

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

# UMI

A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA  
313/761-4700 800/521-0600



THE APPLICATION OF FUZZY LOGIC  
TO THE DIAGNOSIS OF AUTOMOTIVE SYSTEMS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for  
the Degree Doctor of Philosophy in the Graduate  
School of The Ohio State University

By

Ahmed A. Soliman, M.S.

\* \* \* \* \*

The Ohio State University

1997

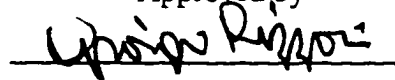
Dissertation Committee:

Professor Giorgio Rizzoni, Adviser

Professor Sheikh Akbar

Professor Stephen Yurkovich

Approved by



Adviser

Department of Mechanical Engineering

UMI Number: 9813350

---

UMI Microform 9813350  
Copyright 1998, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized  
copying under Title 17, United States Code.

---

**UMI**  
300 North Zeeb Road  
Ann Arbor, MI 48103

*Copyright by  
Ahmed A. Soliman  
1997*

## ABSTRACT

Fault diagnosis of a physical plant is crucial for its healthy performance, as it ultimately could prevent catastrophic failure, help comply with environmental regulations, and enhance customer satisfaction. There exist several methods to detect and isolate incipient faults that might cause a plant's performance to deviate from the nominal. Fault detection and diagnostic methods may utilize subjective or objective knowledge as the basis for the diagnostic scheme.

A methodology for the integration of subjective (heuristic) and objective (analytical) knowledge for fault diagnosis and decision-making is proposed in this dissertation. The structure, challenges, and benefits of such integration are explored. When a fault occurs in a dynamic system, the observed outputs deviate from their nominal value thus affecting the performance of the system. These observed *symptoms* are measured using a wide range of sensors and typically these measurements are sampled over different time scales and knowledge from various sources is utilized. The fusion of such sensor measurements, the combination of different time scale data, and the integration of the knowledge sources is a serious challenge that must be overcome by the integration scheme. In this methodology two parallel paths are followed, making use of the system's

input-output data. Along the first path a model-based scheme is applied, where primary residuals are generated using sliding mode nonlinear observers. These residuals are evaluated and processed using fuzzy logic for fault detection and isolation. Meanwhile along the second path a knowledge-based scheme is applied using a combination of fuzzy estimation techniques and fuzzy heuristic rules for fault detection and isolation. The two paths merge together and integration of both schemes is completed in the inference mechanism. Finally a fuzzy decision-making process yields the diagnosis. The philosophy and structure of the diagnostic scheme is explained.

The experimental verification of this methodology is then carried out on a Ford 4.6 L V8 engine controlled by an electric dynamometer housed in the Ohio State University Mechanical Engineering Department and Center for Automotive Research. A hierarchical diagnostic method that explores the functional structure of the automotive engine subsystems and components is applied. The application and results in the context of automotive engine exhaust emission controls system diagnostics are demonstrated. Fault detection and isolation of induced faults was completed successfully.

Dedicated to my small and big family



## ACKNOWLEDGMENTS

No words can express my thanks to my advisor, Professor Rizzoni, for his valuable advise, continuous guidance, and stimulating intellectual support, he is a scholar who pursues theoretical research without losing sight of its practical application.

I am grateful to all my colleagues in the Powertrain Control and Diagnostics (PCAD) Laboratory for their stimulating discussions and help in carrying my experimental work, especially Mr. Yong-Wha Kim. Also, many thanks go to Mr. José Candau of the University of Valladolid in Spain for his help with the fuzzy evaluation algorithms. The PCAD laboratory setup would not have been possible without the technical contribution of Mr. Joseph West and Mr. Michael Bowlus. Their expertise and help with the hardware was beyond the call of duty.

The support from the Center of Industrial Sensors and Measurements (CISM) and its director, Professor Akbar, is greatly appreciated. Also, I would like to acknowledge the assistance of Professor Yurkovich and Professor Passino in understanding fuzzy logic and its application.

Last but not least I thank my wife Maria Soliman for her tolerance, support, and understanding during my roller coaster period. Also, I am in deep gratitude to my mother and father who never lost faith in my capabilities and for their unlimited support. To all my friends who are seeking precise answers for fuzzy problems, thank you.

## VITA

April 4, 1957..... Born- Cairo, Egypt

1979..... B.S.M.E, Ain Shams  
University,  
Cairo, Egypt

1986..... M.S.M.E, Ain Shams  
University,  
Cairo, Egypt

1990-present..... Graduate Research  
Associate  
The Ohio State University,  
Columbus, Ohio

## FIELDS OF STUDY

Major Field: Mechanical Engineering

## TABLE OF CONTENTS

	<u>Page</u>
Abstract .....	ii
Dedication .....	iv
Acknowledgments .....	v
Vita .....	vi
List of Tables .....	x
List of Figures .....	xi
Chapters	
1. INTRODUCTION .....	1
1.1. LITERATURE REVIEW .....	6
1.1.1. Federal Regulations and the Automotive Industry .....	7
1.1.2. Knowledge-Based Diagnosis .....	12
1.1.3. Pattern Recognition Approaches .....	16
1.1.4. Model-Based Diagnosis .....	18
1.1.5. Design of a Hybrid Diagnostic System .....	27
1.2. RESEARCH OBJECTIVES .....	31
1.3. CONTRIBUTIONS OF THIS WORK .....	32
1.4. SUMMARY .....	33

2. BACKGROUND.....	34
2.1. INTERNAL COMBUSTION ENGINE EMISSIONS .....	34
2.2. EMISSION CONTROLS.....	39
2.3. DIAGNOSTIC REGULATIONS .....	43
2.4. DIAGNOSTIC METHODS .....	48
2.4.1. Dynamic Model-Based Methods.....	49
2.4.2. Knowledge-Based Methods .....	70
2.5. HIERARCHICAL STRUCTURE.....	74
2.6. SUMMARY .....	76
3. EXPERIMENTAL SETUP .....	78
3.1. EXPERIMENTAL FACILITIES.....	78
3.1.1. Engine Test Cell Instrumentation.....	79
3.2. DESIGN OF EXPERIMENTS.....	82
3.2.1. Sample Size.....	83
3.2.2. Test Repeatability.....	86
3.2.3. Test Cycle.....	87
3.2.4. Induced Faults .....	89
3.2.5. Experimental Procedures.....	90
3.3. BASELINE EXPERIMENTAL DATA.....	91
3.4. SUMMARY .....	94
4. DIAGNOSTIC METHODOLOGY .....	96
4.1. FAULT DETECTION AND ISOLATION (FDI).....	97
4.1.1. Model-Based FDI:.....	103
4.2. RESIDUAL AND SYMPTOM EVALUATION.....	109
4.2.1. Knowledge-Based FDI.....	117
4.3. FUZZY ESTIMATION AND IDENTIFICATION .....	122
4.3.1. Fuzzy C-Means Clustering.....	123
4.4. DIAGNOSTIC INTEGRATION .....	126
4.5. FUZZY DECISION-MAKING.....	127
4.6. SUMMARY .....	129
5. APPLICATION AND RESULTS.....	131
5.1. ENGINE MODEL.....	132
5.1.1. Nonlinear Observers.....	135
5.2. FUZZY EVALUATION .....	141
5.3. FUZZY ESTIMATION.....	147
5.4. RULE INTEGRATION .....	168
5.5. FUZZY DECISION-MAKING.....	170

5.6. SUMMARY .....	172
6. CONCLUSION AND FUTURE WORK.....	174
6.1. CONCLUSION .....	174
6.2. FUTURE WORK .....	175
BIBLIOGRAPHY .....	176
APPENDICES.....	188
A. FUZZY SYSTEMS .....	189
B. ENGINE SENSORS.....	196
C. ENGINE MODEL .....	201

## LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 2.1 OBD-II Monitored Systems .....	45
Table 2.2 California Tailpipe Emission Standards (grams/mile) .....	47
Table 2.3 Generic Categories of Knowledge-Based Methods Applications .....	71
Table 3.1 Engine Specifications .....	80
Table 3.2 Induced Faults .....	89
Table 4.1 System Variables and Faults .....	103
Table 4.2 Fault Isolation Logic .....	112
Table 4.3 Fuzzy Evaluation and Isolation Logic .....	114
Table 4.4 Fuzzified Symptoms .....	116
Table 4.5 Fuzzy Decision-Making Table .....	128
Table 5.1 List of Observed Quantities .....	137
Table 5.2 Fault Signature Table .....	142
Table 5.3 Fuzzy Isolation Decision Table .....	146
Table 5.4 Fuzzy Singleton Output for Fault Isolation .....	164
Table C.1 List of Symbols for Engine Model .....	206
Table C.2 List of Engine Model Parameters .....	207

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1.1 General Structure of the Experimental Knowledge-base.....	15
Figure 1.2 The Integrated Diagnostic Model.....	15
Figure 1.3 A Dynamic Model-based Diagnostic System .....	20
Figure 1.4 Relationship between Combustion Pressure and Crankshaft Speed.....	21
Figure 1.5 A Perturbation Model of Throttle-to-Speed Dynamics.....	25
Figure 1.6 A Model for Fueling Dynamics of Spark Ignition Engine.....	25
Figure 1.7 Engine Block Diagram.....	27
Figure 1.8 Hybrid FDI System .....	30
Figure 2.1 Emission Levels vs A/F Ratio.....	36
Figure 2.2 Catalytic Converter Efficiency .....	40
Figure 2.3 Engine Emission Control System .....	42
Figure 2.4 Block Diagram of a Model-based Residual Generation Scheme.....	49
Figure 2.5 General System and Fault Representation .....	53
Figure 2.6 Fault Model.....	54
Figure 2.7 Computational Form of the RG .....	57
Figure 2.8 Parity Space Directionality .....	58

Figure 2.9 Observer Structure .....	61
Figure 2.10 Multiple Observer Scheme .....	64
Figure 2.11 Conventional State Observer vs. Unknown Input Observer .....	65
Figure 2.12 Hierarchical System Structure .....	75
Figure 3.1 Experimental Setup.....	82
Figure 3.2 HC Concentration Variability .....	85
Figure 3.3 NO Concentration Variability .....	85
Figure 3.4 CO Concentration Variability .....	86
Figure 3.5 Test Cycle Dynamometer Speed .....	88
Figure 3.6 Test Cycle Dynamometer Torque .....	88
Figure 3.7 TPS Sensor Output.....	91
Figure 3.8 MAF Sensor Output.....	92
Figure 3.9 UEGO Sensor Output .....	92
Figure 3.10 Calculated Fuel Mass.....	93
Figure 3.11 HEGO Sensor Output .....	93
Figure 3.12 Engine Speed .....	94
Figure 4.1 Architecture of a Model- and Knowledge-Based.....	98
Figure 4.2 General Structure of a Diagnostic System .....	99
Figure 4.3 Diagnostic Scheme.....	102
Figure 4.4 General Dynamic System and Fault Representation.....	103
Figure 4.5 Model-Based FDI.....	105
Figure 4.6 Nonlinear Observer Residual Generator .....	107



Figure 4.7 Residual Evaluation Scheme.....	110
Figure 4.8 Residual Fuzzy Evaluation .....	111
Figure 4.9 Fuzzy Evaluation Membership Function .....	113
Figure 4.10 Implied Fuzzy Set for Normal or High Residuals.....	115
Figure 5.1 Dynamic Engine Model .....	132
Figure 5.2 Low Frequency Dynamic Engine Model .....	134
Figure 5.3 Air/Fuel Ratio Subsystem Model.....	135
Figure 5.4 Observer Configuration .....	138
Figure 5.5 Nonlinear Observer Performance.....	140
Figure 5.6 Fuzzy Residual Evaluation Membership Functions.....	142
Figure 5.7 MAF Calibration Fault.....	143
Figure 5.8 MAF Residual for MAF Fault .....	143
Figure 5.9 MAP Residual for MAF Fault .....	143
Figure 5.10 TPS Residual for MAF Fault .....	143
Figure 5.11 TPS Calibration Fault .....	144
Figure 5.12 MAF Residual for TPS Fault .....	144
Figure 5.13 MAP Residual for TPS Fault .....	144
Figure 5.14 TPS Residual for TPS Fault.....	144
Figure 5.15 MAF Fault Isolation.....	146
Figure 5.16 TPS Fault Isolation .....	146
Figure 5.17 HC Membership Functions .....	149
Figure 5.18 NO Membership Functions.....	149

Figure 5.19 CO Membership Functions .....	149
Figure 5.20 $\lambda$ Membership Function .....	149
Figure 5.21 Heuristic Fuzzy System for $\lambda$ Estimation .....	150
Figure 5.22 Heuristic Fuzzy Estimation of Air/Fuel Ratio .....	152
Figure 5.23 Fuzzy C-Means Estimation of Air/Fuel Ratio .....	152
Figure 5.24 Estimated Air/Fuel Ratio Variability .....	153
Figure 5.25 Estimated Air/Fuel Ratio for MAF Fault.....	153
Figure 5.26 Estimated Air/Fuel Ratio for TPS Fault.....	153
Figure 5.27 UEGO vs Fuzzy Estimator.....	154
Figure 5.28 UEGO vs Fuzzy Estimator.....	154
Figure 5.29 UEGO vs Fuzzy Estimator.....	154
Figure 5.30 UEGO vs Fuzzy Estimator.....	154
Figure 5.31 EGR Stuck Open Fault Species' Concentration Level .....	156
Figure 5.32 EGR Stuck Closed Fault Species' Concentration Level.....	157
Figure 5.33 Injector Stuck Closed Fault Species' Concentration Level .....	158
Figure 5.34 Misfire Fault Species' Concentration Level .....	159
Figure 5.35 MAF Fault Species' Concentration Level .....	160
Figure 5.36 TPS Fault Species' Concentration Level .....	161
Figure 5.37 Intake Leak Fault Species' Concentration Level .....	162
Figure 5.38 Fuzzy Output for MAF Fault Isolation .....	164
Figure 5.39 Fuzzy Output for TPS Fault Isolation .....	165
Figure 5.40 Fuzzy Output for Intake Leak Fault Isolation .....	165

Figure 5.41 Fuzzy Output for Injector Fault Isolation.....	166
Figure 5.42 Fuzzy Output for EGR Open Fault Isolation .....	166
Figure 5.43 Fuzzy Output for EGR Closed Fault Isolation.....	167
Figure 5.44 MAF Fault Decision .....	171
Figure 5.45 TPS Fault Decision .....	171
Figure A.1 Fuzzy System .....	190
Figure A.2 Some Typical Membership functions .....	192
Figure B.1 Mass Air Flow Sensor (MAF).....	197
Figure B.2 Throttle Position Sensor (TPS) .....	197
Figure B.3 Engine Coolant/Air Temperature Sensor .....	198
Figure B.4 Fuel Injector Pulse Width.....	198
Figure B.5 Idle Air By Pass Sensor.....	199
Figure B.6 Fault Induction Circuit .....	199
Figure B.7 Misfire Circuit.....	200

## CHAPTER 1

### INTRODUCTION

In the early 1950's smoggy skies over Los Angeles, caused by the exhaust emissions created by the intense traffic [1]<sup>1</sup> persuaded the federal government to set standards to lower the levels of unburned hydrocarbons (HC), nitrogen oxides (NO<sub>x</sub>) and carbon monoxide (CO) emitted by mobile sources.

To implement the air quality standards, the California Motor Vehicle Pollution Control Board (CMVPCB) was created in 1960. Also, California and the federal government used a driving cycle to certify 1966 vehicles and newer models, which was referred to as either the California cycle or the Federal Test Procedure (FTP).

The first major Clean Air Act, which established the Environmental Protection Agency (EPA), was adopted by the Congress in 1970. The new agency had the responsibility of regulating motor vehicle pollution, and also identified Inspection and Maintenance (I/M) programs as an alternative for improving the air quality. Calculated

---

<sup>1</sup> Numbers between brackets are references listed in the Bibliography.

mass emissions, rather than just concentration, formed the basis of the 1970-1971 FTP; where exhaust emissions over a three phase driving cycle are collected in three sample bags in accordance with the different test phases. The constant volume sampling (CVS) method, was developed in 1972 as a true mass emission concentration measurement which employed variable dilution sampling, and also different gas sampling bags were used.

The above mentioned regulations led to the appearance of the charcoal canister, exhaust gas recirculation (EGR) valves and, finally, the catalytic converter in 1975. Moreover, in 1977 amendments of the Clean Air mandated I/M for high-pollution areas. The EPA in 1978 issued its first policy for I/M programs. Furthermore, as the standards grew stricter, 3-way catalysts, on-board computers, and oxygen sensors appeared in 1981. Periodic testing for malfunctioning emission control systems was adopted by 64 cities nationwide by establishing I/M programs in 1983. These programs consist of a start idle-mode, two-speed or loaded steady-state test that measures the tailpipe exhaust gas concentrations with repair-grade analyzers, of lower quality and cost than laboratory equipment.

The growing use of electronic dash panel displays in passenger cars and the increasing use of sophisticated emission and powertrain control systems, led California to adopt the On-Board Diagnostic (OBD) regulations in 1985 for 1988 and later model-year light- and medium-duty vehicles. OBD systems were intended to reduce the time between occurrence of a malfunction and its detection and repair. Their objective is to reduce emissions caused by malfunctions, and also to minimize the damage that could occur to

other vehicle components or systems. This is achieved by incorporating additional software and hardware to collect and analyze data already available to the on-board computer, thus monitoring the entire emission control system.

In 1989, the California Code of Regulations (CCR) known as OBD-II was adopted by the California Air Resources Board (CARB). This code requires the automotive manufacturers to implement new comprehensive on-board diagnostic systems replacing the original OBD (or OBD-I) starting with 1994 model year. OBD-II monitors more components and systems than OBD-I, including the catalyst, the evaporative control system and every emission control system and electronic emission related powertrain component [2]. OBD-II seeks to monitor most emission control system components, such that a malfunction is signaled as the exhaust emissions level exceed 1.5 times the applicable standards. It is also designed to detect engine misfire, which could cause serious damage to the catalytic converter. The vehicle operator is notified at the time when the vehicle begins marginally to exceed emission standards. OBD-II is also intended to provide additional information to technicians for diagnosis and repair of emission related problems. A new vehicle communication system is used to provide the service technicians with detailed information about system performance and detected malfunctions; this is made achievable by storing fault codes that lead technicians to the likely area of the malfunction, and by continuously updating the information for some engine data to help the technician isolate the specific fault. Further, OBD-II includes a "freeze frame" feature, which allows the computer to store in its memory the exact operating conditions when a fault occurred, so intermittent faults can be investigated by

revisiting the same conditions when the problem occurred. A standard access electrical connector which is identical for all vehicle makes is required, which means a single inexpensive generic tool can be used to read out trouble codes.

Although OBD-II requirements reflect state-of-the-art diagnostic system capability, there are limitations which apply to the current techniques for detecting malfunctioning components, as on-board diagnostics do not provide sufficient scope and accuracy [3]. These limitations do not allow OBD-II systems to take the place of the FTP test for measuring vehicle emissions. The reason is that such monitoring systems can detect when components are functioning outside of their operating range, but are limited in their ability to determine whether they are functioning accurately within the range. It has been shown that emissions can exceed the applicable standards if the components are operating improperly within their normal range and before they can be detected to be malfunctioning.

Thus, the EPA is working on implementing new I/M programs. As an example, the test known as the I/M-240, may be mandated for high-pollution areas by amendments to the Clean Air Act. The I/M-240 is analogous to the FTP for emissions certification; I/M-240 includes cycles of acceleration and deceleration under loaded conditions and determines grams per mile for HC, CO, CO<sub>2</sub>, and NO<sub>x</sub>. The urban portion of the FTP is about 30 minutes long, and the I/M-240 covers just 4 minutes; however, much of the normal operating range of a vehicle is included. A cycle of this type will soon be routine during vehicle registration and may involve special "lanes" to efficiently test a large

number of vehicles. These inspection programs are designed to test vehicles on a chassis dynamometer while collecting CVS diluted tailpipe samples.

One of the gray areas in the implementation of regulations limiting the generation of pollutants from mobile sources is the actual effectiveness of the exhaust gas emissions control strategy in vehicles that have been in use for some time. While it is possible today to conduct limited diagnostics with the on-board engine computer by performing periodic checks to verify the validity of the signals measured by the on-board sensors, and to measure tailpipe emissions during routine inspection and maintenance, the task of correlating these measurements with each other to provide an on-line, accurate diagnosis of critical malfunctions has thus far proven to be a very challenging task.

A typical automotive emission control system relies on the measurements provided by various sensors, including for example air charge temperature (ACT), intake manifold absolute pressure (MAP), exhaust gas oxygen (EGO) concentration, and engine speed, to adjust the air and fuel mixture as closely as possible to the stoichiometric air-to-fuel ratio (AFR). Maintaining the AFR close to stoichiometry allows the catalytic converter to operate in a high conversion efficiency region, and to significantly reduce the emission of the three major pollutants (for gasoline engines), that is, CO, HC, and NO<sub>x</sub>. Improper operation of any of the sensors, as well as actual component malfunctions (e.g., a fouled spark plug or fuel injector, or a vacuum or compression leak), may sufficiently upset the delicate balance sought by the electronic controller to cause the mixture to significantly deviate from the desired set point. Such a condition results not only in increased



emissions, but possibly in further damage to components (e.g., the catalytic converter), with significant cost to the consumer or to the manufacturer.

It is clearly important to diagnose malfunctions in the emission control system at the incipient stage. Since in practice the first visible *symptom* of a problem with the emission control system may be the measurement of excessive tailpipe emissions during an inspection and maintenance test, it is very important that the exhaust gas measurements be somehow correlated with the health of the emission control system. This is a particularly challenging problem, because the chemistry of exhaust gas production is very complex, and not easily modeled by conventional dynamic system modeling methods. One approach to understanding the correlation between emission control system malfunctions and measured tailpipe emissions is to generate an extensive data base, and to analyze the data to uncover characteristic patterns that may assist in the diagnosis of specific faults. Fuzzy identification and estimation using the data from the exhaust emissions and sensor data is one of the methods chosen in the present study to provide useful fusion of such diverse measurements.

## **1.1. LITERATURE REVIEW**

In this section, a brief review of recent developments in automotive diagnostics is presented. The section begins with an overview of recent products in the industry then goes further to categorize the different approaches that have been used to diagnose automotive faults. Each category is covered briefly, and finally the use of a hybrid approach is emphasized.

### **1.1.1. Federal Regulations and the Automotive Industry**

Energy and air pollution are two key areas in which the automotive industry has had a major impact on society. To comply with the environmental requirements of better fuel economy and lower emissions level, motor vehicle manufacturers had to re-examine the inter-relationships between fuel economy and emissions involving engine design, operating parameters, fuels, lubricants, and vehicle configuration [4]. The significant dependence on foreign petroleum sources, coupled with the relatively large consumption of petroleum by the transportation sector, has resulted in *Federal Regulations* of the automotive fuel economy and emissions. Moreover the USA EPA, through the *Clean Air Act* , and the State of California Air Resources Board have set stricter emission levels for the protection of the environment, leading to the increasing use of sophisticated electronic emission and powertrain control systems. One of the less desirable consequences of these changes is that the extensive use of electronics has made it difficult for the technicians to detect defective components through the use of visual inspection, and to perform routine diagnostic tests with traditional test equipment. Thus, a new generation of diagnostic equipment and procedures has evolved over the last two decades, leading to two generations of on-board diagnostic requirements (OBD I and OBD II), increasing the number of components and systems to be monitored by the diagnostic tools. Although on-board diagnostics provide the service technicians with detailed information about system performance and detected malfunctions, their performance range is limited, and off-board diagnostic equipment has not yet been and probably will not be completely replaced by on-board equipment.

A number of on- and off-board diagnostic systems have been developed to keep up with continuous advances in automotive electronic and control systems; in spite of the effort put forth by original equipment manufacturers (OEM) and suppliers, the performance of available diagnostic and test equipment is still amenable to further improvement, especially as it pertains to the on-board diagnosis of incipient and intermittent faults.

Different approaches have been used to diagnose automotive faults; these approaches fall into four categories, based on the strategy for diagnosis:

1. Knowledge-based methods including rule-based and expert system.
2. Pattern recognition techniques, including neural networks and vibration analysis.
3. Dynamic model-based methods .
4. Hybrid system as a combination of any of the above.

The development of new automotive test equipment in recent years has certainly been influenced by factors external to the industry, in addition to the natural evolution of technology [5]. A major impact has come from exhaust emissions control regulations, which have forced the re-design of existing equipment and the development of new equipment (e.g., exhaust gas analyzers), as service by substitution is no longer an acceptable solution [6]. As an example, the Sun Test-Link is an application developed by Sun to service automotive on-board computer systems and works in conjunction with a conventional automotive engine analyzer [5]. It has the ability to collect, store, and analyze a variety of signals through external data ports.

Automotive electronic diagnostic equipment should provide a flexible, cost-effective, and reliable tool which eliminates the requirement that the technician be intimately familiar with every function of every system to be serviced [7]. A compromise must be made between the tester cost and the completeness of its tests. A complete tester should have provisions for making the necessary measurements to isolate a fault after it has been identified on the vehicle.

On-board vehicular diagnostic systems exploit computer-based engine control systems, and work with the assistance of electronic dash panel displays in passenger cars. The choice between on-board and off-board will depend on system design constraints, corporate repair objectives, computer memory and hardware costs, and dealership/manufacturer willingness to invest in long term equipment needs [8].

A natural candidate for more specialized diagnostics is the exhaust gas emissions control system. A predecessor of today's advanced diagnostic systems was introduced in 1983 [9], with the aim of diagnosing this particular subsystem. The diagnostic instrument described in [9] consisted of 4-bit microcomputer which was able to monitor the essential control functions. Four years later (1987), Ford Motor Co. released the Electronic Engine Control (EEC-IV) monitor and an eight-channel recorder box which not only allowed the technician to observe and capture the microprocessor signals, but also provided acceptable values and limits for each type of application being tested [10].

With the advent of compact diagnostic microcomputers, two avenues of service assist became available: on- and off-board diagnostic systems. Off-board diagnostics provide the maximum flexibility, as changes can be introduced at any time [11], and

specific test sequence can be prescribed. On the other hand, on-board diagnostics can capture "difficult" faults, such as those that occur intermittently, or during special operating conditions. General Motors (GM), for example, has proposed that a mix between on- and off-board diagnostics is the best approach to service vehicles with the best outcome [11]. The GM service research center developed TECH 1 in 1986, which is a low cost, hand-held interactive diagnostic tester [11]. This tester is a compact, microprocessor based unit designed using a Motorola CMOS 6303 microprocessor. Different diagnostic cartridges were developed to update the tester. A set of specialized diagnostics could be performed using a programmed Functional Test Director (FTD), while up to twenty tests could be performed including: oxygen sensor, coolant sensor, thermostat, etc. This interactive instrument conducts tests on both the electronic system key components and the engine elements that are hard to diagnose. TECH 1 showed that the combination of both on- and off-board diagnostics could provide a powerful tool to increase the accuracy and speed of automotive repairs.

Intermittent faults are the most time consuming faults leading to customer dissatisfaction and distrust, and ultimately to customer loss. An approach to overcome this difficulty is to employ a "Blackbox", similar in concept to those commonly used in aircraft, which is used to store data and to capture control system inputs and outputs while the fault exists. This approach was exploited by General Motors with the Vehicle Service Monitor (VSM), a portable microcomputer capable of storing 32K bytes of data and of updating the information by pushing off the stack of the oldest information, thus easily catching intermittent shorts and opens. A similar concept was proposed in [12]: a

"Blackbox" was equipped with eight analog and eight digital inputs, plus a reference input (ignition in cylinder one). The memory used in this device was capable of storing 10 seconds of engine data, and was kept up to date by overwriting existing data.

In 1985 the On-Line Automated Service Information System (OASIS), was introduced by Ford to enhance the technician's ability to solve vehicle problems and to minimize the amount of information-sorting and base vehicle measurement required [13]. In the following year 1986 GM introduced the Computerized Automotive Maintenance System (CAMS), an automated service tool which provides the service technician with automated testing, isolation of failures, and also presents repair procedures, allows for automatic re-testing and documents the repair actions [14]. The GM-CAMS terminal is an IBM PC-AT computer with a touch screen allowing communication with the technician and two probe systems connected to the vehicle. One probe is connected to the serial bus that provides diagnostic information to the service technician through the flashing check engine light or with the use of scanners, this connection is to determine the state of the engine control unit and to read data that is present in the serial data stream. The second probe is used to isolate the fault and to determine the faulty part that needs repair, it also tests the wiring integrity and the engine control module performance. Both systems can receive service information and warranty data.

The development of such diagnostic systems is continuing in the 1990's, and this was a brief review. There has also been a major effort to standardize diagnostic practices in the industry by establishing standard fault codes and transmission procedures [15]. Further, recent legislative activities in California and in the Northeast of the United States

have significantly raised the importance of on-board diagnostics. These continuing developments will serve as further stimulus to an already exciting sector of the automotive industry.

### **1.1.2. Knowledge-Based Diagnosis**

Many of the past applications involving diagnosis have been rule based. That is, simple production rules are used to provide a mapping between the possible causes and the inputs of the system and the possible faults. Survey of various knowledge based diagnosis strategies can be found in [16] and [17]. New, powerful techniques based on Artificial Intelligence (AI) concepts give us the ability to build and reason about deep models and involve a wide range of information, such as learning from experience, probabilistic information, and learning from examples. Different diagnostic systems can be developed by focusing on different levels of knowledge representation. The set of malfunctions and the relations between the observations and the malfunctions are the basic knowledge required for diagnosis.

One of the methods used for diagnosis is an expert system. An expert system is defined as a computer program which models the decision-making process of a human expert [18]. Expert systems are different from conventional computer programs as the knowledge base is a separate entity of the program and can be revised and expanded as new knowledge is acquired. Also, expert systems allow uncertainty to be considered in the modeling of expert behavior, and can explain the line of reasoning leading to recommended actions [18]. Vehicle diagnosis is a promising application of expert systems if these can be combined with some quantitative techniques. Repairs could be

significantly improved by using expert systems which combine the expertise of several different mechanics. Expert systems are best at solving problems which require an understanding of how the components relate to each other. Expert systems can also be consistent, and may provide insights that the expert might overlook [19]. However, a great deal of knowledge must be assembled to build an expert system that approaches the performance of an expert mechanic. The principal drawbacks of expert systems as summarized in [19]:

*A common misconception that people have about artificial intelligence and expert systems, in particular, is that these somehow provide magical approaches to solving complex problems. What is offered by researchers from this field is a set of tools that can aid in the solution of some problems. The construction of an expert system involves a time consuming interaction between the expert and a knowledge engineer.*

An example of a diagnostic expert system was developed by [20], with their Integrated Diagnostic Model (IDM). The IDM is implemented in Lisp and runs on a symbolic 3600 Lisp machine. It uses multiple knowledge bases, of deep and shallow knowledge, to implement rules and represent knowledge in an appropriate fashion. Two types of knowledge, "experimental" and "functional", are represented. Meanwhile, a three level semantic network, shown in Figure 1.1, is used as a basis for the experimental knowledge, that models a human expert diagnostic and repair performance. These levels are:



1. Information level.
2. Hypothesis level.
3. Solutions level.

The information is examined, hypotheses are proposed, and then they are verified or rejected on the basis of more information until one is verified and a solution is suggested. The cost of verification, cost of repair, and the likelihood of failure are the criteria used to rate the information and the hypothesis. The device under diagnosis is represented by a qualitative model based on its functional knowledge, which provides a hierarchical representation of the system. The diagnosis begins by pinpointing the malfunctioning subsystem and moving further to the components within the subsystem. Figure 1.2. shows the overall expert system, where the two knowledge sources and their inference engines are integrated and controlled by the executor to create a unified problem solving system. The executor determines which part of the expert systems, the experimental inference engine, or the functional inference engine will handle the problem and provides for communication between the two systems while it manages the user interface. This concept was tested in [20] on the diagnosis of an automotive electrical system.

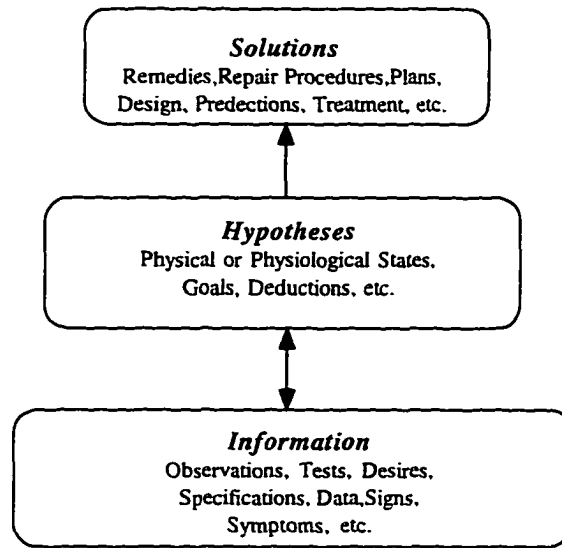


Figure 1.1 General Structure of the Experimental Knowledge-base [20]

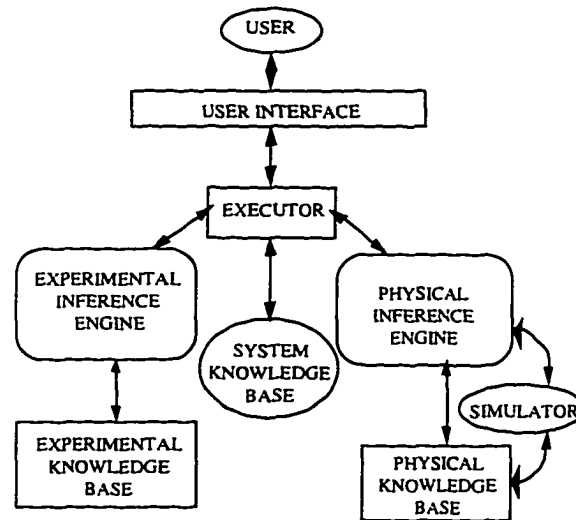


Figure 1.2 The Integrated Diagnostic Model [20]

### 1.1.3. Pattern Recognition Approaches

A significantly different approach to the diagnosis of faults in engineering systems is the use of *pattern recognition* and *signature analysis methods*. This field has been well studied, and a rather comprehensive review of various techniques is presented in [34]. An example of pattern recognition using neural networks is adopted in [35]. The objective of this study was to exploit the natural classification properties of neural network to diagnose faults during an end-of-assembly cold test, with a principal focus on mechanical faults. Statistical classifier systems and neural networks were used together to obtain optimum performance for the diagnostic system. The engine used in the study was a 4.0 liter 6 cylinders, motored at about 150 rev/min by an electric motor with an in-line torque transducer to measure the dynamic crankshaft torque. Pressure transducers monitored both the intake and exhaust manifold pressures, the crankcase air pressure, and the oil pressure, with an actual data acquisition time of less than 1 second. A conventional backpropagation (BP) neural network with 350 input nodes, 50 hidden nodes and 29 output nodes was first used. Dimensionality reduction was achieved using Principal Component Analysis (PCA), which consists of the projection of vectors of input samples onto a new set of orthogonal axes, chosen to represent the largest variance in the sample of the data represented. With the reduction provided by PCA the network training time was reduced by a factor of 100. Actual data from over 1000 different pre-production engines were used as a case study to train a neural network. Technician analysis of the data was compared with the results produced by the network, with good agreement in most of the cases.

A different type of signature analysis, based on vibration signals was used in [36], in which study of vibration analysis of a 350 horsepower diesel truck engine was performed as a means of detecting faults in the individual cylinders and in the crankshaft bearings. Two sources of engine vibration were considered, the cylinder pressure and the main bearing forces. Cepstral windowing of a single channel signal obtained by measuring the vibration of a head bolt vibration signal permitted reconstruction of the cylinder pressure. To estimate the cylinder pressure waveform from the block acceleration signals a single-input three-output linear inverse filter was used, while the method of conditioning partially coherent sources was applied to the bearing force prediction. This was achieved by subtracting the total cylinder pressure contribution from the block acceleration, and then inverse filtering. A multi-channel monitoring system was also used to obtain signals simultaneously from an array of transducers on the engine.

A similar approach is used by R.H. Lyon Co. (Cambridge, Mass.), as reported in [37], in the development of a system called the Diesel Engine Unit Condition Evaluator (DEUCE), which will be incorporated in an expert diagnostic system. The basic concept behind the DEUCE system is that different components emit their own defined signature during normal operation, so that by comparing these baseline signatures with the sampled signals, a deviation exceeding a predetermined limit will indicate a change in the performance of the measured component. Accelerometers and photo diodes are used in the DEUCE system to acquire the signals for processing, while cepstral windowing is used to separate overlapping signals in conjunction with inverse filtering. Development of an expert diagnostic system using the DEUCE processing techniques is taking place at

General Electric's Transportation Systems division, incorporating temperature, speed and pressure sensing along with vibration processing to troubleshoot diesel engines used in locomotives. Faults which result in a change in the combustion, such as faulty injectors and leaks in the intake manifold, are introduced in the laboratory, then the DEUCE system is used to monitor the combustion anomalies and to indicate the source of the fault. This can lead to the early detection of the faults that affect the combustion and to avoid serious damage, thus improving reliability of the engine. Other references can be found in [38].

A characteristic of pattern recognition and signature analysis methods is that in order for reliable signatures to be established, faults must be injected into the system; although this is usually possible for many mechanical components, some of the more subtle malfunctions that can occur in electronic control systems may not be easily predicted or simulated in the physical system. A different class of diagnostic methods that can take such incipient or "soft" faults into account is described in the next section.

#### **1.1.4. Model-Based Diagnosis**

In spite of the fact that expert systems can diagnose malfunctions, they are not well suited to detect "soft calibration" changes in sensors and actuators [39]. Also, the size of the expert system increases as the vehicle and its subsystems under diagnosis get more complicated. On the other hand, methods based on mathematical models of the engine dynamic behavior may be capable of detecting very small calibration changes in any sensor or actuator, as well as locating totally failed components.

It is important at this point to observe that two kinds of models can be distinguished:

1. Models where an underlying mathematical description is used in the form of an algorithm.
2. Models where a representation of the causal sequences by means of which the agent understands the device's function is used for simulation [16].

In this section we will be primarily concerned with the first class of models, leading to *algorithmic model-based diagnostic methods*, or Model Based Diagnostic Systems (MBDS).

Algorithmic approaches to data interpretation for diagnostic purposes are relatively well studied. A general functional description of the approach is shown in Figure 1.3. An algorithmic model based FDI system utilizes known (or approximate) mathematical relationships between the inputs and outputs of a given system. Using this analytical knowledge (i.e., an objective model of the system), the FDI system anticipates the system output, given the inputs. Any fault in the system will generate a discrepancy between the actual system outputs and those predicted by the MBDS. This difference between actual and predicted system performance forms the basis of this diagnostic method. Normally, it is assumed that the underlying process can be modeled via differential equations (or difference equations), with input vector  $u(t)$ , output vector  $y(t)$ , and state vector  $x(t)$ . Based only on the inputs and outputs that can be sensed (in general, the state  $x(t)$  is not measurable), an algorithm must be developed that will characterize the performance of

the system and provide diagnostic information. The interested reader will find excellent surveys of these methods in [40]- [44].

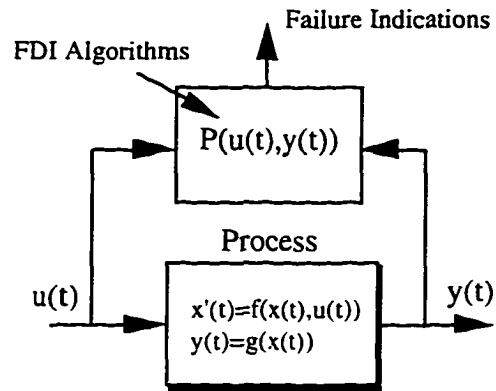


Figure 1.3 A Dynamic Model-based Diagnostic System

As reported in [47], the techniques just described can be effective in detecting and isolating engine faults, provided that a suitable model has been validated. The starting point in the development of a model appropriate for an FDI strategy is the choice of a suitable model. Further, it is necessary to carry out appropriate model reduction in order to produce algorithms which may be implemented in real time, either in a microprocessor or in a custom integrated circuit. In controlled systems, this design cycle cannot be separate from the formulation of the control strategy, as the control algorithms affect the dynamic behavior of the plant, and many sensors can be shared by the controller and monitoring system. Models are a key to the successful application of these methods. In

the case of engine diagnostics, for example, it is possible to capture the essential diagnostic information in an IC engine by considering three subsystem models, described below.

A model relating combustion pressure to crankshaft speed fluctuations [49] is depicted in Figure 1.4 in block diagram form. The model describes the relationship between the combustion pressure in each cylinder,  $P_i$ , the net torque applied at the crank throw by each cylinder,  $T_e$ , and the instantaneous fluctuations of the speed of the crankshaft,  $\omega$ . The model permits the estimation of instantaneous torque for each cylinder by means of a relatively inexpensive measurement of crankshaft position, and has been validated in a number of engines in the laboratory and in-vehicle [49] and [50]. Road tests have also validated the usefulness of this model for the purpose of diagnosing engine misfire [50].

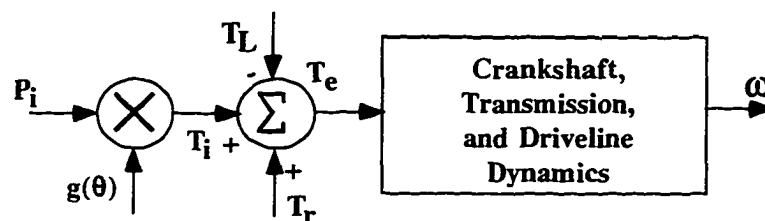


Figure 1.4 Relationship between Combustion Pressure and Crankshaft Speed



The following variables are defined with reference to Figure 1.4

$P_i$  = Indicated cylinder pressure

$g(\theta)$  = Geometric function for each stroke

$T_L$  = Load torque

$T_i$  = Indicated torque

$T_r$  = Reciprocating inertia torque

$T_e$  = Net torque

$\omega$  = Crankshaft angular speed

A second subsystem model that has found diagnostic applications consists of a perturbation model of the throttle-to-crankshaft speed dynamics of a spark ignition engine, shown in Figure 1.5. This model has been exploited to provide partial sensor fault detection capabilities in electronically controlled engines [49]. In spite of its relative simplicity it has proven useful in diagnosing malfunctions in the throttle position, mass air flow (or manifold pressure), and crankshaft speed sensors. The fault detection algorithms based on this model have been tested on several production engines, both in-vehicle and in the laboratory [49]. This model has the inherent advantage of being expressed in terms of a number of measured variables: throttle position, mass air flow or intake manifold absolute pressure, engine speed, spark angle and fuel quantity are known, either as measured or control variables in the engine control strategy. The model represents the throttle-to-crankshaft speed dynamics of the engine and can be reduced to a relatively simple linearized perturbation form. However, the range of validity of the perturbation model is rather restricted, unless some of the critical nonlinearities are taken

into account. Among these, the inherent nonlinearity of the intake manifold flow dynamics, and the dependence of the engine volumetric efficiency on engine speed and load pose severe limitations on the range of applicability of the model. The linear perturbation model used of Figure 1.5 can be augmented by judicious modeling of the nonlinear effects mentioned above to provide useful diagnostic functions over a greatly extended range of operation. A study conducted on a 4-cylinder engine [51] has indicated that the dynamic relationship between throttle angle and intake manifold absolute pressure can be characterized by a linear model over the entire range of engine operation (0-100% throttle, idle to 5000 rev/min), with a maximum error of approximately 5%, provided that the measured input (throttle position) is appropriately pre-warped by a static function, representing the nonlinear characteristics of the pressure-flow characteristics. This result is particularly important, considering that in order to design algorithms that can be implemented in real time it is advantageous (though not strictly necessary) to use linear models. Similarly motivated static pre-warping characteristics can be applied to the spark and fuel inputs, resulting in extended range of operation of the model. It should be remarked that, although it is possible to rely strictly on static relationships (engine maps) to perform some useful diagnostics, the additional information embedded in dynamic relationship may make it possible to differentiate between faults that have similar effects on the measured outputs. As an example, consider the model of Figure 1.5. Modeling the dynamics of the intake manifold makes it possible to distinguish between a fault in the manifold pressure sensor, a fault in the throttle position sensor, or a vacuum leak in the

manifold, in spite of the fact that each of these faults affects the same measured variable, namely, intake manifold absolute pressure.

The third model discussed in this review is of particular interest from the standpoint of emission control system diagnostics. This model, shown in Figure 1.6, has been reported by [52], and describes the dynamics of the intake manifold, of the air-to-fuel ratio (A/F) controller (typically a PI controller), of the exhaust gas oxygen ( $O_2$ ) sensor, and of the fuel injection, combustion, and exhaust mixing processes. The model was successfully employed in developing an individual cylinder fueling strategy based on a single  $O_2$  sensor. The key aspects of this model are that it explicitly relates control variables to exhaust emissions performance through the  $O_2$  sensor, the controller and the fuel injectors. Thus, in principle, it is possible to diagnose the degraded performance of any of these components from the available measurements, although the model has not been explicitly utilized for diagnostic purposes in the reported work. In addition to the flow-pressure nonlinearity described earlier, and to the presence of delays, the subsystem of Figure 1.6 is also characterized by a hard nonlinearity in the output, due to the switching characteristic of the  $O_2$  sensor. The dynamics of the exhaust mixing process are also likely to be load and speed dependent. The application of such powertrain nonlinear models towards fault detection is covered in [136]-[138].

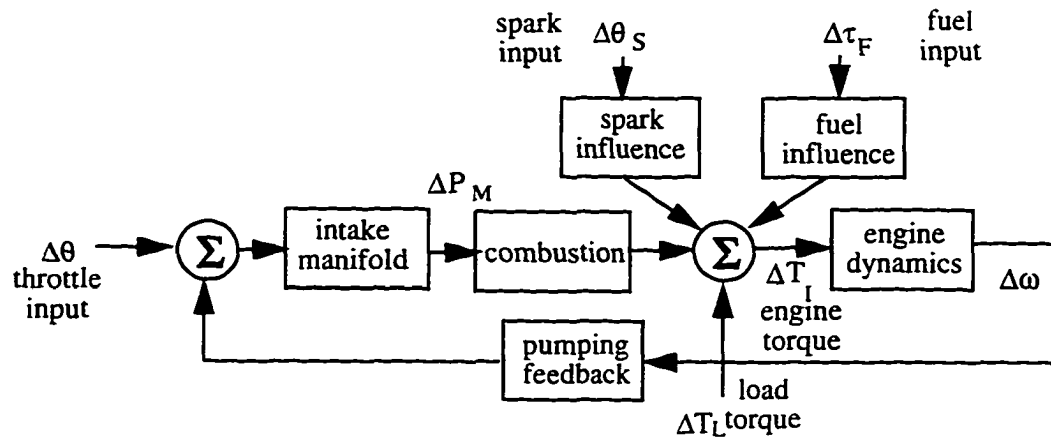


Figure 1.5 A Perturbation Model of Throttle-to-Speed Dynamics

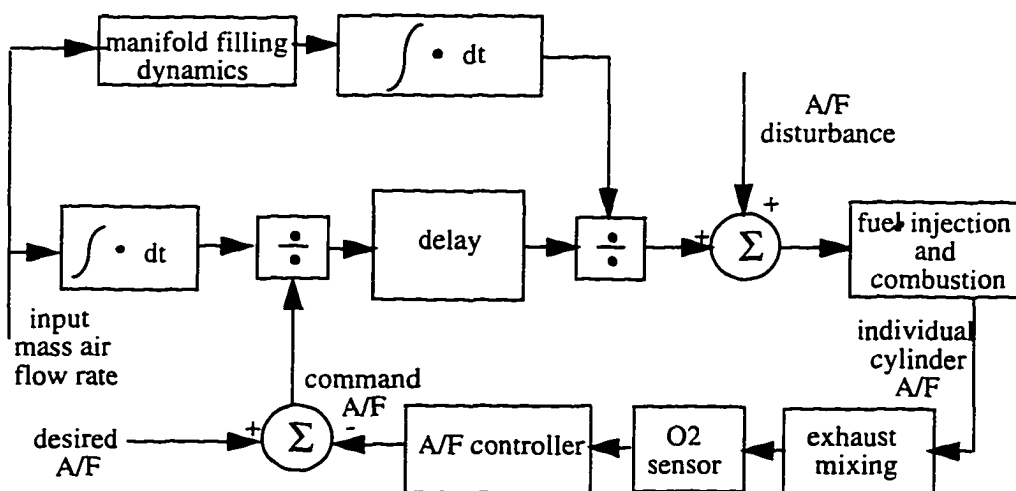


Figure 1.6 A Model for Fueling Dynamics of Spark Ignition Engine

The use of physical models to gain physical insights into the powertrain systems enhances the effectiveness of the diagnostic system. Although this may not always be possible for on-board implementation, it is useful to have a hierarchy of models plus algorithms to accomplish the same basic objectives with varying level of accuracy (i.e., on-board diagnostics vs service bay). Model-based diagnostic methods do rely in part on the data acquired over the engine's normal operating conditions [47] and [48], if for no other reason to identify the optimal values of the model parameters.

A model which captures the physical processes in an engine [45] is shown in Figure 1.6, where the most important phenomenon have been modeled and included in the block diagram. The model of Figure 1.7 is an amalgamation of the many different types of models used by the differing engine research groups plus some new features that do not appear in any control-oriented models. This model is sufficient as a fault simulator and for control objective. The model has been used in [136] and [137] for fault diagnosis in automotive systems.

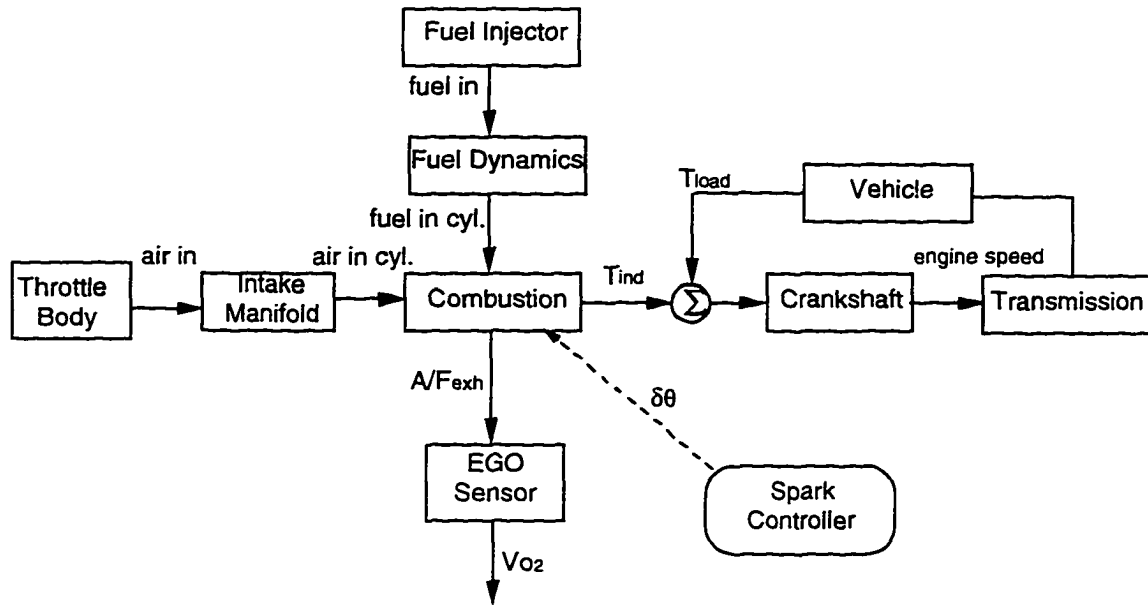


Figure 1.7 Engine Block Diagram

#### 1.1.5. Design of a Hybrid Diagnostic System

Three phases may be identified in the design of the diagnostic strategy:

- i. The optimal use of analytical models to enhance the *diagnosability* of the plant, for example by suitable selection of an optimal sensor set to guarantee sufficient analytic redundancy at minimum hardware costs.
- ii. The reduction of the same analytical models to well understood subsystems, permitting the implementation of simpler monitoring strategies for each subsystem (“Divide and Conquer”).

- iii. The integration of the subsystem models by means of a supervisory monitoring system. The development of a design methodology according to this sequence has the potential to result in more efficient local and global monitoring strategies.

In a practical engineering situation, the diagnostic problem is characterized by two distinctive features:

- i) The diagnostician typically has good engineering knowledge of the plant.
- ii) There are certain system specific features to every diagnostic problem which make it difficult to translate a given solution to other systems.

These facts often lead to *ad hoc* diagnostic strategies, which satisfy the engineer's intuition, but are not necessarily as effective as those which might be afforded by a more systematic study.

A combination of model-based fault detection strategies, which exploit known or measured dynamic relationships between system parameters and variables to assess dynamically the performance of a plant, in conjunction with a set of logical rules, representing a compilation of available diagnostic and design experience, may provide a useful framework for the design of diagnostic instruments [39].

This could be approached, for example, by using fuzzy logic rules, representing both qualitative and quantitative knowledge, in conjunction with model based diagnosis. Such an approach could be the answer to overcome the complexity of expert systems and the limitations of the FDI. Fuzzy rules can be expressed logically using simple linguistics rules, and using every day language so it is easy to understand by the machine operator

who can easily interpret the effect or outcome of each rule. As it is a human concept, it has the potential to perform well in many tasks which currently require human intuition and experience and this relates perfectly to the diagnostic problems.

Each of the two approaches has its own merits and problems. For instance, the dynamic model based methods to FDI can often be implemented in real time, but they suffer from providing an incomplete picture of the system performance and failure status, and normally do not provide a user-friendly interface. The rule-based methods are often difficult to implement in real-time and have difficulties in interpreting time varying sensor data, but provide for higher-level reasoning about failures that includes considerations about system structure, and provide for a user-friendly interface.

Development of a *hybrid* approach to FDI involves determining the best way to combine one or more model based approaches with one or more rule-based approaches. Generally speaking, the decision-making component of such a hybrid system is provided with the results of the algorithmic system and the system input/output data. The real-time data interpretation is thus delegated to the algorithmic system which in turn passes its results to the rule-based system. With this additional information, the rule-based system's diagnosis task becomes more manageable. The functional operation of a hybrid FDI system is depicted in Figure 1.8.

It is in fact the case that the trend in the rule-based approaches described above is to include some numerical preprocessing of the system sensor data before performing a diagnosis task. There have been, however, specific hybrid approaches developed for FDI. One such hybrid approach to FDI was developed in [52]- [55] for aircraft. In this work



the authors show how to develop low level actuator fault detection systems for aircraft then show how the research in the cognitive modeling of pilots at diagnostic tasks can be utilized to implement an expert system to interpret the results of the low level algorithms. Essentially the hybrid approach was composed of a deep knowledge expert system approach combined with a Beard-Jones fault detection filter. Many other possibilities exist for combining algorithmic and AI-based approaches.

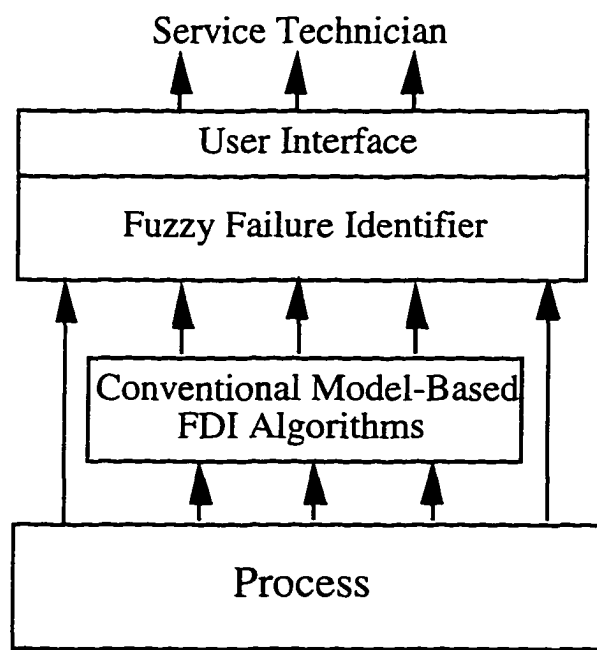


Figure 1.8 Hybrid FDI System [127]

Furthermore, a combination of analytical/ fuzzy model-based fault diagnosis concept is applied in [58] to the high-pressure-preheater line of a power plant, where the residuals generated from analytical knowledge of the plant is evaluated using fuzzy residual evaluation. The process of transforming quantitative knowledge (residuals) into qualitative knowledge (faulty symptoms) is covered in [59]. Also, in [91] the diagnosis of automotive faults was achieved by a combination of fuzzy estimators and FDI algorithms.

For this research work we adopt the fuzzy rule base approach as the diagnostic decision making scheme, and as the method for representing the qualitative knowledge. Fuzzy rules can be expressed logically using simple linguistics rules, and using every day language so it is easy to understand by the machine operator who can easily interpret the effect or outcome of each rule. As it is a human concept, fuzzy logic has the potential to perform well in many tasks which currently require human intuition and experience and this relates perfectly to the diagnostic problem at hand.

## **1.2. RESEARCH OBJECTIVES**

The objectives of the present study are listed below.

1. To generate a general methodology for the diagnosis of complex systems.
2. To integrate model-based FDI algorithms with knowledge-based algorithms for the diagnosis of engine malfunctions.
3. To provide a methodology for the development of a diagnostic tools that can integrate the information supplied by conventional tailpipe inspection programs with on-board diagnostics to provide fast and reliable fault diagnosis.

4. To understand the correlation between emission control system malfunctions and measured tailpipe emissions.
5. To generate an extensive data base to analyze, and to uncover characteristic patterns that may assist in the diagnosis of specific faults
6. To generate rules from heuristic knowledge that can be used as a fuzzy knowledge-base.
7. To fine tune the fuzzy rule-base to detect faults and to make decisions about which system, subsystem, or component is malfunctioning.

### **1.3. CONTRIBUTIONS OF THIS WORK**

- A literature survey of the diagnostic methods and procedures over the last decade was completed along with pinpointing the gap that can be filled in the area of diagnostics.
- A computer instrumented engine test cell for high accuracy engine data acquisition was developed, that allowed for engine exhaust emissions and actuator/sensor data collection.
- A methodology that integrates model-based fault detection methods, which exploit known or measured dynamic relationships between system parameters and variables to assess dynamically the performance of a plant, in conjunction with knowledge-based methods, representing a compilation of available diagnostic and design experience, was formulated which provides a useful framework for the design of diagnostic instruments.

- An extensive data base was generated and analyzed to uncover characteristic patterns that correlates emission control system malfunctions and measured tailpipe emissions which assists in the diagnosis of specific faults.
- A series of MATLAB<sup>®</sup> programs were written in order to analyze the comprehensive data base of emissions and sensor data which lead to more systematic methods to provide useful fusion of such diverse measurements in a diagnostic scheme.
- A novel application of fuzzy logic systems for the diagnosis of I.C. engines malfunctions was demonstrated. Moreover, the generation of rules using the concentration levels of the tailpipe emissions as the antecedent, and the fault as the consequent was achieved.

#### **1.4. SUMMARY**

In this chapter the problem at hand was investigated, where the literature available was reviewed. Also, an overview of the emission and diagnostic regulations relevant to the automotive emission control system was presented. The different approaches and methods used for diagnostics were explored and the equipment used was surveyed leading to the gap that can filled with the result of this dissertation work. The objectives and the contributions of this work were stated, in the following chapter the background behind this work is covered.

## CHAPTER 2

### BACKGROUND

The I. C. engine is a complex system, with many subsystems and components interacting to deliver a required output. In this chapter an overview of engine emissions and their control strategy, diagnostic regulations, and diagnostic methods is covered. Moreover, a diagnostic hierarchical method for detecting and isolating Emission Control System (ECS) faults is introduced with the overall goals of this research work.

#### 2.1. INTERNAL COMBUSTION ENGINE EMISSIONS

Internal combustion (I. C.) engines convert chemical energy contained in the fuel (primarily hydrocarbons) into mechanical energy. The chemical energy consists of the fuel enthalpy, combustion air sensible enthalpy, and cooling air sensible energy. The fuel-air mixture is the working fluid which is combusted in the cylinder, while the burned products after combustion are the constituents of the working fluid that provides the desired power output. A control volume constructed around the engine will have the fuel-air mixture and the cooling air as inputs and the exhaust gases, the cooling air, and the useful work as outputs.

Engines can be classified into categories [57], one of which is according to the method of ignition. In spark ignition engines, the combustible mixture is considered a homogeneous mixture of fuel vapor (hydrocarbons) and air [4]. The quality of the combustion, fuel economy, power output, and emission levels are all dependent on the mixture (Air/Fuel) ratio. The fuel is normally mixed with air in the engine intake system to produce a combustible homogeneous air/fuel mixture [105]. The air/fuel ratio can vary from lean to rich, where it deviates from the chemically correct mixture known as stoichiometric mixture. In a complete combustion process fuel (hydrocarbons) and air (oxygen and nitrogen) are ignited to produce CO, water and nitrogen, but the typical engine combustion produces unburned hydrocarbons (HC), nitrogen oxides ( $\text{NO}_x$ ), carbon monoxide (CO), carbon dioxide ( $\text{CO}_2$ ) and water ( $\text{H}_2\text{O}$ ). These are the exhaust pollutants regulated by the EPA. A hydrocarbon fuel can be completely oxidized if sufficient oxygen is available. Insufficient oxygen (air) in the mixture (a rich mixture) leads to an incomplete combustion which produces higher levels of HC and CO. Furthermore, excess oxygen in the mixture (a lean mixture) leads to higher combustion temperatures that lead to increase in the  $\text{NO}_x$  emission levels. Figure 2.1 shows the change in the emission levels as the Air/Fuel mixture deviates from the stoichiometric ratio.

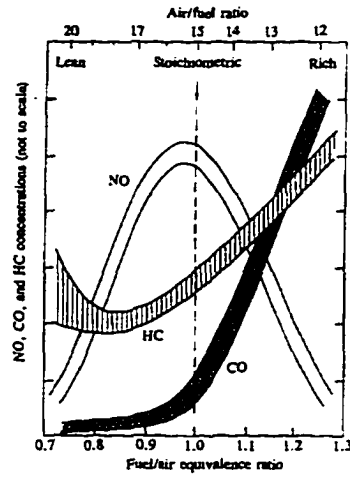
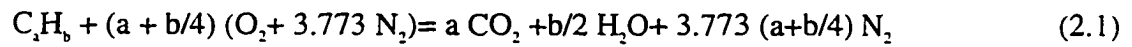


Figure 2.1 Emission Levels vs A/F Ratio [57]

The complete combustion equation of a general hydrocarbon fuel of average composition  $C_xH_y$  with air, [57] is:



The previous equation defines the stoichiometric (or chemically correct, or theoretical) proportions of fuel and air. If  $y = a/b$ , the stoichiometric air to fuel ratio depend on the fuel composition;

$$A/F = 34.56(4+y)/12.011 + 1.008y. \quad (2.2)$$

The ratio of air to fuel for Indolene, often used for emission testing, where  $y=1.86$  is 14.58/1, when less air is available the mixture is known as rich and if there is excess air it is known as lean mixture. Now we define the following variables of interest:

$$\text{The equivalence ratio } \phi = \frac{(F / A)_{actual}}{(F / A)_{stoich}} \quad (2.3)$$

$$\text{The relative air/fuel ratio } \lambda = \frac{(A / F)_{actual}}{(A / F)_{stoich}} \quad (2.4)$$

Leading to the three mixture conditions below:

Rich  $\phi > 1, \lambda < 1$

Stoichiometric  $\phi = 1, \lambda = 1$

Lean  $\phi < 1, \lambda > 1$

The following is a summary of the mechanism of formation of exhaust emission as described in [4]:

#### **Hydrocarbons (HC):**

They are a resultant of unburned or partially burned fuel molecules due to incomplete combustion. The formation of HC has been attributed to the following:

- i. Wall quenching as the combustion flame front propagates near a wall or into a crevice.
- ii. Oil film absorption and desorption.



- iii. Deposit storage and release.
- iv. Incomplete combustion of the charge.
- v. Excessive charge stratification.

**Nitrogen Oxides ( $\text{NO}_x$ ):**

Nitrogen and oxygen atoms in the air under the high pressure and temperature during combustion react to form nitrous oxide ( $\text{NO}$ ) and nitrogen dioxide ( $\text{NO}_2$ ). The higher the burned gas temperature, the higher the rate of  $\text{NO}$  formation.

**Carbon Monoxide ( $\text{CO}$ ):**

Carbon in the fuel is partially oxidized due to incomplete combustion. With rich fuel-air mixtures, there is insufficient oxygen to burn fully all the carbon in the fuel to  $\text{CO}_2$ . The sources of  $\text{CO}$  formation are:

- i. Rich mixture combustion.
- ii. Slow kinetics during expansion.
- iii. Wall and crevice effects
- iv. Partial HC oxidation.

**Carbon Dioxide ( $\text{CO}_2$ ):**

It is a product of complete and incomplete combustion.

## 2.2. EMISSION CONTROLS

To comply with the more stringent emission regulations and the increased emphasis on fuel economy, the air/fuel ratio in today's automobiles is controlled close to the stoichiometric ratio to take advantage of the catalytic converter. This is achieved by a closed-loop control system that uses an oxygen sensor as a feedback signal to a microprocessor that outputs commands to both of the fuel injectors and the spark ignition system to optimize the engine performance and to minimize emissions.

In conjunction with the oxygen sensor, a three way catalytic converter is used to reduce engine emission levels of the three major pollutants (HC, NO<sub>x</sub>, and CO). The catalytic converter maximum efficiency can be reached in a narrow window close to stoichiometry where  $\lambda = \phi = 1$ . As shown in Figure 2.2, the three pollutants' concentration rises substantially once the mixture air/fuel ratio deviates from stoichiometry. The catalytic converter conversion efficiency for a given component x is determined by the following equation [57]:

$$\eta_{cat} = \frac{\dot{m}_{x,in} - \dot{m}_{x,out}}{\dot{m}_{x,in}} \quad (2.5)$$

Where  $\dot{m}$  represents the mass flow rate of the component.

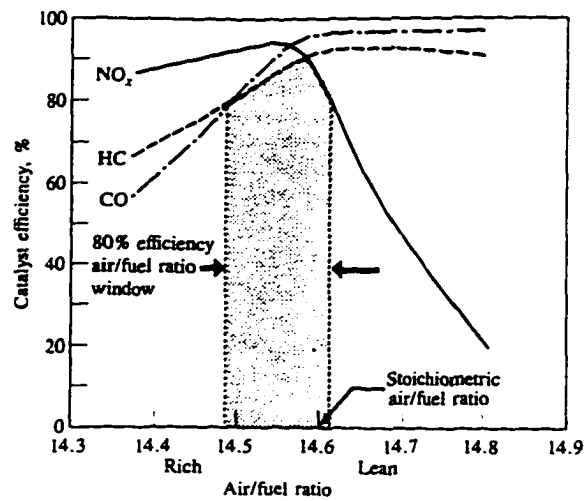


Figure 2.2 Catalytic Converter Efficiency [57]

Closed-loop control of ignition timing and fuel injectors aids in the reduction of the emission levels. The excess air factor lambda ( $\lambda$ ) is measured by means of an oxygen sensor that is located in the exhaust system. The oxygen, or lambda, sensor contains a ceramic material that has an electrical response to changes in the oxygen partial pressure in the exhaust stream relative to ambient; this sensor becomes active when it reaches a temperature  $T > 300\text{ }^{\circ}\text{C}$  [105]. The concentration of the amount of oxygen relative to ambient contained in the exhaust is related to the air/fuel ratio. Thus, the voltage output is a measure of  $\lambda$ . The sensor voltage is compared to a reference voltage; the output of the comparator circuit then switches between high and low values which correspond to rich and lean mixtures respectively. This voltage is processed by the Engine Control Unit

(ECU), which outputs a signal to the fuel injectors (pulse width) adjusting the amount of fuel injected according to the input voltage.

The Nernst equation [57] describes the potential across the ceramic active element as a function of the oxygen partial pressures:

$$V_o = \frac{RT}{4F} \ln\left(\frac{P''O_2}{P'O_2}\right) \quad (2.6)$$

In (2.6),  $R$  is the universal gas constant,  $T$  is the temperature,  $F$  is, and  $P''$ , and  $P'$

An additional method that is used to reduce  $NO_x$  is by recirculating fraction of the exhaust gas into the intake manifold through the Exhaust Gas Recirculation (EGR) valve. This EGR valve recirculates the exhaust gas back to the combustion chamber which means adding an inert gas to the mixture and therefore reducing the combustion temperature leading to the reduction of  $NO_x$  formation. In addition HC emissions which originate from evaporative sources (i.e. evaporation of the fuel in the tank) are reduced by the use of an evaporative emission control system which is designed to store and to dispose of the fuel tank vapors. This evaporative system consists of an active-charcoal canister that stores the HC vapors; these vapors are then purged into the intake manifold to be burned with the mixture at operating conditions that require additional enrichments.

Furthermore, open-loop control using other available sensors, e.g. Throttle Position Sensor (TP), Manifold Air Pressure (MAP), Mass Air Flow (MAF), is used at operating conditions which require a richer or leaner mixtures that these ideally in the case of closed-loop control; cold start and heavy acceleration are few conditions among

others which require the open-loop control strategy to take over. Figure 2.3 depicts a system's view of an engine emission control system.

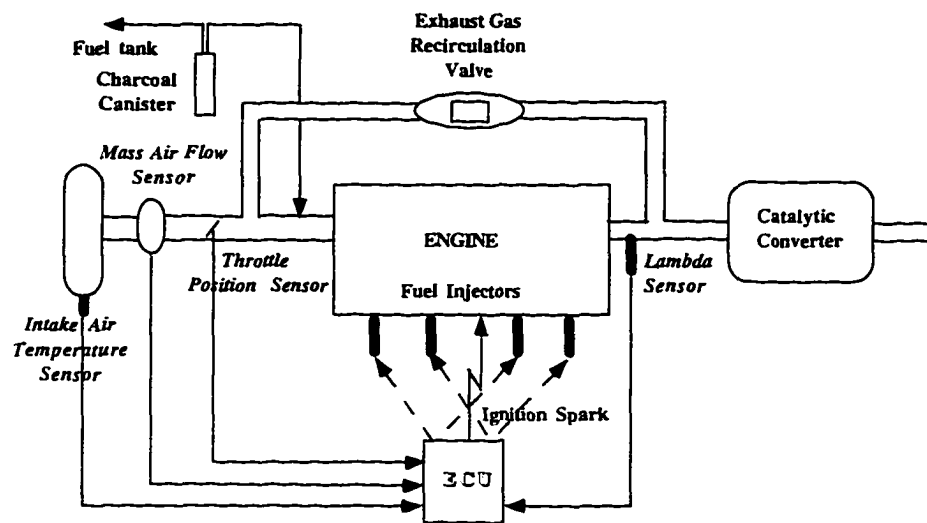


Figure 2.3 Engine Emission Control System

The air/fuel ratio control problem is not as simple as it might appear from this description. Manufacturing variations in system components, transient excursions in load and temperature, and malfunctions in any of the components of Figure 2.3 add complexity to the control problem.

### **2.3. DIAGNOSTIC REGULATIONS**

As mentioned in Chapter 1, in 1975 stringent emission regulations for passenger vehicle tailpipe led to the appearance of the charcoal canister, Exhaust Gas Recirculation (EGR) valve, and finally the catalytic converter. Moreover, in 1977 amendments of the Clean Air mandated Inspection and Maintenance (I/M) programs for high-pollution areas. In 1978 the EPA issued its first policy for I/M programs. Furthermore, as the standards grew stricter, three-way catalysts, on-board computers, and oxygen sensors appeared in 1981.

The growing use of electronic dash panel displays in passenger cars and the increasing use of sophisticated emission and powertrain control systems, led California to adopt the on-board diagnostic (OBD) regulations in 1985 for 1988 and later models light and medium-duty vehicles. OBD systems were designed to reduce the time between occurrence of a malfunction and its detection and repair. The objective of OBD is to reduce emissions caused by the malfunction and also minimize the damage that could occur to other vehicle components or systems. This is achieved by incorporating additional software and hardware to collect and analyze data already available to the on-board computer, thus monitoring the entire emission control system.

In 1989, California Code of Regulations (CCR), known as OBD-II was adopted by the California Air Resources Board (CARB). OBD-II requires the manufacturers to implement new comprehensive on-board diagnostic systems, replacing OBD-I, starting with the 1994 model year. OBD-II monitors more components and systems than OBD-I, including:

- Catalyst
- Evaporative control system
- Emission control system
- Electronic emission related powertrain components.
- Detection of engine misfire.

OBD-II systems are intended to be capable of detecting most malfunctions when performance of a component or a system deteriorates to the point that the vehicle emission exceed a threshold value tied to the applicable emission standard, Table 2.1 shows the systems monitored by OBD-II.

The vehicle operator is notified at the time when the vehicle begins to marginally exceed emission standards, by Malfunction Indication Light (MIL). OBD-II also requires that additional information be provided to technicians for diagnosis and repair of emission related problems.

ITEM	LEGAL REQUIREMENTS	DIAGNOSTIC TECHNIQUE
Catalyst Diagnostics	Light MIL when HC conversion efficiency of catalyst falls to 60%.	Two O <sub>2</sub> Sensors
Misfire Monitoring	Light MIL on detecting Misfire.	Measuring change in engine speed
Fuel system Monitoring	Light MIL when abnormality in the fuel system exceeds 1.5 times the emission control level.	Value of $\lambda$ with O <sub>2</sub> sensor.
O <sub>2</sub> Sensor Monitoring	Light MIL when output voltage or response speed exceeds 1.5 times the emission control level.	Response time of two O <sub>2</sub> Sensors
EGR Monitoring	Light MIL when EGR flow rate is too large or small exceeding 1.5 times the emission control level.	Measuring temperature change in EGR flow path with a detector.
Secondary Air System Monitoring	Light MIL when Secondary air flow rate is below 1.5 times the emission control level.	Reading O <sub>2</sub> sensor signals by providing a secondary air jet in the upstream of the O <sub>2</sub> sensor.
Evaporative leak monitoring	Light MIL when evaporative system causes HC vapor to leak into the atmosphere is below 1.5 times the emission control level.	Evaporative check pressure sensor
Monitoring of the other emission related components	Light MIL when any component or system related to emission performance malfunctions. Vehicle speed sensor, crank angle sensor, throttle sensor, water temperature sensor, neutral sensor, knock sensor, ISC valve, purge valve, VTC valve, electronic control A/T, etc.	Checking wires for breakage or by using an operation mode in which normal operation can be confirmed

Table 2.1 OBD-II Monitored Systems [106]



An additional vehicle communication system is used to provide the service technicians with detailed information about system performance and detected malfunctions, this is made achievable by storing fault codes that leads technicians to the likely area of the malfunction and continuously updating the information for some engine parameters to help them isolate the specific fault. Most components are monitored, except for the catalyst and evaporative system, such that a malfunction is signaled as the emission exceed 1.5 times the applicable standards. OBD-II is designed to detect engine misfire, which could prevent serious damage to the catalytic converter. Further, it also includes "freeze frame", which allows the computer to store in memory the exact operating conditions when a fault occurred, so intermittent faults can be investigated by revisiting the same conditions when the problem occurred. A standard access electrical connector which is identical for all makes is required, which means a single inexpensive generic tool can be used to read out fault codes.

Although OBD-II requirements reflect state-of-the-art diagnostic system capability, there are limitations which apply to the current techniques for detecting malfunctioning components. These limitations do not allow OBD-II systems to take the place of the FTP test for measuring vehicle emissions. The reason is that monitoring systems can detect when components are functioning within there operating range, but are limited with ability to determine whether they are functioning accurately within the range.

The tailpipe mass emission standards are getting more stringent. CARB has adopted a low-emission vehicles/clean fuel program in California during mid- and late-1990s, where four vehicle types are defined, these are Transitional Low Emission Vehicle

(TELV), Low Emission vehicle (LEV), Ultra Low Emission Vehicle (ULEV), and Zero Emission Vehicle (ZEV). Table 2.2 shows the emission level standards [107];

	<i>CV</i>	<i>TLEV</i>	<i>LEV</i>	<i>ULEV</i>	<i>ZEV</i>
NMOG	0.25	0.125	0.075	0.04	0.0
CO	3.4	3.4	3.4	1.7	0.0
NO <sub>x</sub>	0.4	0.4	0.2	0.2	0.0

Table 2.2 California Tailpipe Emission Standards (grams/mile) [107]

These new standards make the early detection of emission related faults more important, as their malfunctions lead to a significant increase in the emission levels. Moreover, this means the increase in the complexity of the emission control system, and as a consequent the complexity OBD system. All this require a reliable diagnostic system that can detect the incipient failures and avoid further degradation of the performance of the emissions control system. The growing demand for a powerful diagnostic off-board tool that can integrate all the of the available information about the automotive systems dictates a new approach towards fault detection in such systems. An overview of diagnostic methods used will be covered in the following sections

## **2.4. DIAGNOSTIC METHODS**

Fault diagnosis of a physical plant is crucial for its healthy performance, as it ultimately could prevent catastrophic failure, help comply with environmental regulations, and enhance customer satisfaction. There exist several methods to detect and isolate incipient faults that might cause the plant performance to deviate from the nominal. Fault detection and diagnostic methods may utilize physical or analytical redundancy as a basis for the diagnostic scheme. Physical, or hardware, redundancy schemes require the installation of extra hardware components which are used to monitor their counterparts [108]. The simplest example of this scheme are double- or triple-redundant sensors found in high performance aerospace systems [56]. On the other hand, analytical redundancy scheme requires a model of the plant which generates signals that are processed logically to monitor the plant. The models in turn require knowledge of the plant; this knowledge might be subjective (expert systems, etc.) or objective knowledge (mathematical models, etc.) [109]. Different diagnostic methods have evolved that are based on objective knowledge of the plant (mathematical models), while other are based on subjective knowledge (expert systems). These diagnostic methods can be grouped as follows:

1. Dynamic model-based methods
2. Knowledge-based methods

The characteristics of the preceding methods are exploited in the following sections.

### 2.4.1. Dynamic Model-Based Methods

Model-based fault detection methods use knowledge of a system model to generate an estimate of the system states (or outputs) from the measured inputs. The general form of a model based diagnostic scheme is displayed in Figure 2.4. The inputs to the residual generator are the commanded inputs to, and the measured outputs of, the plant. The residual generator then produces as its output a vector of residuals which are then processed by a decision maker that draws inferences about the presence or absence of faults.

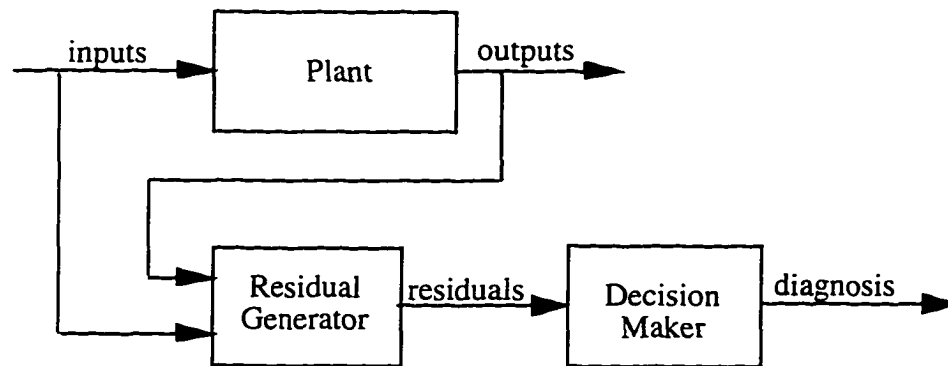


Figure 2.4 Block Diagram of a Model-based Residual Generation Scheme

The system models can be either nonlinear or linear, the (more general) nonlinear systems equations in discrete time are as follows,

$$x(k+1) = f(x(k), u(k)) \quad (2.7)$$

$$y(k) = h(x(k), u(k)) \quad (2.8)$$

where,  $x \in \mathcal{R}^{n \times 1}$  is the state vector,  $f \in \mathcal{R}^{n \times 1}$  is the system dynamics vector function,  $u \in \mathcal{R}^{m \times 1}$  is the input vector,  $y \in \mathcal{R}^{p \times 1}$  is the output vector, and  $h \in \mathcal{R}^{p \times 1}$  is the output vector function. While a linear discrete time system is represented in the state space form as follows,

$$x(k+1) = Ax(k) + Bu(k) \quad (2.9)$$

$$y(k) = Cx(k) + Du(k) \quad (2.10)$$

where  $x \in \mathcal{R}^{n \times 1}$  is the state vector,  $A \in \mathcal{R}^{n \times n}$  is the dynamics matrix,  $B \in \mathcal{R}^{n \times m}$  is the input distribution matrix,  $y \in \mathcal{R}^{p \times 1}$  is the vector of outputs,  $u \in \mathcal{R}^{m \times 1}$  is the vector of inputs,  $C \in \mathcal{R}^{p \times n}$  is the output distribution matrix and  $D \in \mathcal{R}^{p \times m}$  is the direct term matrix.

The residuals, which are quantities that represent the difference between the measured and estimated values, are generated using different methods [40]. The residual generator is designed to produce residuals that have different properties under different fault conditions, in order to distinguish among the various faults. Typically, the residual generator is designed such that one or more elements of the residual vector are nonzero in the presence of a fault, while the remaining elements are zero. The methods for generating residuals include, according to [40] the following:

1. Parity equations or consistency relations

2. Diagnostic observers
3. Kalman filters
4. Parameter estimators

While the fault diagnosis process includes the following steps:

- i. Detection (presence)
- ii. Isolation (location)
- iii. Identification (size)

These may be performed sequentially or in parallel, and are affected by: disturbances, noise, and modeling errors. These discrepancies may lead to false alarms, missed detection, and misclassification. In this background section we shall not consider noise and modeling error to permit a simpler overview of the main ideas. The desired properties of a diagnostic algorithm are listed:

1. Insensitivity to disturbance
2. Noise suppression
3. Robustness to modeling error
4. Sensitivity to faults

The basic idea is to comparing actual plant behavior (measured) with predictions based on mathematical model. Clearly, while a scalar residual suffices for detection, we need a vector residual for isolation.

## Faults, Disturbances, and Modeling Errors

In the following pages, a general model relating faults, disturbances, and other errors to the plant dynamics is briefly reviewed. *Faults* are the cause of failure; we seek to recognize the performance degradation resulting from a fault. A well accepted relationship between faults and failures, which can have a recursive interpretation, is shown below.

$$\textit{Fault} \rightarrow \textit{Malfunction (error)} \rightarrow \textit{Failure}$$

Faults may be represented as *additive* (i.e. can be modeled as an extra input acting on the system), or *multiplicative*, i.e., corresponding to a change in some plant parameters. If at all possible we try to model faults as additive, because this model can be handled more easily by FDI algorithms.

*Disturbances* are undesired inputs; there may be no difference in some cases between an additive fault and a disturbance. The distinction is philosophical: ignore disturbances and detect faults. Both faults and disturbances are modeled as unspecified functions of time (sometimes we specify “drift”, “jump”, “intermittent”). *Modeling errors* correspond to uncertainty in parameters of modeled system and lead to multiplicative faults. One of the great challenges in FDI is the distinction between parameter changes induced by faults and model uncertainty. The reader may find a more in-depth treatment of this subject in [41]. A general system and fault representation is shown in Figure 2.5.

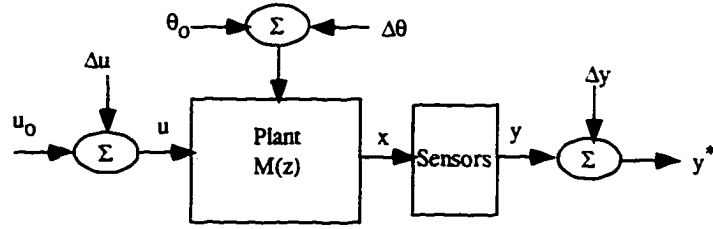


Figure 2.5 General System and Fault Representation

In Figure 2.6,.  $u_0$  is the commanded input vector,  $\Delta u$  is input fault vector, and  $u$  is the actual plant input.  $\theta_0$  is the nominal parameter vector,  $\Delta\theta$  is the parameter (component) fault,  $x$  is the state vector  $M$  is the (matrix) transfer function,  $y$  is the actual output vector,  $\Delta y$  is the output fault vector, and  $y^*$  is the measured output vector. A general representation of the system can be in the form

$$\sum_{k=0}^N a_k y(n-k) = \sum_{k=0}^N b_k u(n-k),$$

where  $y \in \mathcal{H}^M$ , and  $u \in \mathcal{H}^K$ .

Let the  $\mathcal{Z}$ -transform of  $u(k)$  is  $U(z) = \mathcal{Z}(u(k))$ , and similarly for  $y$ . Then a multi-input multi-output (MIMO) system can be represented in the  $z$ -domain by

$$Y(z) = M(z)U(z).$$

In Figure 2.6

$u_c \subset u$  denotes controlled inputs,



$u_m \subset u$  denotes measured inputs.

If  $u^{(i)} \notin u_c$  and  $\notin u_m$ , then it is treated as disturbance.

Assume command values are known for  $u_c$ , and measured values are known for  $u_m$ , and  $y$ .

Then, additive faults may be represented as follows:

$\Delta u_c(k), \Delta U_c(z)$  Input actuator fault

$\Delta u_m(k), \Delta U_m(z)$  input sensor fault

$\Delta y(k), \Delta Y(z)$  output sensor fault

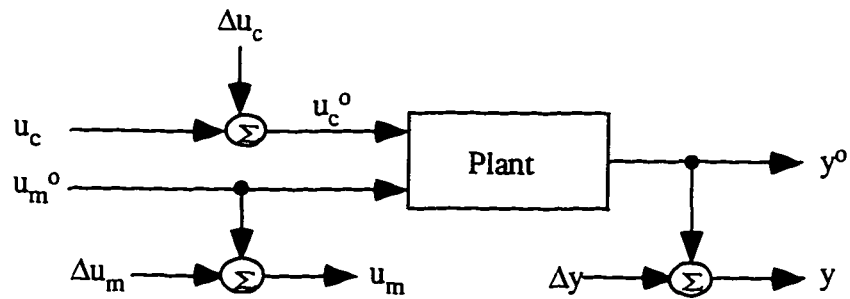


Figure 2.6 Fault Model [96]

Let the actual inputs and outputs of the system be:  $u_c^o$ ,  $u_m^o$ ,  $y^o$ , and let the measured/commanded inputs and outputs be represented by:  $u_c$ ,  $u_m$ ,  $y$ .

$$u_c^o = u_c + \Delta u_c$$

$$u_m^o = u_m - \Delta u_m$$

$$y^o = y - \Delta y$$

The nominal model  $Y(z) = M(z) U(z)$  actually applies to  $U_c^o$ ,  $U_m^o$ ,  $Y^o$ .

Thus,

$$U(z) = \begin{bmatrix} U_c(z) \\ U_m(z) \end{bmatrix} \text{ and } M(z) = [M_c(z) : M_m(z)],$$

and we can represent the system as shown below, to include the effects of faults.

$$\begin{aligned} Y(z) - \Delta Y(z) &= Y_o(z) = [M_c(z) : M_m(z)] \begin{bmatrix} U_c(z) \\ U_m(z) \end{bmatrix} \\ &= M(z) u(k) + M_c(z) \Delta U_c(z) - M_m(z) \Delta U_m(z) \end{aligned} \quad (2.11)$$

Finally,

$$Y(z) = M(z) U(z) + S_f(z) P(z) \quad (2.12)$$

where  $P(z) = [\Delta U_c' : -\Delta U_m' : \Delta Y']'$ , and  $S_f(z) = [M_c(z) : M_m(z) : I]$ .

### Parity Equation Residual Generation (PERG)

In our context a residual generator (RG) is a linear discrete time algorithm that generates a residual with the following properties;

$$r(k), R(z) = 0 \quad \text{No fault present}$$

and (2.13)

$$r(k), R(z) \neq 0 \quad \text{Fault present}$$

The generic form of a RG in the frequency domain is:

$$R(z) = V(z) Z(z) + W(z) Y(z) \quad (2.14)$$

Applying equation (2.13) to (2.14) we obtain

$$V(z) U(z) + W(z) M(z) U(z) = 0 \quad \forall U(z),$$

or

$$V(z) = -W(z) M(z) \quad (2.15)$$

Thus, we can write

$$R(z) = W(z) [ Y(z) - M(z) U(z) ] \quad (2.16)$$

We call the above expression the *computational form* of the RG as shown in Figure 2.7.

The design of a PERG consists of selecting  $W(z)$  in such a way as to obtain the desired fault isolation (and noise rejection and modeling error robustness) properties. Gertler discusses the design procedures for PERG in great detail in [41].

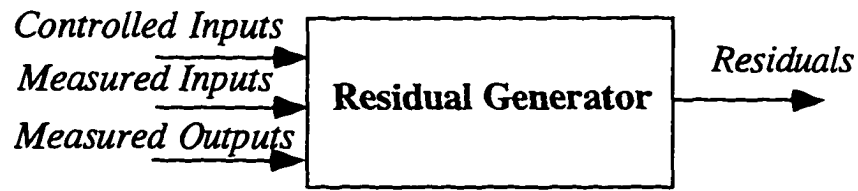


Figure 2.7 Computational Form of the RG

If we express  $Y(z) - M(z)U(z)$  from equation (2.12)

$$Y(z) - M(z)U(z) = S_F(z) P(z) \quad (2.17)$$

we can write

$$R(z) = W(z) [S_F(z) P(z)] \quad (2.18)$$

This is the *internal form* of the RG, which explicitly shows the effects of the faults.

The residuals obtained as shown above are called primary. primary residuals may be enhanced to facilitate the process of isolating faults. Two methods have been proposed to accomplish isolation:

1. Structured residuals
2. Directional residuals

*Structured Residuals* lead to a *fault signatures* or *codes*. One can then apply a threshold test to the *ith*  $r_i(k)$  to obtain

$$\varepsilon_i(z) = 0 \quad \text{if } |r_i| < \eta_i$$

$$\varepsilon_i(z) = 1 \quad \text{if } |r_i| \geq \eta_i \quad i = 1, \dots, c.$$

Then,

$\varepsilon = [\varepsilon_1 \dots \varepsilon_c]$ , is the fault code signature.

*Directional Residuals* are so called because the residual vector is confined to lie in one-dimensional subspace at all times, including during transients.

$$R(z | P_j) = \varphi_j \psi_j(z) P_j(z)$$

where  $\varphi_j$  is in the direction of the  $j$ th fault,  $\psi_j(z)$  is the transfer function (matrix) from the  $j$ th fault to the residual.

Note that if each  $r_i$  responds only to one  $P_j$ , the residual  $r$  is both structured and directional. Figure 2.8 depicts the appearance of a directional residual in a hypothetical three-dimensional case.

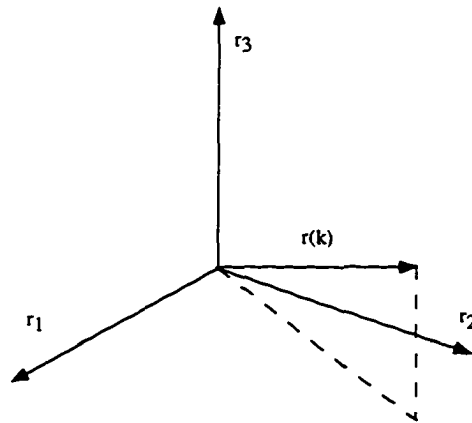


Figure 2.8 Parity Space Directionality

The basic concepts outlined above apply to all residual generators. More details may be found in [41]. In the next pages we shall review residual generation methods based on observers.

### Observer Based Residual Generation

One method for obtaining a residual is to estimate the inputs and outputs of a system using one or more observers. Observers are algorithms based on the state space form of a system's representation. For the purpose of this background we shall assume a Linear Time Invariant (LTI) system representation in discrete time:

$$x(k+1) = A x(k) + B u(k) \quad (\text{state equation})$$

$$y(k) = C x(k) + D u(k) \quad (\text{output equation})$$

Let

$\nu$  = number of states

$\kappa$  = number of inputs

$\mu$  = number of outputs,

then

$A$  is  $\nu \times \nu$      $B$  is  $\nu \times \kappa$      $C$  is  $\mu \times \nu$      $D$  is  $\mu \times \kappa$

For the present development we assume  $D=0$  if  $D \neq 0$ , we use  $y' = y - Du$  as our output vector.

### Observer Design

The observer design problem may be stated as follows:

“Given  $y(k)$ ,  $u(k)$  we wish to estimate  $x(k)$ ”. This objective can be achieved using a model of the dynamic system in the form

$$\hat{x}(k+1) = A_c \hat{x}(k) + L y(k) + z(k) \quad (2.19)$$

where

$\hat{x}(k)$  = estimated state.

Let  $e = x - \hat{x}$ , then

$$e(k+1) = Ax(k) - A_c \hat{x}(k) + L y(k) + B u(k) - z(k) \quad (2.20)$$

Let  $z(k) = Bu(k)$  and  $y(k) = Cx(k)$ , then

$$e(k+1) = (A - LC) x(k) - A_c \hat{x}(k)$$

If we choose  $A_c = A - LC$  we have

$$e(k+1) = (A-LC) e(k) \quad (2.21)$$

i.e., the estimation error approaches to 0 if  $A-LC$  is stable ;  $\hat{x}(t) \xrightarrow{t \rightarrow \infty} x(t)$ .

If  $(A, C)$  is observable, it is always possible to find  $L$  that permits arbitrary assignment of eigenvalues of  $(A-LC)$ , i.e.: we can arbitrarily specify the response of the observer. The structure of the observer is shown in Figure 2.9.

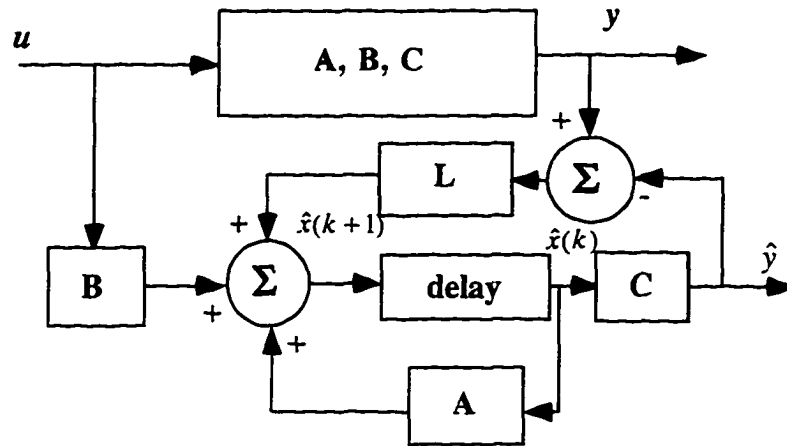


Figure 2.9 Observer Structure

Usually, the observer is represented in the following form.

$$\hat{x}(k+1) = A \hat{x}(k) + B u(k) + L(y(k) - \hat{y}(k)) \quad (2.22)$$

Note that this structure also matches the structure of the standard RG, since the output estimation error

$$r(k) = y(k) - \hat{y}(k)$$

may be interpreted as a primary residual.

A number of observer design procedures exist; in essence they all amount to pole placement (eigenvalue assignment) in the matrix  $(A - LC)$ .

Diagnostic Observers use the available design freedom to accomplish special tasks: decoupling, structuring, giving direction to residuals. We shall briefly review:

1. Observer design by Eigenstructure Assignment



## 2. The Unknown Input Observer (UIO)

### Eigenstructure Assignment for Fault Detection

Assume we can model the faults to be detected by additive vectors, which could also represent disturbances to be decoupled. A common representation is given in equation (2.23).

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k) + Eq(k) + Kp_a(k) \\ y(k) &= Cx(k) + Du(k) + P_s(k)\end{aligned}\tag{2.23}$$

In equation (2.23)

$p_a$  = vector of actuator faults

$p_s$  = vector of sensor faults

$q$  = vector of disturbance

$x \in R^n, u \in R^r, y \in R^m$ .

The above representation only includes faults and disturbances that can be modeled as additive ones. We shall not consider multiplicative faults at this time.

The design procedure for the class of diagnostic observers considered here may be summarized as follows. write the observer in the form

$$\begin{aligned}\hat{x}(k+1) &= (A - LC)\hat{x}(k) + (B - LD)u(k) + Ly(k) \\ \hat{y}(k) &= C\hat{x}(k) + Du(k)\end{aligned}\tag{2.24}$$

Define the state estimation error  $e(k)$ :

$$e(k) = x(k) - \hat{x}(k).$$

Then the evolution of  $e(k)$  is governed by equation (2.25).

$$\begin{aligned}
e(k+1) &= (A - LC)e(k) + Eq(k) + KP_a(k) - LP_s(k) \\
&= A_o e(k) + Eq(k) + KP_a(k) - LP_s(k)
\end{aligned} \tag{2.25}$$

Define the  $p$ -dimensional residual as

$$\begin{aligned}
r(k) &= W[y(k) - \hat{y}(k)] \\
&= W[Cx(k) + Du(k) + p_s(k) - C\hat{x}(k) - Du(k)] \\
&= W[Ce(k) + p_s(k)] \\
&= WCe(k) + Wp_s(k)
\end{aligned} \tag{2.26}$$

where  $W$  is a matrix to be selected according to the design specification. In transfer function form we can write

$$\begin{aligned}
e(z) &= [zI - A_o]^{-1} EQ(z) \\
&\quad + [zI - A_o]^{-1} KP_a(z) \\
&\quad - [zI - A_o]^{-1} LP_s(z).
\end{aligned} \tag{2.27}$$

Substituting (2.27) into (2.26), we obtain the expression for the residual vector given in equation (2.28).

$$\begin{aligned}
R(z) &= [W - WC(zI - A_o)^{-1}]P_s(z) \\
&\quad + WC[zI - A_o]^{-1} KP_a(z) \\
&\quad + WC[zI - A_o]^{-1} EQ(z).
\end{aligned} \tag{2.28}$$

To decouple the residual from the effect of the disturbances we wish to set  $WC[zI - A_o]^{-1} E = 0$ .

In [104] algorithms are given for accomplishing such decoupling. Similarly, the observer may be decoupled from the effect of sensor and actuator faults by selecting  $W$  such that  $[W - WC(zI - A_o)^{-1}] = 0$  (sensor faults) or  $WC[zI - A_o]^{-1}K = 0$  (actuator faults). If exact decoupling is not possible, approximate solutions may be obtained using optimization methods [104].

The ability to decouple some of the unknown inputs may be exploited in multiple observer scheme, such as the one depicted in Figure 2.10, where two observers are employed to decouple a disturbance and one of two faults. In the scheme of Figure 2.10 a nonzero element in the residual vector indicates that the unknown input that has not been decoupled is present.

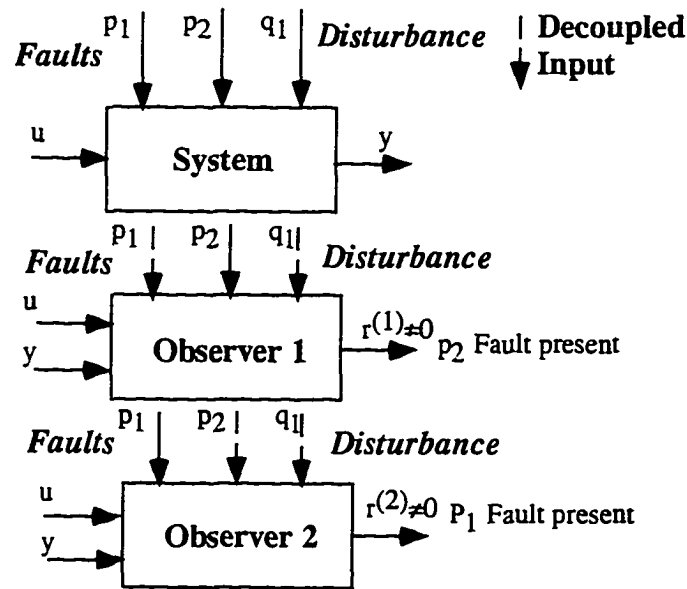


Figure 2.10 Multiple Observer Scheme

## The Unknown Input Observer

A second method that permits the direct estimation of unknown inputs is the unknown-input observer (UIO) [101] and [102]. Consider a linear time-invariant system described by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Dd(k) \\ y(k) &= Cx(k) \end{aligned} \quad (2.29)$$

where  $x \in R^n$ ,  $u \in R^p$ ,  $d \in R^q$ ,  $y \in R^m$  are the state vector, the known input vector, the unknown input vector, and the measurement vector, respectively.  $A$ ,  $B$ ,  $D$  and  $C$  are constant matrices of appropriate dimensions.

The conventional observer estimates unknown states using the available measurements and inputs. The simple block diagram of conventional observer is illustrated in Figure 2.11a.

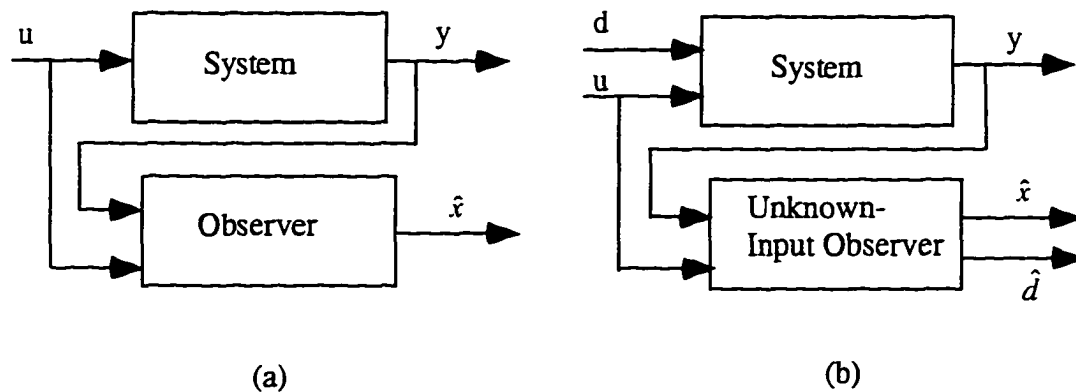


Figure 2.11 Conventional state observer vs. Unknown input observer

As shown in Figure 2.11, the state observer estimates the unknown states  $\hat{x}$  using the measurement  $y$  and input  $u$ . On the other hand, the UIO estimates the unknown states  $\hat{x}$  and the unknown inputs  $\hat{d}$  using the output  $y$  and the known input  $u$ . The simplified block diagram of the UIO is also shown in Figure 2.11b.

### UIO Design Procedure

The UIO for a linear time invariant system can be formulated as follows. This design procedure is mainly developed by [101].

Under the assumption that  $\text{rank } D=q$  (this assumption will be explained later), it is possible to choose a nonsingular matrix such that

$$T = \begin{bmatrix} N & D \end{bmatrix}, \quad N \in R^{n \times (n-q)} \quad (2.30)$$

Using this nonsingular matrix, the system given by equation (2.29) can be transformed as

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u + \bar{D}d \quad (2.31)$$

$$y = \bar{C}\bar{x} \quad (2.32)$$

where

$$x = T\bar{x} = T \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \quad \bar{A} = T^{-1}AT = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix},$$

$$\bar{B} = T^{-1}B = \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix}, \quad \bar{D} = T^{-1}D = \begin{bmatrix} 0 \\ I_q \end{bmatrix},$$

$$\bar{C} = CT = \begin{bmatrix} CN & CD \end{bmatrix},$$

with  $\bar{x}_1 \in R^{n-q}$  and  $\bar{x}_2 \in R^q$ .

In equation (2.31) and (2.32), it is possible to decouple the system into two parts: one directly related to the unknown input, and one which does not see the unknown input.

Then, the latter system can be defined as :

$$\begin{bmatrix} I_{n-q} & 0 \end{bmatrix} \dot{\bar{x}} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \end{bmatrix} \bar{x} + \bar{B}_1 u \quad (2.33)$$

$$y = \begin{bmatrix} CN & CD \end{bmatrix} \bar{x} \quad (2.34)$$

under the assumption that  $\bar{x}_2$  can be measured.

Also, if the matrix  $CD$  has full column rank, then there exists a nonsingular matrix

$$U = \begin{bmatrix} CD & Q \end{bmatrix} \quad \text{with } Q \in R^{m \times (m-q)} \quad (2.35)$$

where

$$U^{-1} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad \text{with } U_1 \in R^{q \times m} \text{ and } U_2 \in R^{(m-q) \times m}.$$

By pre-multiplying both sides of measurement equation (3.32) by  $U^t$ ,

$$U_1 y = U_1 CN \bar{x}_1 + \bar{x}_2 \quad (2.36)$$

$$U_2 y = U_2 CN \bar{x}_1 \quad (2.37)$$

and by substituting (2.36) into (2.33) and combining it with (2.37), it is possible to obtain the system

$$\dot{\bar{x}}_1 = \tilde{A}_1 \bar{x}_1 + \bar{B}_1 u + E_1 y \quad (2.38)$$

$$\bar{y} = \tilde{C}_1 \bar{x}_1 \quad (2.39)$$

where

$$\tilde{A}_1 = \bar{A}_{11} - \bar{A}_{12}U_1CN, \quad E_1 = \bar{A}_{12}U_1,$$

$$\tilde{C}_1 = U_2CN, \quad \bar{y} = U_2y.$$

If the pair  $\{\tilde{A}_1, \tilde{C}_1\}$  is observable or detectable, the conventional Luenberger observer design procedure can be applied :

$$\dot{w} = (\tilde{A}_1 - L\tilde{C}_1)w + \bar{B}_1u + L^*y, \quad w \in R^{n-q} \quad (2.40)$$

where  $L \in R^{(n-q) \times (m-q)}$  and  $\dot{L} = LU_2 + E_1$ . With appropriate choice of  $L$ ,  $w \rightarrow \hat{\hat{x}}_1$  as  $t \rightarrow \infty$  and we have

$$\hat{x} = T\hat{\hat{x}} = T \begin{bmatrix} w \\ U_1y - U_1CNw \end{bmatrix} \quad (2.41)$$

where  $\hat{\hat{x}} \rightarrow x$  as  $t \rightarrow \infty$ .

Furthermore, using the result obtained above, it is possible to estimate the unknown input.

$$\hat{d} = U_1\dot{y} + G_1w + G_2y + G_3u \quad (2.42)$$

where

$$G_1 = U_1CNLU_2CN + U_1CN\bar{A}_{12}U_1CN$$

$$- U_1CN\bar{A}_{11} - \bar{A}_{21} + \bar{A}_{22}U_1CN$$

$$G_2 = -U_1CNLU_2 - U_1CN\bar{A}_{12}U_1 - \bar{A}_{22}U_1$$

$$G_3 = -U_1CN\bar{B}_1 - \bar{B}_2.$$

### UIO existence conditions

The existence condition of UIO for the system described by equation (2.40) can be summarized as

$$(a) \text{ rank } CD = \text{rank } D$$

$$(b) \text{ rank } \begin{bmatrix} sI_{n-q} - \bar{A}_{11} & -\bar{A}_{12} \\ CN & CD \end{bmatrix} = n$$

The proof of these conditions can be found in [101]. Instead of proving the above condition, the physical interpretation of the conditions will be presented in the following paragraphs.

The first condition (a) states the required number of measurements and quantities which needs to be measured. To satisfy this condition, the number of measurements in the output vector  $y$  must be equal or greater than the number of unknown input vector  $d$  and that part of the state which is directly coupled to the unknown inputs must be obtainable from the measurement vector.

The second condition (b) states that the transmission zeros of the triple  $(C, A, D)$  must be stable. This condition stems from the fact that the UIO performs an implicit inversion of the system and the transmission zeros of the system become the poles of the UIO. Therefore, to obtain a stable UIO, the transmission zeros of the system should be stable.

The UIO finds direct application in FDI schemes in that it can be used to estimate unknown (additive faults) inputs. Also, in [108] it is shown that the UIO can be used in a multiple decoupling scheme such as the one depicted in Figure 2.10.



In model-based diagnosis the system model used is, where possible, linear, to minimize the computational complexity. However, certain types of static nonlinearities can be easily incorporated into system models with little increase in computing time. Due to the presence of noise, unmeasured disturbances and uncertainty in the system model, the residuals generated can be nonzero even in the absence of any faults. Thus in order to avoid the possibility of false alarms it is necessary to set some threshold values for the residuals. If the residuals lie within these threshold values then it is assumed that there is no fault. Threshold selection may be done by either analytical or empirical means depending on the complexity of the application. Improper selection of thresholds might lead to false alarms and missed detection. Therefore, knowledge-based diagnosis methods could either replace or complement model-based diagnosis methods [100].

#### **2.4.2. Knowledge-Based Methods**

The uncertainty and imprecision that are inherent in model-based methods can be adequately managed using knowledge-based methods. The basis of this method is the compilation of subjective knowledge about the system under study that helps making inferences about the problem at hand. There are two schemes that will be the focus of this section, these are:

- Expert systems
- Fuzzy systems

#### **Expert Systems**

An expert system is a computer program that captures human expert knowledge and reasoning process in dealing with complex problems in order to solve them expertly [31].

There are many types of expert systems, classified according to their application, knowledge representation mechanism, inference methods, or special features. One classification according to their applications is summarized in [27] as shown in Table 2.3.

<i>Category</i>	<i>Problem Addressed</i>
Interpretation	Inferring situation descriptions from sensor data
Prediction	Inferring likely consequences of given situations
Diagnosis	Inferring system malfunctions from observables
Design	Configuring objects under constraints
Planning	Designing actions
Monitoring	Comparing observations to plan vulnerabilities
Debugging	Prescribing remedies for malfunctions
Repair	Executing a plan to administer a prescribed remedy
Instruction	Diagnosing, debugging, and repairing student behavior
Control	Interpreting, predicting, repairing, and monitoring system behavior

Table 2.3 Generic Categories of Knowledge-Based Methods Applications [27]

Expert systems have had great success in the medical domain. One of the first applications of expert systems in medicine is MYCIN [26]; it deals with infectious disease diagnosis and therapy selection. Other medical applications like NEOMYCIN and CASNET among many others are covered in detail in [33]. Moreover, expert systems that

diagnose malfunctions of technological devices and systems are mentioned in [17] and [33]. Three main parts are used to construct an expert system according to [24]:

- a. Inference engine
- b. Knowledge base
- c. Working memory

While the knowledge-base contains the expert domain knowledge, the working memory stores the information gained from the user and is also used as a scratch pad. Finally, the inference engine uses both the expert knowledge and the user information to provide an expert solution [24]. Rules, semantic networks, and frames are the mostly used methods of knowledge representation [24]. Production rules in the form *IF...THEN...* were used as the scheme for knowledge representation in MYCIN. The other form of knowledge representation is semantic networks, where nodes are linked with arcs of different weights, these were used in CASNET as a general tool for building expert systems for the diagnosis and treatment of diseases whose mechanisms are well understood [32]. While the frames are data structures that represent stereotyped situations and are made up of a set of slots that may contain procedures, data, or pointers to other frames [24].

The last part of the knowledge-base system is the inference engine, which determines what knowledge is to be used and outputs recommendations and procedures to be followed. Backward and forward chaining are the main modes of operation of the inference engine. While forward chaining starts from the data and ends with the

conclusion, backward chaining performs in an opposite fashion. Furthermore, some experts systems use the combination of forward and backward chaining.

Most expert systems deal with the uncertainty and the imprecision in the knowledge base by introducing certainty factors, which are probabilistic. Thus uncertainty in the conclusion is introduced, leading to a final conclusion that is in doubt [21]. Therefore, certainty factors must be included in the design of the expert systems as possibilistic descriptions in the form of fuzzy rather than crisp numbers. For an overview of fuzzy systems, refer to Appendix A.

There are issues that cannot be dealt with in conventional expert systems: fuzziness of antecedent and/or consequent in the rules; partial mismatch between the antecedent of a rule and a fact supplied by the user; and the presence of fuzzy quantifiers in the antecedent and/or consequent [21]. The introduction of uncertainty in expert system using fuzzy logic makes it possible to deal efficiently with these different issues when conventional techniques stand short. The role of fuzzy logic system in dealing with uncertainty and imprecision in expert systems is discussed in detail in the next section.

### **Fuzzy Systems**

Most of the facts and rules in expert systems contain fuzzy antecedents and thus they are fuzzy propositions. Fuzzy logic provides a natural conceptual framework for the analysis and design of expert systems, because its main purpose is to provide a systematic basis for representing and inferring from imprecise rather than precise knowledge. The main features of fuzzy logic which are relevant to the management of the uncertainty in

expert systems is covered in more detail in [21], where these features are summarized as follows:

1. A proposition  $p$  can have an intermediate truth-value over the fuzzy subset of  $T$  ( $T$  finite or infinite truth-value set).
2. The predicates may be crisp or more generally fuzzy, e.g. *low high, tall*, etc.
3. Fuzzy logic allows the use of fuzzy quantifiers exemplified by *most, many, few*, etc.
4. Fuzzy logic provides a method of representing the meaning of both non-fuzzy and fuzzy predicate-modifiers exemplified by *not, very, less*, etc.
5. Three principal modes of qualification are: truth qualification, probability qualification, and possibility qualification.

In practice most of the time we need to associate levels of confidence with rules, antecedents and consequents. The inference mechanism should be able to compute the confidence of the consequent given the confidence values of the antecedent and the confidence of the rule, this is where fuzzy reasoning systems are of great value [17]. This allows for the concise representation of a human expert's knowledge about the diagnosis task. Relevant references in the area of analysis of fuzzy information can be found in [21]- [23]. An overview of fuzzy systems is covered in Appendix A.

## **2.5. HIERARCHICAL STRUCTURE**

In a complex system with many subsystems and components interacting to produce a desired output, the malfunction of any of these subsystems causes the output of the

system to deviate from the desired value; thus affecting the performance of the plant. The diagnostic scheme suggested in this study explores the functional structure of these subsystems and components and how they interact when organizing them for the purpose of diagnostics. In particular, the hierarchical structure shown in Figure 2.12 is used to probe for all potential faults, where the dashed line represents the interacting subsystems or components; this interaction can take place only among subsystems or components at the same level. Therefore, a reduction of the number of entries in a fault-symptom look-up table can be achieved. In this structure the process of searching for the suspect component is simplified. The example of Figure 2.12 is referred to the case of interest in this dissertation, i.e. the diagnosis of an internal combustion engine.

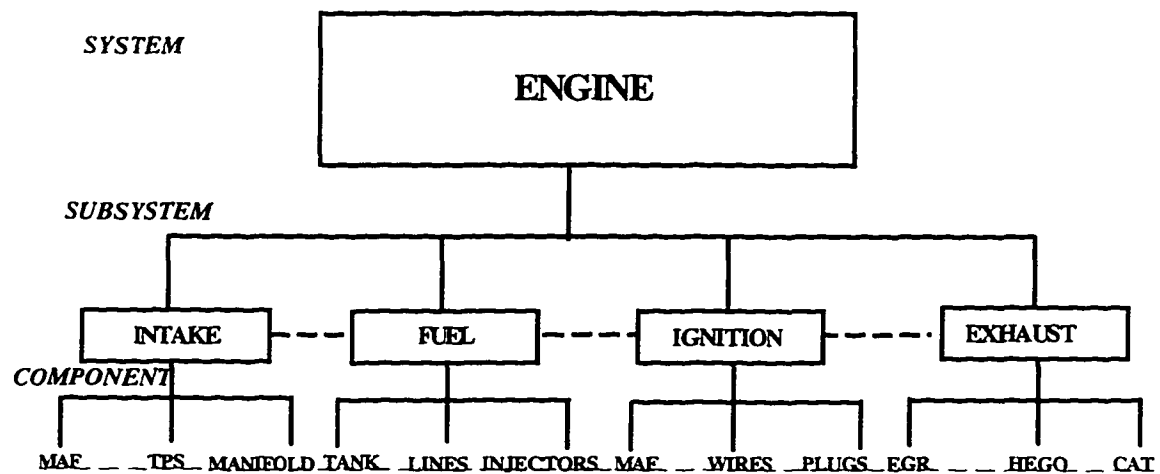


Figure 2.12 Hierarchical System Structure

The automotive powertrain is a very complicated physical processes, especially because it operates in a harsh and rapidly changing environment. There exists no universal powertrain model, which could describe the evolution of the whole process, and there is no unified algorithm which could handle all kinds of faults. Therefore, decentralized diagnostic strategies, based on a variety of models and diagnostic algorithms, must form the basis of current and future generations of integrated powertrain diagnostic system.

Breaking down the system into different physical models or sub-models based on the faults to be diagnosed will increase redundancy and reduce computations. Furthermore, according to different models, different algorithms can be applied to detect faults. In particular, some faults cannot be detected by a single physical model alone, but will affect the output of several small models or measured outputs. Therefore, faulty signals relating to this single fault will all be sent to the higher level, called the supervisory level, which uses fuzzy logic rules to make a final decision.

The methodology used to construct a diagnostic system capable of integrating quantitative models with qualitative reasoning will be discussed in detail in Chapter 4.

## **2.6. SUMMARY**

In the present chapter, an overview has been given of the relevant background, to explain how different diagnostic methodologies may be used in the context of the automotive engine exhaust emission controls system diagnostics and a hybrid approach was proposed as the diagnostic scheme. A hierarchical diagnostic method that explores

the functional structure of the automotive engine subsystems and components was suggested. In the following chapter the experimental work that was carried during this work will be explained in detail.



## CHAPTER 3

### EXPERIMENTAL SETUP

In this chapter the experimental setup and procedures are described in detail, including the engine test cell and the hardware and software used for running the experiments and collecting data. Moreover, analyses are conducted to determine the size of a suitable sample space for the experiments, and a diagnostic test cycle is designed. Sample runs are shown, and the induced faults are tabulated. The baseline experimental data are displayed, setting the stage for the results described in Chapter 5.

#### 3.1. EXPERIMENTAL FACILITIES

The experimental verification of the diagnostic techniques proposed in Chapter 2 was performed in one of the engine test cells at the Center for Automotive Research housed at the Ohio State University. This section describes the facilities, the instrumentation, and the experimental procedures relevant to this study. A Ford 4.6 L V-8 engine was loaded by an electric dynamometer and the throttle was controlled using a digital throttle controller. The engine was equipped with both production and laboratory grade sensors and actuators. A test cycle that covers a driving profile of acceleration and

deceleration, was simulated. Also, a Fourier Transform Infra Red (FTIR) gas analyzer was used to analyze the concentration of the raw exhaust components, and data from the sensors and actuators were acquired simultaneously with the exhaust gas analysis. The following is the experimental set-up in details.

### **3.1.1. Engine Test Cell Instrumentation**

#### **Engine**

A Ford 4.6 Liter engine with a V8 configuration and with the specifications shown in Table 3.1 was under investigation for this work. The following production sensors were available for the experiments; Air Charge Temperature sensor (ACT), Throttle Position Sensor (TPS), Mass Air Flow sensor (MAF), and two Heated Exhaust Gas Oxygen sensors (HEGO) which are located in the exhaust manifold upstream of the catalytic converter. The engine was controlled by the production Electronic Engine Controller (EEC-IV) which outputs the needed injector pulse width and spark timing according to the load and speed conditions. A commercial break-out box was used to link the engine EEC-IV controller to the data acquisition system thus making all the input/output data from the ECC-IV accessible. A laboratory grade manifold absolute pressure sensor (MAP) SENSOTEC model V/1945-02 with a 0-15.0 psi vacuum output range was used to measure the absolute manifold pressure. Also, a Universal Exhaust Gas Oxygen sensor (UEGO) NGK NTK model MO-1000 was installed in the vicinity of the HEGO sensor to give a linear air/fuel ratio output with  $13 < \text{air/fuel} < 16$  measurement range and a response time of 0.1 sec. The engine speed was measured using a digital Magnetic zero-speed sensor (Electro Corporation) model DZ200SLE and a frequency-to-

voltage converter. The fuel mass flow rate was calculated using a custom built circuit that transforms the pulse width to an output voltage proportional to fuel mass flow. Detailed specifications, circuitry, and calibration curves of the previously mentioned sensors are given in Appendix B.

Displacement	4.6L
Number of Cylinders	V8
Bore	$90.2 * 10^{-3} \text{ m}$
Stroke	$90.0 * 10^{-3} \text{ m}$
Firing Order	1-3-7-2-6-5-4-8
Combustion Chamber Volume	$49-51 * 10^{-6} \text{ m}^3$
Connecting Rod Length	$150.7 * 10^{-3} \text{ m}$

Table 3.1 Engine Specifications

### Electric Dynamometer

A General Electric DC dynamometer model TLC-50 with a capacity of 100 HP/3500RPM was controlled by a DyneSystems DYN-LOC-IV dynamometer speed/torque controller in conjunction with a DTC-1 throttle controller. The throttle controller, acting on a pulse width modulated solenoid throttle actuator ACT-148, allows excellent regulation of throttle position by means of closed-loop feedback, permitting simultaneous load and speed control by acting as a slave to the dynamometer controller. A SENSOTEC load cell model 41 with a capacity of 1000 lbs provided the measurement

for average and instantaneous load torque at the dynamometer cradle, and a magnetic pick up sensor, mounted above a 60 teeth gear, provided an angular measurement from which dynamometer speed was calculated by the dynamometer controller.

### **Data Acquisition**

The data acquisition hardware consisted of A National Instruments SCXI-1000 Chassis with a SCXI-1328 Module, which allowed the acquisition of 32 channels simultaneously. The data acquisition Input/Output device is an AT-MIO-16E-2 with a maximum sampling rate of 500 ks/s with 8 digital I/O channels housed in a Pentium 200 MHz based PC. The software used was LabVIEW™ was externally triggered using the cylinder identification pulse (CID) and a scanning rate using the Profile Ignition Pickup (PIP). Moreover, 15000 samples were acquired to cover the test cycle.

### **Gas Analyzer**

A Nicolet REGA 7000 FT-IR analyzer was used for the analysis of the raw exhaust gases which was carried out on a Macintosh 68040 25MHz desktop computer. The data were collected and sampled at a rate of one sample per second real-time. The analyzer is capable of analyzing up to 40 exhaust components per test method and up to 50 user-definable test cycles. Figure 3.1 shows the experimental set-up.

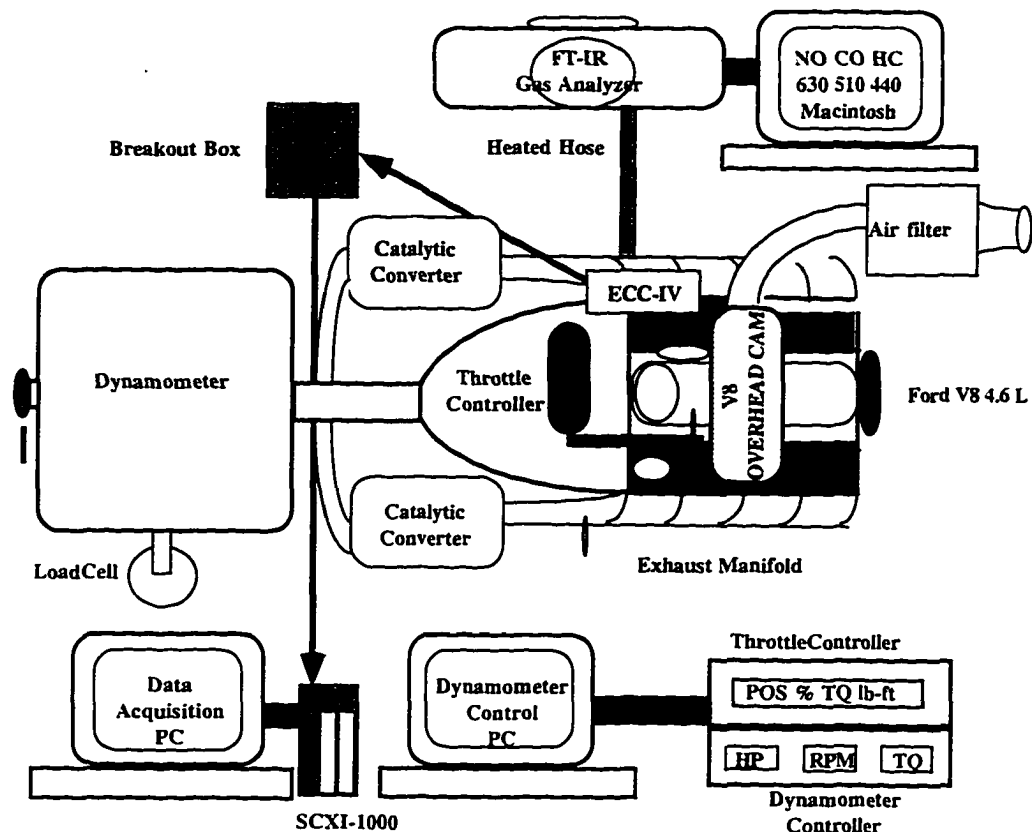


Figure 3.1 Experimental Setup

### 3.2. DESIGN OF EXPERIMENTS

The experiments were carefully designed to satisfy the following objectives:

1. Statistically significant sample size.
2. Repeatability of the tests.
3. Satisfactory test cycle.
4. Real life induced faults.

### 3.2.1. Sample Size

The data acquisition hardware and software described in the previous section allows the operator to collect data for 120 seconds of gas concentration  $C(\theta)$ , where  $\theta$  is the component. This corresponds to 15000 points when using the PIP signal to set the scanning rate. In order to determine the size of the sample that would yield a statistically meaningful result, pilot experiments were run as shown in Figure 3.2-3.4. The objective of the experiment is to identify the baseline levels for the three components. The sources of variation were mainly the ambient conditions and the combustion cyclic variation. The number of observations needed was decided by the use of the chebyshev's inequality [140]. For the random variable  $X$  with a finite second moment, that is, let

$$E[|X|^2] < +\infty, \quad (3.1)$$

then

$$P[|X - E[X]| \geq \varepsilon] \leq \frac{\text{var}(X)}{\varepsilon^2} \quad (3.2)$$

where  $P$  is the probability, and  $E$  is the expected value, equation (3.2) holds for any  $\varepsilon > 0$ .

let

$$X = \frac{1}{N} \sum_{i=1}^N C^{(i)}(\theta) - E[C(\theta_n)] = \hat{C}(\theta_n) - \bar{C}(\theta_n), \quad (3.3)$$

where  $N$  is the sample size and the estimator for  $\bar{C}(\theta_n)$  is the sample mean.

Then, if we interpret  $\varepsilon$  as

$$\varepsilon = k\sigma \quad (3.4)$$

where  $\sigma$  is the standard deviation of  $C(\theta_n)$ , and for any positive constant  $k$  we can write

$$\begin{aligned}
 P\left[|\hat{C}(\theta_n) - \bar{C}(\theta_n)| > k\sigma\right] &\leq \frac{E[\hat{C} - \bar{C}]^2}{k^2 \sigma^2} = \frac{\text{var}[\hat{C}(\theta_n)]}{k^2 \sigma^2} \\
 &= \frac{\text{var}[C(\theta_n)]}{Nk^2 \sigma^2} = \frac{1}{Nk^2}.
 \end{aligned} \tag{3.5}$$

Thus, if we wish our estimate of  $\bar{C}(\theta_n)$  to be within  $k$  standard deviations of the true mean with probability  $x$ , the required sample size will be

$$N = \frac{1}{k^2 x} \tag{3.6}$$

If the original population distribution is not normal and if  $N$  is large ( $N > 30$ ), the distribution for  $\bar{C}(\theta_n)$  is normal (central limit theorem [140]). For a confidence level of 90% and for  $k = 0.5$  with probability of error of estimation 0.1, the number of samples to be collected according to equation (3.5) is  $N = 40$ . In this study the number of a samples  $N$  was chosen to be greater than forty in all cases, except for the case of misfire we chose  $N = 40$  to avoid damaging the catalytic converter. The number of required runs were completed which provided adequate sample size for both healthy and faulty engine conditions. This was achieved by running the test cycle on a daily basis while collecting data using both the exhaust emission analyzer and the data acquisition system simultaneously. Then faults were introduced and the test was conducted several times for each fault induced. The test cycle used for the experiment work will be covered in detail in Section 3.2.3.

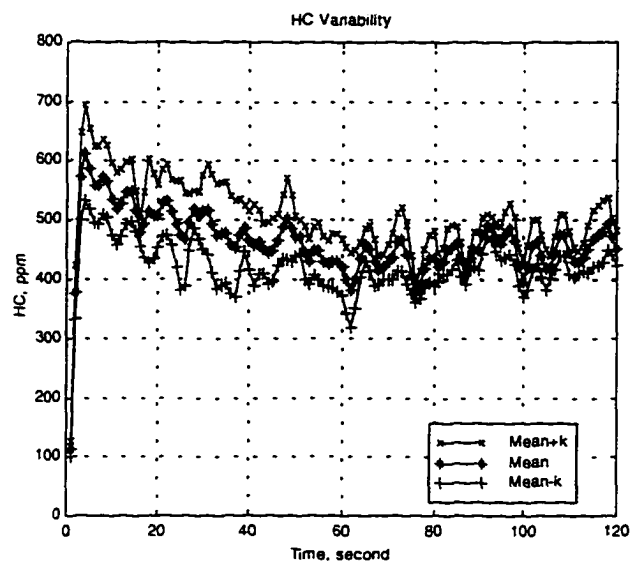


Figure 3.2 HC Concentration Variability

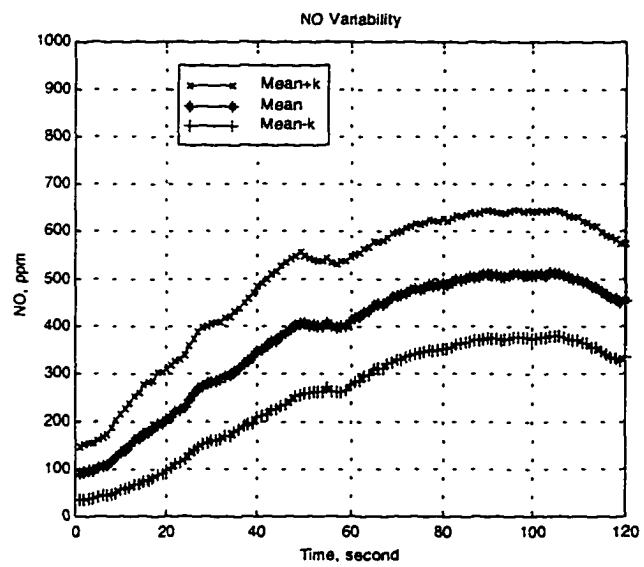


Figure 3.3 NO Concentration Variability



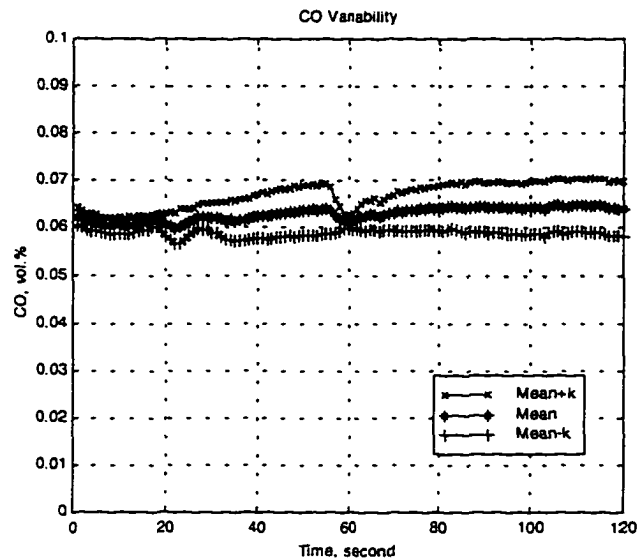


Figure 3.4 CO Concentration Variability

### 3.2.2. Test Repeatability

To achieve a repeatable test, all running conditions were maintained identical over the whole testing period except for the ambient conditions. The engine operating temperature of 190°F was reached before starting the test, and the FT-IR analyzer hose, filter, and gas cell were heated up to the recommended temperatures. The variable ambient environment temperature and relative humidity were taken into account as interfering inputs when setting the baseline data in order to avoid false alarms and missed detection.

### 3.2.3. Test Cycle

The test cycle was chosen in such a way that it represented a typical urban driving profile. The cycle consisted of acceleration, deceleration, and steady state load and speed conditions which covered a wide range of the engine's operating conditions. This cycle, as any other cycle, can be simulated as shown in Figure 3.5, using the DTC to control the throttle position; also, the load profile shown in Figure 3.6 was controlled by the DYNLOC dynamometer controller and was calculated using the road load equation (3.6) [57]:

$$P_r = [2.73C_r M_v + 0.0126C_d A_v S_v^2] S_v * 10^{-3} \quad (3.7)$$

where:

$P_r$  = Road- load power (kW).

$S_v$  = Vehicle speed (km/h).

$C_d$  = Drag coefficient.

$C_r$  = Coefficient of rolling resistance.

$A_v$  = Frontal area of vehicle  $.(m^2)$

$M_v$  = Mass of vehicle (kg).

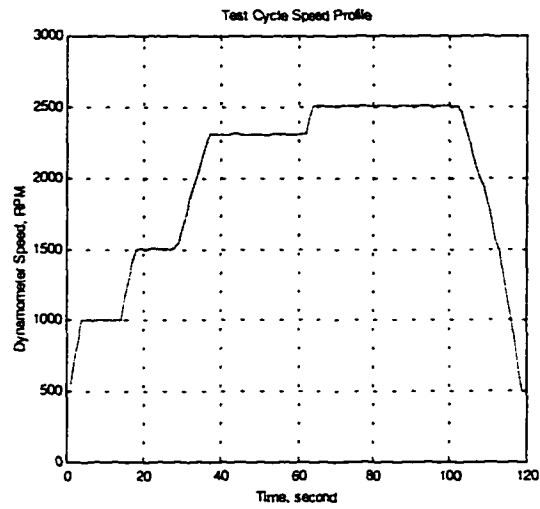


Figure 3.5 Test Cycle Dynamometer Speed

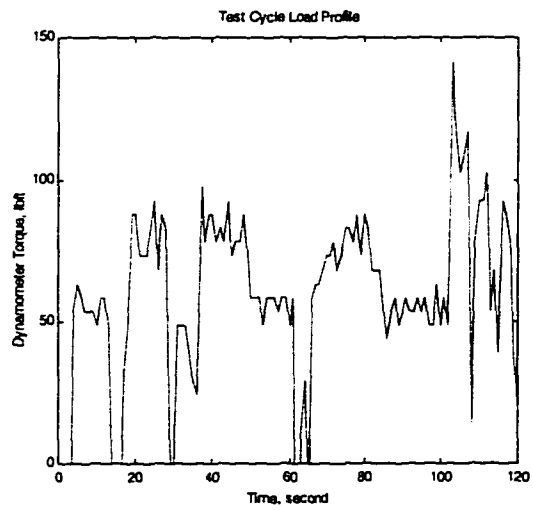


Figure 3.6 Test Cycle Dynamometer Torque

### 3.2.4. Induced Faults

A catalog of real life faults that are typical for the engine under investigation was obtained from the field; some of these faults are tabulated in Table 3.2. Furthermore, intermittent misfire conditions were generated using a custom ignition shut-off circuit. The misfire circuit interrupts the ignition command signal to cylinder number one at an adjustable rate of cycles per misfire. Also, intake manifold air leak was induced. The sensor calibration circuit allowed the induction of multiplicative fault to the sensor under investigation. Appendix B includes more information about these circuits and the nature of their operation.

<i>Sensor/Actuator</i>	<i>Fault</i>
Throttle Position (TPS)	Calibration error (25%-75%)
Mass Air Flow (MAF)	Calibration error (25%-75%)
Intake Manifold	Leak
Exhaust Gas Recirculation (EGR)	Stuck Open or Closed
Injector	Stuck Closed

Table 3.2 Induced Faults

### 3.2.5. Experimental Procedures

The following procedures were followed in all experiments.

1. The engine was warmed-up to its operating temperature ( $88^{\circ}\text{C}$ ).
2. The FT-IR sample hose was heated to a temperature of ( $165^{\circ}\text{C}$ ).
3. The FT-IR sample filter was heated to a temperature of ( $165^{\circ}\text{C}$ ).
4. The FT-IR gas cell was heated to a temperature of ( $165^{\circ}\text{C}$ ).
5. A background of ambient air was collected to set the FT-IR bench.
6. The bench was aligned and was tuned automatically by the Rega 7000 software.
7. The transmission was shifted to the drive gear.
8. Engine dynamometer room temperature was maintained between  $21^{\circ}\text{C}$ - $27^{\circ}\text{C}$ , and relative humidity was in the range of 30%-70%.
9. Faulty conditions were induced.
10. The dynamometer was set in the speed mode, while the throttle controller was set in the position mode.
11. The simulated cycle torque-speed data set was down-loaded from the control computer while simultaneously output data from the different sensors and actuators were acquired, and the raw exhaust gases were analyzed using the Nicolet FT-IR at a sampling rate of one sample per second.

### 3.3. BASELINE EXPERIMENTAL DATA

During the period of study experimental data were acquired on a daily basis to guarantee a representative baseline measurement. This was achieved by acquiring the sensors', actuators', and exhaust gases' data twice per day for the healthy engine, and similarly for the faulty conditions. Figures 3.7-3.12 show the typical output of the sensors, actuators. This was a principal factor in determining the thresholds and membership function values for the applied algorithms.

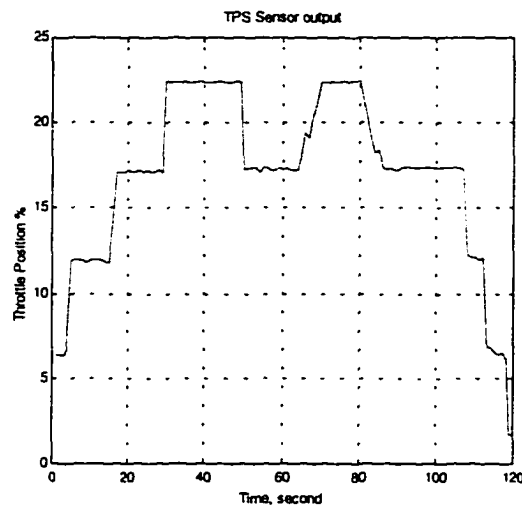


Figure 3.7 TPS Sensor Ouput

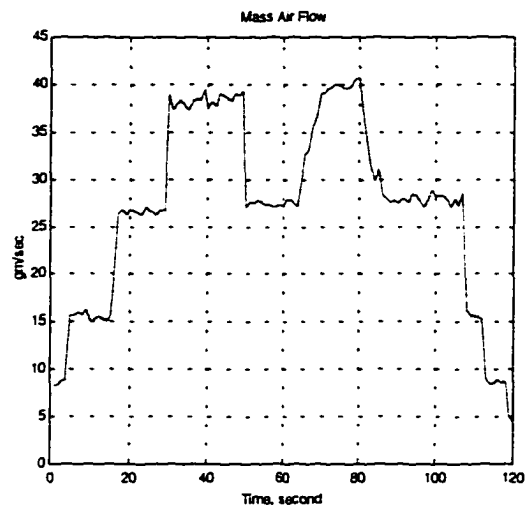


Figure 3.8 MAF Sensor Output

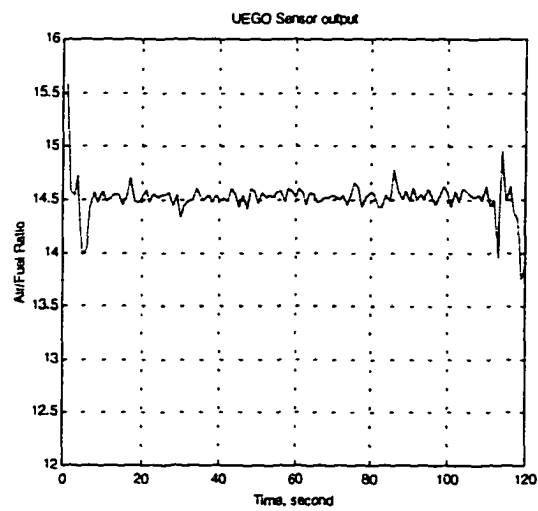


Figure 3.9 UEGO Sensor Output

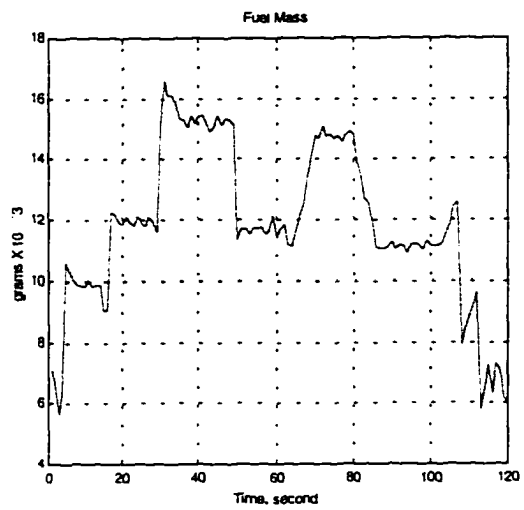


Figure 3.10 Calculated Fuel Mass

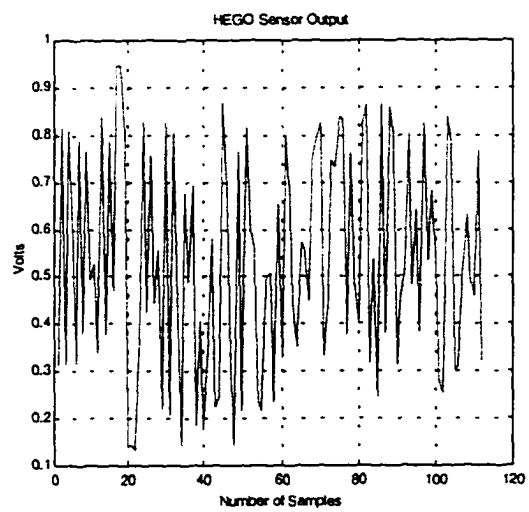


Figure 3.11 HEGO Sensor Output



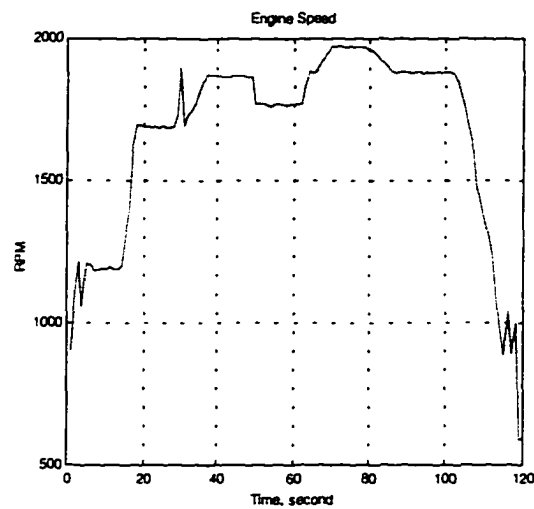


Figure 3.12 Engine Speed

### 3.4. SUMMARY

In this chapter the experimental setup used in this study was covered. The engine cooling, exhaust, and fueling systems were designed and installed by the researcher. Moreover, the engine mounts and output shaft connecting the engine to the dynamometer were modified for the laboratory setup. The test cycle used in this application was designed carefully to cover a wide range of engine load and speed that was adequate for the diagnostic application and comparable to the national and international test profiles used for emission testing. The fault inducing circuit and misfire circuit were designed and built in the electronics laboratory. The data acquisition system was capable of data collection in both the time and the crank-angle domain. Furthermore, the faults induced

and the test procedures were covered in this chapter. The pilot experiments and the choice of the sample size was explained. The experiments were carried on daily basis for a period of 12 months. The methodology used in this study is covered in the next chapter.

## CHAPTER 4

### DIAGNOSTIC METHODOLOGY

A methodology for the integration of subjective (heuristic) and objective (analytical) knowledge for fault diagnosis and decision-making is covered in this chapter. The structure, challenges, and benefits of such integration are explored. When a fault occurs in a dynamic system, the observed outputs deviate from their nominal value thus affecting the performance of the system. These observed *symptoms* are measured using a wide range of sensors and typically these measurements are sampled over different time scales and knowledge from various sources is utilized. The fusion of such sensor measurements, the combination of different time scale data, and the integration of the knowledge sources is a serious challenge that must be overcome by the integration scheme. Furthermore, the choice of the thresholds to account for model error and noise is a crucial issue. These issues are addressed within the frame of integration of different diagnostic methods in the following sections. In this methodology the system's input-output data follow two paths that are in parallel. Along the first path a model-based scheme is applied, where primary residuals are generated using sliding mode nonlinear observers. These residuals are evaluated and processed using fuzzy logic for fault

detection and isolation. Meanwhile along the second path a knowledge-based scheme is applied using a combination of fuzzy estimation techniques and fuzzy heuristic rules for fault detection and isolation. The two paths merge together and integration of both schemes is completed in the inference mechanism. Finally a fuzzy decision-making process yields the diagnosis. The methodology is covered in the following sections, where the model-based scheme is introduced first in Section 4.1., and followed by the fuzzy evaluation technique in Section 4.2. Then the knowledge-based fuzzy estimation scheme follows in Sections 4.3. The integration is covered in Section 4.4 and the decision-making process is explained in Section 4.5.

#### **4.1. FAULT DETECTION AND ISOLATION (FDI)**

Objective knowledge, such as a mathematical model that describes the dynamics of a system, combined with subjective knowledge, such as heuristic rules, are used to detect and isolate the faults that occur in the system under investigation. For complex processes, such combination of quantitative and qualitative knowledge, accounts for the incomplete knowledge of the process [100]. The architecture of a fault diagnosis system that combines both strategies is described in [100] and is shown in Figure 4.1. The main components of this fault detection system are the knowledge base, the data base, the inference engine, and the explanation component. The combination of the heuristic knowledge and the analytical knowledge is performed by the inference engine. The inference engine has access to the analytical knowledge, the heuristic knowledge, and the actual data.

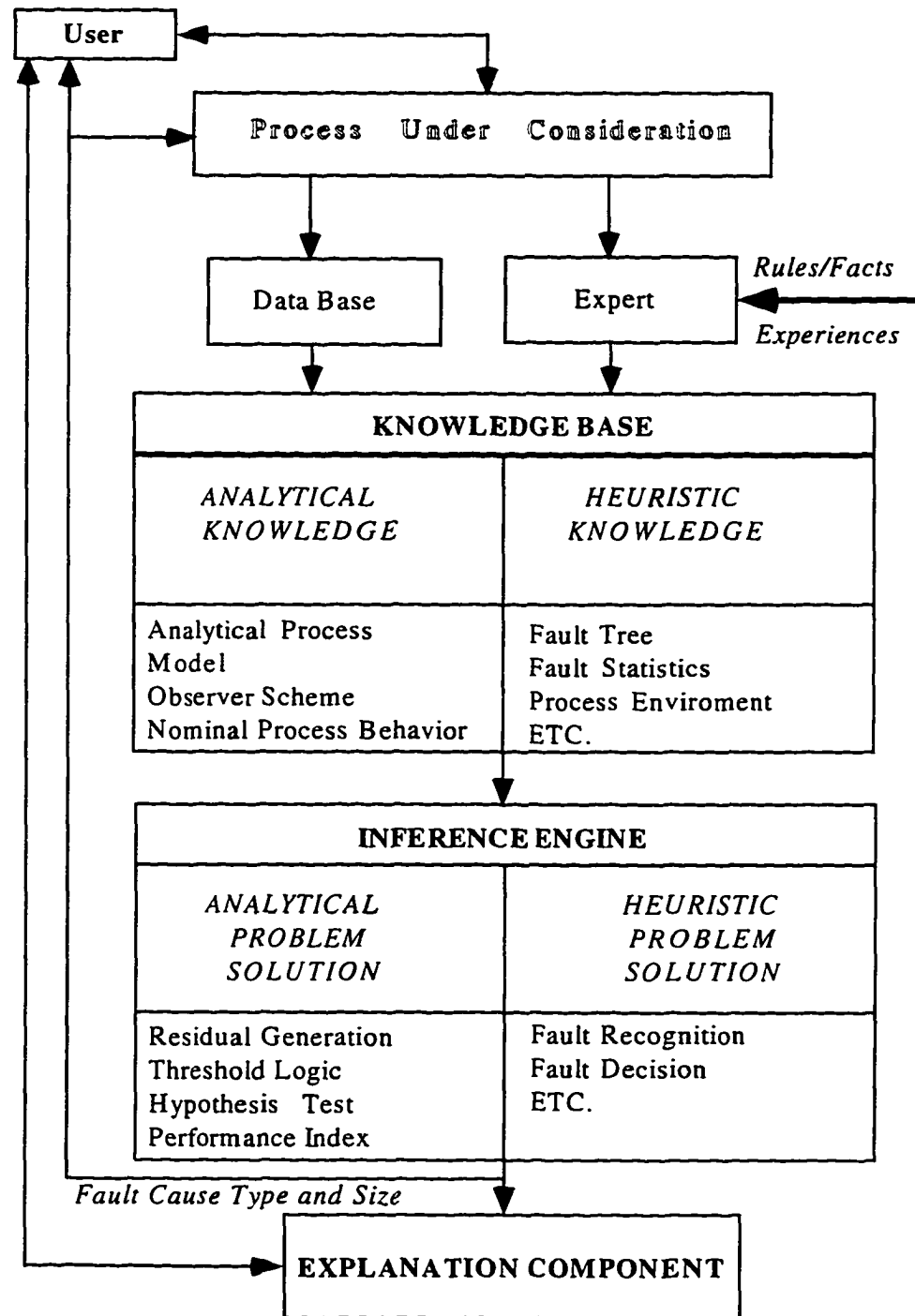


Figure 4.1 Architecture of a Model- and Knowledge-Based Fault Diagnosis System [100]

The general structure of a diagnostic system, that combines model- and knowledge-based models was suggested in [123] and is shown in Figure 4.2. In this structure primary residuals are generated using a model-based residual generator. These primary residuals are then evaluated and secondary residuals are generated. These secondary residuals are processed and the diagnosis is made by the diagnostic supervisor.

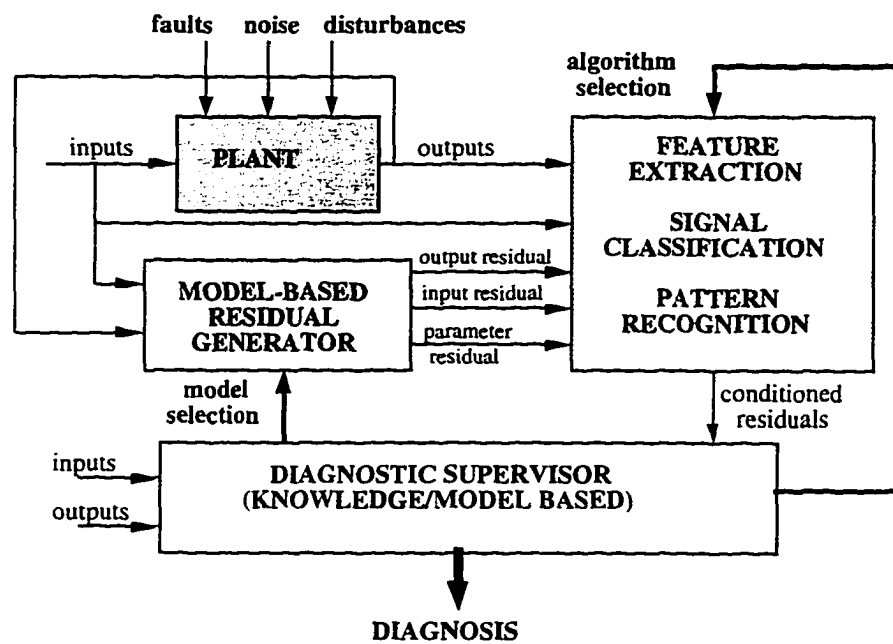


Figure 4.2 General Structure of a Diagnostic System

There are different methods used to process the data available from the distinct forms of knowledge mentioned previously; these are summarized by [72] as follows;

- a. Limit value checking; where thresholds are set *a priori* and the direct measured signals are checked against these thresholds.
- b. Signal analysis; where measured signals are manipulated using correlation functions, frequency spectra, auto regressive moving average models, etc.
- c. Process analysis; where measured signals and mathematical models are used with parameter estimation, observers, or filters to generate residuals that are checked against fixed or adaptive thresholds.

Moreover, subjective knowledge from human experts can be used to produce *heuristic symptoms*; these symptoms can be represented as linguistic variables and vague numbers [72]. When uncertainty is introduced in the diagnostic scheme the use of non-Boolean logic inference becomes evident. The following are the most common non-Boolean reasoning methods as listed in [17]:

- a. Bayesian inference
- b. Fuzzy logic
- c. Theory of confirmation
- d. Shafer-Dempster theory
- e. Theory of usuality

To avoid the effects of uncertainty and disturbance, fuzzy logic is a tool that may be used for fault decisions [119]. The use of an adaptive threshold that varies according to the control activity of the process can reduce the rate of false alarm and missed detection [119].

In the scheme proposed in this study the plant's inputs and outputs are compared against those obtained analytically. The discrepancies between the measured and calculated values (primary residuals) are generated using model-based methods, where a bank of sliding mode observers is used to generate primary residuals. The primary residuals are evaluated using fuzzy evaluation techniques which uses heuristic knowledge about the system. Meanwhile, a fuzzy estimator generates a residual that is used for fault detection. A fuzzy decision-making supervisor, which uses the evaluated residuals from the model-base, the heuristic knowledge, and the plant's inputs and outputs to diagnose the fault. The fuzzy logic system can handle both analytical and heuristic symptoms, which can be embedded in the fuzzy diagnosis system [110]. The diagnostic scheme proposed in this study is shown in Figure 4.3. In the following sections the model- and knowledge-base diagnostic methods applied are covered. Also, the fuzzy residual generator and evaluation method are described in detail.



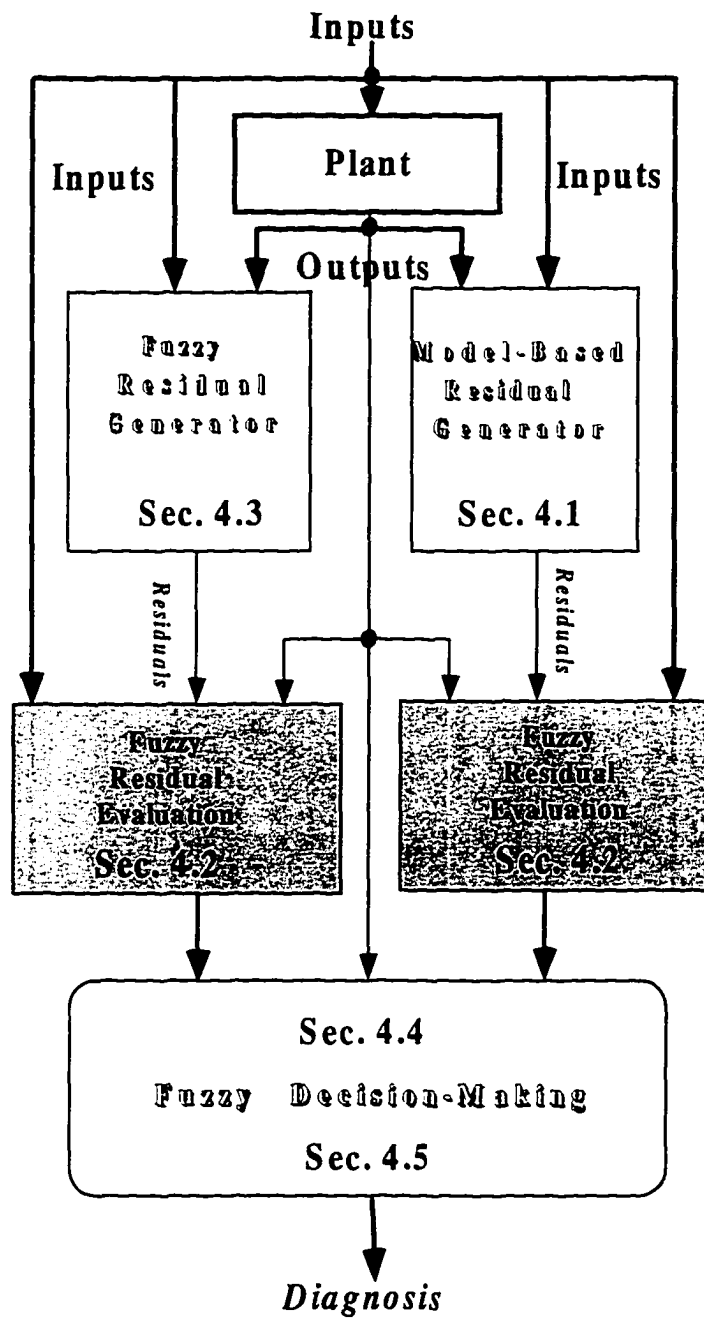


Figure 4.3 Diagnostic Scheme

#### 4.1.1. Model-Based FDI:

The general system and fault representation is shown in Figure 4.4, where a block diagram representation of a dynamic system is depicted. The variables and their description are also listed in Table 4.1.

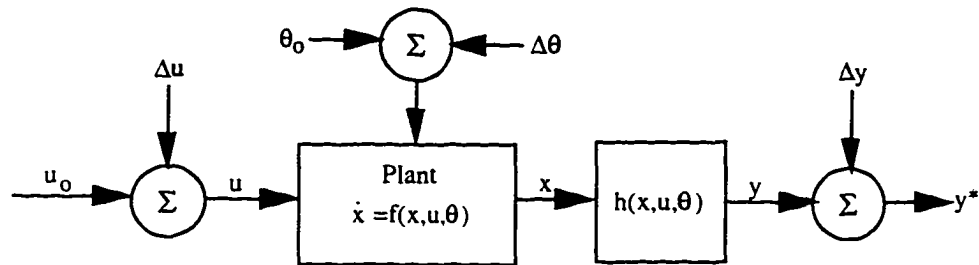


Figure 4.4 General Dynamic System and Fault Representation

No.	Variable	Description
1.	$u_o$	Input vector
2.	$\Delta u$	Input fault vector
3.	$\theta_o$	Nominal parameter vector
4.	$\Delta \theta$	Parameter (component) fault vector
5.	$x$	State Vector
6.	$f(\cdot, \cdot, \cdot)$	State evolution vector field
7.	$h(\cdot, \cdot, \cdot)$	Output measurement function
8.	$y$	Actual output vector
9.	$\Delta y$	Output fault vector
10.	$y^*$	Measured output vector

Table 4.1 System Variables and Faults

In the description of Figure 4.4, the system is described by the state equations

$$\dot{x} = f(x, u, \theta)$$

$$y = h(x, u, \theta).$$

In the above equations, faults are represented by deviations in the nominal values of commanded or measure inputs (i.e.:  $u \rightarrow u + \Delta u$ ), measured outputs ( $y \rightarrow y + \Delta y$ ), and system parameters ( $\theta \rightarrow \theta + \Delta \theta$ ).

In Chapter 2 a distinction was made between additive and multiplicative faults. In the more general nonlinear case, such a distinction is not as clear, and we will not attempt to make it. In the present chapter we make use of an approach developed in [136] and [133] based on nonlinear observers, and in particular those of the sliding mode type, to generate residuals that will then see further processing in the way of fuzzy evaluation.

### **Residual Generation**

Objective knowledge, such as that provided by model-based FDI methods can be described by the block diagram shown in Figure 4.5. Given a mathematical model of the plant that captures the dynamics of the system, the observations acquired through sensory measurements are compared to those values analytically obtained for the same variables. The resulting discrepancy between the measured and calculated variable is called the primary residual. This step generates a residual which reflects the system's behavior, and which has nominally zero value under healthy conditions.

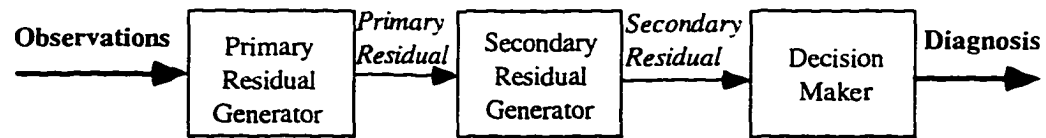


Figure 4.5 Model-Based FDI

In practice the primary residual deviates from zero due to noise and model error that are unavoidable in model-based methods. The presence of any significant noise requires an additional step in which the primary residual is converted to a logical pattern or a signature which reflects the condition of the plant. This signature can be obtained by simply filtering the primary residual, by statistical testing, or by more complicated means such as frequency domain analysis. The result of the processing of the primary residual is referred as the secondary residual. Also, any model error, which is unavoidable in practice, can be at least in part accommodated by similar approaches. Finally, the secondary residual (logical pattern, or signature) is analyzed to isolate the fault and its cause. Such analysis can be performed through the comparison of a set of patterns known to belong to simple failure or by more complex logical procedures, such as knowledge-based inference. The last two steps, namely secondary residual generation and decision-making, are intended to detect and isolate the fault including time and location.

### Nonlinear Observers for FDI

FDI for nonlinear systems can be achieved by generating residuals using nonlinear observers. These observers have been proposed in [115]-[118] among others. Recently, a nonlinear FDI scheme that is based on constructing inverse dynamic models of nonlinear systems was proposed by [122]. The nonlinear observer based residual generation problem can be formulated as follows; consider a nonlinear system given by

$$\dot{x} = f(x, u) \quad (4.1)$$

$$y = h(x, u) \quad (4.2)$$

Under the assumption of a smooth system [121], and assuming that  $H$  defined below is full rank, a simple form of a nonlinear estimator is

$$\dot{\hat{x}} = f(\hat{x}, u) + H(\hat{x}, u)(y - \bar{y}) \quad (4.3)$$

$$\hat{y} = h(\hat{x}, u) \quad (4.4)$$

where

$$H(\hat{x}, u) = \left. \frac{\partial \hat{f}}{\partial y} \right|_{\hat{x}, u} \quad (4.5)$$

is a time-varying observer gain matrix. The state estimation error equation then becomes

$$\dot{e}(t) = \left[ \frac{\partial f}{\partial y} - H(x, u) \frac{\partial h}{\partial x} \right]_{\hat{x}, u} e(t) \quad (4.6)$$

and the output estimation error is

$$e(t) = y(t) - \hat{y}(t) = h(x, u) - h(\hat{x}, u) \quad (4.7)$$

$e(t)$  is the primary residual. The nonlinear system residual generator using a nonlinear observer is illustrated in Figure 4.6. Diagnostic observers for the nonlinear system have been reported in [116], [117], and [118].

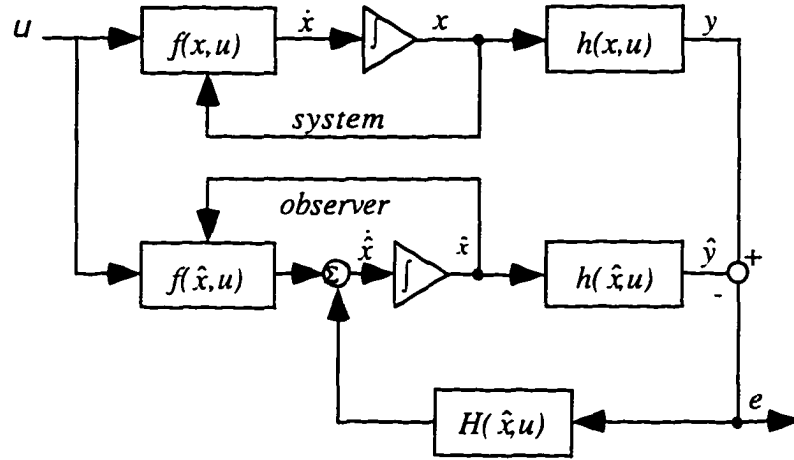


Figure 4.6 Nonlinear Observer Residual Generator

In this study, sliding mode estimation techniques are used as they are suitable for application to systems that are significantly nonlinear, and exhibit good robustness properties with respect to modeling uncertainty and parameter variation. The sliding mode approach to identification and estimation in dynamic systems consists of constructing an observer with a discontinuous auxiliary function, and of enforcing sliding

mode such that the model and plant outputs coincide. Then, the average value of the auxiliary function depends on the unknown states, parameters and disturbances and may be used for their evaluation. For sliding mode observer additional references on design and applications refer to [129]-[135].

The sliding mode observer design idea may be illustrated as follows; let a system be represented by the equation

$$\dot{x} = f(x, t, a) \quad (4.8)$$

with known state vector  $x$  and an unknown parameter vector  $a$ .

The observer equation is of the form

$$\dot{\hat{x}} = f_o(\hat{x}, t, a_o) + V \quad (4.9)$$

with  $\hat{x}$ ,  $f_o$  and  $a_o$  being estimates of  $x$ , and the nominal values of  $f$  and  $a$ , respectively. The auxiliary input  $V$  is defined as,

$$V = -M \text{sign}(s),$$

where  $V$  is a discontinuous function of the mismatch, and

$$s = \hat{x} - x.$$

To reach the sliding manifold  $s$ , a gain  $M$  is selected to satisfy the reaching condition, that is:  $s\dot{s} < 0$ , where

$$\dot{s} = \dot{\hat{x}} - \dot{x} = f_o(\hat{x}, t, a) + V - f(x, t, a) = f_o(\hat{x}, t, a) - f(x, t, a) - M \text{sign}(s). \quad (4.10)$$

If

$$M > |f_0(\hat{x}, t, a_0) - f(x, t, a)|,$$

then

$$s\dot{s} < 0.$$

Once the reaching condition is satisfied, sliding mode will occur after a finite time-interval. When sliding occurs  $s \equiv 0$ , and therefore  $\hat{x} \rightarrow x$ .

Once the primary residual is generated using the nonlinear observer scheme mentioned previously, it has to be further processed to generate a secondary residual that can be used for diagnosis. In the next section the method of evaluation adopted in this study is explained.

## 4.2. RESIDUAL AND SYMPTOM EVALUATION

The proposed diagnostic scheme shown in Figure 4.3 in the previous section utilizes a fuzzy logic residual evaluation scheme for fault detection. Sensitivity to faults and low false alarm rates are the ultimate goal of robust residual evaluation [62]. Residuals are generated in such a way that a fault can be distinguished from the others using one or more residuals. A decision is made on the condition of the system by evaluating these residuals. The evaluation of the residuals can be approached using different methods [58], e.g. pattern recognition, adaptive thresholds, statistical evaluation, neural networks, or fuzzy logic. Residuals are generated using nonlinear observers which require analytical models of subsystems. Residual evaluation for fault detection depends on the robustness



of the models used and on the quality of the acquired information [63]. A general scheme for residual evaluation is shown in Figure 4.7.

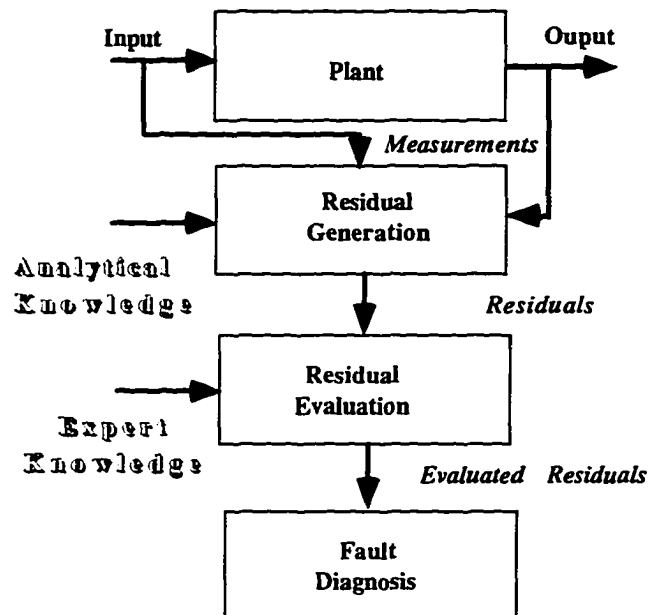


Figure 4.7 Residual Evaluation Scheme [58]

Residuals generated using quantitative knowledge are evaluated using qualitative knowledge that can deal with imprecision. This reduces the possibility of false alarm and missed detection. The primary residuals are fuzzified and the heuristic knowledge about the symptoms is used for the evaluation, this output is weighted information in terms of linguistic variables; and the membership functions are chosen to decouple the possible faults, as shown in Figure 4.8.

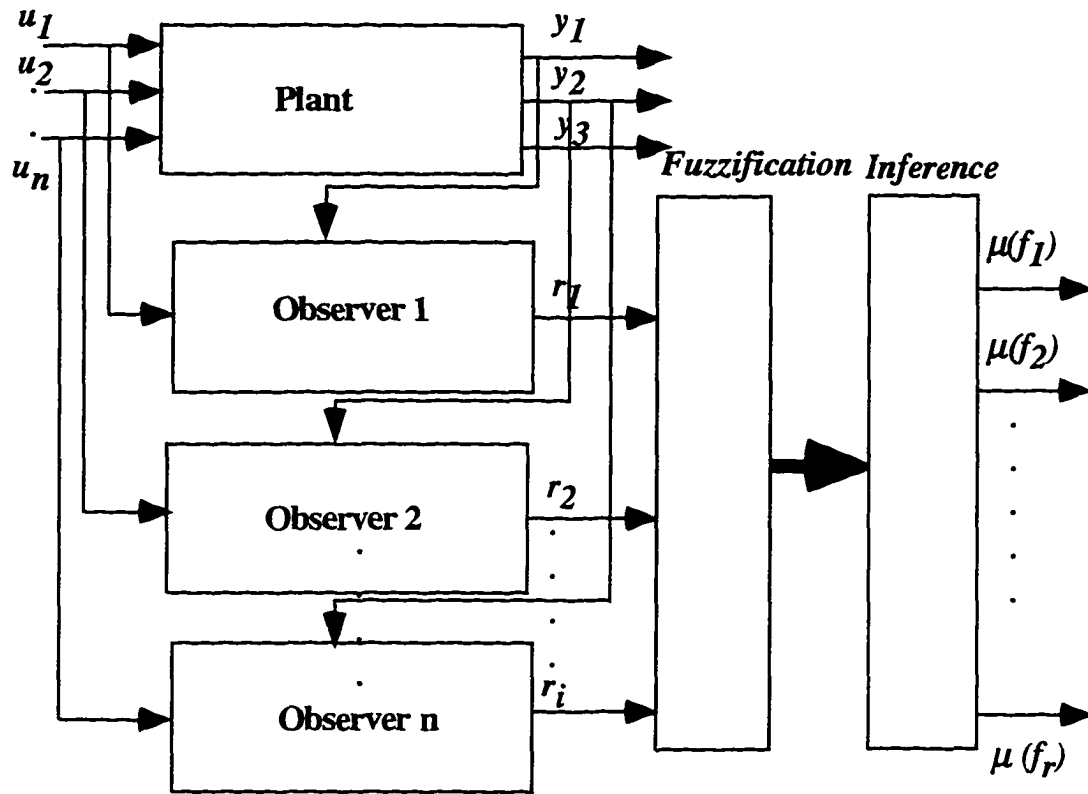


Figure 4.8 Residual Fuzzy Evaluation

In Figure 4.8,  $u_i, i=1,2,\dots,n$  are the inputs,  $y_i, i=1,2,\dots,m$  are the outputs,  $r_i, i=1,2,\dots,m$  are the residuals, and  $\mu(f_r)$  is the possibility of each of the  $r$  faults that have been hypothesized. The threshold is defined using fuzzy membership functions to increase the sensitivity to faults and to lower the false alarm rate. The following are the steps followed for the fuzzy residual evaluation:

1. Primary residuals are generated using nonlinear observers.

2. Optimal thresholds are calculated for the nominal process.
3. Fuzzy membership functions are generated for residual evaluation.
4. Fuzzy rules are formulated for fault detection.
5. Fuzzy inference is made for fault diagnosis.

Classical residual evaluation for fault detection and isolation is approached by using a logic tables such as the one shown in Table 4.2. In Table 4.2,  $A$  and  $B$  are the components monitored and  $f_1, f_2, f_3$  are the *symptoms* in the case of a fault. If  $B$  is faulty, for example, then the first symptom will be equal to one and all the other are zero.

	$A$	$B$
$f_1$	0	1
$f_2$	1	0
$f_3$	1	0

Table 4.2 Fault Isolation Logic [63]

In this case the thresholds are used to generate symptoms  $f_i$  which are used for isolation. Fuzzy linguistic replaces the logical approach by expressing the effect of the *symptom* [63]. For example, with reference to Table 4.2, we can generate fuzzy rules such as:

*If  $B$  is faulty then symptom  $f_1$  is high or normal*

and symptom  $f_2$  is low

and symptom  $f_3$  is low.

The symptom is evaluated using the membership function;

$$\mu_i(f_r): x \in X \rightarrow [0,1],$$

where  $X$  is the universe of discourse  $X = \{x_1, x_2, \dots, x_p\}$ , and  $\mu_i(f_r)$  is the possibility of symptom  $f_r$  is  $\{x_1, x_2, \dots, x_p\}$ .

The membership functions can be represented as shown in Figure 4.9 with  $x_1$ : low,  $x_2$ : normal, and  $x_3$ : high.

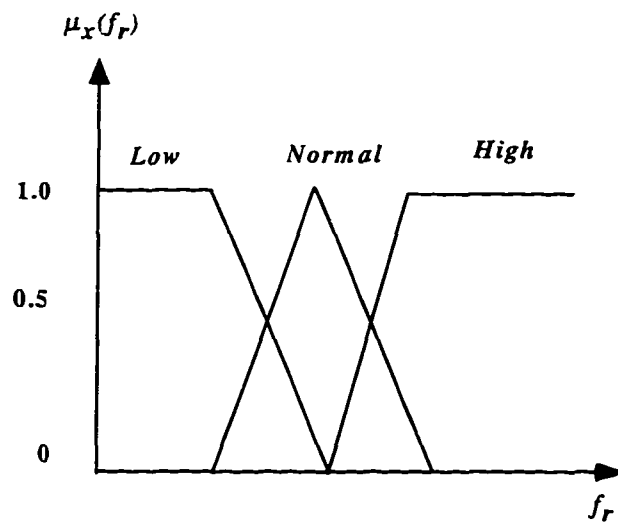


Figure 4.9 Fuzzy Evaluation Membership Function

The universe of discourse  $X = \{low, normal, high\}$ , and the logic represented in Table 4.2

could be replaced by the logic in Table 4.3 ;

	<i>A</i>	<i>B</i>
$f_1$	low or normal	high or normal
$f_2$	high	low
$f_3$	high	low

Table 4.3 Fuzzy Evaluation and Isolation Logic

In fuzzy logic we can define

$$\mu(x \cap y)(f_r) = \max \{ \mu_x(f_r), \mu_y(f_r) \},$$

and

$$\mu(x \cup y)(f_r) = \min \{ \mu_x(f_r), \mu_y(f_r) \}.$$

The membership function for  $f_1$  is *high or normal* is shown in Figure 4.10.

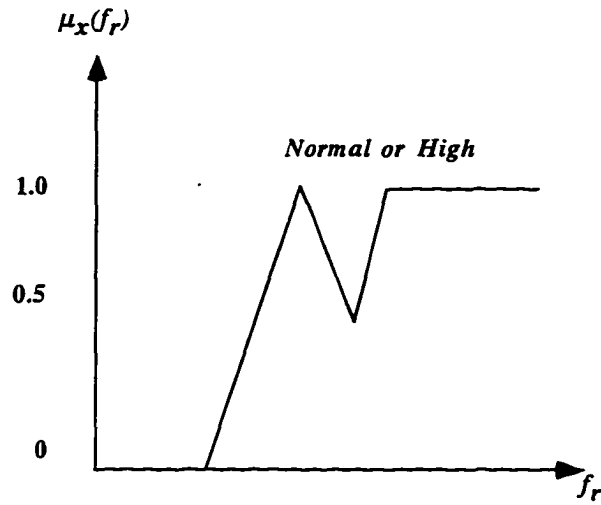


Figure 4.10 Implied Fuzzy Set for Normal or High Residuals

Let  $E$  and  $F$  represent the set of elements in a process and the set of symptoms, respectively:

$$E = \{A_1, A_2, \dots, A_j, \dots, A_n\}, \quad j = 1, \dots, n, \text{ and } F = \{f_1, \dots, f_i, \dots, f_m\}, \quad i = 1, \dots, m.$$

In this fuzzy system the fuzzy linguistic value vector  $X^r = [x_1^r \ x_2^r \ x_m^r]^T$  is associated with each element  $r = 1, \dots, n$ . The symptoms  $F$  are matched against  $X^r$  and a fuzzy rule is generated which is in the following form:

*IF  $f_1$  is  $x_1^r$  and  $f_2$  is  $x_2^r$ , ..., and  $f_m$  is  $x_m^r$  THEN  $A_r$  is faulty*

The possibility of the fault is computed as follows:

$$\mu_{\text{faulty}}(A_r) = \min \{ \mu_{x_1^r}(f_1), \mu_{x_2^r}(f_2), \dots, \mu_{x_m^r}(f_m) \},$$

$A_r$  is declared faulty when  $\mu_{faulty}(A_r)$  reaches a maximum value which can be chosen by the user. In Table 4.4 an example that gives the fuzzy symptoms for  $A$ , and  $B$  is shown. The isolation is performed using the fuzzy rules previously mentioned in Table 4.3. The result of the possibility calculation is given below.

	<i>Low</i>	<i>Normal</i>	<i>High</i>
$f_1$	0.8	0.2	0
$f_2$	0	0.2	0.9
$f_3$	0	0.2	0.6

Table 4.4 Fuzzified Symptoms

$$\mu_{faulty}(A) = \text{Min} \{ \text{Max}(0.8, 0.2), 0.9, 0.6 \} = 0.6$$

$$\mu_{faulty}(B) = \text{Min} \{ \text{Max}(0.2, 0), 0, 0 \} = 0$$

The outcome of the fault possibility calculation indicates that  $A$  is the faulty component. The residuals are generated using a mathematical model of the system and the symptoms are generated by fuzzy evaluation of the residuals. Moreover, the fuzzy inference engine uses the heuristic rules acquired from the expert to isolate the faulty component. The use of knowledge-base fuzzy rules is covered in the following section.

#### **4.2.1. Knowledge-Based FDI**

Subjective knowledge approaches to FDI such as knowledge-based expert systems use qualitative models based on available knowledge of the system [100] for fault detection and isolation. The main components of knowledge-base systems, as shown in Figure 4.1, are listed in [100] as follows:

1. Knowledge base
2. Data base
3. Inference engine
4. Explanation component

The heuristic reasoning acquired from an expert in the field under investigation is used for fault diagnosis. The detection and isolation of a fault is achieved by the evaluation of the heuristic knowledge. When the system can be represented linguistically, and low and high level of processing is required, fuzzy logic can be used to represent the rules in the knowledge base [68]. The advantage of using fuzzy rule based systems in the knowledge-based methods are listed by [68];

1. Fuzzy logic is a well defined theory.
2. Common sense, vague, and ambiguous knowledge can be represented easily.
3. Process uncertainty can be accommodate .
4. The ability to tune the fuzzy system to suitably represent the plant.
5. Different inference methods are possible.
6. Fuzzy rules can approximate any continuous function to a great degree of accuracy.



Conditional statements are the fundamental component of the rules in a rule based expert system [78] and may be of the form,

“if  $V_1$  is  $A$  then  $V_2$  is  $B$ ”.

A possibility distribution is generated by such propositions [21]; for example the statements

if  $X$  and  $Y$  are the base sets for  $V_1$  and  $V_2$  respectively then

if  $V_1$  is  $A$  then  $V_2$  is  $B$

induce a conditional possibility distribution  $\pi_{V_1/V_2}$  over  $X \times Y$  such that

$$\Pi_{V_1/V_2}(x,y) = \text{Min} (1, 1-A(x)+B(y)) \quad (4.11)$$

or

$$\Pi_{V_1/V_2}(x,y) = \text{Max} (1-A(x), B(y)) \quad (4.12)$$

Thus, a possibility distribution is introduced by data statements and rules. The possibility distribution for more complex forms of rules is covered in [78], and can be captured in the fuzzy rule-base. Moreover, to use fuzzy compositional inference adds new information to the global data base. This is possible because fuzzy compositional inference combines conjunction and projection principles [78]. Also, the use of linguistic quantifiers in the rules improves the performance of the expert system as it can provide a better representation of the expert's knowledge. Furthermore, qualified statements about the certainty of the information provided can be included.

## **Fuzzy Logic and Knowledge Representation**

Knowledge representation using the conventional approaches, such as classical probability theory and first order logic, stands short when it comes to representing the fuzzy quantifiers and common sense knowledge [79]. Fuzzy logic, on the other hand, has the following characteristics, which make it well suited to represent uncertainty.

- i. In approximate reasoning exact reasoning is a limiting case.
- ii. Everything is a matter of degree.
- iii. Fuzzification of any logical system is possible.
- iv. Fuzzy (elastic) constraints are used to represent the knowledge available.
- v. Propagation of elastic constraints are used for inference.

The principal modes of reasoning in fuzzy logic are the following [79];

1. Categorical reasoning; where no fuzzy quantifiers and no fuzzy probabilities are contained in the premises.
2. Syllogistic reasoning; where fuzzy quantifiers are included in the premises.
3. Dispositional reasoning; where preponderantly (but not necessarily always true) propositions are the premises.
4. Qualitative reasoning; where if-then rules with linguistic variables in the antecedent and consequent.

In fuzzy logic, the predicates are fuzzy rather than crisp. There exist a variety of predicate modifiers that act as hedges, e.g., *very*, *more* or *less*. Also, the quantifiers are interpreted as fuzzy numbers and probabilities can be represented by linguistic or fuzzy probabilities. While possibility in classical logic is bivalent, in fuzzy logic it is graded. A collection of

propositions establish the knowledge in a general fuzzy setting [79]. Test-score semantics are the basis for knowledge representation [80], where a proposition is regarded as a collection of elastic (fuzzy) constraints. When representing the meaning of a proposition,  $p$ , through the use of test-score semantic, the following steps are considered [79]:

1. The variables  $X_1, \dots, X_n$  whose values are constrained by the proposition are identified.
2. The constraints  $C_1, \dots, C_m$  which are induced by  $p$  are identified.
3. Each constraint,  $C_i$ , is characterized by describing a testing procedure which associates with  $C_i$  a test score  $\tau_i$ , representing the degree to which  $C_i$  is satisfied, where  $\tau_i$  is expressed as a number in the interval  $[0, 1]$ .
4. The partial test scores  $\tau_1, \dots, \tau_m$  are aggregated into smaller number of test scores  $\bar{\tau}_1, \dots, \bar{\tau}_k$ , which are represented as an *overall vector test score*  $\tau = \bar{\tau}_1, \dots, \bar{\tau}_k$ . In most cases  $k=1$ .

The meaning of a proposition,  $p$ , is represented not by the overall test-score  $\tau$  but by the procedure which leads to it. Thus, by dealing with uncertainty and imprecision, a higher level of expressive power is achieved by test-score semantics. This allows for representing the meaning of much wider variety of propositions in natural language [79].

In knowledge-based methods the production rule, in a rule-based expert systems, is the primary knowledge structure [81]. The structure of these rules is

If  $V_1$  is  $A_1$  and  $V_2$  is  $A_2$  . . . and  $V_n$  is  $A_n$  then  $U$  is  $B$

When the values of  $A_i$ 's as well as the knowledge about the  $V_i$ 's are imprecise, the application of approximate reasoning is apparent. A simple rule, where we assume  $V$  is a variable which takes as its value an element in the set  $X$ , will be of the form:

$$V \text{ is } A,$$

where  $A$  is a fuzzy subset of  $X$ . This rule indicates that the value of  $V$  is an element of  $A$ .

The rule gets translated into a possibility distribution  $\Pi_v$  on  $X$ , as shown in [81], so that:

$$\Pi_v(x) = A(x),$$

where  $\Pi_v(x)$  indicates the possibility the  $V$  has assumed the value  $x$ . For the more general rule (If  $V_1$  is  $A_1$  and  $V_2$  is  $A_2$  . . . and  $V_n$  is  $A_n$  then  $U$  is  $B$ ), it gets translated into

$$\Pi_{v_1, v_2, v_3, \dots, v_n, u} \text{ on } X_1 \times X_2 \dots \times X_n \times Y$$

so that:

$$\Pi_{v_1, v_2, v_3, \dots, v_n, u}(x_1, x_2, \dots, x_n, y) = 1 \wedge (1 - (A_1(x_1) \wedge A_2(x_2) \dots \wedge A_n(x_n)) + B(y))$$

where  $\wedge$  stands for min, where the smaller of the operands of  $\wedge$  is taken.

The rules in a fuzzy expert system can be derived from the expert, a model of the process, or from training input-output data. Heuristic rules that are generated by the expert can be used to make inferences about the system under consideration. A model of the process under investigation can be used to derive the rules, using simulated data. Moreover, numerical input-output data obtained from experiments can be used to identify a fuzzy system and to construct rules [87]. This fuzzy system can serve as a parameter estimator, and similarly as a predictor. In the following section the estimation and training techniques adopted in this work are explained.

### 4.3. FUZZY ESTIMATION AND IDENTIFICATION

A rule-base can be constructed from input-output data for the system for which parameter estimation must be performed. Different methods have been developed to construct fuzzy systems for estimation and identification. One method is given in [73] for strongly nonlinear systems, where two sources of information are available: 1) human experience in the form of "IF-THEN" rules, and, 2) input-output numerical data. The approach developed in [73] is rather general, and combines both sources of information into a common fuzzy rule base to solve the problem at hand. Other methods for estimation are covered in more detail in [90] and [87]. In this section we present the fuzzy clustering technique in which cluster centers are identified where each center represents the center point of the input membership function. These clusters can be used to generate a Takagi-Sugeno fuzzy inference system that models the data behavior and in which the output is a crisp linear function. Takagi-Sugeno's fuzzy system is composed of fuzzy *IF-THEN* rules of the form

$$IF H^j THEN \hat{y}^j = c_o^j + c_1^j x_1 + \dots + c_n^j x_n \quad (4.13)$$

$j = 1, \dots, N$  ( $N$  is the number of rules) where input fuzzy set

$$H^j = \{(\underline{x}, \mu_{H^j}(\underline{x})) : \underline{x} \in U_1 \times \dots \times U_n\}, U_i \text{ is the } i^{th} \text{ input universe of discourse,}$$

$\underline{x} = [x_1, \dots, x_n]^T$  is the vector of fuzzy system inputs,  $\mu_{H^j}(\underline{x})$  is the membership function

for  $H^j$ ,  $\hat{y}^j = \underline{c}_j^T \hat{\underline{x}}$  is a crisp output that is a linear function of the inputs given by the real-

valued parameters  $\underline{c}_j = [c_o^j, \dots, c_n^j]^T$ , and  $\hat{\underline{x}} = [1, \underline{x}^T]^T$ . The output of the Takagi-Sugeno's

fuzzy system is a weighted average of  $\hat{y}^j$  for  $j = 1, \dots, N$  given by;

$$\hat{y} = g(\underline{x}; \underline{\theta}) = \frac{\sum_{j=1}^N \mu_{H'}(\underline{x}) \hat{y}^j}{\sum_{j=1}^N \mu_{H'}(\underline{x})} \quad (4.14)$$

where  $\underline{\theta}$  is a vector of parameters which characterize the fuzzy system.

$$\underline{\theta} = [\underline{v}_1^T, \dots, \underline{v}_N^T, c_1^T, \dots, c_N^T]^T \quad (4.15)$$

We use the algorithm developed in [92] to construct a Takagi-Sugeno fuzzy system. A systematic approach to automating a fuzzy system that approximates the functional mapping represented in the output-input data  $F$  was proposed in [92]. This was achieved by estimating the appropriate fuzzy relation patches (fuzzy clustering) and identifying  $N$  cluster centers  $\underline{v}_j \in \mathcal{R}^n, j = 1, \dots, N$  where each cluster center represents the center points for input membership functions. A weighted least squares computation, that uses the output portion of the training data set to calculate the coefficients  $\underline{c}_j$  for  $N$  linear functions, was used to determine the consequent of the fuzzy rule. Fuzzy C-Means is a clustering technique that can be applied to perform fuzzy clustering in order to obtain the antecedent of the rules in the fuzzy system. An overview of fuzzy clustering technique is given in the following section in addition to the technique used to obtain the consequent part of the rules.

#### 4.3.1. Fuzzy C-Means Clustering

In order to construct the desired fuzzy system from input-output data, “Fuzzy C-Means” clustering algorithm is employed to specify the antecedent of the fuzzy system as shown in [93]-[95]. This algorithm partitions the data into different patches (clusters),

and a degree of membership of each point in the space is assigned using fuzzy clustering. For a training data set  $F$  with input portion  $\underline{x}_i$ , and for fuzzy c-means we wish to minimize the objective function given by

$$J = \sum_{i=1}^{m_F} \sum_{j=1}^N (\mu_{ij})^m (\underline{x}_i - \underline{v}_j)^T (\underline{x}_i - \underline{v}_j) \quad (4.16)$$

where  $m > 1$  is a design parameter,  $N$  is the number of clusters,  $m_F$  is the number of training data points,  $\underline{x}_i$  for  $i = 1, \dots, m_F$  is the input portion of the training data set  $F$ ,  $\underline{v}_j$  for  $j = 1, \dots, N$  are the cluster centers, and  $\mu_{ij}$  is the grade of membership for  $\underline{x}_i$  in the  $j^{\text{th}}$  cluster. To minimize (4.23) a set of necessary conditions are given in [93] for  $\mu_{ij}$  and  $\underline{v}_j$ , when  $\|\underline{x}_i - \underline{v}_j\|_2^2 > 0$  for all  $j$ . This is given by

$$\underline{v}_j = \sum_{i=1}^{m_F} (\mu_{ij})^m \underline{x}_i / \sum_{i=1}^{m_F} (\mu_{ij})^m \quad (4.17)$$

for each  $j = 1, \dots, N$ , and

$$\mu_{ij} = \left[ \sum_{k=1}^N \left( \frac{\|\underline{x}_i - \underline{v}_k\|_2^2}{\|\underline{x}_i - \underline{v}_j\|_2^2} \right)^{m-1} \right]^{-1} \quad (4.18)$$

for each  $i = 1, \dots, m_F$  and  $j = 1, \dots, N$ . However, if  $\|\underline{x}_i - \underline{v}_j\|_2^2 = 0$  for some  $j$  then  $\mu_{ij}$  are nonnegative numbers satisfying  $\sum_{j=1}^N \mu_{ij} = 1$  and  $\mu_{ij} = 0$ , if  $\|\underline{x}_i - \underline{v}_j\|_2^2 \neq 0$ . The Fuzzy C-Means algorithm is based on (4.23), (4.24), and (4.25), and is an iterative algorithm. The steps followed to generate these clusters are given in [90]. Clustering with optimal output

predefuzzification, which is a weighted least square approach, is used to choose the consequent parameters. Minimizing the squared error, weighted by the value of the cluster membership between the output portion and the training data set and a parameterized function of the input portion of the training data, can be achieved by minimizing the cost function  $J_j$  given by

$$J_j = \sum_{i=1}^{m_j} \mu_{ij} (\underline{c}_j^T \hat{\underline{x}}_i - y_i)^2 \quad (4.19)$$

for  $j = 1, \dots, N$  where  $\mu_{ij}$  is the membership of the input portion of the  $i^{\text{th}}$  training data point for the  $j^{\text{th}}$  cluster,  $y_i$  is the output for the  $i^{\text{th}}$  training data point, and the multiplication of  $\underline{c}_j^T$  and  $\hat{\underline{x}}_i$  define the output associated with the  $j^{\text{th}}$  rule and the  $i^{\text{th}}$  training point. The solution to the minimization problem can be found in [87] and [90]. The training data are used in conjunction with optimal output defuzzification to calculate the linear function which defines the crisp output for each rule  $y^j(\underline{x}) = \underline{c}_j^T \hat{\underline{x}}$  [90]. This yields a fuzzy system of the form;

$$g(\underline{x}; \underline{\theta}) = \frac{\sum_{j=1}^N y^j(\underline{x}) \mu_{H^j}(\underline{x})}{\sum_{j=1}^N \mu_{H^j}(\underline{x})} \quad (4.20)$$

A Multi-Input Single-Output (MISO) fuzzy system can be constructed by combining the optimal output predefuzzification and fuzzy clustering. The spread of the input-output data, and the richness of information contained in the data pairs, are few of the issues that have to be considered when choosing the data structure [87]. Moreover, the fuzzy system



generated from heuristic knowledge about the plant can be incorporated with that obtained from the training data.

In the diagnostic scheme proposed in Figure 4.3, the input-output data follow two paths. The first path applies model-based methods, where a mathematical model of the system and nonlinear observers are utilized to generate primary residuals. These residuals are evaluated using fuzzy adaptive thresholds. The second path applies knowledge-based methods, where heuristic knowledge incorporated with the fuzzy estimation techniques explained previously are utilized to generate an estimate of the output. So far various forms of knowledge are available to the diagnostic system, the integration of this knowledge is the final step towards the decision-making. In the following section the diagnostic integration criteria is covered.

#### 4.4. DIAGNOSTIC INTEGRATION

A crucial step towards making the diagnostic decision is to combine the residuals and symptoms generated by the different diagnostic methods, i.e. the analytical residuals and the heuristic symptoms (see Figure 4.3). Along the first path, the primary residuals are generated using the model-based algorithm, where observers generate residuals for each subsystem of the plant. These residuals are evaluated using fuzzy adaptive thresholds, where quantitative knowledge (residual) is transformed into qualitative knowledge (symptom). The residuals are fuzzified, where crisp numbers are mapped into fuzzy sets, and evaluated as shown in Section 4.2. A fault event has to be assigned for the set of *symptoms*  $F=\{f_1, \dots, f_m, \dots\}$ ,  $i = 1, \dots, m$ , and a fuzzy production rule is expressed in the following form:

$$\text{IF } f_1 \text{ is } x_1^r \text{ and } f_2 \text{ is } x_2^r, \dots, \text{ and } f_m \text{ is } x_m^r \text{ THEN fault is } A_r \quad (4.21)$$

Where the fuzzy linguistic vector  $X^r = [x_1^r \ x_2^r \ x_m^r]^T$  and is associated with each element  $r = 1, \dots, n$ .

On the other hand, along the second path heuristic symptoms are available where knowledge-based methods are used to generate fuzzy production rules in the same form as that stated in (4.28). It is apparent that in this diagnostic scheme both paths lead to a unified representation of the analytical and heuristic symptoms.

The integration of the model- and knowledge-based knowledge is achieved at the inference mechanism, where the rules representing the same symptom can be combined into one rule. Then the fault can be diagnosed and a decision can be made by analyzing the relation between the symptoms and the faults. In the following section the decision-making process is covered.

#### 4.5. FUZZY DECISION-MAKING

Robustness is a principal goal of the proposed diagnostic scheme. A robust diagnostic scheme is intended as one where false alarms and missed detection are minimized. A simple logic table as shown in Table 4.2 is classically used to isolate faults. In the proposed diagnostic scheme, a fuzzy decision table is used instead to isolate the fault. The decision-making Table 4.5 implements fuzzy logic rules where linguistic variables and values are specified to isolate the fault. The linguistic variables associated with the inputs are the *symptoms*  $F = \{f_1, \dots, f_m, \dots\}$ ,  $i = 1, \dots, m$ . The linguistic values of the input are  $X^r = [x_1^r \ x_2^r \ x_m^r]^T$ ,  $r = 1, \dots, n$ , and the consequent of the rule is the fault. The consequent is

chosen to be a zero-order Takagi-Sugeno (singleton) fuzzy system of the general form of equation (4.20) to enhance the efficiency of the defuzzification process; the fuzzy rule has the form

$$\text{IF } f_1 \text{ is } x_1^r \text{ and } f_2 \text{ is } x_2^r, \dots, \text{ and } f_m \text{ is } x_m^r \text{ THEN } A_r = k,$$

where  $k$  is a crisply defined constant, and each fault is assigned a crisp number.

The membership functions are chosen *a priori* for the input universe of discourse. The fuzzy system used *singleton* fuzzification, *Gaussian* membership functions, product for the premise and implication, and center of gravity defuzzification are used. The input symptoms are compared to the premises of all rules to determine which rule is applicable. The fault is determined using the rule that is applicable at the time of detection. An example of such rules is tabulated in Table 4.5 where each element of this table is a rule in the form

$$\text{IF } f_1 \text{ is } \textit{high} \text{ and } f_2 \text{ is } \textit{low}, \dots, \text{ and } f_m \text{ is } \textit{normal} \text{ THEN } \textit{fault} = 1$$

		<i>Symptom</i>			
		$f_1$	$f_2$	...	$f_m$
<i>Fault</i>	1	<i>high</i>	<i>low</i>	...	<i>normal</i>
	2	<i>high</i>	<i>high</i>	...	<i>low</i>
	⋮	⋮	⋮	⋮	⋮
	$n$	<i>low</i>	<i>low</i>	...	<i>high</i>

Table 4.5 Fuzzy Decision-Making Table

The rule-base derived from Table 4.5 involves thresholds that are determined using fuzzy logic for evaluation, as explained in Section 4.2. Moreover, the fuzzification of the residuals generated by the model-based method allows a gradual transition from a normal condition to a faulty condition as opposed to the crisp thresholds, where an abrupt change occurs as the residual crosses the preset threshold. The analytical and heuristic knowledge redundancy available at the decision-making stage provides additional information that is included in the rule-base to produce a more infallible decision. The output is represented in the form of a histogram displaying the fault number of occurrences, or the degree of compatibility of the possible faults; then a human operator can make the final decision. Ultimately an automated fault alarm could be implemented.

#### **4.6. SUMMARY**

In this chapter the methodology for integrating model- and knowledge-based diagnostic methods has been explained. There exist two paths; one path for each of the diagnostic methods. The first path implements model-based methods, where a mathematical model of the plant under investigation is utilized in conjunction with nonlinear observers to generate primary residuals. The residuals are evaluated using fuzzy logic to accommodate the uncertainty and model error that are inherent in such analytical methods. The outcome of this fuzzy residual evaluation is secondary residuals (symptoms) that are fuzzified and included in the premise of production rules of the form IF-THEN, while the fault source is included in the consequent. Along the other path knowledge-based methods are implemented simultaneously, where heuristic rules are generated. These fuzzy rules in conjunction with fuzzy estimation techniques using input-

output data from the plant, are used to detect faults. Furthermore, the knowledge available from both methods is integrated at the inference mechanism, as they are represented in unified form of IF-THEN fuzzy production rules. Finally a fuzzy decision-making table is generated from which fuzzy rules can be extracted where the symptoms are the antecedent of the rule and the consequent is the isolated fault. The validity of the proposed methodology is demonstrated in the next chapter where its application towards fault diagnosis in an internal combustion engine and results are introduced.

## CHAPTER 5

### APPLICATION AND RESULTS

In this chapter the methodology proposed in Chapter 4 is applied to the problem of fault detection and isolation in an internal combustion engine. The available measured data is processed along the two paths of Figure 4.3 as explained in Chapter 4. Along the first path a model-based scheme is applied, which utilizes a mathematical engine model. The engine model is a Low Frequency Dynamic Engine Model (LFDEM) which is subdivided into systems, subsystems, and components. A bank of nonlinear observers processes the input/output data and generates the primary residuals. These residuals are evaluated using fuzzy logic thresholds. Along the second path, a knowledge-based scheme is applied. The knowledge-base scheme utilizes heuristic rules in conjunction with fuzzy estimation techniques to detect engine faults using the engine-out (or tailpipe) exhaust gas emissions. The integration of the knowledge available from both methods is achieved in the inference mechanism. A fuzzy decision-making scheme is constructed and was capable of detecting and isolating the faults induced in the engine. The results and procedures followed for such application are explained in detail in this chapter.

## 5.1. ENGINE MODEL

The engine model shown in Figure 5.1 depicts a simplified dynamic model, based on the understanding of the physical phenomena taking place within the engine systems, subsystems, and components, to derive the dynamic equations of each block. These equations are listed in Appendix C. This model contains different systems of the engine, these are subdivided into subsystems, and furthermore, to the component level. This hierarchical organization mentioned in Section 2.5. is a key factor in the “Divide and Conquer” strategy that was adopted throughout this study.

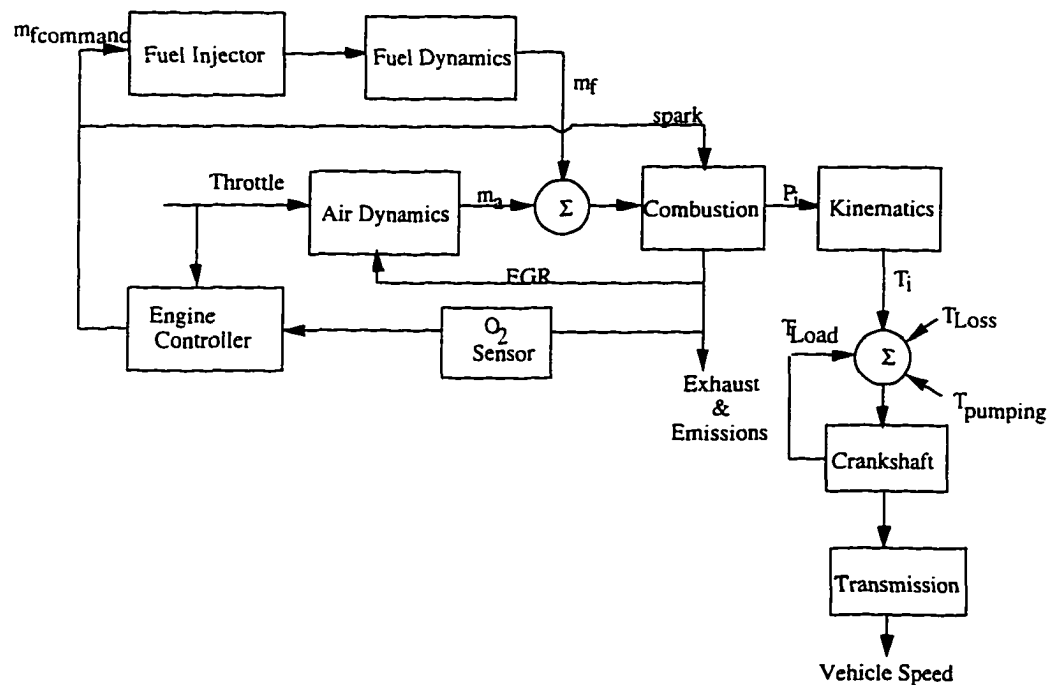


Figure 5.1 Dynamic Engine Model

A LFDEM which describes the engine dynamics with mean values rather than the instantaneous values is shown in Figure 5.2. The subsystems to be modeled are the throttle, the intake manifold dynamics (including pumping); the fuel dynamics (a wetting model); and the oxygen sensor dynamics with exhaust delay. Breaking down the system into different physical models or sub-models, based on the faults to be diagnosed, increases the redundancy and reduces computations, as the input/output relation for each subsystem is less complicated. Also, according to different models, different algorithms can be applied towards fault detection. In particular, some faults cannot be detected by a single physical model by itself, as this fault might affect the output of several sub-models or measured outputs. To illustrate the concept, Figure 5.3 depicts the air, fuel and exhaust subsystems. This is an especially important set of subsystems for diagnostic purposes, because improper operation of these subsystems results in increased exhaust emissions. The faults considered in this chapter all pertain to the subsystems of Figure 5.3. The equations describing the air, fuel and exhaust subsystems are listed in Appendix C. A discrete-crank-angle version of these equations was used in the realization of the observers.

In the remainder of the chapter we shall follow the structure of Figure 4.3 to highlight how the methodology of Chapter 4 was applied to the problem of engine diagnostics.



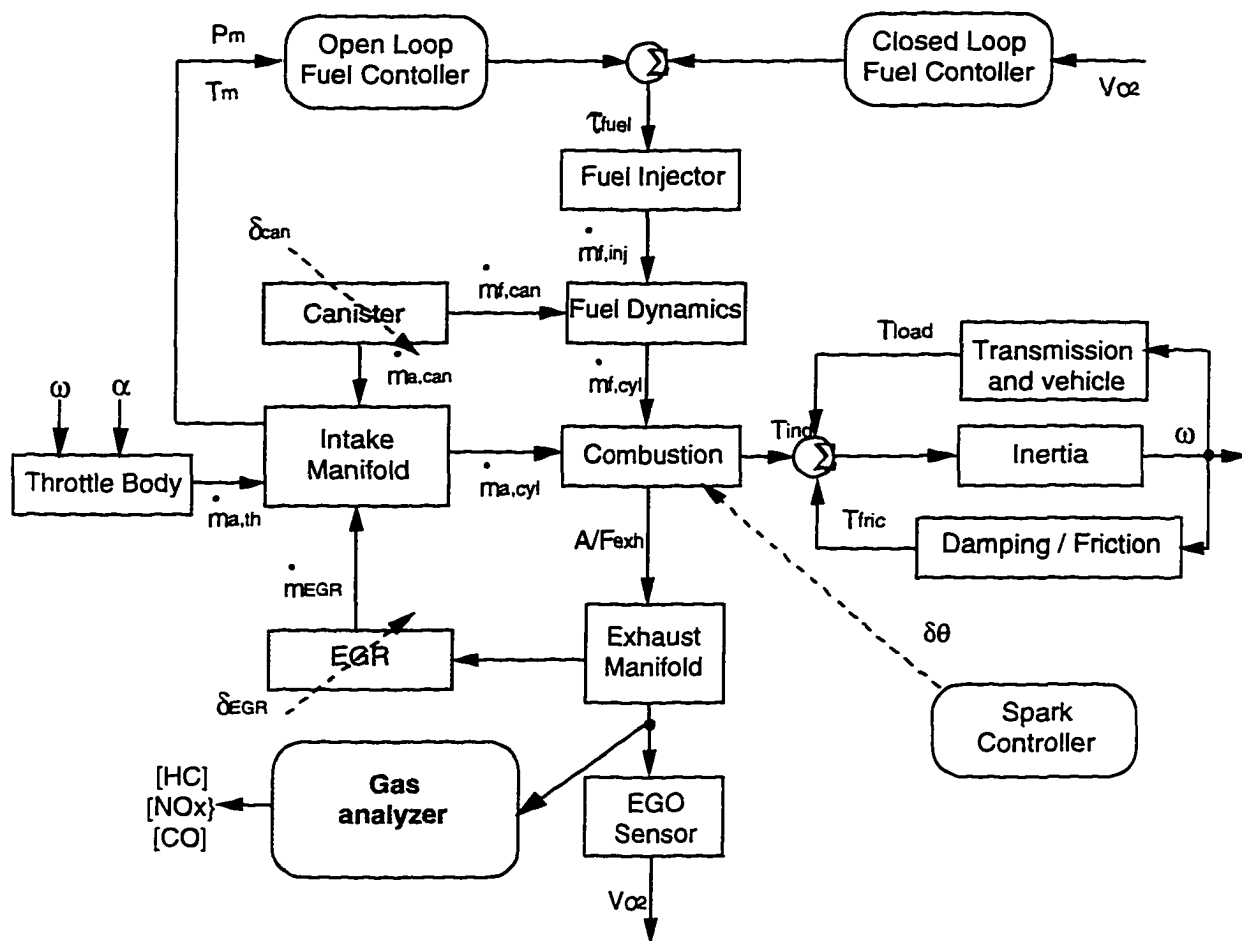


Figure 5.2 Low Frequency Dynamic Engine Model

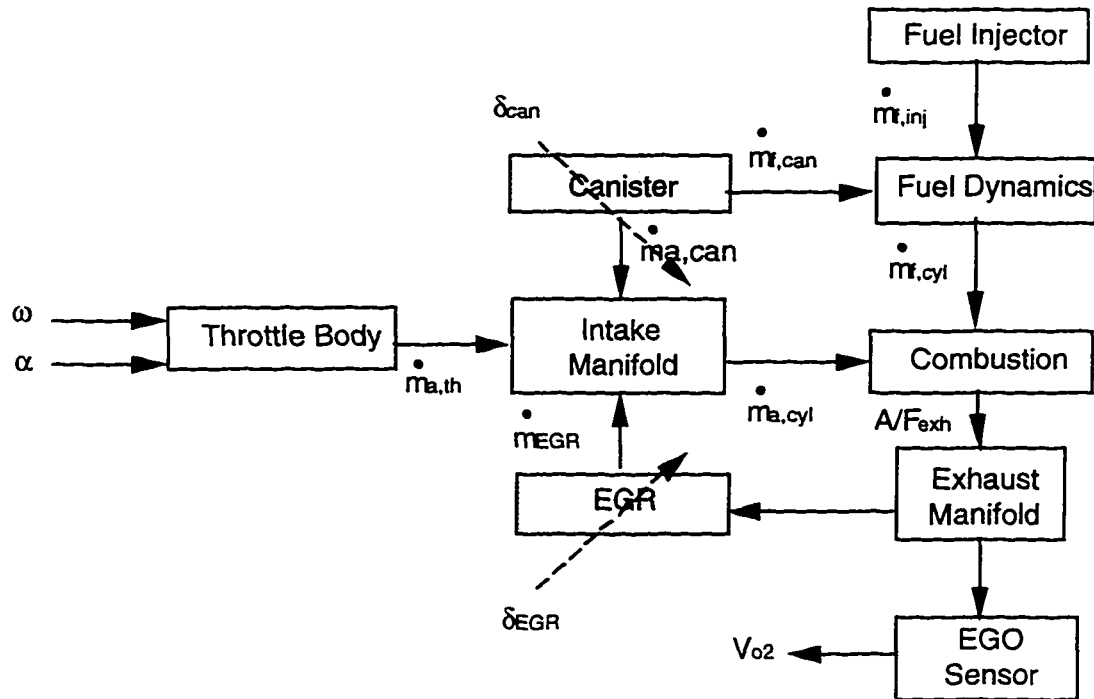


Figure 5.3 Air/Fuel Ratio Subsystem Model

A discrete crank-angle domain engine model was used towards the nonlinear observer design. In the next section the observer structure and design is covered.

### 5.1.1. Nonlinear Observers

#### Discrete Sliding Observer Design

In Section 4.1.1. the design background of nonlinear sliding observer was introduced in broad, general concepts. However, in this application practical considerations dictated the use the use of discrete crank-angle domain powertrain models and of discrete sliding

mode estimators. The design of the discrete sliding estimators was performed using the delta approximation; i.e., it is assumed that the following equation holds;

$$X(k+1) - X(k) = \Delta T \delta X(k) \quad (5.1)$$

where  $\delta X(k)$  approximates the derivative of the states at the sample instant  $k$  and  $\Delta T$  is the sampling interval. The difference equation describing the system is

$$X(k+1) = f(X(k)) \quad (5.2)$$

$$y = h(X) \quad (5.3)$$

We define

$$H = [h_1 \quad h_2 \quad \dots \quad h_n]^T \quad (5.4)$$

where

$$h_1(X) = f \circ h(X) \quad (5.5)$$

$$h_2(X) = f \circ f \circ h(X) = f^2 \circ h(X) \quad (5.6)$$

$$h_i(X) = f^i \circ h(X), \quad i = 2, \dots, \quad (5.7)$$

Then the observer can be constructed as follows:

$$\hat{X}(k+1) = \hat{X}(k) + \left( \frac{\partial H}{\partial X} \right)^{-1} (L - H(\hat{X})) \quad (5.8)$$

where:

$$L = [L_1 \quad L_2 \quad \dots \quad L_n]^T \quad (5.9)$$

$$L_1 = h_1(\hat{X}) + M_1 \text{sign}(y - h_1(\hat{X})), \text{ and} \quad (5.10)$$

$$L_i = h_i(\hat{X}) + M_i \text{sign}(L_{(i-1)eq} - h_i(\hat{X})), \quad i = 2, \dots, n$$

where  $M_i > 0$  represent the gains of the observer and  $L_{ieq}$  are low pass filtered versions of  $L_i$ . The proof of convergence for the observer, under the assumption that the delta approximation holds, is similar to the one for the continuous case, and is given in [134]. This discrete sliding mode estimator design procedure was applied to the problem of powertrain state and input estimation. Input estimation was performed through augmenting the list of states with the input that is to be estimated. Three observers were designed, as shown in Figure 5.4, to estimate different sets of variables for the purposes of residual generation for fault detection. The observers were designed as listed in Table 5.1 to estimate different sets of variables for the purposes of residual generation for fault detection the list of observers are displayed.

No.	Measured Variable	Estimated Variables
1.	$P_m$	$P_m, \alpha$
2	$P_m$	$P_m, \dot{m}_{ath}$
3.	$\phi_s, \alpha$	$P_m$

Table 5.1 List of Observed Quantities

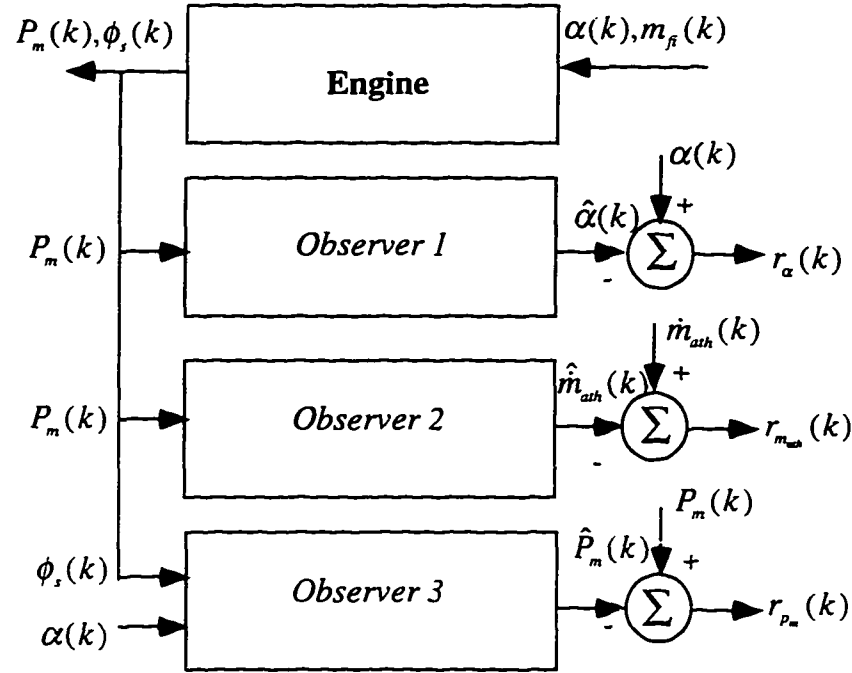


Figure 5.4 Observer Configuration

Note that in the above figure (^) denotes estimated variables. The generated residuals are  $r_\alpha$  from observer 1,  $r_{math}$  from observer 2, and  $r_{Pm}$  from observer 3. The residual elements are grouped into a single residual vector  $r=[r_\alpha \ r_{math} \ r_{Pm}]$ . In Figure 5.4  $\alpha$  is throttle position in degrees,  $\phi_s$  is the measured equivalence ratio, and

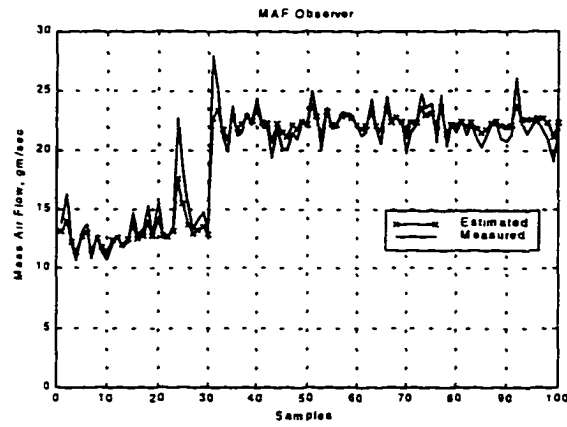
$P_m$  = intake manifold pressure [Pa]

$m_{fi}$  = fuel mass from injector [kg]

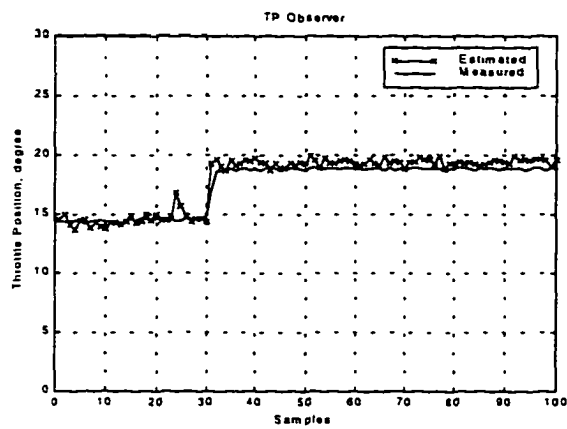
$\dot{m}_{ath}$  = air mass flow rate from throttle [kg/sec]

The performance of the observers, is summarized in Figure 5.5, where the mass air flow, the throttle position, and the intake manifold pressure are show in (a), (b), and (c),

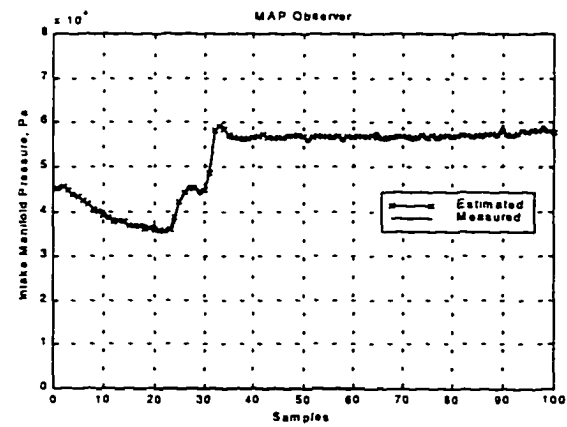
respectively, during a tip-in condition. It can be seen that the observers are capable of tracking the system's states and inputs in an acceptable manner, and are therefore well suited to serve as primary residual generators. In the next subsection we illustrate how fuzzy evaluation of these primary residuals is conducted.



(a) Mass Air Flow



(b) Throttle Position



(c) Intake Manifold Pressure

Figure 5.5 Nonlinear Observer Performance

## 5.2. FUZZY EVALUATION

The primary residuals are generated using the model-based scheme mentioned in the preceding section. These primary residuals can be non-zero not only due to the presence of faults, but also due to the presence of noise, model error, and disturbances. In binary threshold logic, where the threshold is crisp, a trade off between false alarms and missed detection is encountered. As explained in Section 4.2., the use of fuzzy logic to evaluate residuals can enhance the performance of the diagnostic scheme. This technique was applied to the experimental system described in Chapter 3. Two faults from Table 3.2 were induced: a calibration fault each of the MAF and TPS sensors, induced using the circuit described in Appendix A. Data for the no fault, and faulty sensors were acquired using the data acquisition system of Section 3.1.1. A section of the test cycle shown in Figure 3.5 (Throttle position 15% -20%) was chosen for validation purposes. The calibration fault was gradually induced covering a range from 0%-50% calibration fault of the input. The residuals from the observers used were generated for both the healthy and faulty condition. The mean and the standard deviation of the residual generated for the healthy condition were used to determine the membership functions for the fuzzy evaluation system. Two saturated triangular membership functions (low and high) were generated. The low threshold is calculated as the mean +  $2 \times$ (standard deviation), while the high is mean +  $4 \times$ (standard deviation); this was chosen to minimize the noise effects. These membership functions are shown in Figure 5.6. The vector of thresholds is evaluated using these membership functions, and is then fuzzified with min/max operators (as explained in Section 4.2) to produce symptoms  $F$ . Making the assumption



that faults occur one at a given time, and observing the structure of the residual generated, the fault isolation logic listed in Table 5.2 was used.

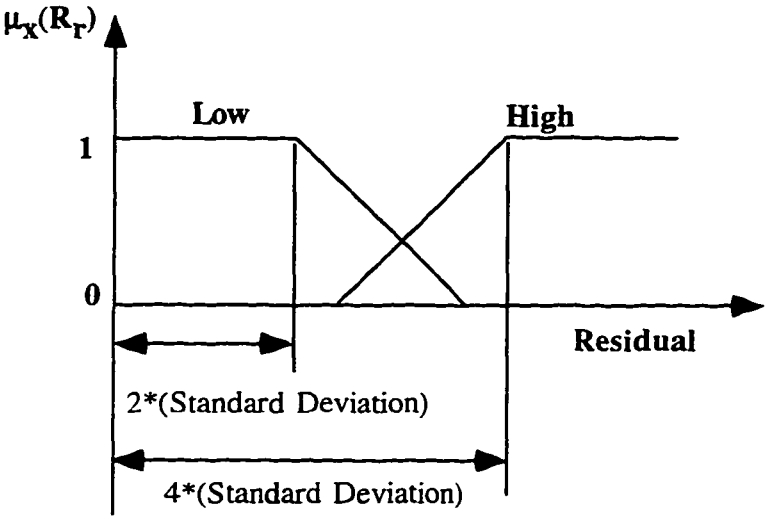


Figure 5.6 Fuzzy Residual Evaluation Membership Functions

<i>Fault Location</i>	<i>Symptoms F</i>
Throttle sensor	high, high, high
Mass flow meter	low, high, high
Manifold pressure sensor	low, low, high

Table 5.2 Fault Signature Table

The MAF 50% calibration fault is shown in Figure 5.7; Figure 5.8-5.10 show the residuals. Figure 5.11-5.14 show the corresponding results for the TPS calibration fault.

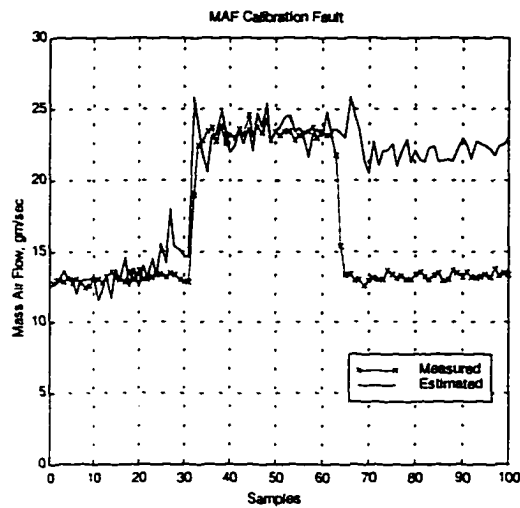


Figure 5.7 MAF Calibration Fault

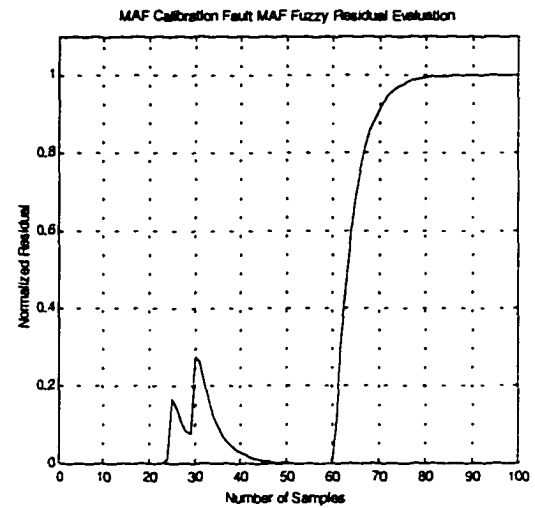


Figure 5.8 MAF Residual for MAF Fault

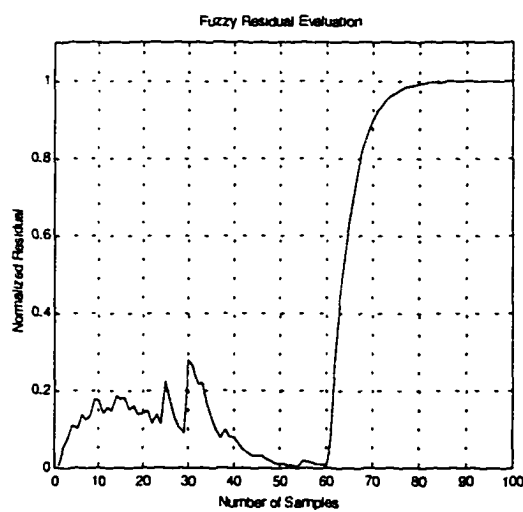


Figure 5.9 MAP Residual for MAF Fault

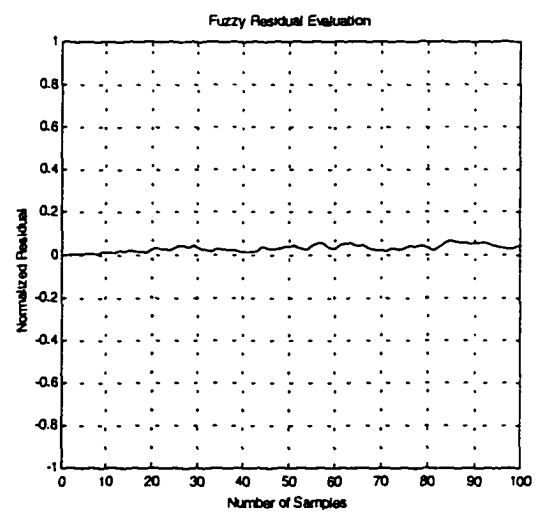


Figure 5.10 TPS Residual for MAF Fault

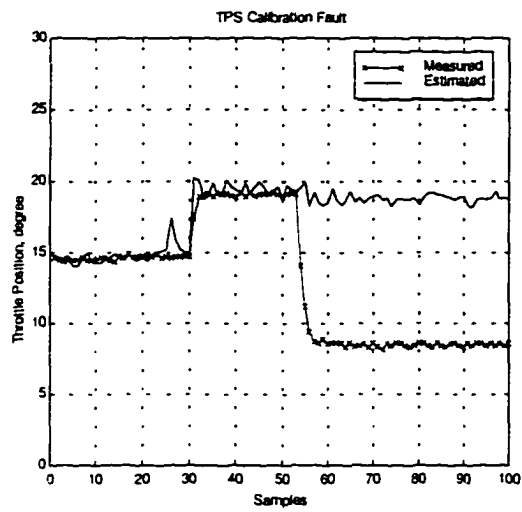


Figure 5.11 TPS Calibration Fault

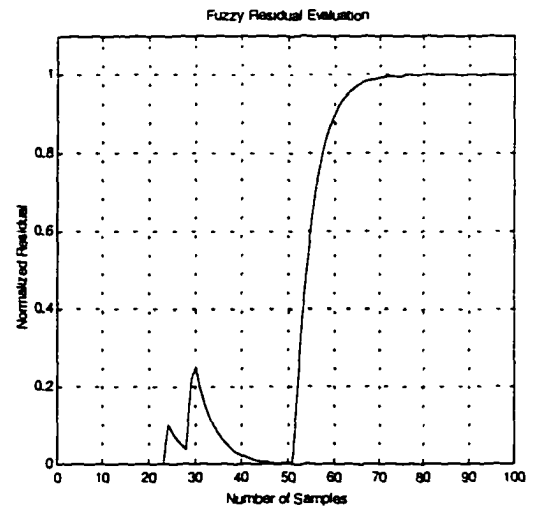


Figure 5.12 MAF Residual for TPS Fault

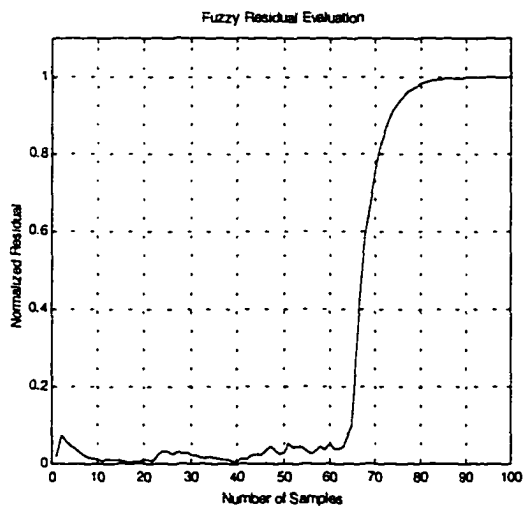


Figure 5.13 MAP Residual for TPS Fault

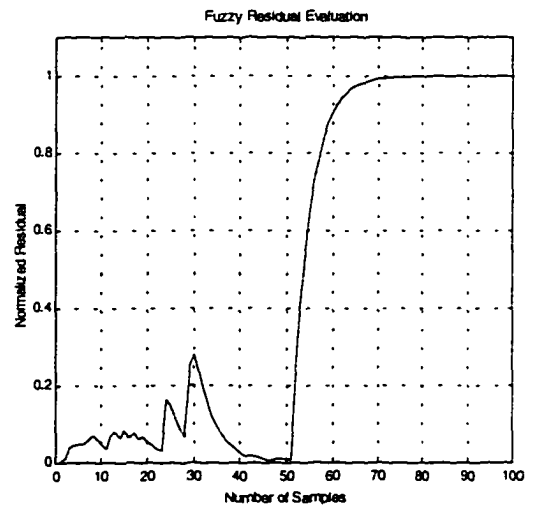


Figure 5.14 TPS Residual for TPS Fault

Once the symptom's vector  $F$  is generated, the next step is to isolate the fault. Fault isolation was made possible by the use of fuzzy logic. Fuzzy rules were generated using Table 5.3, where the antecedent is the level of the residual and the consequent is the fault. A crisp number  $k$ , was assigned to each fault. The rules for the MAF, and TPS fault are as follows;

If  $f_1$  is low and  $f_2$  is high and  $f_3$  is high Then fault is MAF=1

If  $f_1$  is high and  $f_2$  is high and  $f_3$  is high Then fault is TPS=2

where  $f_1$ ,  $f_2$ , and  $f_3$  are the TPS, MAF, and MAP residuals, respectively.

The same methodology can be used to generate five more rules, with a total of seven rules, one rule per fault induced. The outputs of these rules are displayed in Figure 5.15 and Figure 5.16, where the residuals are displayed as well as the histogram of the residuals. The histogram gives a clear indication of the isolated fault. The results shown in this section indicate that the model-based approach can be quite effective in isolating specific faults that would not be detected by the commercial OBD-II system. However, vehicle aging, vehicle-to-vehicle variability, and differences in operating conditions may severely degrade the performance of such a diagnostic system. The parallel path approach, based on fuzzy estimation from measured exhaust gas emissions, serves therefore a dual purpose: it reinforces the model-based diagnosis, and it provides integration between the engine control systems and the engine-out emissions.

		<i>Symptom</i>		
		$f_1$	$f_2$	$f_3$
<i>Fault</i>	<i>MAF=1</i>	<i>low</i>	<i>high</i>	<i>high</i>
	<i>TPS=2</i>	<i>high</i>	<i>high</i>	<i>high</i>

Table 5.3 Fuzzy Isolation Decision Table

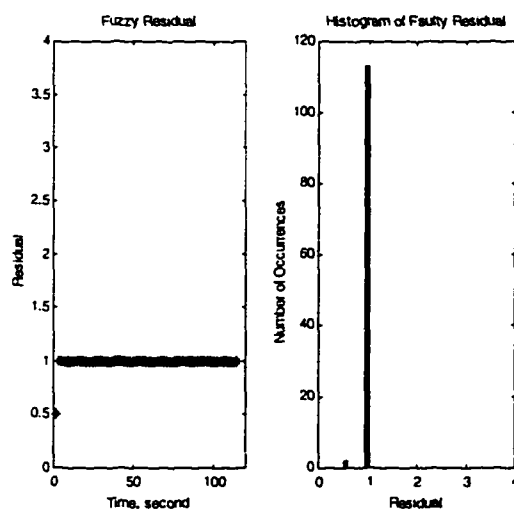


Figure 5.15 MAF Fault Isolation

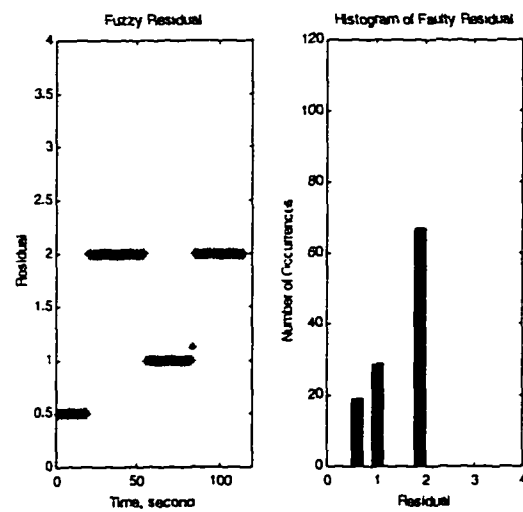


Figure 5.16 TPS Fault Isolation

### 5.3. FUZZY ESTIMATION

The system and subsystems in an internal combustion engine are strongly nonlinear. One of these subsystems is the combustion block of Figure 5.1, which is a challenge to model, especially when it comes to relating the charge air/fuel ratio with the exhaust components' levels. The change in the chemical balance of these species are the first symptom of a malfunction. Heuristic knowledge of the combustion process can be used to estimate the air/fuel ratio from the concentration level of the engine-out exhaust emissions. This knowledge can be in the form of fuzzy production *IF-THEN* rules. Moreover, input-output data can be used to generate additional rules that can be combined with the heuristic rules to form a rule-base for fault diagnosis. Fuzzy c-means clustering with optimal predefuzzification as covered in Section 4.3. was used to generate fuzzy rules. The three components regulated by the EPA,  $\text{NO}_x$ , CO, and HC, from the input data, while the air/fuel ratio measured by the UEGO sensor mentioned in Section 3.1.1. provides the output.

The heuristic rules are based on the relation between the air/fuel ratio and the three previously mentioned components which is shown in Figure 2.1. Fifteen linguistic rules that map the inputs to the output were generated and they are as follows:

1. *If* (HC is low) *then* (lambda is lean)
2. *If* (HC is normal) *then* (lambda is Stoichiometric)
3. *If* (HC is high) *then* (lambda is rich)
4. *If* (HC is very high) *then* (lambda is very rich)

5. *If* (HC is very low) *then* (lambda is very lean)
6. *If* (NO is very high) *then* (lambda is very lean)
7. *If* (NO is high) *then* (lambda is lean)
8. *If* (NO is normal) *then* (lambda is Stoichiometric)
9. *If* (NO is very low) *then* (lambda is very rich)
10. *If* (NO is low) *then* (lambda is rich)
11. *If* (CO is very high) *then* (lambda is very rich)
12. *If* (CO is high) *then* (lambda is rich)
13. *If* (CO is normal) *then* (lambda is Stoichiometric)
14. *If* (CO is very low) *then* (lambda is very lean)
15. *If* (CO is low) *then* (lambda is lean)

The linguistic variables that describe the input  $\tilde{u}_i \in U_i$  are the three exhaust species;  $\tilde{u}_1 =$  "HC",  $\tilde{u}_2 =$  "NO", and  $\tilde{u}_3 =$  "CO". The output linguistic variable  $\tilde{y}_i \in Y_i$  is the relative air/fuel ratio  $\tilde{y}_1 =$  "Lambda". The linguistic values for the inputs  $\tilde{A}_i = \{\text{very high, high, normal, low, very low}\}$ , and for the output  $\tilde{B}_i = \{\text{very rich, rich, stoichiometric, lean, very lean}\}$ . The universe of discourse for the inputs  $U_i$  and for the output  $Y_i$  is determined from the experiments. This fuzzy system has multi-input single-output MISO where  $n = 3$  inputs and  $N_i = 5$  membership functions on each universe of discourse. Therefore, the rule-base contains 125 possible rules according to  $\prod_{i=1}^n N_i = N_1 \cdot N_2 \cdots N_n$  [87]. The membership functions are Gaussian, for computational simplicity and are shown in Figures 5.17-5.20.

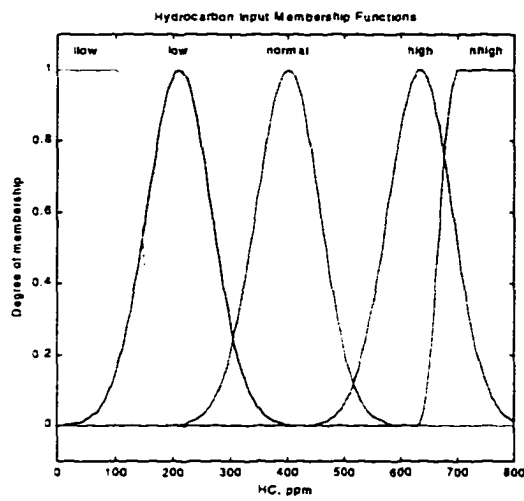


Figure 5.17 HC Membership Functions

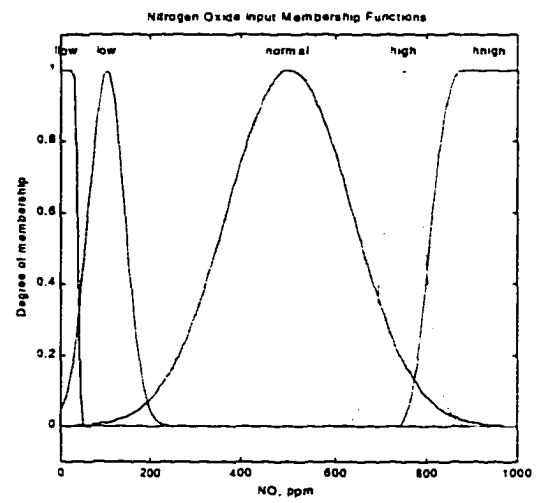


Figure 5.18 NO Membership Functions

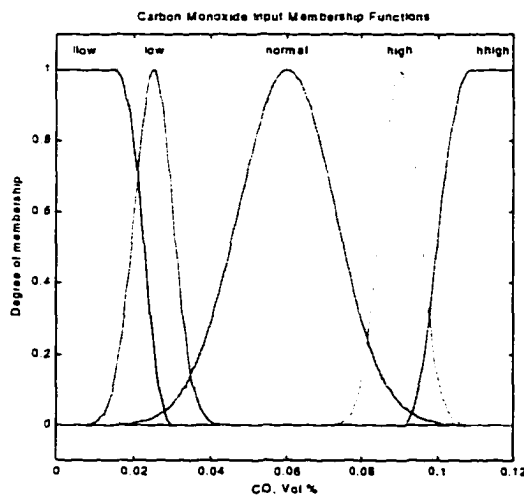


Figure 5.19 CO Membership Functions

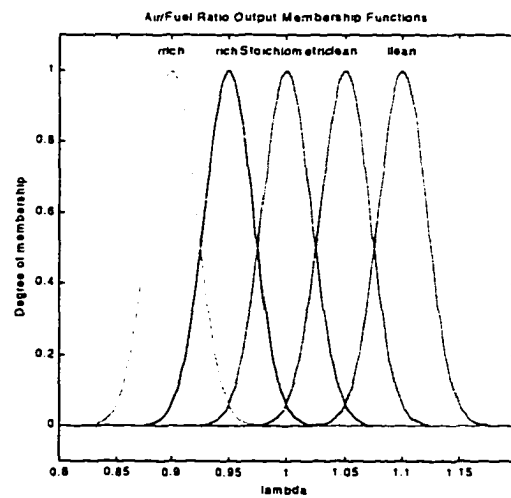
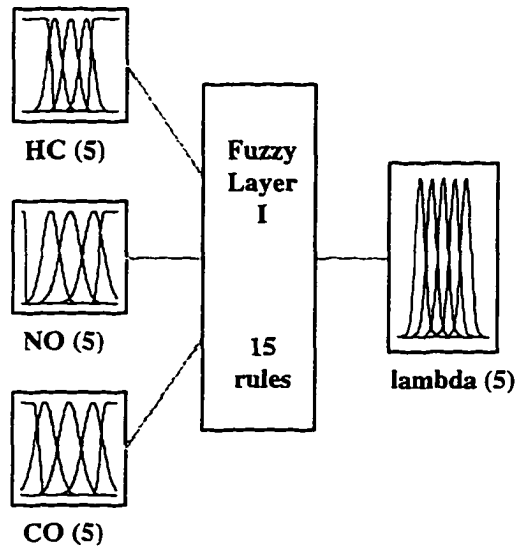


Figure 5.20  $\lambda$  Membership Function



The MATLAB<sup>®</sup> fuzzy logic toolbox [128] and a series of specially written programs were used to build the fuzzy system used in this study and shown in Figure 5.21. The fuzzy system used in this study has the following properties; “Gaussian fuzzification”; “and” = min; “or” = max; implication = min; aggregation = max; and defuzzification = centroid; the definitions of these terms are given in Appendix A.



**Fuzzy Layer I: 3 inputs, 1 outputs, 15 rules**

Figure 5.21 Heuristic Fuzzy System for  $\lambda$  Estimation

The data acquired from the FT-IR gas analyzer for the engine running at a healthy condition over the test profile of Figure 3.5 were used as the input for the fuzzy system. The baseline tests allowed for the tuning of the membership functions. The output is the estimated  $\lambda$  (the stoichiometric air/fuel ratio for this engine is 14.54). In Figure 5.22, the

estimated and measured values of the air/fuel ratio are shown. Furthermore, the input-output data were used to generate a fuzzy system using fuzzy c-means clustering with optimal predefuzzification as covered in Section 4.3.1. and the result of the air/fuel estimation for the same data used for the heuristic system is shown in Figure 5.23. This shows that the estimator generated using input-output data performed better than the one generated using heuristic rules.

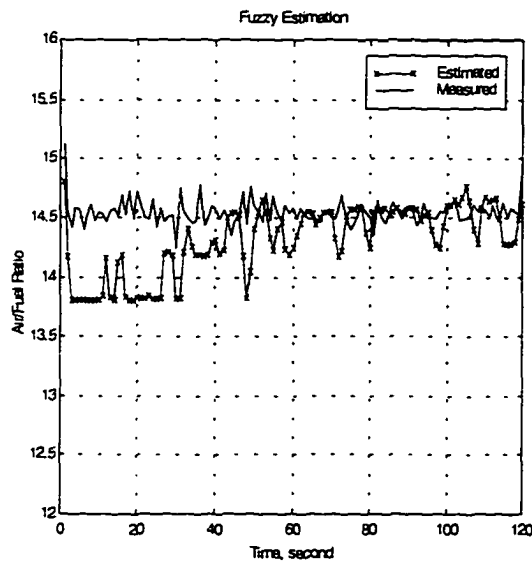


Figure 5.22 Heuristic Fuzzy Estimation of Air/Fuel Ratio

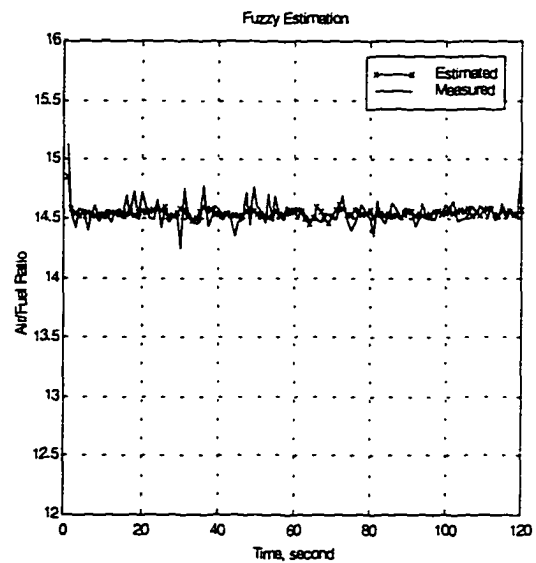


Figure 5.23 Fuzzy C-Means Estimation of Air/Fuel Ratio

The healthy engine data set was used to train the fuzzy system and to estimate the air/fuel ratio for all the tests representing the sample space, with 720 input-output data pairs,  $N=20$  clusters, cluster center's range of influence of 0.05, fuzziness factor  $m=1$ . The

5.24, with a maximum root mean square error (RMSE) = 0.07. The value 0.07 for the RSME was the threshold set to trigger a flag declaring a malfunction. Training the fuzzy system with healthy data enhanced the sensitivity of the estimator to faults. The faults listed in Table 3.2, were induced and the exhaust data was fed to the fuzzy system, the RMSE error increased substantially to a value  $\geq 0.1$ . The results from a run with a MAF calibration fault of 30% is shown in Figure 5.25 with RMSE = 0.1, and for a TPS 30% calibration fault is shown in Figure 5.26 with RMSE = 0.16. Moreover, the UEGO sensor output was compared to the estimated air/fuel ratio and is shown in Figures 5.27, and 5.38 for the intake and EGR faults, respectively. It was observed that for these faults the estimator gave values that represent the faulty condition, while the UEGO sensor was less sensitive to these faults. On the other hand some faults had the opposite effect on the UEGO sensor as shown in Figure 5.28, where it was more sensitive to these faults than the fuzzy estimator. This can be related to the fact that the UEGO sensor is designed in such a way that it responds to the oxygen partial pressure, and different faults had a distinct effect on the oxygen content in the exhaust emissions. Furthermore, since the fuzzy estimator was trained with healthy input-output data, the combined effect of different species on the estimator output when subjected to faulty data is the subject of future research, where the estimator could be trained using faulty data.

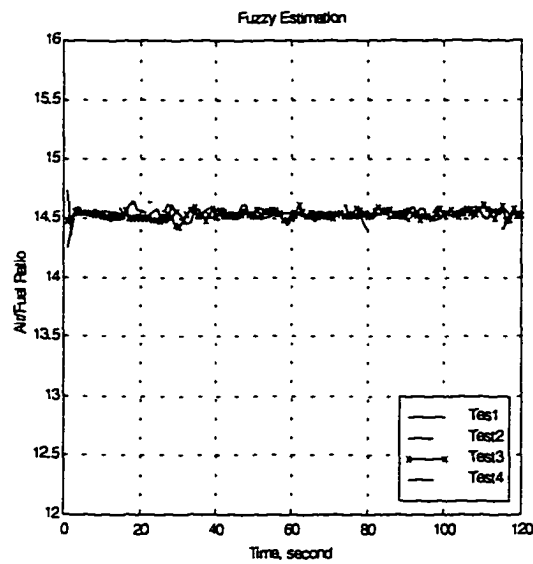


Figure 5.24 Estimated Air/Fuel Ratio Variability

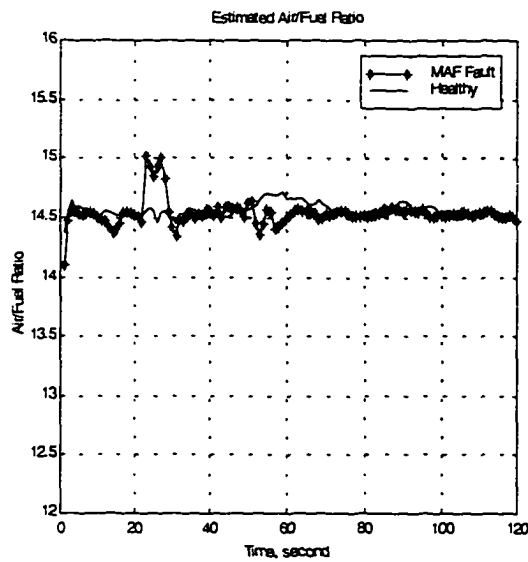


Figure 5.25 Estimated Air/Fuel Ratio for MAF Fault

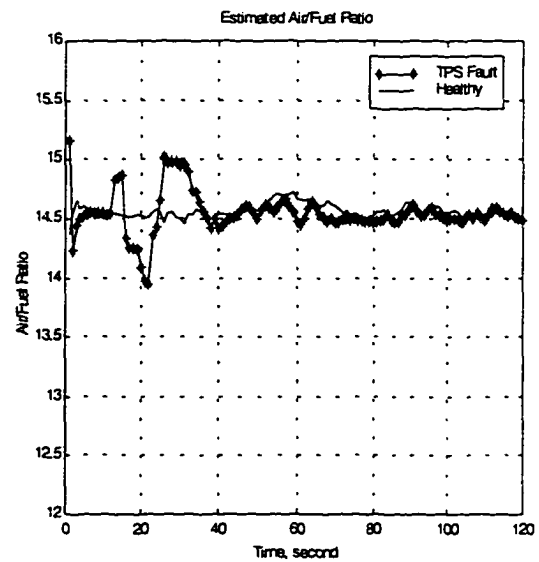


Figure 5.26 Estimated Air/Fuel Ratio for TPS Fault

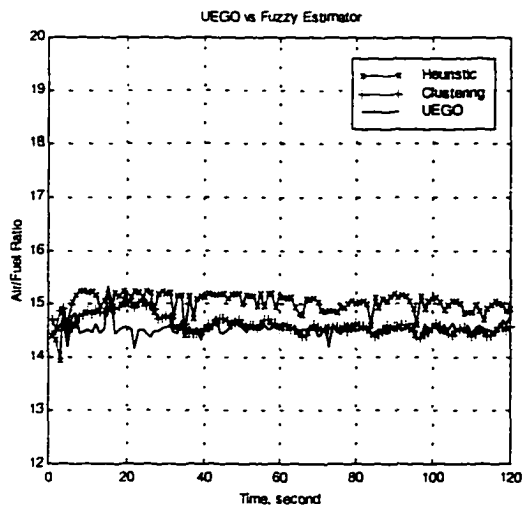


Figure 5.27 UEGO vs Fuzzy Estimator  
Intake Leak Fault

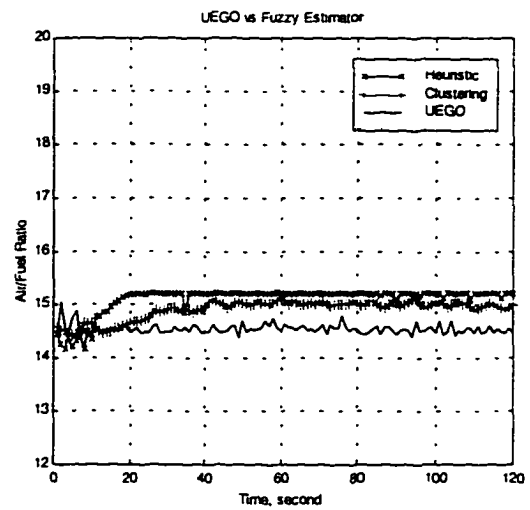


Figure 5.28 UEGO vs Fuzzy Estimator  
EGR Stuck Closed Fault

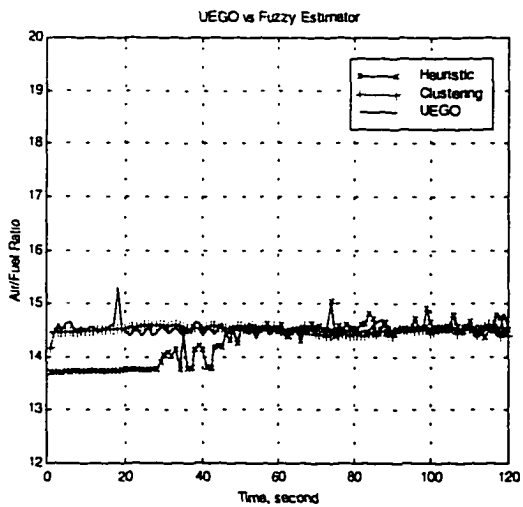


Figure 5.29 UEGO vs Fuzzy Estimator  
TPS Fault

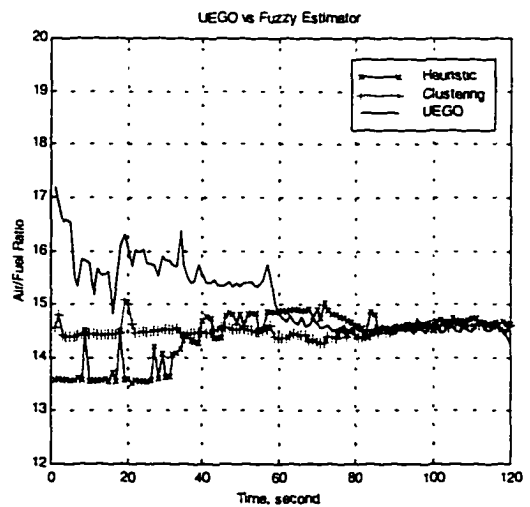
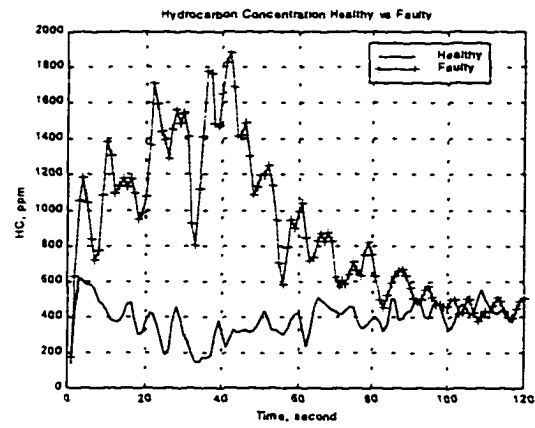


Figure 5.30 UEGO vs Fuzzy Estimator  
MAF Fault

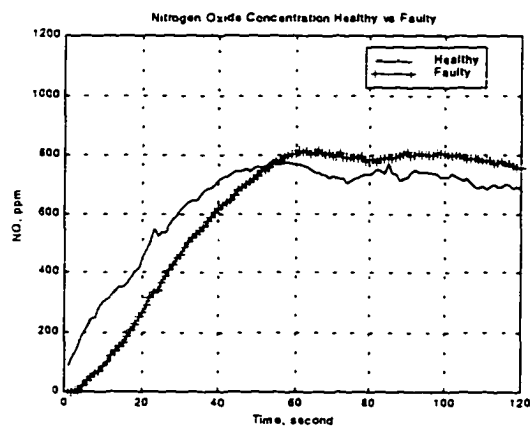
The fuzzy system that was used to estimate the air/fuel was tested for all the faults listed in Table 3.2 and proved to be robust to changes in ambient conditions. In spite of the fact that all the faults induced were detected, the source of the fault was yet to be determined. The fault isolation process was based on the way the exhaust emissions reacted to each fault independently. The fault isolation process required inducing one fault at a time and acquiring the exhaust data. The difference between the healthy and faulty condition in concentration level for each of the species was monitored, and fuzzy heuristic rules were generated to isolate the fault. The result of these tests are shown in Figure 5.31-Figure 5.37; were (a) HC concentration, (b) NO concentration, and (c) CO concentration respectively. The membership functions are chosen *a priori* for the input universe of discourse. The same membership functions generated heuristically for the air/fuel ratio estimation were used to fuzzify the inputs. The output of these rules was chosen to be a zero-order Takagi-Sugeno fuzzy system of the general form of equation (4.16) to enhance the efficiency of the defuzzification process. The fuzzy rule has the following form;

$$\text{IF } f_1 \text{ is } x_1^r \text{ and } f_2 \text{ is } x_2^r, \dots, \text{ and } f_m \text{ is } x_m^r \text{ THEN } A_r = k$$

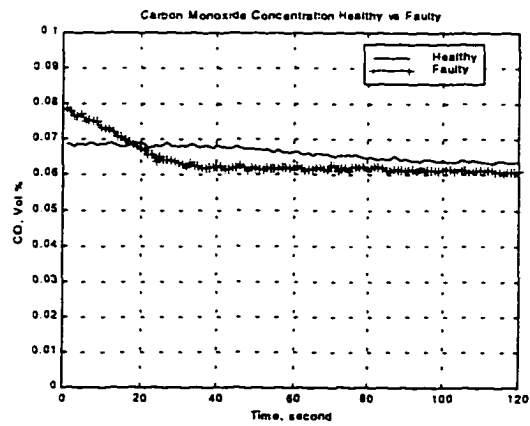
where  $k$  is a crisply defined constant, and each fault is assigned a crisp number as listed in Table 5.4.



(a)

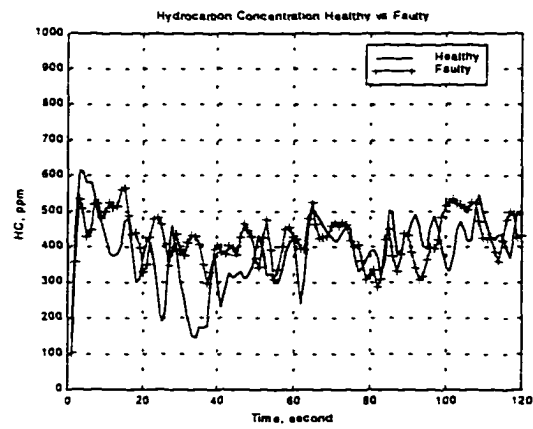


(b)

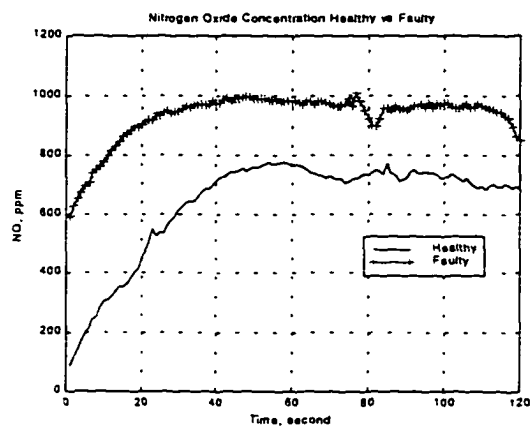


(c)

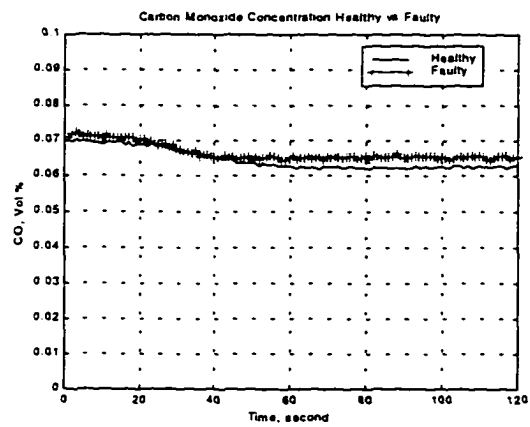
Figure 5.31 EGR Stuck Open Fault Species' Concentration Level



(a)



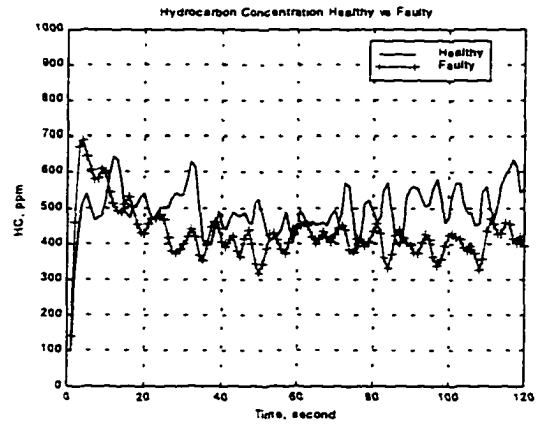
(b)



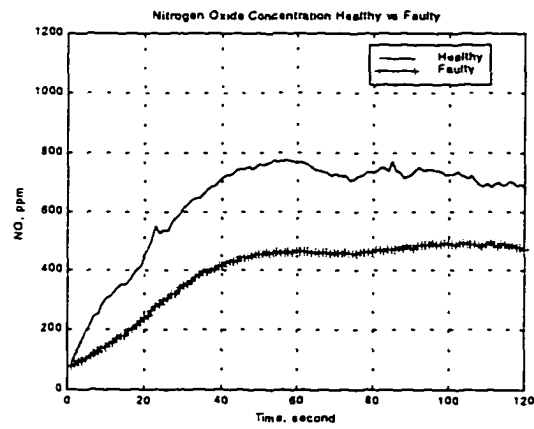
(c)

Figure 5.32 EGR Stuck Closed Fault Species' Concentration Level

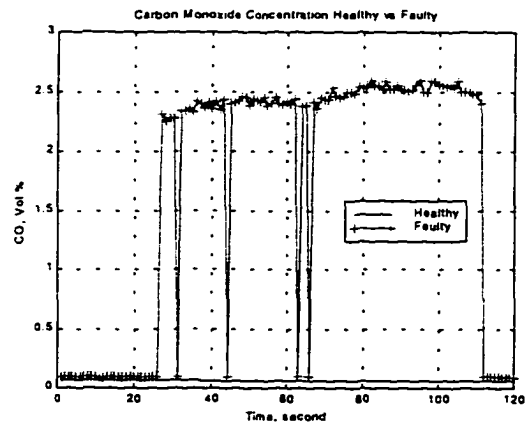




(a)

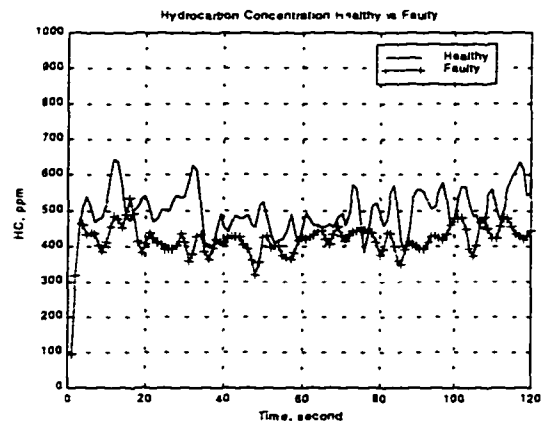


(b)

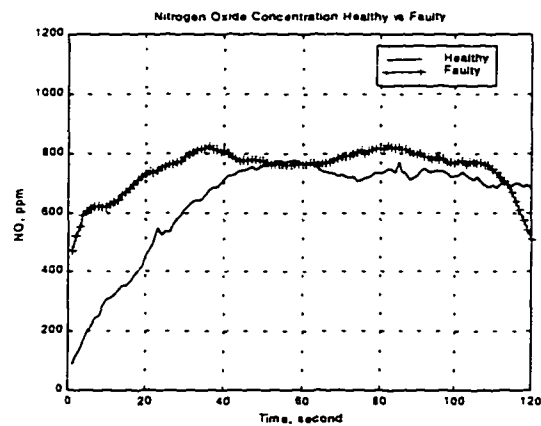


(c)

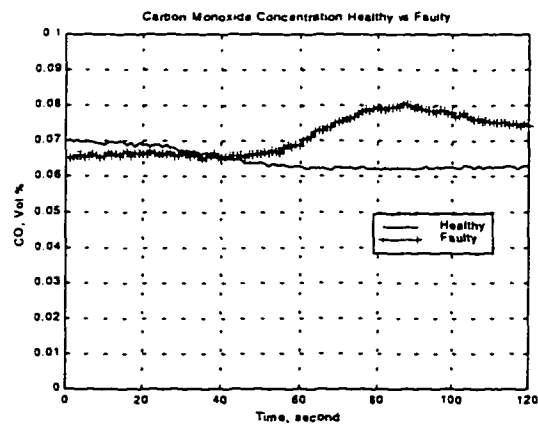
Figure 5.33 Injector Stuck Closed Fault Species' Concentration Level



(a)

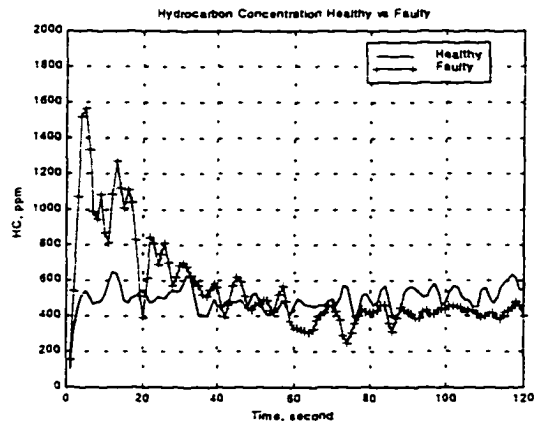


(b)

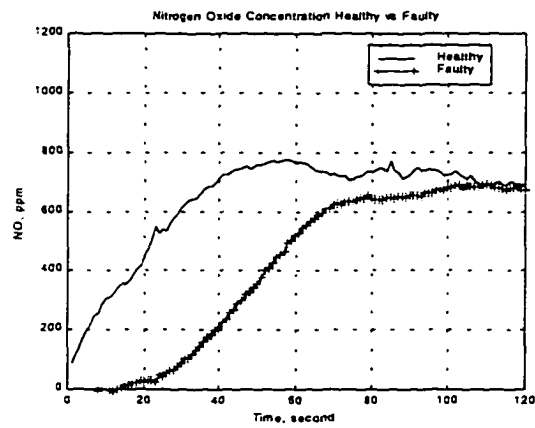


(c)

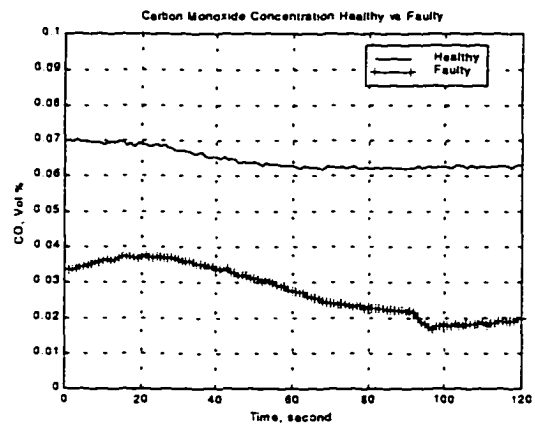
Figure 5.34 Misfire Fault Species' Concentration Level



(a)

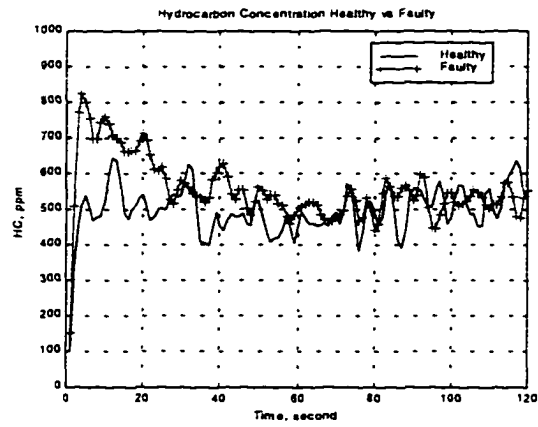


(b)

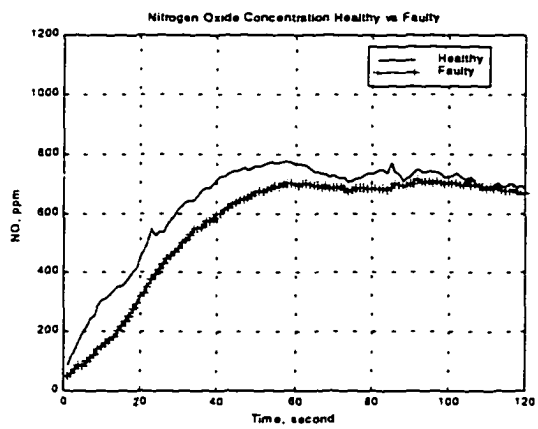


(c)

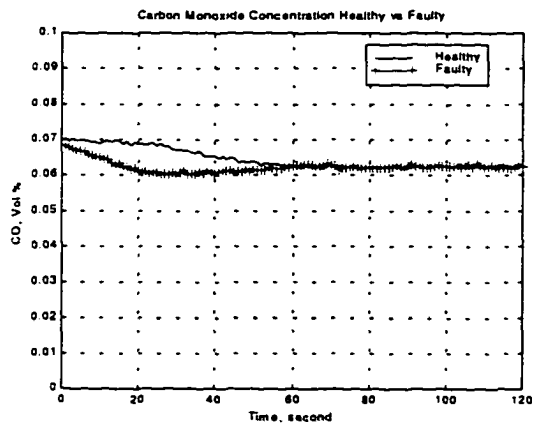
Figure 5.35 MAF Fault Species' Concentration Level



(a)

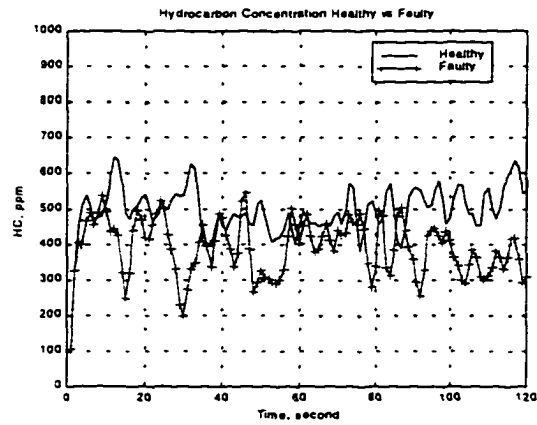


(b)

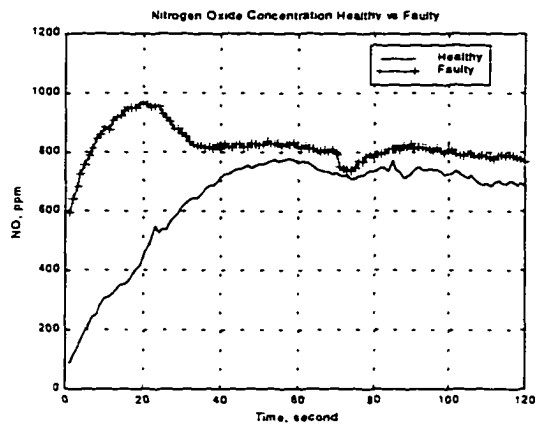


(c)

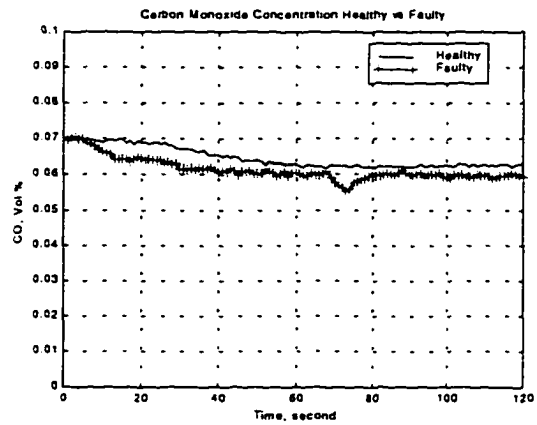
Figure 5.36 TPS Fault Species' Concentration Level



(a)



(b)



(c)

Figure 5.37 Intake Leak Fault Species' Concentration Level

The rules were generated as explained in Section 4.3. from the species' concentration levels (shown in Figures 5.31-5.37); eight rules were generated and they are as follows;

1. If (HC is very high) and (NO is very low) and (CO is low)  
then (Fault is MAF)
2. If (HC is high) and (NO is low) and (CO is normal) then (Fault is TPS)
3. If (HC is low) and (NO is high) and (CO is normal) then (Fault is Intake)
4. If (HC is normal) and (NO is low) and (CO is very high)  
then (Fault is Injector)
5. If (HC is normal) and (NO is high) and (CO is high) then (Fault is Misfire)
6. If (HC is very high) and (NO is low) and (CO is normal) then (Fault is EGRO)
7. If (HC is normal) and (NO is very high) and (CO is normal)  
then (Fault is EGRC)
8. If (HC is normal) and (NO is normal) and (CO is normal) then (Nofault)

Where EGRO and EGRC, are EGR opened and closed respectively.

The fuzzy system properties are: singleton fuzzification; Gaussian membership functions are used, product for the premise and implication; and center of gravity defuzzification.

The input symptoms are compared to the premises of all rules to determine which rule is applicable. The fault is determined using the rules that are fired at the time of detection.

Table 5.4 shows the faults and the assigned  $k$ , and the isolation output for the induced faults is shown in Figures 5.38-5.43, where the residuals and their histogram are displayed.

Fault	k
No Fault	0
MAF	1
TPS	2
Intake Leak	3
Injector	4
Misfire	5
EGR Stuck Open	6
EGR Stuck Closed	7

Table 5.4 Fuzzy Singleton Output for Fault Isolation

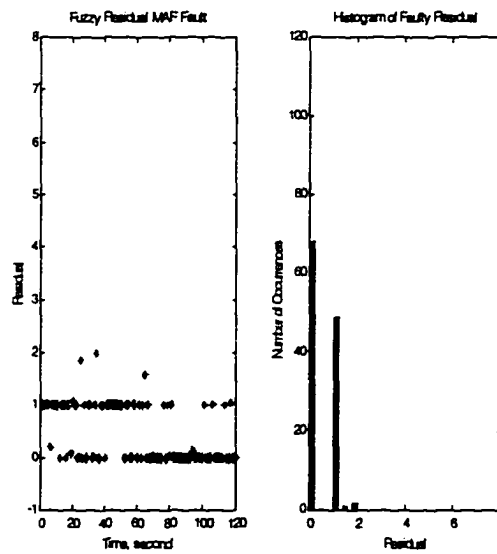


Figure 5.38 Fuzzy Output for MAF Fault Isolation

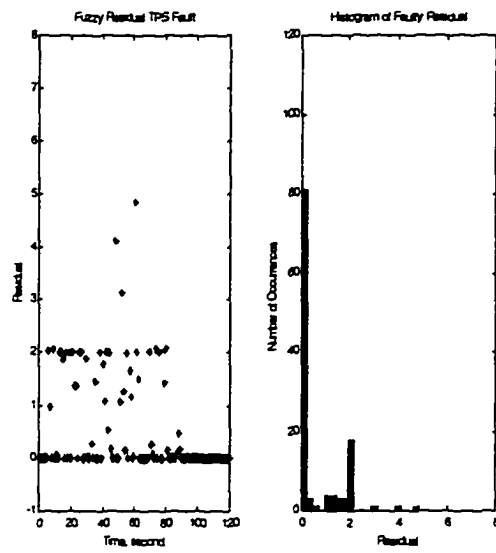


Figure 5.39 Fuzzy Output for TPS Fault Isolation

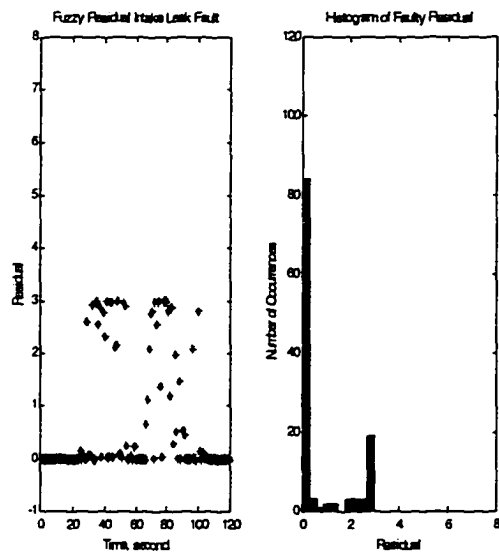


Figure 5.40 Fuzzy Output for Intake Leak Fault Isolation



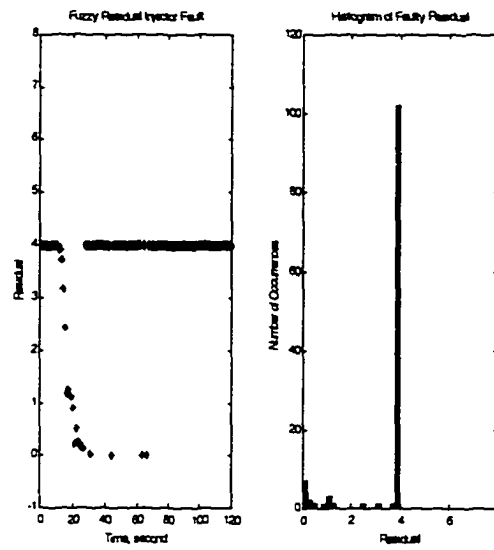


Figure 5.41 Fuzzy Output for Injector Fault Isolation

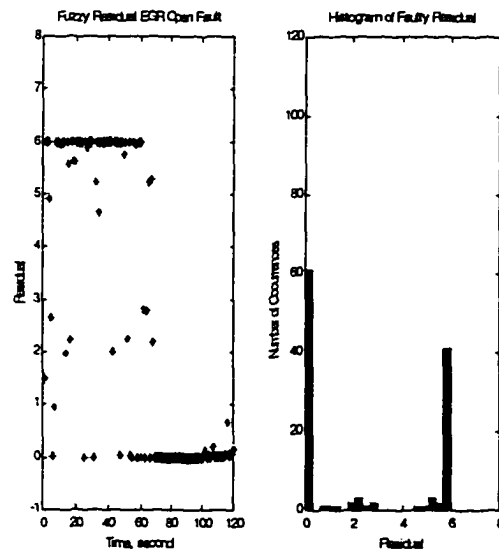


Figure 5.42 Fuzzy Output for EGR Open Fault Isolation

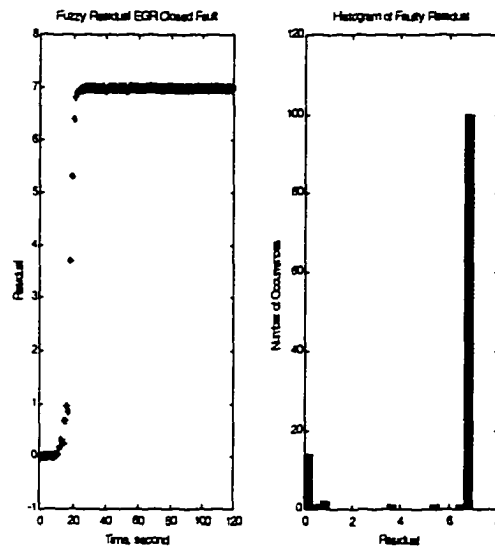


Figure 5.43 Fuzzy Output for EGR Closed Fault Isolation

Once the fault is isolated by both the fuzzy and the model-based schemes, the next step is to integrate the results for fault diagnosis. This can be achieved in the inference mechanism as covered in Section 4.4. This is possible because of the unified representation of all symptoms as a degree of membership. The rule-base integration scheme is covered in the next section.

#### 5.4. RULE INTEGRATION

The integration of the results obtained for the model- and knowledge-base schemes can be achieved at the inference mechanism. For each fault to be isolated there are two methods that can be used to achieve such integration;

- i. Combining the premise of the rule from the model- and knowledge-base methods in the rule-base.
- ii. Combining the defuzzified output of the rule from the model- and knowledge-base methods in the rule-base

The first method is possible in such cases where all *symptoms* have the same vector size and the data is acquired in the same domain (time or event based). Let the symptoms be  $F = \{f_1, \dots, f_r, \dots\}$ ,  $i = 1, \dots, m$ , and the linguistic values of the input be  $X^r = [x_1^r \ x_2^r \ x_m^r]^T$ ,  $r = 1, \dots, n$ , and the consequent of the rule be the faults. Then the symptoms  $F$  are matched against  $X^r$  and a fuzzy rule which is in the following form;

$$IF f_1 \text{ is } x_1^r \text{ and } f_2 \text{ is } x_2^r, \dots, \text{ and } f_m \text{ is } x_m^r \text{ THEN Fault is } A_r = k$$

In this application a typical rule is as follows;

$$IF \text{ HC is very high and NO is very low and CO is low and } f_1 \text{ is high and } f_2 \text{ is low and } f_3 \text{ is high THEN Fault is MAF} \approx 1$$

The fault can be isolated and the time of occurrence can be decided. The rule-base for this method included a large number of rules, as there are six inputs and six outputs.

The second method allows for the integration of the data available from different domains. The defuzzified output vector from both domains can be augmented in one vector, which is the symptom, and the premise for the production rule. This was possible

because the analytical and heuristic symptoms are represented in a unified way as proposed in [75]. The fuzzy system used “singleton” fuzzification and defuzzification which is covered in Appendix A, this simplified the process of combining the fuzzy sets that represent the inputs with the rule premises as suggested in [87]. The premise of the rule is the symptom and the consequent is the possible fault. The defuzzified outputs from both analytical and heuristic knowledge for the MAF and TPS calibration fault were tested. The applied rules were as follows;

*IF symptom is TPS THEN fault is TPS=100*

*IF symptom is MAF THEN fault is MAF=200*

Now as the integration is complete, the final step is the decision-making. Fuzzy systems are well suited for problems where solutions are problem dependent. The following section covers the fuzzy decision-process.

## 5.5. FUZZY DECISION-MAKING

Once the integration is complete, the linguistic rule-base is available for the inference mechanism. Multiple rule activation from all symptoms to all faults or sequential rule activation from level to level can be used for decision-making as proposed in [75]. The latter was used in this work as the isolation was performed by each scheme separately and symptoms can be partitioned. The input vector was made of the symptom  $f_i$ . The input was discretized to seven regions with five boundaries associated with the real number line. For the studied MAF and TPS fault the input was discretized as:

- MAF=1
- TPS=2

The output is the fault and for the MAF and TPS the region assigned for these faults are ;

- MAF=100
- TPS=200

The fuzzy system used “singletons” for the inputs and were positioned at 1 and 2 respectively. Also, for the output “singletons” were positioned at 100 and 200 respectively. Weighted average was used for defuzzification, and minimum was used to quantify the implication. The rules were as follows:

*IF* symptom is TPS *THEN* fault is TPS=100

*IF* symptom is MAF *THEN* fault is MAF=200

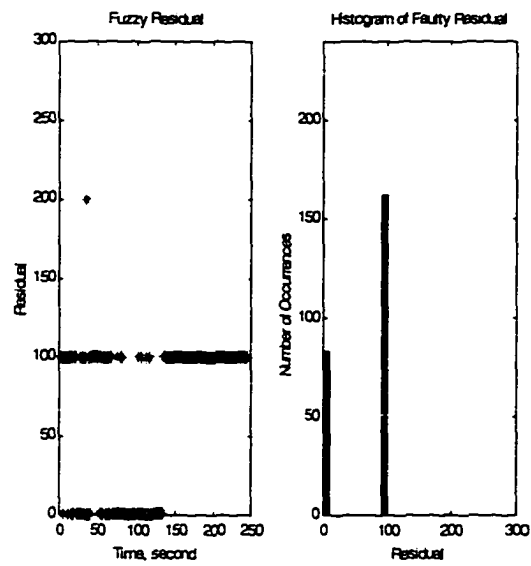


Figure 5.44 MAF Fault Decision

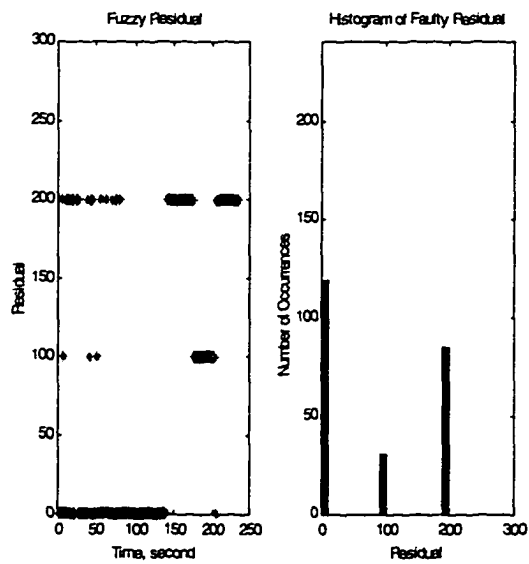


Figure 5.45 TPS Fault Decision

A vector which combined the outputs obtained from the heuristic fuzzy logic scheme and that of the analytical scheme was constructed. The symptom vector for one possible fault at a time was the input to the fuzzy decision-making system. The fault number “100” displayed in Figure 5.44 indicated a MAF fault, while the fault number “200” displayed in Figure 5.45 indicated a TPS fault.

## **5.6. SUMMARY**

In the application and results demonstrated in this chapter two faults were introduced separately. A MAF and a TPS calibration fault was induced using the fault induction circuit of Figure B.6. The input-output data from the engine and the exhaust gases were acquired simultaneously and processed by the diagnostic scheme of Figure 4.3. Along the first model-based path, the observers generated primary residuals as explained in Section 4.1., these residuals were evaluated using the fuzzy logic evaluation technique suggested in Section 4.2. This allowed for detection and isolation of the induced fault and a number was assigned to each fault. The second path applied the fuzzy logic estimation techniques suggested in Section 4.3. to the exhaust gas emissions. When the estimated air/fuel ratio indicated a malfunction, the next step was to isolate the fault, this was achieved by a fuzzy logic system that was able to isolate the fault and assign a number to the fault which matched the fault isolated by the model-based method. The integration was possible as the fault and its symptoms were represented in a unified form (Section 4.4.). The decision-making process was the result of the defuzzification of such integrated rules. The fuzzy diagnostic scheme was successful in detecting and isolating both faults

induced. More faults were detected and isolated using the knowledge-based method, yet the model-based method required the construction of more nonlinear observers.



## CHAPTER 6

### CONCLUSION AND FUTURE WORK

#### 6.1. CONCLUSION

The study presented in this dissertation has accomplished the goal of integrating model- and knowledge-based diagnostic methods in one diagnostic scheme to enhance the performance of fault diagnosis tools in complex nonlinear systems. This research work had demonstrated that fuzzy logic is a suitable method to integrate both types of knowledge. This integration has been accomplished by the amalgamation of two parallel paths. In the first, mathematical model-based methods are applied; in the second knowledge-based methods are used. The combined effect of the two methods, which are integrated at the decision-making stage, resulted in a more robust diagnosis.

The major results contributed by this dissertation is listed below.

- (1) Survey of model- and knowledge-based diagnostic methods for nonlinear systems.
- (2) Formulation of a methodology that integrates subjective and objective knowledge in one diagnostic scheme.

- (3) Development of a computer instrumented engine test cell for high accuracy engine data acquisition.
- (4) Verification of the diagnostic methodology in the context of an internal combustion engine emission control system.
- (5) Demonstration of the novel application of fuzzy logic as a decision-making tool in the automotive field.
- (6) Correlation of exhaust emission levels with engine malfunctions and including this heuristic information in the diagnostic scheme.

## **6.2. FUTURE WORK**

The work presented in this dissertation can be further extended to real-time implementation. Also, the completeness of the rule base will be of great value. Evaluation of the diagnostic performance in terms of recognized statistical detection criteria (e.g.: receiver operating characteristics) will result in a lower false alarm and miss-detection rate. Moreover, the implementation of such algorithm with real world hardware will enhance the diagnostic method practical value (e.g.: the field of Inspection and Maintenance (I/M) of vehicles).

## BIBLIOGRAPHY

- [1] A. J. Haagen-Smith, "Chemistry and Physiology of Los Angeles Smog," *Ind. Eng. Chem.*, vol.44, p. 1342, 1952.
- [2] M. H. Costin, "On-board diagnostics of vehicle emission system components: review of upcoming government regulations," in *Proceedings IFAC SAFEPROCESS symposium*, Baden-Baden, Germany, Sept. 1991, pp. 497-501.
- [3] M. Gray, "Overcoming the Limitations of the System Architecture of On-Board Vehicle Diagnostics," SAE paper no. 940432, presented at The International Congress and Exposition, Feb. 28-March 3, 1994, Detroit, MI.
- [4] N. A. Henein and D. J. Patterson, "Combustion Engine Economy, Emissions and Controls," University of Michigan Engineering Summer Conferences, June 1989.
- [5] K. Johnson, "Changes in Diagnostic Equipment as a Result of Technological Advances in Automotive Systems," SAE paper no. 861107, presented at the West Coast Int'l meeting, Universal City, CA, August 4-7, 1986.
- [6] M. J. Walker, "Portable Multi-Mode Analyzer for Electronic Ignition," SAE paper no. 800819 presented at SAE Passenger Car Meeting, Dearborn, MI, June 9-13, 1980.
- [7] R. Carp, R. Wakeman, and H. Weissler, "A Low-Cost Interactive Diagnostic Tester," SAE paper no. 850426, presented at SAE International Congress & Exposition, Detroit, MI, February 25- March 1, 1985.
- [8] R. Van Beylen, T. Dunn, and F. Maaseidvaag, "The Pragmatics and Development of a Diagnostic Aid for Automotive Electronic Systems," SAE paper no. 850423, presented at SAE International Congress & Exposition, Detroit, MI, February 25-March 1, 1985.
- [9] J. Troutner, "An On-Board Microprocessor Vehicle Diagnostic Instrument," SAE paper no. 830276, presented at SAE International Congress & Exposition, Detroit, MI, February 25- March 1, 1985.

- [10] S. Rachedi and L. Tedesco, "Development of Diagnostic Tools in Automotive Electronics," SAE paper no. 871582, presented at SAE Future Transportation Technology Conference & Exposition, Seattle, Washington, August 10-13, 1987.
- [11] J. Zaleski and C. Perkins, "TECH 1 Interactive Diagnostic Tester," SAE paper no. 861108, presented at the West Coast International meeting, Universal City, CA, August 4-7, 1986.
- [12] P. Malaterre, "Transient Fault Recorder," ISATA paper no. 88034, 1988.
- [13] L. Tedesco, "Service Bay Diagnostic System," SAE paper no. 861030, presented at SAE International Congress & Exposition, Detroit, MI, February 1986.
- [14] S. Hamilton, "Computerized Automotive Maintenance System (CAMS)," 16<sup>th</sup>. ISATA paper no. 87054A, proceeding Vol.2, Florence, Italy, May 1987.
- [15] SAE, "On-Board Diagnostics for Light and Medium Duty Vehicle Standards Manual," SAE HS-3000, 1997.
- [16] R. Milne, "Strategies for Diagnosis," IEEE Transactions On Systems, Man, And Cybernetics, Vol. SMC-17, no. 3, May/June 1987.
- [17] S. G. Tzafestas, "System Fault diagnosis Using Knowledge-based Methodology," Chapter 15 of "Fault Diagnosis in Dynamic Systems," Patton et al., Prentice Hall Book, 1989.
- [18] R. Klausmeier, "Using A. I. in the Vehicle Diagnostic Systems," West Coast International Meeting, Universal City, CA , August 4-7, 1986.
- [19] D. Dankel II, "Expert Systems: Misconception and Reality," SAE paper no. 860334, presented at SAE International Congress & Exposition, Detroit, MI, February 1986.
- [20] P. K. Fink and J. C. Lusth, "A Second Generation Expert System for Diagnosis and Repair of Mechanical and Electrical Devices," SAE paper no. 860337, presented at SAE International Congress & Exposition, Detroit, MI, February 1986.
- [21] L. A. Zadeh, "The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems," Part I of, "Approximate Reasoning In Expert Systems," M. M. Gupta et al. eds, Elsevier Science Pub. Co., Inc., NY, 1985.
- [22] J. C. Bezdek, "Analysis of Fuzzy information," Volume II Artificial Intelligence and Decision Systems, CRC Press, Inc., 1987.
- [23] O. Hall and A. Kandel, "Designing Fuzzy Expert Systems," TUV Rheinland, 1986.
- [24] A. Kandel, "Fuzzy Expert Systems," CRC Press, Inc., Florida USA, 1992.

- [25] A. Di Nola et al., "Fuzzy Relation Equations and their Application to Knowledge Engineering," Kluwer Academic Publishers Book, 1989.
- [26] E. H. Shortliffe, "Computer-Based Medical Consultations: *MYCIN*," Elseiver, NY, USA 1976.
- [27] F. Hayes-Roth, D. A. Waterman, and D.B. Lenat, "Building Expert System," Addison-Wesley Publishing Co, Inc MA, USA, 1983.
- [28] J. Kriz, "Knowledge-Based Expert System in Industry," Ellis Horwood Limited, West Sussex, England, 1987.
- [29] I. Graham and P. L. Jones, "Expert Systems Knowledge, Uncertainty, and Decision," Chapman and Hall Ltd, NY, USA, 1988.
- [30] C. V. Negoita, "Expert Systems and Fuzzy Systems," Benjamin/Cummings Publishing Co, Inc, CA, USA 1985.
- [31] S. M. Weiss and C. A. Kulikowski, "A Practical Guide to Designing Expert Systems," Rowman and Allanheld, Publishers, NJ, USA 1984.
- [32] S. M. Weiss and C. A. Kulikowski, "Representation of Expert Knowledge for Consultation: The CASNET and EXPERT Projects," in Artificial Intelligence in Medicine, Szolovits, P., Ed., AAAS Symp. Series, Westvie Press, Boulder, CO, USA, 1982.
- [33] L. Johnson and E. T Keravnou, "Expert Systems Architecture," Kogan Page Limited, London, England 1988.
- [34] Simon Braun, Ed., Mechanical Signature Analysis, Academic Press, 1986.
- [35] K. Marko, B. Bryant, and N. Soderborg, "Neural Network Application To Comprehensive Engine Diagnostics," paper presented at IEEE Systems, Man, And Cybernetics conference, Chicago, October 22, 1992.
- [36] R. DeJong, R. Powell, and J. Manning, "Engine Monitoring Using Vibration Signals," SAE paper no. 861246, presented at International Off-Highway & Powerplant Congress & Exposition, September 8-11, 1986, Milwaukee, WI, USA.
- [37] L. O'Connor, "Diagnosing Diesel Engines ," Mechanical Engineering, March 1992, vol.114/No.3.
- [38] A. Sood et al., "Engine Fault Analysis: Part I- Statistical Methods," IEEE Transactions on Industrial Electronics, Vol.IE-32, No.4, November 1985.

- [39] W. B. Ribbens, "Failure Diagnosis For Automotive Applications," 18<sup>th</sup> ISATA paper no. 88034, 1988.
- [40] J. Gertler, "A Survey of Fault Detection and Identification," proceedings of IFAC, IMACS SAFEPROCESS Symposium, Baden- Baden, Germany, September 1991.
- [41] J. Gertler, "Fault Detection and Diagnosis in Engineering Systems," Marcel Dekker, New York, 1998.
- [42] R. Patton and C. Chen, "Model based fault detection and identification methods," Proceedings of IFAC, IMACS SAFEPROCESS Symposium, Baden Baden, Germany, September 1991.
- [43] R. Isermann, "Fault Detection and Identification by Parameter Estimation," Proceedings of IFAC, IMACS SAFEPROCESS Symposium, Baden Baden, Germany, September 1991.
- [44] G. Rizzoni, "Fault Diagnosis in Mechatronic Systems," ME 874, graduate course at the Dept. of Mechanical Engineering, the Ohio State University.
- [45] G. Rizzoni and K. Srinivasan, "Powertrain Dynamics," ME 971, graduate course at the Dept. of Mechanical Engineering, the Ohio State University.
- [46] G. Rizzoni and P.S. Min, "Application of Fault Detection Filters to the Diagnosis of Sensor Faults in Electronically Controlled Engines," IEEE Transactions on Vehicular Technology, May, 1991.
- [47] J. Gertler, et al., "Model-based on-board fault detection and diagnosis for automotive engines," in Proc. IFAC SAFEPROCESS symposium, Baden-Baden, Germany, Sept. 1991, pp. 503-508.
- [48] J. Gertler, et al., "Model-based on-board fault detection and diagnosis of automotive engines," Control Engineering Practice, vol. 1, No.1 pp. 3-17, Jan. 1993.
- [49] G. Rizzoni, "Estimate of Indicated Torque from Crankshaft Speed Fluctuations: A Model for the Dynamics of the IC Engine," IEEE Transactions on Vehicular Technology, August, 1989.
- [50] F. Connolly and G. Rizzoni, "Real-time Estimation of Engine Torque for the Detection of Engine Misfires," in Proceedings of 1991 ASME Symposium on Advanced Automotive Technologies, Atlanta, GA, December, 1991.
- [51] G. Rizzoni, P. Azzoni, and G. Minelli, "On-board diagnosis of emission control system malfunctions in electronically controlled internal combustion engines," in Proceedings of TOOLDIAG '93, Toulouse, France, April 3-5, 1993.

- [52] J. Grizzle, K. Dobbins, and J. Cook, "Individual Cylinder Air-to-Fuel Ratio Control Using a Single EGO Sensor," in *IEEE Transactions on Vehicular Technology*, August 1991.
- [53] K. M. Passino and P. J. Antsaklis, "Fault Detection and Identification in an Intelligent Restructurable Controller," *Journal of Intelligent and Robotics Systems*, Vol. 1, No. 1, pp. 145-161, June 1988 .
- [54] P. J. Antsaklis, K. M. Passino, S. J. Wang, "Towards Intelligent Autonomous Control Systems: Architecture and Fundamental Issues, *Journal of Intelligent and Robotics Systems*, Vol. 1, No. 4, pp. 315-342, 1989.
- [55] P. J. Antsaklis and K. M. Passino, eds., "An Introduction to Intelligent and Autonomous Control," Kluwer Academic Publishers, Norwell, MA, 1993.
- [56] W. A. Kwong, K. M. Passino, E. G. Lukonen, and S. Yurkovich, "Expert Supervision of Fuzzy Learning Systems for Fault Tolerant Aircraft Control," *Proceedings of the IEEE*, Vol. 83, No. 3, March 1995 .
- [57] J. B. Heywood, "Internal Combustion Engine Fundamentals," McGraw-Hill Company Book, New York, 1988.
- [58] B. Köppen-Seliger, N. Kiupel, H.S. Kellinghaus., and P.M. Frank, "A Fault Diagnosis Concept for a High-Pressure -Preheater Line," 34<sup>th</sup> CDC, paper no. 295085.
- [59] N. Kiupel and P.M. Frank, "Process supervision with the aid of fuzzy logic," *Proc. IEEE International Conference on Systems, Man and Cybernetics*, Le Touquet, France, October 1993.
- [60] P. M. Frank and N. Kiupel, "Fuzzy supervision and application to lean production," *International Journal of Systems and Science*, vol. 24, No. 10, pp 1935-1944, 1993.
- [61] C. J. Kim and R. B. Don, "A Structure of Fuzzy Decision-Making System for Power System Production," *IEEE International Conference on Fuzzy logic* pp. 998-1003 1993.
- [62] H. Schneider and P.M. Frank, "Observer-Based Supervision and Fault Detection in Robots Using Nonlinear and Fuzzy Logic Residual Evaluation," *IEEE Transactions on Control Systems Technology* , vol. 4, No.3, pp.274-282, May 1996.
- [63] C. Auburn, D. Sauter, H. Noura, and M. Robert, "Fault Diagnosis and reconfiguration of systems using fuzzy logic: application to a thermal ," *International Journal of Systems and Science*, vol. 24, No. 10, pp 1945-1954, 1993.

- [64] J. P. Rosselot, J. F. Balmat, and A. Gut, "A Fuzzy Logic Knowledge-Based System in Naval Decision-Support Aids," IEEE International Conference on Fuzzy logic pp. 833-838, 1993.
- [65] L. A. Zadeh, "Fuzzy logic," IEEE Computer, vol.21, pp. 83-93, April 1988.
- [66] L. A. Zadeh, "Fuzzy sets," Information and Control, vol.8, pp. 383-353, 1965.
- [67] P. Ralston, and T. Ward, "Fuzzy Logic Control of Machining," Manufacturing Review, vol. 3, no. 3, September 1990.
- [68] N. Kasabov, "Hybrid Fuzzy Connectionist Rule-Based Systems and the Role of Fuzzy Rules Extraction," IEEE International Conference on Fuzzy Systems, vol.1, pp.49-56, Japan, March 1995.
- [69] Z.X. Yang, K. Susuki, Y. Shimada, and H. Sayama, "Fuzzy Fault Diagnostic System based on Fault Tree Analysis," IEEE International Conference on Fuzzy Systems, vol.1, pp.165-170, Japan, March 1995.
- [70] P. Chen, T. Toyota, and Y. Sasaki, "Fuzzy Diagnosis and Fuzzy Navigation for Plant Inspection and Diagnosis Robot," IEEE International Conference on Fuzzy Systems, vol.1, pp.185-192, Japan, March 1995.
- [71] P.M. Frank, "Fault Diagnosis in Dynamic Systems using Analytical and knowledge-based Redundancy- A Survey and Some Results," Automatica, vol.26, no. 3. pp. 459-474, 1990.
- [72] R. Isermann and P. Ballé, "Trends in the Application of Model based fault Detection and Diagnosis of Technical Processes," IFAC 13<sup>th</sup> Triennial World Congress, pp. 2-12, San Francisco, USA 1996.
- [73] L. X. Wang and J. Mendel, "Generating Fuzzy Rules by Learning from Examples," IEEE Transactions on Systems, Man and Cybernetics, vol.22, no.6, pp. 1414-1427, 1992.
- [74] B. D Liu and C. Y. Huang, "Systematic Design Approach for Multivariable Fuzzy Expert System," Proceedings of IEEE International Conference on Fuzzy Systems, vol. 3, pp. 2094-2099, 1994.
- [75] R. Isermann and T. Pfeufer, "Integration Model Based Fault Detection and Diagnosis with Application to Electromechanical Drives," IFAC 13<sup>th</sup> Triennial World Congress, pp. 85-90, USA 1996.
- [76] E. C. Payne and R. C. McArthur, "Developing Expert System," John Wiley and Sons, Inc. N.Y, USA, 1990.



- [77] K. Lecot and D. Parker, "Customizable Inexact Reasoning," IEEE Proceedings of Western Conference on Knowledge-Based engineering and Expert Systems, pp.95-101, June 1986.
- [78] R. Yager, "Expert Systems Using Fuzzy Logic," in "An Introduction to Fuzzy Logic Applications in Intelligent Systems," R. Yager, and L. Zadeh eds., pp. 27-44, Kluwer Academic Publishers, Massachusetts, USA, 1992.
- [79] L. Zadeh, "Knowledge Representation in Fuzzy Logic," in "An Introduction to Fuzzy Logic Applications in Intelligent Systems," R. Yager, and L. Zadeh eds., pp. 1-25, Kluwer Academic Publishers, Massachusetts, USA, 1992.
- [80] L. Zadeh, "Test-Score Semantics as a Basis for a Computational Approach to Representation of Meaning," Literary and Linguistic Computing 1, pp. 24-35, 1986.
- [81] R. R. Yager, "On the Representation of Relational Production Rules in Expert Systems," in "Fuzzy Expert Systems," A. Kandel, Ed., pp. 55-67, CRC Press, Inc., Florida USA, 1992
- [82] A. Jones, A. Kaufmann, and H. Zimmermann, "Fuzzy sets Theory and Applications," D. Reidel Publishing Co., Holland, 1985.
- [83] D. Dubois and H. Prade, "Fuzzy sets and Systems: Theory and Applications," Academic Press Inc., N.Y., USA 1980.
- [84] A. Kandel, "Fuzzy Mathematical Techniques with Applications," Addison-Wesley Publishing Co., Inc., USA, 1986.
- [85] A. Geyer-Schulz, "Fuzzy Rule-Based Expert Systems and Genetic Machine Learning," Physica-Verlag Heidelberg, Germany 1995.
- [86] E. Mamdani and B. Gaines eds, "Fuzzy Reasoning and its Applications," Academic Press Inc. N.Y., USA 1981.
- [87] K. M. Passino and S. Yurkovich, "Fuzzy Control," Addison-Wesley Longman, Inc., Menlo Park, California, USA, 1998.
- [88] S. Yurkovich, "A Control Engineer's Perspective on Fuzzy Control," Presented at European Control conference, Brussels, Belgium, July 1997.
- [89] K. M. Passino and S. Yurkovich, "Fuzzy Control," In The Control Handbook, W. S. Levine, ed., Boca Raton: CRC Press, 1996.
- [90] E. Laukonen, "Training Fuzzy Systems to Perform Estimation and Identification," Master's Thesis, Dept. of Electrical Engineering, The Ohio State University, 1994.

- [91] E. G. Laukonen, K. M. Passino, V. Krishnaswami, G-C. Luh, and G. Rizzoni, "Fault Detection and Isolation for an Experimental Internal Combustion Engine via Fuzzy Identification," *IEEE Trans. on Control Systems Technology*, 3(30), pp. 347-355, September 1995.
- [92] S.-K. Sin and R. J. P. deFigueiredo, "Fuzzy System Design Through Fuzzy Clustering and Optimal Predefuzzification," in proceedings of the second IEEE conference on Fuzzy Systems, vol. 1, pp. 190-195, San Francisco, California, 1993.
- [93] J. C. Bezdek, "A convergence theorem for the fuzzy ISODATA clustering algorithms," in *Fuzzy Models for Pattern Recognition*, J. C. Bezdek, and S. K. Pal, eds., pp. 130-137, IEEE Press, 1992.
- [94] M. P. Windham, "Geometrical fuzzy clustering," in *Fuzzy Models for Pattern Recognition*, J. C. Bezdek, and S. K. Pal, eds., pp. 123-129, IEEE Press, 1992.
- [95] J. C. Bezdek, R. J. Hathaway, M. J. Sabin, and W.T. Tucker, "Convergence theory for fuzzy c-means: Counterexamples and repairs," in *Fuzzy Models for Pattern Recognition*, J. C. Bezdek, and S. K. Pal, eds., pp. 138-142, IEEE Press, 1992.
- [96] J. Gertler and D. Singer, "A new structural framework for parity equation based failure detection and isolation," *Automatica*, 26, pp. 381-388, 1990.
- [97] J. Gertler, "Analytical redundancy methods in fault detection and isolation," *IFAC/IMACS Symposium on Fault Detection Supervision & Safety for Technical Processes*, Baden-Baden, Germany, pp. 9-21.
- [98] E. Y. Chow and A. S. Willsky, "Analytical redundancy and the design of robust failure detection systems," *IEEE Transactions on Automatic Control*, AC-29, pp. 603-614, 1984.
- [99] N. Viswanadham and R. Srichander, "Fault Detection Using Unknown Input Observers," *Control Theory and Advanced Technology*, Vol.3, No.2, pp.91-101, June 1987.
- [100] P. M. Frank, "Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy - a Survey and Some New Results," *Automatica*, Vol.26, No.3, pp.459-474, 1990.
- [101] M. Hou and P.C. Müller, "Design of Robust Observers for Fault Isolation," *Proceedings of the IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes - SAFEPROCESS '91*, Baden-Baden, Germany, pp. 265-270, September 1991.

- [102] V. L. Symos, "Computational Observer design Techniques for Linear Systems with Unknown Inputs Using the Concept of Transmission Zeros," IEEE Trans. on Automatic control, vol. AC-38, pp. 790-794, 1993.
- [103] R. V. Beard, "Failure accommodation in linear systems through self-reorganization," Report MVT-71-1, Man Vehicle lab, MIT, Cambridge, Massachusetts, 1971.
- [104] R. J. Patton and J. Chen, "Robust Fault Detection Using Eigenstructure Assignment: A Tutorial Consideration and Some New Results," Proceedings of the 30th IEEE Conference on Decision and Control, Brighton, England, pp. 2242-2247, December, 1991.
- [105] Bosch Automotive Handbook, 2<sup>nd</sup> Edition, Robert Bosch GmbH, Stuttgart, Germany, 1986.
- [106] T. Magi, "Prospects for Failure Diagnostics of Automotive Electronic Control Systems," Proceedings of the International Congress on Transportation Electronics, paper no. 94C055, Dearborn, Michigan 1994 .
- [107] M. Ross, R. Goodwin, R. Watkins, M. Wang, and T. Wenzel, "Real-World Emissions from Model Year 1993, 2000 and 2010 Passenger Cars," distributed by the American Council for an Energy-efficient economy, USA, November 1995.
- [108] R. J. Patton, P. M. Frank, and R. Clark, "Fault Diagnosis in Dynamic Systems, Theory and Applications," Prentice Hall, 1989.
- [109] J. Mendel, "Fuzzy Logic Systems for Engineering: A Tutorial," Proceedings of IEEE, vol.83, no.3, pp. 345-377, March 1995.
- [110] M. Ayoubi, "Neuro-Fuzzy Structure for Rule Generation and Application in the Fault Diagnosis of Technical Processes," Proceedings of the American Control Conference, vol.4, pp. 2757-2761, Seattle, Washington, June 1995.
- [111] Y. Chen, "A Fuzzy Decision System Based on Statistical Learning for Fault Classification," Proceedings of IEEE International Conference on Fuzzy Systems, vol. 2, pp.1459-1463, 1994.
- [112] L.-X. Wang, "Training of Fuzzy Logic Systems Using Nearest Neighborhood Clustering," Proceedings of IEEE International Conference on Fuzzy Systems, vol. 1, pp. 13-17, San Francisco, California, March 1993.
- [113] T. W. Cheng, D. B. Goldgof, and L.O. Hall, "Fast Clustering with Application to Fuzzy Rule Generation," Proceedings of IEEE International Conference on Fuzzy Systems, vol. 4, pp. 2289-2295, Yokohama, Japan, March 1995.

- [114] R. N. Clark, "A Simplified Instrument Failure Detection Scheme," IEEE Transactions on Aerospace and Electronic Systems, AES-14, pp. 558-563, 1978.
- [115] D. Hengry and P.M. Frank, "Component failure detection via nonlinear state observers," IFAC Workshop, September 1986, Kyoto, Japan.
- [116] R. Seliger and P.M.Frank, "Robust Component Fault Detection and Isolation in Nonlinear Dynamic Systems Using Nonlinear Unknown Observers," Proceedings of the IFAC/IMACS symposium on Fault Detection, Supervision and Safety for Technical Process-SAFEPROCESS 1991, pp. 313-318 Baden-Baden, Germany.
- [117] J. Wünnenberg, "Observer-based fault detection in dynamic systems, Ph.D. thesis, University of Duisburg, 1990, Germany.
- [118] T. Himmelspach, "Model-Based Failure Detection and Isolation (FDI) for Deterministic Nonlinear Dynamic Systems," Ph.D. Dissertation, University of Michigan Ann Arbor, 1990.
- [119] R.J. Patton, "Robust Model-Based Fault Diagnosis: The State of The Art," Proceedings of the IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes - SAFEPROCESS '94, Espoo, Finland, 1994.
- [120] R. Isermann and T. Pfeufer, "Integrated Model Based Fault Detection and Diagnosis with Application to Electromechanical Drives," IFAC 13<sup>th</sup> Triennial World Congress, pp. 85-90, San Francisco, USA 1996.
- [121] A. Isidori, "Nonlinear Control System: An Introduction," Springer-Verlag, New York, 1989.
- [122] V. Krishnaswami and G. Rizzoni, "Nonlinear parity equation residual generation for fault detection and Isolation," Proceedings of the IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes - SAFEPROCESS '94, pp. 317 - 322, Espoo, Finland, 1994.
- [123] A. Soliman, V. Krishnaswami, and G. Rizzoni, "Integrated Powertrain Diagnostic System (IPDS): Linking On- and Off-Board Diagnostic Strategies," SAE paper no. 960621, presented at The International Congress and Exposition, Feb.1996, Detroit, MI.
- [124] A. Soliman, V. Krishnaswami, and G. Rizzoni, "The Application of Fuzzy Logic to the Diagnosis of Automotive Systems," SAE paper no. 970208, presented at The International Congress and Exposition, February 1997, Detroit, MI.
- [125] A. Soliman and G. Rizzoni, "The effects of various engine control system malfunctions on exhaust emissions levels during the I/M 240 cycle," SAE paper

no. 940448, presented at The International Congress and Exposition, March 1-5, 1994, Detroit, MI.

- [126] A. Soliman, G. Rizzoni, and V. Krishnaswami, "The Effect of Engine Misfire on Exhaust Emission Levels in Spark Ignition Engines," SAE Paper no. 950480, presented at The International Congress and Exposition, Feb.1995, Detroit, MI.
- [127] A. Soliman, G. Rizzoni, and K. M. Passino, "A Survey of Automotive Diagnostic Equipment and Procedures," SAE Paper no. 930769, presented at The International Congress and Exposition, March.1993, Detroit, MI.
- [128] J.-S. R. Jang and N. Gullrey, "Fuzzy Logic Toolbox User's Guide," The MathWorks, Inc. Natick, Mass. USA, January 1995.
- [129] V. Utkin, "Sliding Modes in Control and Optimization," Springer Verlag, 1992.
- [130] S. V. Drakunov, "Sliding Mode Observer Based on Equivalent Control Method," Proceedings of 31-st CDC, Dec. 16-18, Tuscon, Arizona, pp. 2368-2369, 1992.
- [131] H. Hashimoto et. al., "VSS Observer for Linear Time Varying Systems," Proceedings of 16-th IECON, Nov. 27-30, Pacific Grove, California, Vol. 1, pp.34-39, 1990.
- [132] E. A. Misawa and J. K. Hedrick, "Nonlinear Observer A State-of-the-Art survey," ASME J. Dynamic Systems, Measurement and Control, v.111, No.3, 1989.
- [133] V. Krishnaswami and G. Rizzoni, "Design of Sliding Mode Observers for Vehicle Steering Diagnostics and Control," Proceedings of the 1995 American Control Conference, ACC '95, pp. 2113 - 2117, Seattle, U.S.A., 1995.
- [134] V. Krishnaswami and G. Rizzoni, "Vehicle steering system state estimation using sliding mode observers," in Proceedings of the 34<sup>th</sup> Conference on Decision and Control, CDC 34, pp. 3391 - 3396, New Orleans, LA, December, 1995.
- [135] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding Mode Observers for Fault Detection," in Proceedings of IFAC Symposium on Fault Detection, Supervision, and Safety for Technical Processes, SAFEPROCESS '97, pp. 522-527, Hull, England, August 1997.
- [136] Y. W. Kim, G. Rizzoni, A. Soliman, P. Azzoni, and D. Moro, "Powertrain Diagnostics Using Nonlinear Sliding Mode Observers," in Proceedings of IFAC Symposium on Fault Detection, Supervision, and Safety for Technical Processes, SAFEPROCESS '97, pp. 825-830, Hull, England, August 1997.
- [137] A. Soliman, G. Rizzoni, and Y. W. Kim, "Diagnosis of Automotive Emission Control System Using Fuzzy Inference," in Proceedings of IFAC Symposium on

Fault Detection, Supervision, and Safety for Technical Processes, SAFEPROCESS '97, pp. 726-730, Hull, England, August 1997.

- [138] V. Krishnaswami and G. Rizzoni, "Robust Residual Generation for Nonlinear System Fault Detection and Isolation," in Proceedings of IFAC Symposium on Fault Detection, Supervision, and Safety for Technical Processes, SAFEPROCESS '97, pp. 154-159, Hull, England, August 1997.
- [139] V. Krishnaswami, A. Soliman, and G. Rizzoni, "A New Generation of On-Board and Service Diagnostics for Passenger Vehicles," in Proceedings of FISITA, Beijing, China 1995.
- [140] W. B. Davenport, JR., "Probability and Random Processes," McGraw-Hill Publishing Company, Inc., USA, 1970.
- [141] J. Dawson, "Real-time Engine Models for Hardware-in-the-Loop Control Development," Ph. D dissertation, The Ohio State University, 1998.

## APPENDICES

## APPENDIX A

### FUZZY SYSTEMS

The theory of fuzzy logic has its roots back in 1965 when Zadeh presented his ideas of fuzzy sets [66]. An overview of some of the fundamental concepts in fuzzy systems has been presented here to provide background into some of the notation and concepts used in this dissertation. Most of the definitions given in this appendix have been paraphrased from [87].

A fuzzy system is shown in Figure A.1, which is static nonlinear mapping between inputs and outputs. The inputs are  $u_i \in U_i$  where  $i = 1, 2, \dots, n$ , and outputs  $y_i \in Y_i$ , where  $i = 1, 2, \dots, m$ . The outputs and inputs are *crisp*- that is real numbers, not fuzzy sets. These crisp inputs are mapped into fuzzy sets by the fuzzification block, this is needed in order to activate rules which are in terms of linguistic variables, which have fuzzy sets associated with them [109]. The inference mechanism produces conclusions using the fuzzy rules in the rule-base. Crisp outputs are obtained from the defuzzifier block.



## Universe of Discourse

The *crisp* sets  $U_i$  and  $Y_i$  are called the *universe of discourse* for  $u_i$  and  $y_i$ , respectively.

Generally the universes of discourse are simply the set of real numbers or some interval or subset of real numbers.

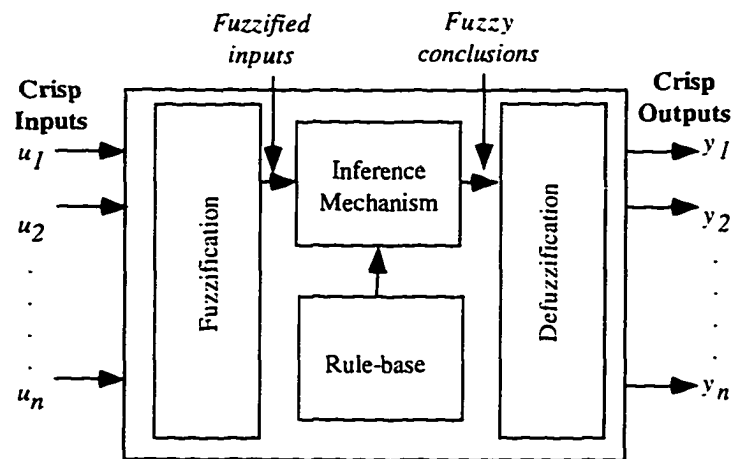


Figure A.1 Fuzzy System

In classical set theory, an element of any universe can be either a member of the set or not. Fuzzy sets, however, are characterized by the fact that an element of the universe of discourse has a so-called degree of membership, determined by a membership function; i.e. an element can not only belong or not belong to a set, but belong more or less to it. This fuzziness is also characteristic for human beings when they are asked to classify

certain elements. The procedure of determining the degree of membership of a crisp input, which is an element of an universe of discourse, is called fuzzification.

### **Linguistic Variables**

These are variables whose values are not number but words or sentences in a natural or artificial language to describe fuzzy system inputs and outputs. Where  $\tilde{u}_i$  is the *linguistic variable* that describes the inputs  $u_i$ . Similarly  $\tilde{y}_i$  is the linguistic variable that describes the output  $y_i$ .

### **Linguistic Values**

Linguistic variables  $\tilde{u}_i$  and  $\tilde{y}_i$  take on *linguistic values* that describe the characteristics of the variable. The set of linguistic values  $\tilde{A}_i = \{\tilde{A}_i^j : j = 1, 2, \dots, N_i\}$  where  $\tilde{A}_i^j$  denotes the  $j^{th}$  linguistic value of the linguistic variable  $\tilde{u}_i$ . Similarly  $\tilde{B}_i = \{\tilde{B}_i^p : p = 1, 2, \dots, M_i\}$ , where  $\tilde{B}_i^p$  denotes the  $p^{th}$  linguistic value of the linguistic variable  $\tilde{y}_i$ .

### **Linguistic Rules**

A set of *condition*  $\rightarrow$  *action* rules, or in *modus ponens* (If-Then) rule maps the inputs to the outputs

**If** antecedent **Then** consequent

Usually, the inputs to the fuzzy system are associated with the antecedent, and the outputs are associated with the consequent. for the multi-input single-output (MISO) the standard rule form is

**If**  $\tilde{u}_1$  is  $\tilde{A}_1^j$  and  $\tilde{u}_2$  is  $\tilde{A}_2^j$  and,.....,  $\tilde{u}_n$  is  $\tilde{A}_n^j$  **Then**  $\tilde{y}_q$  is  $\tilde{B}_q^p$

This can be in the form of multi-input multi-output (MIMO). Generally the rules in the rule-base are distinct.

### Membership Functions

The membership functions  $\mu(u_i)$  are subjectively specified in an ad hoc (heuristic) manner, they are associated with the terms that appear in the antecedent and consequent. Many shapes of the membership function are possible (e.g., triangular, trapezoidal shapes), each will provide a different meaning for the linguistic variable. Some Typical membership functions are shown in Figure A.2.

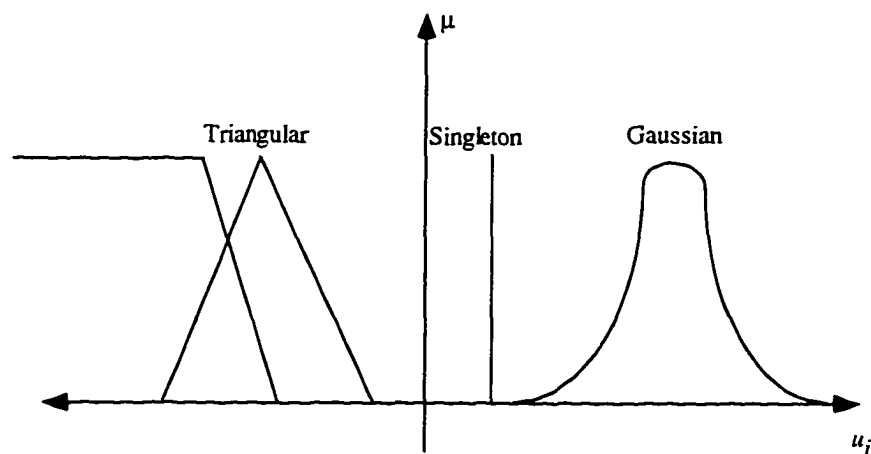


Figure A.2 Some Typical Membership functions

## Fuzzy Sets

Simply a fuzzy set is a crisp set of elements of the universe of discourse paired and coupled with their associated membership values.  $A_i' = \{(u_i, \mu_{A_i'}(u_i)) : u_i \in U_i\}$ .

## Fuzzification

Fuzzification transforms  $u_i$  to a fuzzy set defined on the universe of discourse  $U_i$ . This transformation is produced by operator  $f$  defined by

$$f : U_i \rightarrow U_i^*$$

where

$$f(u_i) = \hat{A}_i^{fuz}, \hat{A}_i^{fuz} \text{ is the fuzzy set}$$

quite often singleton fuzzification is used a fuzzy set  $\hat{A}_i^{fuz} \in U_i^*$  with a membership function defined by

$$\mu_{\hat{A}_i^{fuz}}(x) = \begin{cases} 1 & x = u_i \\ 0 & \text{otherwise} \end{cases}$$

Any fuzzy set with this form for its membership function is called a *singleton*.

## Fuzzy Intersection (AND)

Two methods to define the membership function that represents the intersection of

$A_i^1$  and  $A_i^2$ .

1. *Minimum*,  $\mu_{A_i^1 \cap A_i^2} = \min\{\mu_{A_i^1}(u_i), \mu_{A_i^2}(u_i) : u_i \in U_i\}$
2. *Algebraic Product*,  $\mu_{A_i^1 \cap A_i^2} = \{\mu_{A_i^1}(u_i) \mu_{A_i^2}(u_i) : u_i \in U_i\}$

### Fuzzy Union (OR)

Two methods to define the membership function that represents the union of

$A_i^1$  and  $A_i^2$ .

1. *Maximum*,  $\mu_{A_i^1 \cup A_i^2} = \max\{\mu_{A_i^1}(u_i), \mu_{A_i^2}(u_i) : u_i \in U_i\}$

2. *Algebraic Sum*,  $\mu_{A_i^1 \cup A_i^2} = \{\mu_{A_i^1}(u_i) + \mu_{A_i^2}(u_i) - \mu_{A_i^1}(u_i)\mu_{A_i^2}(u_i) : u_i \in U_i\}$

### Fuzzy Implications

It is the fuzzy quantification of the linguistic rule. The implication method shapes the consequent based on the antecedent. The fuzzy quantify the terms in the antecedent and consequent of the If-Then rule, to make a *fuzzy implication* (a fuzzy relation).

### Aggregation

It is combining the output fuzzy sets into a single fuzzy set in preparation for defuzzification.

### Defuzzification

It is a means to choose a crisp output based on the implied fuzzy sets. Some defuzzification techniques include “centroid”, and “center of gravity”. In the centroid a crisp output is chosen based on the individual implied fuzzy sets and the point of maximum for each output membership function

$$y = \frac{\sum_{j=1}^M \bar{y}^j \mu_{A \circ R_j}(\bar{y}^j)}{\sum_{j=1}^M [\mu_{A \circ R_j}(\bar{y}^j)]}$$

where  $A \circ R_j$  is a single implied fuzzy set for the  $j^{th}$  fuzzy implication, and  $\bar{y}^j$  is the consequent portion of the linguistic rule  $R_j$ .

In the center of gravity a crisp output is chosen based on the center of area of each of the implied fuzzy set.

## APPENDIX B

### ENGINE SENSORS

In this Appendix the engine sensors' calibration curves are shown in Figures B.1-B.5. While the misfire circuit is shown in Figure B.7, this circuit allowed the interruption of the spark command to cylinder #1 at a preset rate that can be dialed in. Also, the fault induction circuit is shown in Figure B.6 which allowed the induction of calibration faults to the sensor/actuator under investigation. This was possible through a potentiometer that covered the range 0%-100% calibration fault.

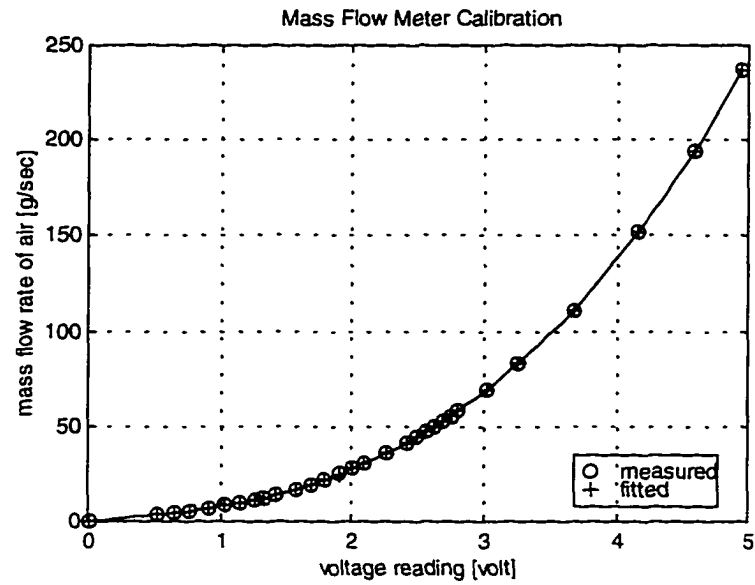


Figure B.1 Mass Air Flow Sensor (MAF)

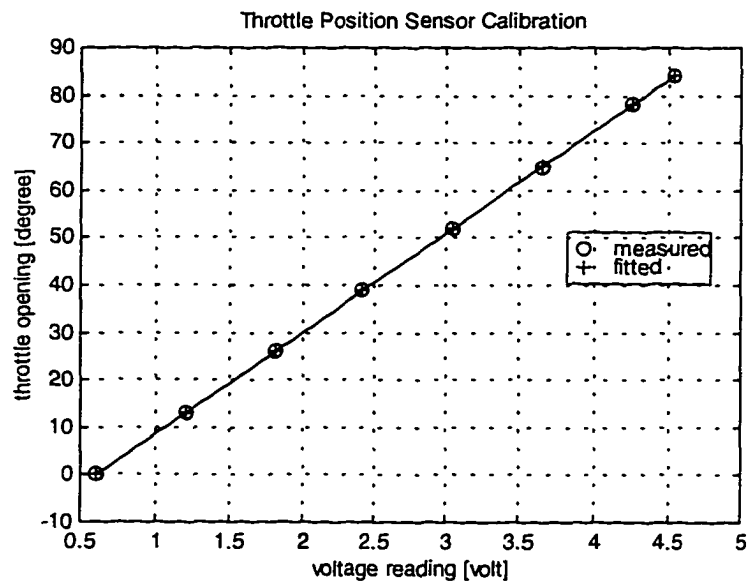


Figure B.2 Throttle Position Sensor (TPS)



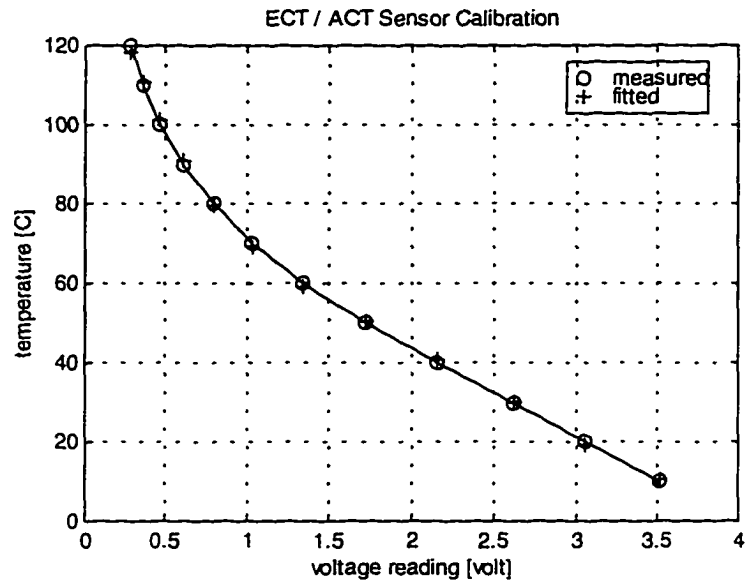


Figure B.3 Engine Coolant/Air Temperature Sensor

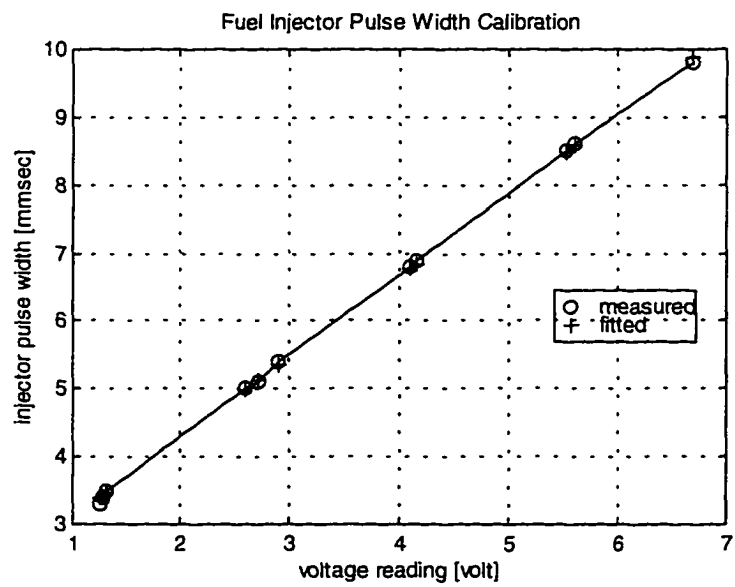


Figure B.4 Fuel Injector Pulse Width

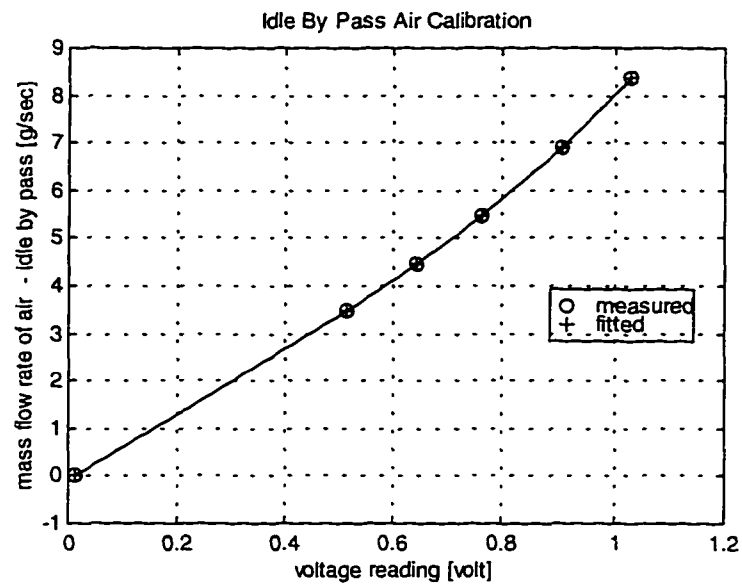


Figure B.5 Idle Air By Pass Sensor

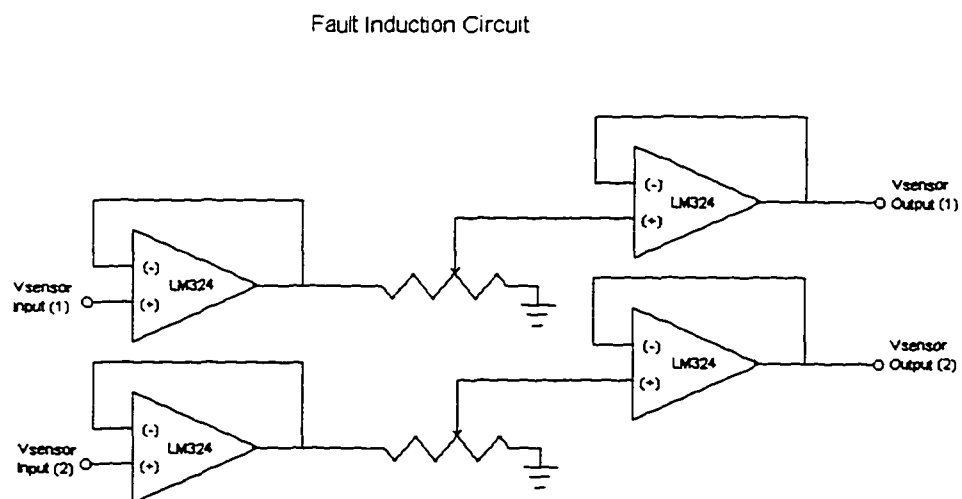


Figure B.6 Fault Induction Circuit

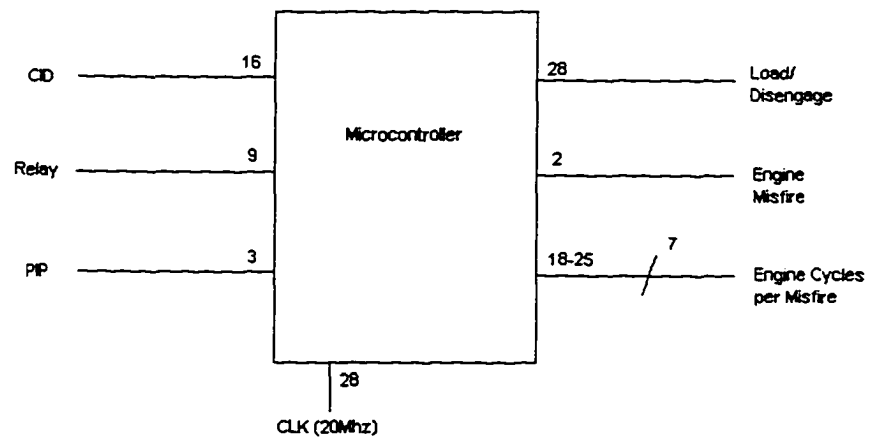
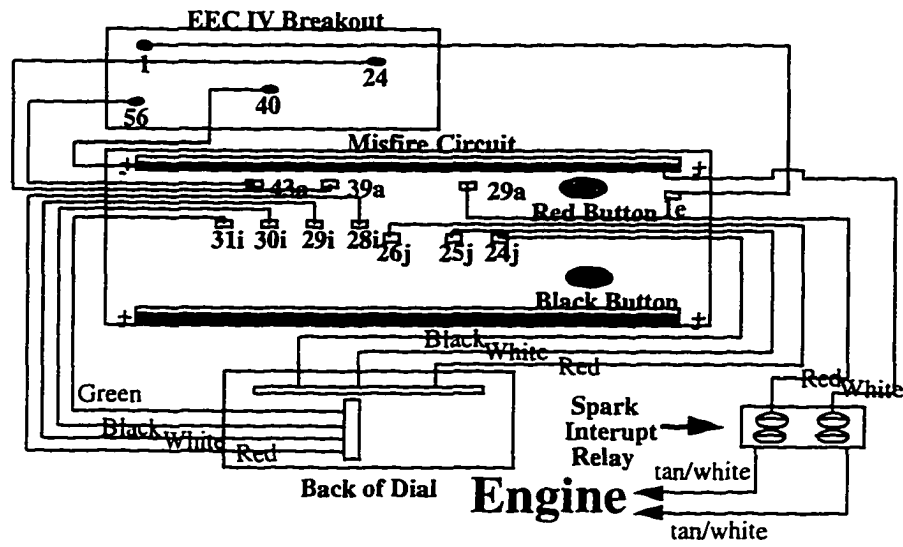


Figure B.7 Misfire Circuit

## APPENDIX C

### ENGINE MODEL

A simplified dynamic engine model was developed by [141], that uses the understanding of the physical phenomena taking place within the engine systems, subsystems, and components, to derive the dynamic equations of each block. This model was identified and verified in the Powertrain Control and Diagnostics Laboratory. The following are the engine model equations.

#### C.1. Engine Model Equations

##### Throttle Area Equation

$$\frac{4A_{th}}{\pi D^2} = \left(1 - \frac{\cos \psi}{\cos \psi_o}\right) + \frac{2}{\pi} \left[ \frac{\frac{a}{\cos \psi} (\cos^2 \psi - a^2 \cos^2 \psi_o)^{1/2}}{-\frac{\cos \psi}{\cos \psi_o} \sin^{-1} \left( \frac{a \cos \psi_o}{\cos \psi} \right) - a(1 - a^2)^{1/2} + \sin^{-1} a} \right] \quad (C.1)$$

### Throttle Area Equation

$$A_{th} = f(\psi, \psi_o) \quad (C.2)$$

### Throttle Mass Flow Rate Equations

$$\text{If } \frac{p_m}{p_o} > \left[ \frac{2}{\gamma + 1} \right]^{\gamma/(\gamma-1)} \quad (\text{unchoked flow}),$$

$$\text{then } \dot{m}_{th} = \frac{C_{D,th} A_{th} p_o}{\sqrt{RT_o}} \left( \frac{p_m}{p_o} \right)^{1/\gamma} \left\{ \frac{2\gamma}{\gamma-1} \left[ 1 - \left( \frac{p_m}{p_o} \right)^{(\gamma-1)/\gamma} \right] \right\}^{1/2} \quad (C.3)$$

$$\text{While, if } \frac{p_m}{p_o} \leq \left[ \frac{2}{\gamma + 1} \right]^{\gamma/(\gamma-1)} \quad (\text{choked flow}), \text{ then}$$

$$\dot{m}_{th} = \frac{C_{D,th} A_{th} p_o}{\sqrt{RT_o}} \gamma^{1/2} \left( \frac{2}{\gamma + 1} \right)^{(\gamma+1)/2(\gamma-1)} \quad (C.4)$$

### Intake Manifold Equations

$$\frac{dm_{a,m}}{dt} = \dot{m}_{a,th} - \sum \dot{m}_{a,cyl} \quad (C.5)$$

$$\sum \dot{m}_{a,cyl} = \frac{\eta_v \rho_{a,m} V_d N}{2} \quad (C.6)$$

$$p_m V_m = m_a R T_m \quad (C.7)$$

$$\frac{dp_m}{dt} + \frac{\eta_v V_d N}{2 V_m} p_m = \dot{m}_{a,th} \frac{RT_m}{V_m} \quad (C.8)$$

### Fueling Equations

$$\dot{m}_{fv} = (1 - X)\dot{m}_{fi} \quad (C.9)$$

$$\ddot{m}_f = \frac{1}{\tau_f} (-\dot{m}_f + X\dot{m}_{fi}) \quad (C.10)$$

$$\dot{m}_f = \dot{m}_{fv} + \dot{m}_f \quad (C.11)$$

### Compression/Expansion Equation

$$p_{cyl} V_{cyl}^\gamma = Const. \quad (C.12)$$

### Combustion Equations

$$\frac{dp_{cyl}}{d\theta} = \frac{\gamma - 1}{V} \left[ \frac{dm_{f, burn}}{d\theta} Q_{LHV} - \frac{dQ_{ht}}{d\theta} \right] - \frac{\gamma}{V} \frac{dV}{d\theta} p_{cyl} \quad (C.13)$$

$$\frac{dm_{f, burn}}{d\theta} = m_{f, total} \frac{a(m+1)}{\Delta\theta} \left( \frac{\theta - \theta_o}{\Delta\theta} \right)^m \exp \left[ -a \left( \frac{\theta - \theta_o}{\Delta\theta} \right)^{m+1} \right] \quad (C.14)$$

### Intake/Exhaust Valve Equations

If  $\frac{p_T}{p_o} > \left[ \frac{2}{\gamma + 1} \right]^{\gamma/(\gamma-1)}$  (unchoked flow), then

$$\dot{m}_{valve} = \frac{C_{D,valve} A_{valve} p_o}{\sqrt{RT_o}} \left( \frac{p_T}{p_o} \right)^{1/\gamma} \left\{ \frac{2\gamma}{\gamma-1} \left[ 1 - \left( \frac{p_T}{p_o} \right)^{(\gamma-1)/\gamma} \right] \right\}^{1/2} \quad (C.15)$$

While, if  $\frac{p_T}{p_o} \leq \left[ \frac{2}{\gamma+1} \right]^{\gamma/(\gamma-1)}$  (choked flow), then

$$\dot{m}_{valve} = \frac{C_{D,valve} A_{valve} p_o}{\sqrt{RT_o}} \gamma^{1/2} \left( \frac{2}{\gamma+1} \right)^{(\gamma+1)/2(\gamma-1)} \quad (C.16)$$

#### Exhaust Pipe Flow Equations

$$\Delta t = \frac{R_{l,n} p_{ex,n} A_{run}}{RT_{ex,n} \dot{m}_{run,n}} \quad (C.17)$$

#### Ideal Gas Law

$$pV = \frac{m}{MW} \bar{R}T$$

$$pV = mRT \quad (C.18)$$

#### Cylinder Mass Equation

$$m_{cyl} = \frac{p_{o,cyl} M_{o,cyl} V_{o,cyl}}{\bar{R}T_{o,cyl}} + \int_{IVO}^{IVC} \dot{m}_{cyl} \frac{1}{w} d\theta \quad (C.19)$$

$$m_{cyl} = \frac{p_{o,cyl} M_{o,cyl} V_{o,cyl}}{\bar{R}T_{o,cyl}} - \int_{EVO}^{EVC} \dot{m}_{cyl} \frac{1}{w} d\theta \quad (C.20)$$

### Engine Dynamics

$$\frac{d\omega}{d\theta} = \frac{\sum_{i=1}^{numcyl} T_{cyl,i} - T_{load}}{J\omega} - \frac{B}{J} \quad (C.21)$$

### Engine Geometry

$$r_c = \frac{V_d + V_c}{V_c} \quad (C.22)$$

$$V_d = \frac{\pi B^2}{4} L$$

$$V = \frac{\pi B^2}{4} \left[ \frac{L}{r_c - 1} + l + a - a \cos \theta - (l^2 - a^2 \sin^2 \theta)^{1/2} \right] \quad (C.23)$$

$$\frac{dV}{d\theta} = \frac{\pi B^2}{4} \left[ a \sin \theta + \frac{a^2 \sin \theta \cos \theta}{(l^2 - a^2 \sin^2 \theta)^{1/2}} \right] \quad (C.24)$$

The list of symbols used in the engine model equations are listed in Table C.1, while the parameters are listed in Table C.2.



No.	Symbol	Description [Units]
1.	$P_m$	intake manifold pressure [Pa]
2.	$P_a$	ambient pressure [Pa]
3.	$T_a$	ambient temperature [K]
4.	$C_{d,th}$	throttle discharge coefficient
5.	$\dot{m}_{a,th}$	air mass flow rate at throttle [Kg/sec]
6.	$R$	ideal gas constant [J/ (kg·K)]
7.	$\dot{m}_{a,cyl}$	air mass flow rate into cylinder [kg/sec]
8.	$V_m$	intake manifold volume [m <sup>3</sup> ]
9.	$V_d$	displacement volume [m <sup>3</sup> ]
10.	$\eta_v$	volumetric efficiency
11.	$\dot{m}_{ff}$	fuel flow rate from film [kg/sec]
12.	$\dot{m}_{fi}$	fuel flow rate from injector [kg/sec]
13.	$\tau_f$	fuel evaporation time constant [sec]
14.	$X$	fraction of injected fuel enter into the film
15.	$\dot{m}_{fc}$	fuel mass flow rate entering the cylinder
16.	$\tau_m$	fuel flow rate from film [kg/sec]
17.	$t_c$	cycle delay [/sec]
18.	$t_t$	transportation delay [/sec]
19.	$\phi_m$	measured equivalence ratio
20.	$\theta_s$	crank-angle sampling interval ( $\pi/2$ radians)

Table C.1 List of Symbols for Engine Model

No.	Parameter	Description
1.	$C_{d,th}$	Throttle discharge coefficient
2.	$\eta_v$	Volumetric Efficiency
3.	$\tau_f$	fuel evaporation constant
4.	$X$	direct entry fraction of fuel
5.	$\theta_c + \theta_t$	transport+EGO sensor delay
6.	$\tau_m$	EGO sensor time constant

Table C.2 List of Engine Model Parameters