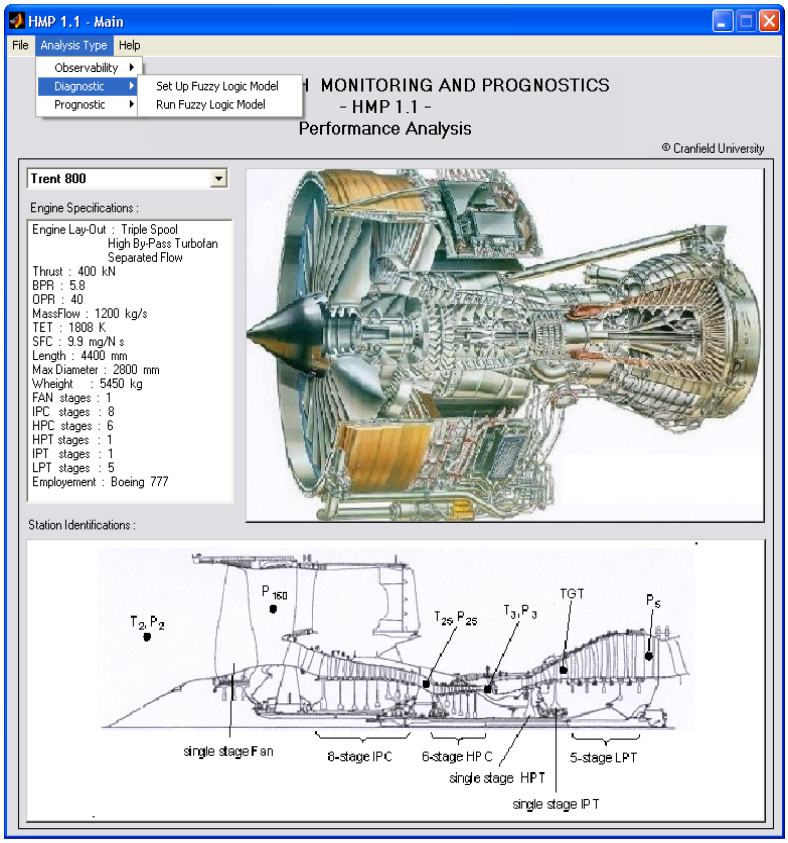


Figure 126: HMP 1.1 Main window, analysis type selection, diagnostics options, image courtesy of Rolls-Royce

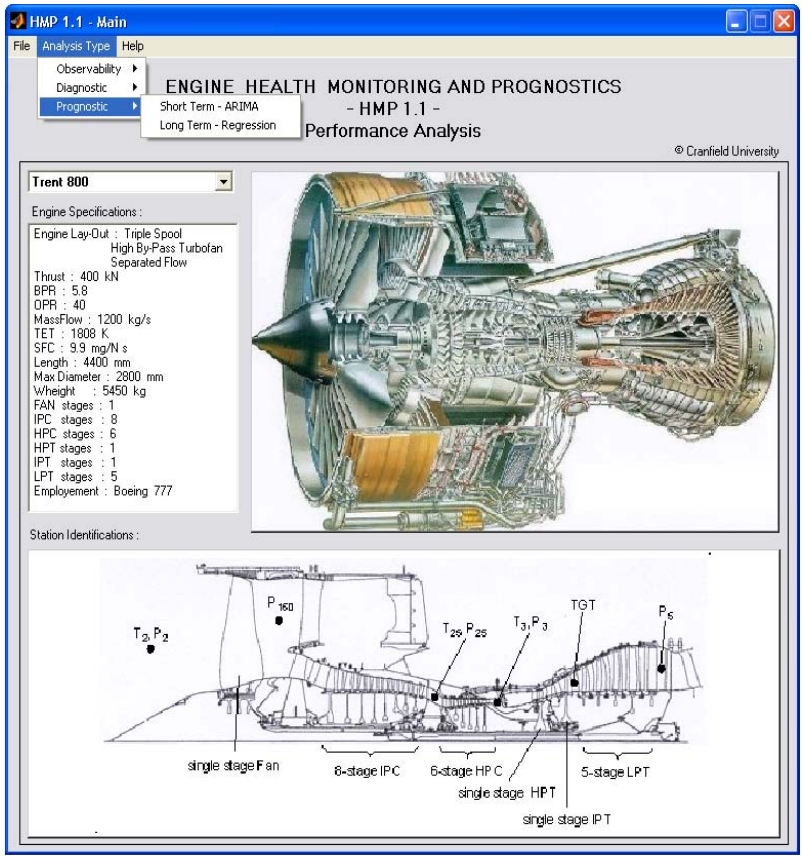


Figure 127: HMP 1.1 Main window, analysis type selection, prognostics options, image courtesy of Rolls-Royce

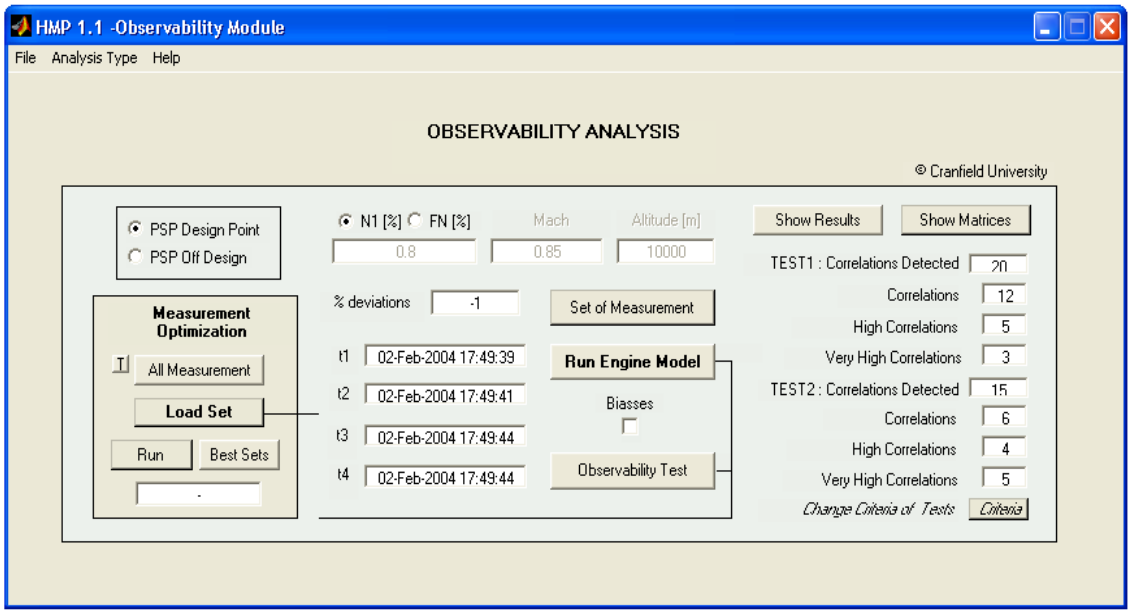


Figure 128: HMP 1.1 Observability module

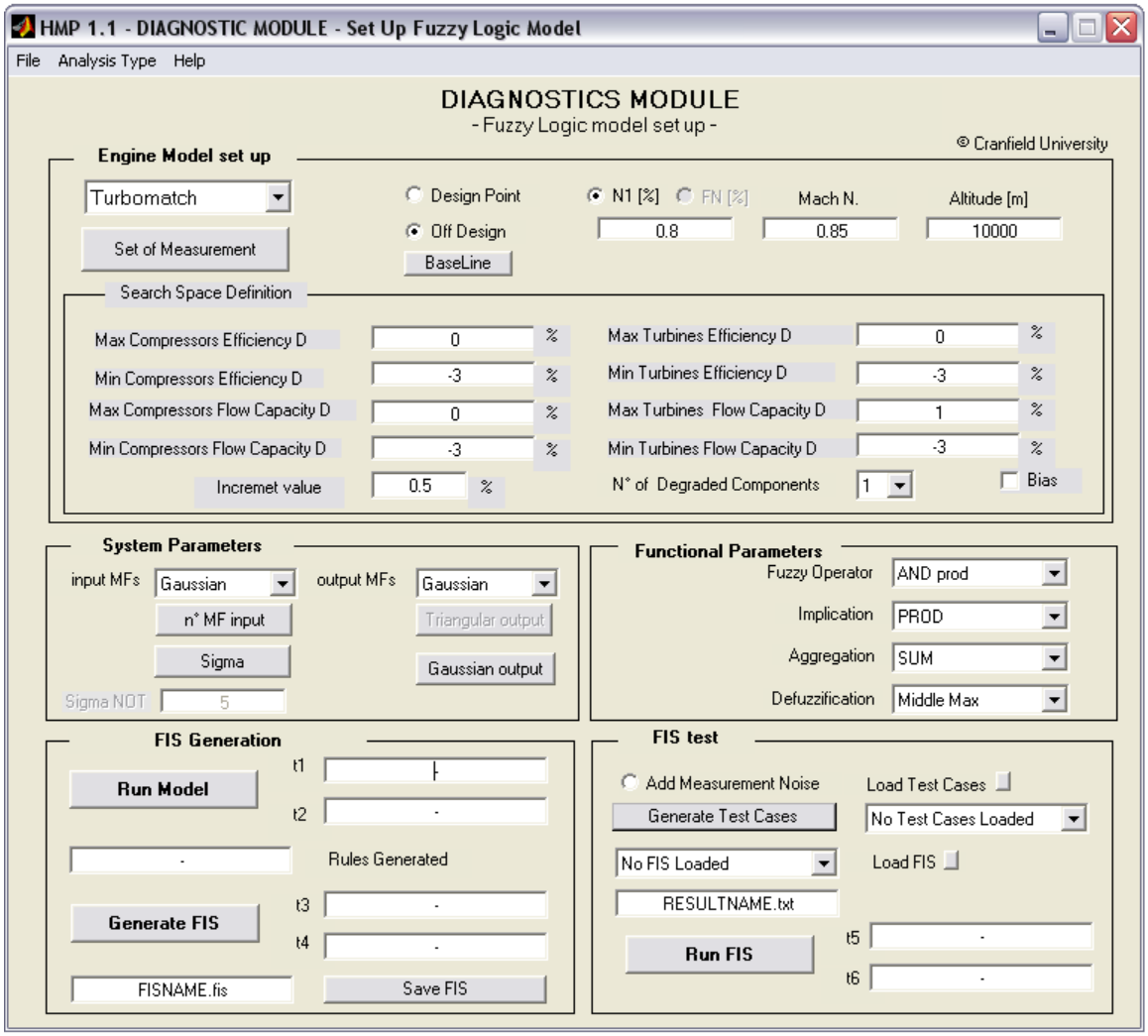


Figure 129: HMP 1.1 Diagnostics module, GUI to set-up a fuzzy logic model

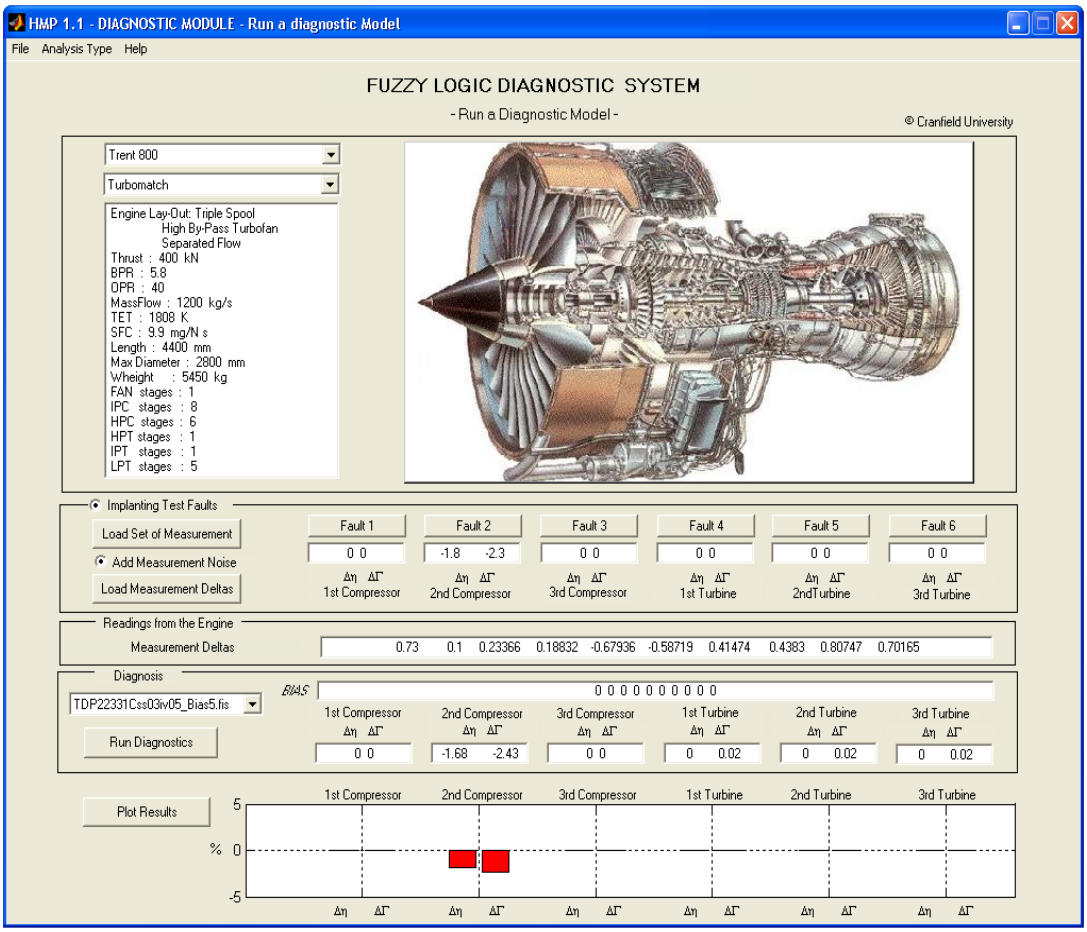


Figure 130: HMP 1.1 Diagnostics module, GUI to operate a fuzzy logic model, image courtesy of Rolls-Royce

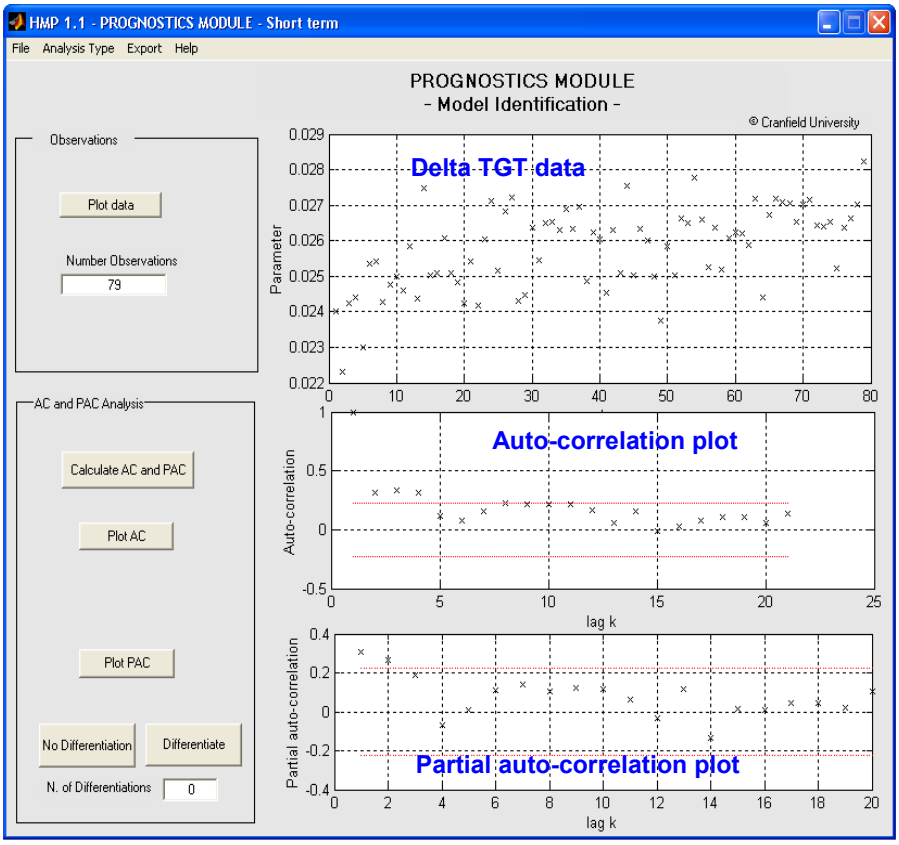


Figure 131: HMP 1.1 Prognostics module, short-term investigations, model identification

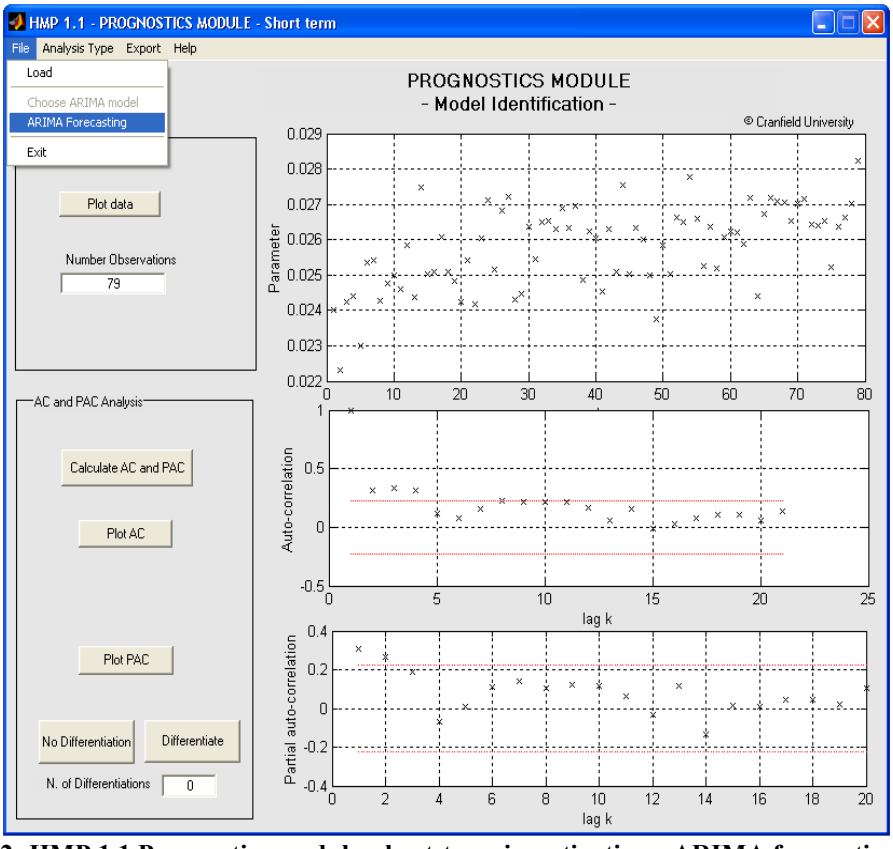


Figure 132: HMP 1.1 Prognostics module, short-term investigations, ARIMA forecasting selection

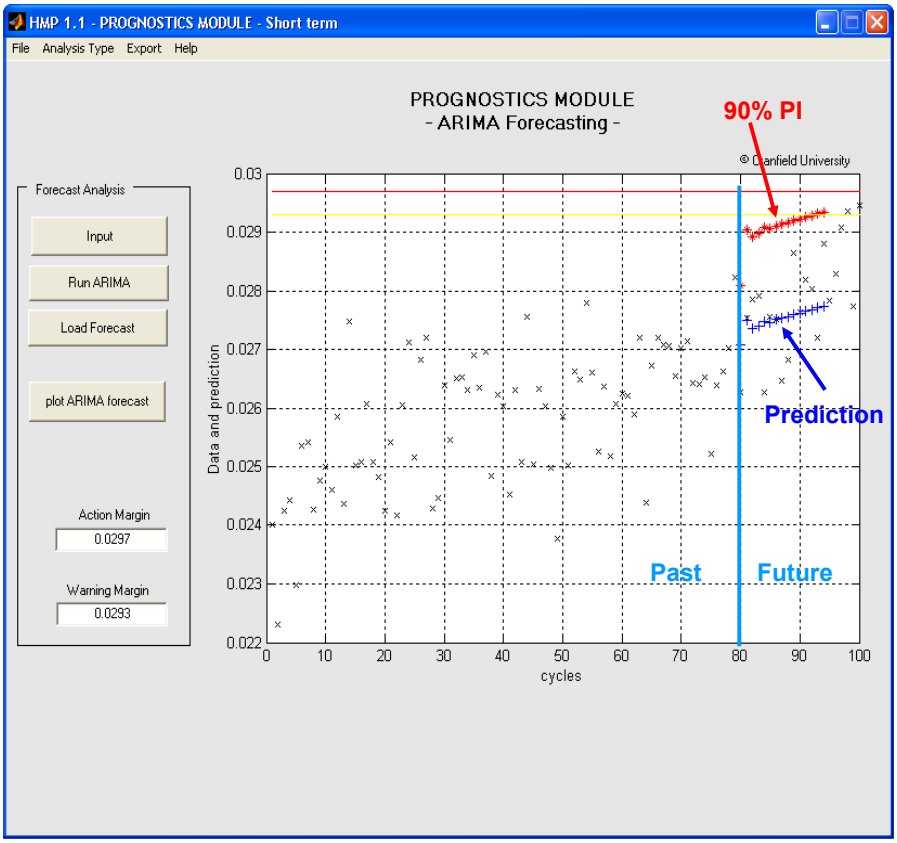


Figure 133: HMP 1.1 Prognostics module, short-term investigations, ARIMA forecasting output

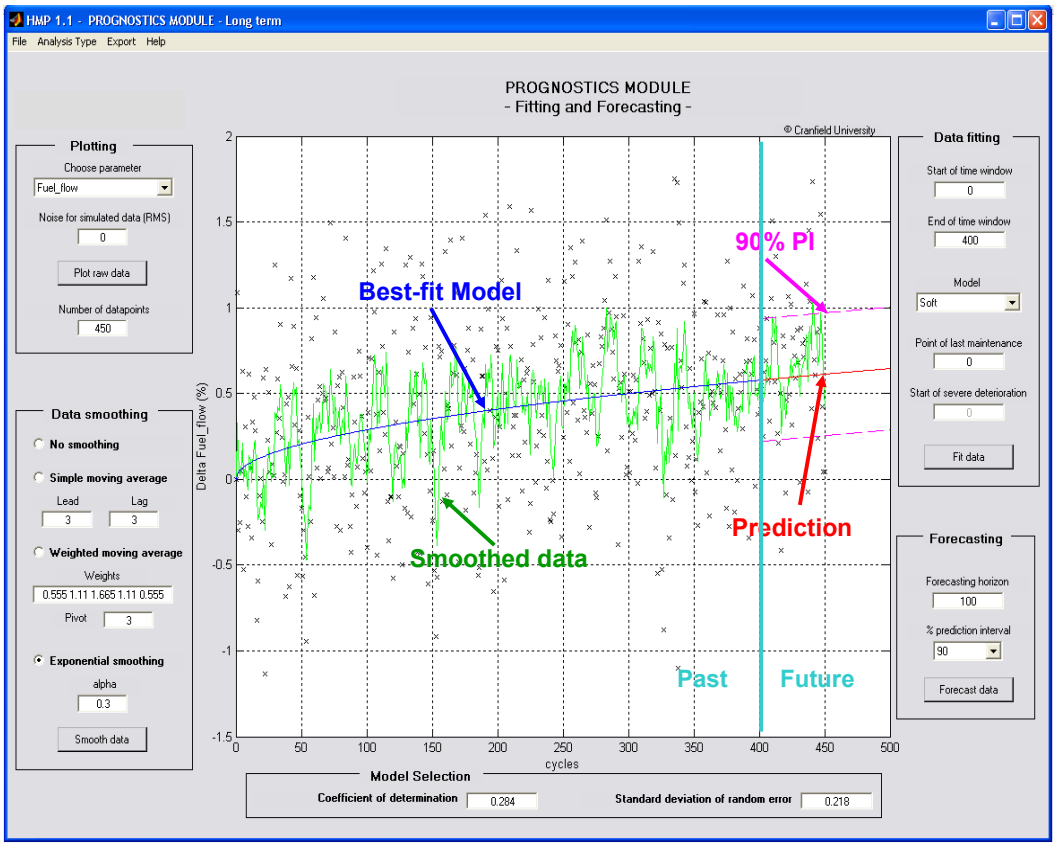


Figure 134: HMP 1.1 Prognostics module, long-term investigations, GUI capable of fitting and forecasting with regression models

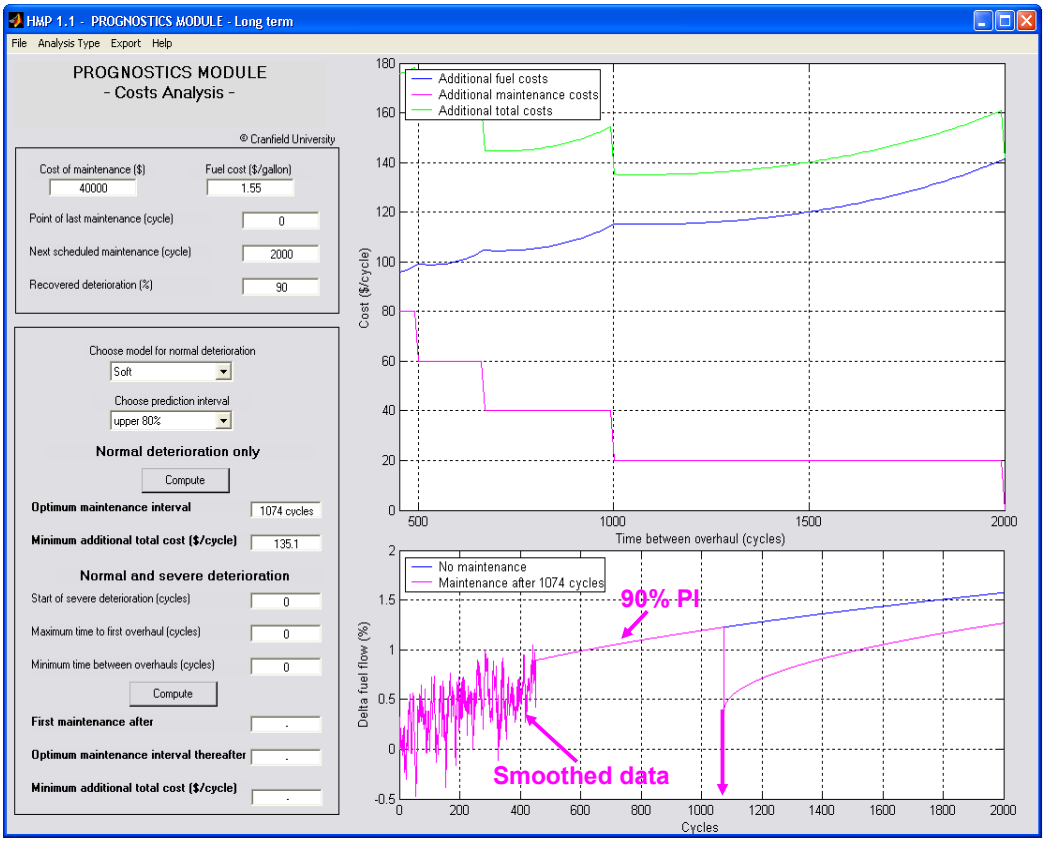


Figure 135: HMP 1.1 Prognostics module, long-term investigations, GUI to undertake cost analyses