Fault Diagnosis and Control of A Diesel Power Generator Using Sliding Mode Techniques

Thesis submitted for the degree of

Doctor of Philosophy
At the University of Leicester

by

K.B. Goh

B. Eng (Honours), MIEE

Department of Engineering
University of Leicester
Leicester, United Kingdom

2004

Fault Diagnosis and Control of A Diesel Power Generator Using Sliding Mode Techniques

K. B. Goh

Abstract

This thesis is concerned with fault diagnosis and control of diesel engines. It is vital that the diesel engine operates in a healthy condition and is appropriately controlled in order to maintain efficiency, robustness and reliability. Early detection of possible minor faults may prevent possible catastrophic faults.

The sliding mode concept, which is based on the notion of a variable structure system, has inherent advantages of robustness and performance specification and is here applied to the problems of fault diagnosis and control of the diesel engine. A real-time condition monitoring and fault diagnosis system and also control system using sliding mode techniques are described. The designs are assessed through tests on a particular Perkins diesel-electric power generator.

A model-based approach incorporating a non-linear sliding mode observer scheme is proposed for fault monitoring. The diesel engine coolant system is considered. The system parameters are monitored using the concept of the equivalent injection signal which is required to maintain the sliding mode. The proposed diagnostic scheme is shown to be robust in estimating component parameters. The approach is applicable to many automotive engine problems and is cost effective as only low cost temperature sensors are involved in the implementation.

The engine control strategy investigates both model-based and model-free sliding mode control techniques. The development of an engine model appropriate for speed control has been considered using closed-loop event-based system identification. Several control algorithms are proposed. Real-time speed control systems have been designed and implemented using Matlab/Simulink/dSPACE. The proposed model-based and model-free controllers show good tracking performance and disturbance rejection properties. The proposed model-free controller is shown to be an appropriate candidate for industrial control of the diesel engine system. The established gain-tuning algorithms allow non-experts to maintain and tune the resulting control schemes.

Acknowledgements

I would like to thank a number of people who have assisted me throughout the ups and downs research period and off research period. First and foremost, to both of my supervisors, my hearty thanks go to Sarah Spurgeon and Barrie Jones for their advice, guidance and encouragement. I deeply appreciate the valuable time they spent with me at the hot engine test bed despite 'the department lift having broken down' and also concern and support for my financial situation. I also thank them for their patience and belief in me.

My sincere thanks go to Engineering Department, Perkins, TRW, EPSRC (ref: GR/L42018) and ORS (ref: ORS/2000024011) for supporting on this research. Special thanks go to my colleague John Twiddle with whom I worked closely on this research. Special thanks go to Chris Edwards for passing on the application letter/form, also to industrial collaborators, P. Scotson, P. Ladlow and M. Scaife. To the department staffs Dipak, Peter, Dave, Paul, Tony, Graham, Geoff etc. I also express my gratitude to Royan, Andy P. W., Daljeet, Anita, Steve, Chinmay, Edwin, Kien Ling, Kien Kiat, Bin, Joe, Hans, Law, Ryan, YuHua, the department's cricket team, badminton members and many more for making life in Leicester more memorable and lively.

My heart-felt thanks go to my family and family-in-law: grandma, father, Ah Ee, parents-in-law, sister (family) & in-law, brothers (family) & in-law and my wife, Whui Cheng for their constant love and support over the years of my studies in England. Not forgetting to my loving mother who did not live long enough to witness this work. If she were around, she would be the happiest person. Lastly, I thank my wife again who has never stopped believing in me and never blamed me on resigning from an engineering job with Hewlett Packard to attempt a PhD. It would have been impossible without her continuous encouragement, constructive criticism and love.

This thesis is dedicated to the memory of my mother,

Led Wah

to all my family members and to my wife

Whui Cheng

List of Publications

This section describes the published work and work under consideration for publication, resulting from the thesis.

Published Papers

1. M.K. Khan, K.B. Goh and S.K. Spurgeon, 2003, "Second order sliding mode control of a diesel engine," *Asian Journal of Control* – Variable Structure System Control, New Designs and Applications (Special Issue), to appear.
2. K.B. Goh, S. K. Spurgeon and N. B. Jones, 2003, "Higher Order Sliding Mode Control of a Diesel Generator Set," *Journal of Systems and Control Engineering*, Proceedings of the Institution of Mechanical Engineers Part I, vol. 217(13), pp. 229-241.
3. K.B. Goh, S. K. Spurgeon and N. B. Jones, 2002, "Fault diagnostics using sliding mode techniques *Control Engineering Practice*, Vol. 10 (2), pp. 207-217.
4. Spurgeon S, Goh, K.B. and Jones B., 2002, "An application of higher order sliding modes to the control of a diesel generator set (Genset*)," 7th International Workshop on Variable Structure Systems, Sarajevo.*
5. N. B. Jones, S. K. Spurgeon, M. J. Pont, J. A. Twiddle, C.L. Lim, C. R. Parikh and K.B. Goh, 2000, "Diagnostic Schemes for Biomedical and Engineering Systems" *MEDSIP 2000 Conference*. No. 476, Page 1 - 8
6. N. B. Jones, S. K. Spurgeon, M. J. Pont, J. A. Twiddle, C.L. Lim, C. R. Parikh and K.B. Goh, 2000, "Aspects of diagnostic schemes for biomedical and engineering systems," *IEE Proc.-Sci. Meas. Tech.* Vol. 147, No. 6, P. 357- 362.

Paper or book chapter under consideration for publication

1. K.B. Goh, S. K. Spurgeon and N. B. Jones, "The application of sliding mode control algorithms to a Diesel Generator Set," Invited contribution to Variable Structure Systems: from principles to implementation, IEE Book Series.
2. K.B. Goh, S. K. Spurgeon, N. B. Jones and J.A. Twiddle, "Model identification and sliding mode control of a diesel electric power generator," Submitted to *IEE Proceedings – Control Theory and Applications*.

Contents

List of Symbols

The symbols are arranged in alphabetical order.

$A_{block\text{-}to\text{-}coolant}$	Heat transfer area
A_{rad}	Heat transfer area of radiator
A_m	State matrix of ideal model
A_p	State matrix of plant
A_{rad}	Surface area of radiator
A_{reg}	Regular form of state matrix
A_i	Controller class/category as defined in the British Standard $(i = 0,1,2)$
B_m	Input matrix of ideal model
B_p	Input matrix of plant
B_{reg}	Regular form of input matrix
B_i	Controller class/category as defined in the British Standard $(i = 1,2)$
$c_{block\text{-}to\text{-}coolant}$	Specific heat of coolant in block
c	Specific heat capacity
c_{rad}	Specific heat capacity of coolant
C_m	Output matrix of ideal model
C_p	Output matrix of plant
C_{reg}	Regular form of output matrix
D	System direct transmission matrix
CO_2	Carbon dioxide
e_i	Speed error $(i = spd$ for speed error, $i = y$ for output error$)$ Error between the model state and estimated state, $(i = 2, 2a$ and $3)$
$\dot{e}_i$	Derivative of observer error, $(i = 2, 2a$ and $3)$
e_{track}	Speed tracking error
f	External disturbance
$f(t,x,u)$	System uncertainties
F	non-singular design matrix
$g\_L_{x_{lat}}$	Scaling gain
g_L_r	Scaling gain
$g\_L_{\dot{r}}$	Scaling gain
G_{fwd}	Feedforward term

G_l	Observer design gain
G_L	Controller design matrix
h_{rad}	Radiator heat transfer coefficient
h_{bc}	Coolant side heat transfer coefficient
h_L	Gain of the linear term in the reachability condition
$\hat{h}$	Estimated heat transfer coefficient
H_2O	Water
I, I_o	Light intensity (measurement of current)
J	Cost function
k	Matrix rank number
k_i	Parameter representation, ($i = 1, 2, 2a, 3$ and 4)
K	State feedback matrix
K_d	Derivative controller gain
K_i	Integral controller gain
K_{int}	Integral controller gain
K_p	Proportional controller gain
K_n	Gain of the discontinuous function, ($n =$ $2, 2a$ and 3)
K_{obs}	Observer feedback matrix
l	Sliding equation/line
L_{add}	Linear error feedback gain
m	Number of inputs
m	Mass of coolant (Chapter 4)
$\dot{m}_B$	Mass rate flow through engine block
$\hat{m}$	Estimated coolant mass
$\dot{m}$	Mass flow rate of coolant
M	Design matrix
n	System order
N	Opacity %
N_{data}	Length of data sequence
N_{mea}	Measured engine speed
N_{nom}	Nominal or reference engine speed
P_{cr}	Critical period
N_2	Nitrogen
$\mathrm{N}(\cdot)$	Null space

P_{transf}	Transformation state matrix
P	Symmetric positive definite matrix
P_{obs}	Symmetric positive definite matrix
$\dot{q}$	heat transfer rate
Q	Positive definite symmetric matrix
r	Reference speed
$r(t)$	Demand vector
R	Constant demand vector
R	Positive definite symmetric matrix
$\mathrm{R}(\cdot)$	Range space
$R(z)$	Rosenbrock's system matrix
s	Sliding variable
$\dot{s}$	First derivative of sliding variable
$\ddot{s}$	Second derivative of sliding variable
s_o	Boundary layer around sliding surface
$s(x)$	Switching function
S	Sliding surface matrix
T or T_{per}	Period of the PRBS sequence
T_1	Engine block inlet temperature
T_2	Engine outlet temperature and radiator inlet temperature
$\hat{T}_2$	Estimated engine outlet temperature and radiator inlet temperature
T_3	Radiator outlet temperature
$\hat{T}_3$	Estimated radiator outlet temperature
T_B	Engine block temperature
T_{amb}	Ambient temperature
T_r	Orthogonal matrix
t	time
u_{eq}	Equivalent control
u_{low}	Input signal after filtering through low pass filter.
u_L	Linear term of system input
u_N	Non-linear term of system input
u_I	Integral control term
u_2	Discontinuous function of sliding variable

v_{cont}	Switching function for controller
v_n	Equivalent injection control term $(n = 2, 2a \text{ and } 3)$
v_{obs}	Switching function for observer
$V(\bullet)$	Lyapunov function
V	Right eigenvector matrix
$\dot{V}_B$	Coolant volume flow-rate through engine block
$\dot{V}_{rad}$	Coolant volume flow-rate at radiator inlet
w	Reference model state
w_{eig}	Left eigenvectors of the closed loop system
W_{HOSM}	Variable controller parameter
$W_{oscillation}$	Value of W when engine speed begins to oscillate
x	System state
$\hat{x}$	Estimated system state
$\dot{x}$	First derivative of the state
$\ddot{x}$	Second derivative of the state
y	System output
z_i	Partitioned states, $(i = 1, 2)$
α	Thermostat valve opening fraction
$\hat{\alpha}$	Estimated thermostat valve opening
α_{vss}	Positive constant
α_i	Positive constant $(i = 1, 2, \ldots, n)$
β	Filter coefficient
δ_{cont}	Smoothing factor for controller
δ_{obs}	Smoothing factor for observer
λ_i	Eigenvalue $(i = 1, 2, \ldots, n)$
λ_{HOSM}	Variable controller parameter
$\lambda_{oscillation}$	Value of λ when engine speed begins to oscillate
ρ	Coolant density
ρ_{cont}	Design gain for the discontinuous term of controller
ρ_{obs}	Design gain for the discontinuous term of observer
ρ_{HOSM}	Variable controller parameter
ϖ	Positive constant

κ	Condition number
ℓ_{IASM}	Switching line
ℓ_{MFSM}	Switching line
ε	Surface emissivity
Σ	Relative emissivity
σ	Steffen-Boltzmann constant
Γ	Stable design matrix
Γ_m	Lower positive bound on the smooth uncertain function, γ
Γ_M	Upper positive bound on the smooth uncertain function, γ
Φ	Stable design matrix
Φ_{HOSM}	Positive norm bound on the smooth uncertain function, ϕ
ϕ	Smooth uncertain function
γ	Smooth uncertain function
Δ	Matrix/parameter perturbation symbol
Δt	Discrete interval of time

List of Abbreviations

A2D	Analog To Digital conversion
ADC	Analog to Digital Converter
AMC	Manufacturer And Customer
ANN	Artificial Neural Network
ARMAX	Auto-Regressive Moving Average eXogenous
BDC	Bottom dead centre
BMEP	Brake Mean Effective Pressure
BSI	British Standard Institution
CI	Compression Ignition
CLDI	Closed-Loop Dynamic Identification
CMFD	Condition Monitoring Fault Diagnosis
dSPACE	Digital Signal Processing And Control Engineering
D2A	Digital to Analog conversion
DAC	Digital to Analog Converter
DC/d.c.	Direct Current
DI	Direct Identification
DSP	Digital Signal Processing
EBDI	Event-Based Dynamic Identification
EGR	Exhaust Gas Recirculation
EPG	Electrical Powered Governor
EPSRC	Engineering and Physic Science Research Council
FDI	Fault Detection and Isolation
GAC	Governor America Corporation
HOSM	Higher Order Sliding Mode
Hz	Hertz
IASM	Integral Action Sliding Mode
II	Indirect Identification
IC	Integrated Circuit
I/O	Input Output
ISO	International Standard Organisation
JREI	Joint Research Equipment Initiative
kW	Kilo Watts
KBM	Knowledge Based Method

LED	Light Emitting Diode
LQ	Linear Quadratic
LQG	Linear Quadratic Gaussian
LQI	Linear Quadratic and Integral
LTI	Linear Time Invariant
MBC	Model-Based Control
MFSM	Model Following Sliding Mode
MIMO	Multi Input Multi Output
MIT	Massachusetts Institute of Technology
MRC	Model-Reference Control
MSFV	Mean Square Fit Value
OLDI	Open-Loop Dynamic Identification
PEM	Prediction Error Method
PI	Proportional-Integral
PID	Proportional-Integral-Derivative
PRBS	Pseudo Random Binary Sequence
PRT	Platinum Resistance Thermometer
PWM	Pulse-Width Modulation
rpm	Revolution Per Minute
SISO	Single Input Single Output
SMC	Sliding Mode Control/Controller
SMO	Sliding Mode Observer
TBDI	Time-Based Dynamic Identification
TBM	Time Based Method
TDC	Top Dead Centre
TF	Transfer Function
VSS	Variable Structure System
WGC	Woodward Governor Corporation

List of Tables

List of Figures

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Appendix

Chapter 1

1 Introduction

This chapter describes problems of interest in the condition monitoring and control of diesel engines. The thesis explores the use of sliding mode techniques to solve these particular engine problems. Issues such as model-based fault diagnostic systems using sliding mode observer schemes, closed-loop event-based system identification, sliding mode controller design and implementation are covered.

Section 1.1 describes engine problems of interest in the research project. Section 1.2 presents the motivation for using sliding mode techniques to solve such industrial problems. A review of previous work, in the area of fault diagnosis and control system, is presented in Section 1.3. The approaches and research project objectives are clearly stated in Section 1.4. The research contribution to industrial applications is highlighted in Section 1.5, followed by the organisation of this thesis in Section 1.6.

1.1 Problem of Interest

With the development of technology, our dependency on engines is increasing. Engines are used in many areas. The construction of the engine is becoming more complex with the new technologies. The ever changing demand and higher specifications from customers have led to a challenging and competitive engine industry. The outcomes motivate the development of a better specified engine, not only with higher efficiency, safer, more robust and reliable, but also meeting the confines of tight legislation.

The complexity of the engine design means that maintenance work becomes more complicated and requires highly skilled operators to service the engine. A complex engine also means that a longer service time is required and sometimes it is so complex that the engines have to be sent back to the manufacturing plant for service. The engine downtime for service may not always be feasible for many manufacturing industries due to the halt in production activities and associated financial losses. In the event of an engine fault, the repair work will be far more difficult. If minor engine faults are not corrected, catastrophic faults can result. Consequences are not just that engines overheat and are damaged, but massive property damage and loss of life can ensue.

For example, aircraft accidents claim the lives of many people and many such accidents are the result of instrument failure [Gertler, 1998]. The space shuttle *Challenger* explosion in 1986 resulted in destruction of the spacecraft and the death of its seven crewmembers [Nasa, 1986]. Undetected faults contributed to these accidents.

Another area of interest is focused on engine control systems. Power generators can be seen operating in remote areas, as a standalone power supply and backup supply unit. Most diesel generators operate at a single speed where the speed here is a measurement of the number of crankshaft revolutions per unit time (i.e. revolutions per minute). The crankshaft drives the attached alternator which generates the alternating current power supply at fixed frequency at constant speed. It is important to maintain a constant speed in order to maintain the desired electrical frequency. However, due to varying electricity demand from the consumer, the generator speed will tend to vary due to load changes. The problem of varying frequency can affect some electrical instruments which are sensitive to the varying frequency and may cause damage. Irregular changes in speed can increase the rate of wear and tear in the engine components and reduce the overall engine life expectancy. Thus, the speed controller must be capable of tracking the nominal generator speed.

Idle speed control problems have been frequently addressed in the literature. Although several robust control strategies have been proposed, the level of acceptance and application are low. It is believed that the industrial community still has scepticism about the complexity associated with many new control algorithms. There is thus a big gap between the control research community and industry. The proportional-integral-derivative (PID) controller is a popular controller which has been used in industry for more than five decades. PID controllers are usually individually tuned at different process conditions and every part of the operation envelope must be considered. Bhatti [Bhatti 1998] has also mentioned that for coupled multi-input multi-output processes, decoupling cannot be easily achieved through PID tuning.

1.2 Motivation

Systems for the early detection and diagnosis of faults are important in the operation of modern diesel engines. Effective fault diagnosis has numerous benefits. A condition monitoring and fault diagnosis system enhances engine safety, which is important throughout the process industry and paramount in safety-critical situations. The constant monitoring of system may give an indication of the occurrence of faults before they become serious. This can improve plant maintenance schedules and thus minimise downtime which incurs high financial cost. Early indicators allow action to be taken to maintain the safe, continued and economic operation of an engine.

Control systems perform like a human brain; they react to external changes and send out the command signals required to control the plant under consideration. The controllers should maintain the engine speed within operating limits and reject external disturbances such as load changes. It is desirable that the controller can show robustness working over a wide operating envelope at different reference speed settings.

A model is a set of equations in a specific form to represent some physical component, process or system. The model representation is increasingly important for fault diagnosis, simulation, control system design, product prototype design and process prediction. Model based fault diagnostic systems can predict the trend of possible faults in the system. Also control system design based on a plant model can reduce the design time, allow simulations to be carried out and facilitate better prediction.

In the report of the Panel considering *Future Directions in Control Theory: A Mathematical Perspective* [Fleming 1988], Fleming classifies the process of controlling a dynamical system into the following four fundamental issues:

1. *Modelling* of the system based on either physical laws or identification.
2. *Signal processing* of the output by filtering, prediction, state estimation, etc;
3. *Synthesizing the control input* using mathematical control theories and carrying out simulation tests.
4. *Implementation* of the controller onto a test rig. This process includes developing real-time algorithms and computer code.

This has been the motivation towards the use of a model based approach for fault diagnostics and control systems design. The choice of a mathematical model for a given system depends on the chosen control strategy.

The improvement in computer technology has certainly benefited the model-based design process both for fault diagnosis and control. Large and complex systems can now be designed and simulated on a computer. Technology also allows real-time simulation, implementation and on-board control.

1.3 Previous Work

A general survey of fault diagnostic methods can be found in several papers [i.e. Isermann 1984, Frank 1990, Gertler 1988]. These authors emphasise model-based methods to solve the fault diagnosis problem. Isermann suggested using a continuous-time model for plant parameter estimation. Frank outlined the principles and the most important techniques of model-based residual generation using parameter identification and state estimation methods. Gertler also surveyed the use of model-based techniques to generate the residuals. These methods employed observer schemes to estimate plant parameters which are non-measurable. Various solutions to the fault diagnostic problem are the parity equation residual generation method [Krishnaswami 1995, Gertler 1995], Neural Network [Jones 2000a, Parikh 2001] and fuzzy logic method [Twiddle 2001, Miguel 1996]

Speed control problems can be traced back to the eighteenth century when James Watt invented the centrifugal governor for speed control of a steam engine [Ogata 1997, Challen 1999]. Other techniques like PID methods [Tao 1995], Optimal control – such as the Linear Quadratic approach [Butts 1999], Fuzzy logic [Thornhill 1999], Sliding mode technique [Bhatti 1999a] have all been used for speed control of an engine. A review of the techniques for both the fault diagnosis and speed control will be presented in detail in the next chapter.

1.4 The Approach of this Research

Engine faults may be due to the malfunction of components. These faults may propagate to the neighbouring parts of the engine and thus the initial faults may worsen. This raises a significant issue in monitoring the health of the engine at an early stage. Keeping the engine in good working condition is vital to maintaining the overall efficiency.

Condition Monitoring and Fault Diagnosis (CMFD) is a valuable measure to ensure that the engine stays in good health. Early detection brings benefits not only in preventing possible catastrophic faults but also in reducing the manufacturing downtime and sustaining the productivity level. The model-based fault diagnosis approach has received much attention. It is based on the concept of analytic redundancy, as opposed to physical redundancy, and uses signals generated by a mathematical model of the system being considered. An advantage of the model based technique is that no additional hardware components are needed in order to realise a CMFD algorithm.

The approaches of the research project for both fault diagnosis and speed control are stated in the research objectives as follows:

1. To develop an *on-board* fault diagnostic system using sliding mode techniques, particularly with application to the diesel generator cooling system. The on-board system means that the diagnostic system must be able to continuously monitor the engine for its day-to-day operating conditions.

2. The fault diagnosis system uses an observer scheme and a model-based approach. The use of additional instrumentation is to be minimised, and the instrumentation should be robust and cost effective to be economically applied within industry.

3. To develop an engine model for isochronous speed control system design. Explore a novel closed-loop event-based system identification method.

4. Design real-time speed control systems using sliding mode techniques. Explore both model-based and model-free sliding mode techniques. The designed controller is required to be such that non-experts can maintain and tune the controller conveniently.

5. Assess the robustness and performance of the controllers developed by comparing them with alternative candidates.

6. To develop a control tuning algorithm which may be more transparent to the end user than those currently available.

1.5 Contributions of the Thesis

The main contribution of this thesis is that the proposed sliding mode techniques can be applied in various applications involving a diesel engine. A list of detailed contributions is as follows:

(a) A new approach to model-based fault diagnostic systems design and implementation using a sliding mode observer scheme is developed. The technique was implemented on cooling system components. The diagnostic method uses cheap, readily available sensors which means that the application is simple and economical to implement. It is demonstrated that sophisticated parameters can be estimated using the sensor set. The *on-board* implementation shows the proposed fault diagnosis scheme has the potential to diagnose particular engine coolant system faults.

(b) A novel way to model engine dynamics for idle speed control has been demonstrated – closed-loop event-based system identification with event-injection of excitation signals.

(c) An integral action sliding mode observer-controller pair design scheme is applied to the proposed closed-loop event-based linear model. The controller has speed tracking capability. A tuning procedure is proposed.

(d) A model-following sliding mode observer-controller pair is applied to the same engine model. A tuning procedure is proposed.

(e) A new method of implementation of a higher order sliding mode controller with modifications to the design algorithm proposed by [Levant 1993, Khan 2001] is considered. The modified control algorithm improves performance and is shown to work over a wide operating envelope. A higher order sliding mode controller tuning algorithm is proposed which is suitable for non-experts to apply.

(f) The performance of the designed controllers is assessed by comparing them with commercial and PI controllers.

1.6 Organisation of the Thesis

Chapter 2 is a preliminary chapter providing a brief history of the diesel engine and its important role in various applications in the world today. The concept of operation is described as well as the test rig diesel power generator. Several fault diagnosis issues are discussed as well as possible faults in the cooling system. Controller benchmarking, specific controller assessment and hardware calibration for the experimental tests are also covered in this Chapter.

In Chapter 3 a detailed literature survey of the sliding mode technique is presented, along with the background of the concept. Both sliding mode observer and controller design is described. An example is used to illustrate the design algorithm. The properties of sliding mode schemes and the robustness of the concept are addressed.

Chapter 4 provides the design procedure for the fault diagnostic system. The model of the particular cooling system under consideration is derived from thermal heat balance equations. The sliding mode observer scheme is employed in the design, along with the description of residual generation using a non-linear model. The off-line test and on-board implementation are discussed.

The model identification for the idle speed control is described in Chapter 5. The closed-loop event-based technique is discussed in detail as well as the hardware design for the data acquisition and event-triggering system. An ARX identification method is addressed in this Chapter. This Chapter ends with the identification procedure, model verification and model analysis.

The design of the idle speed controller is divided into two categories, namely model-based and model-free control. Chapter 6 will demonstrate the design of two different types of model-based sliding mode controllers. The first controller design is based on the integral action speed tracking technique using the identified model of Chapter 5. The second controller uses the model-following concept for the design. Both control system designs involve the use of an observer scheme to estimate the system states. Implementation results are discussed.

Chapter 7 describes the application of a higher order sliding mode control strategy to the speed control problem of a diesel power generator. This approach is regarded as a model-free technique. A modification of the super-twisting algorithm is established and justified by the robustness of the controller. A novel controller-tuning algorithm is presented and the resulting controller performance is compared with that of a commercial controller.

Chapter 8 presents an overall controller performance assessment. The performance tests include the starting of the engine (i.e. transient response), fuel consumption, exhaust emission, robustness assessment by changing the resistive load. Other performance tests include changing of reference speed and steady state performance evaluation.

The final Chapter describes further work that can be developed beyond the scope of this thesis and indicates further research directions on the engine using related techniques. References and Appendices conclude the thesis.

Chapter 2

2 An Overview of the Diesel Engine System and related Fault Diagnosis and Control Techniques

The introduction to this chapter outlines the history of diesel engines. A particular diesel-electric generator is introduced. The cooling system of the diesel engine and the cooling system components are described in detail. Several possible cooling system faults at the component level are addressed in this chapter, and then followed by the engine speed control system. A brief introduction to the commercially available speed controllers is presented and controller performance test criteria are listed in detail. The selection of proper transducers for the test design is described. This chapter also addresses the requirements of fault diagnosis and speed control; instrumentation, data acquisition and instrument calibration are discussed. The techniques used for the design of the fault diagnostic system and speed control system are briefly described. Lastly, a PI controller setup is presented.

2.1 Diesel Engine Systems

The history of the diesel engine can be traced back over a hundred years to the late 19th century. Rudolf Diesel (1858-1913), is the pioneer who invented the diesel engine and the engine was then named after him. He obtained the first patent for the internal combustion (diesel) engine on 28 February 1892. The first fully operational diesel engine was only introduced in 1897 [Moon 1974]. Since then, it has been the most widely used industrial source of power. His original concept allows a maximum amount of work to be obtained from a given heat source which has yet to be bettered commercially.

The diesel engine is a sparkless compression-ignition engine, which transforms the energy stored in fuel into motion and useful work. The compression-ignition (CI) principle employed by the diesel engine distinguishes it from the petrol engine which uses a spark-ignition (SI) principle. In diesel engines, the air is compressed and the fuel is then injected at an appropriate point in the cycle once the air is well above the ignition temperature. In the SI engine, a spark plug is used to ignite the fuel mixture.

Despite competition, the use of the diesel engine is increasing and over the years it has become steadily more efficient, fuel economic and more adaptable. Diesel engines are the most efficient internal combustion engines and can achieve over 50% efficiency [Guzzella 1998]. They are one of the most likely prime movers for efficient cars of the future (the "80 miles per gallon" vehicle).

Advances in technology bring engine design and construction to a higher level of robustness, reliability, increased compression ratios and improved air/fuel mixing. The application of such techniques as turbo-charging has increased the power output of the diesel engine. Improvements in lubrication and cooling techniques also contribute to greater efficiency. The tight emission legislation led to the application of techniques like exhaust gas recirculation (EGR) and electronic fuel control with the aim of reducing the levels of exhaust emission.

In general, diesel engines can be classified into two categories, two-stroke and four-stroke engines. In the two-stroke engine, combustion occurs in the region of top dead centre (TDC) of every revolution and gas exchange at every revolution at bottom dead centre (BDC). TDC is the position where the cylinder reaches its maximum upward movement and vice versa for the BDC. The position of TDC and BDC are shown in Figure 2-1. In the four-stroke engine, there are four cycles namely, intake, compression, expansion, and exhaust. The combustion only occurs once every two revolutions. The main advantage [Challen 1999] of the four-stroke cycle is that it provides a longer period for the gas exchange process which results in purer trapped charge. This allows time for sufficient fresh air to be drawn into the combustion chamber to mix with the diesel and the exhaust can be drawn out of the combustion chamber. It also lowers the thermal loading associated with engine internal components like pistons, cylinder heads and liners. The application of the two-stroke engine can be seen in use in marine and stationary applications while four-stroke engines are used in the majority of other applications.

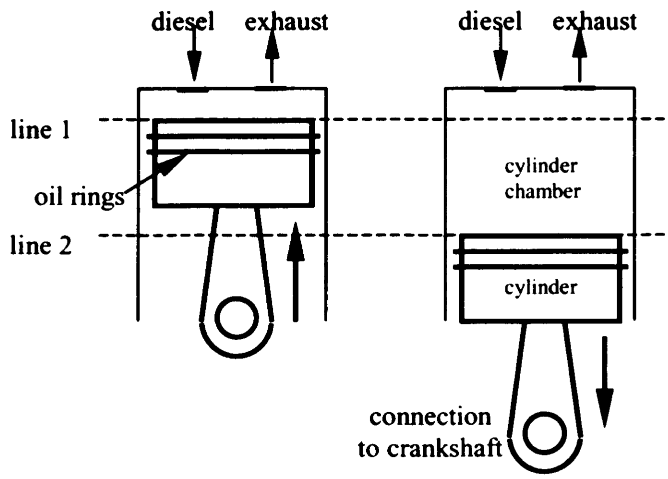


Figure 2-1: Schematic shows: line 1- top-dead-centre and line 2 - bottom dead centre.

2.2 Test-Bed Diesel Power Generator Set

The research diesel engine is the prime mover in an electricity generator set (genset), a Perkins 1000 series, four litre, four-cyclinder, turbo charged diesel engine. It was purchased by a Joint Research Equipment Initiative grant [JREI (GR/M307777)] provided by the EPSRC and Perkins Engine Company Ltd. The engine drives an alternator at a constant speed of 1500 revolution per minute (rpm). The alternator generates maximum power of 65-kilowatts (kW) and dissipates the generated power via an electrical resistor load bank. The diesel generator is shown in Figure 2-2.

2.3 Diesel Engine Cooling System

A cooling system on an engine is used to remove excess heat from the engine to keep the engine operating at its most efficient temperature and to get the engine up to the correct temperature as soon as possible after starting. Ideally, the cooling system keeps the engine running at its most efficient temperature no matter what the operating conditions are. Figure 2-3 shows a schematic of the engine cooling system. The cooling system consists of a thermostat valve, a coolant pump, a radiator and a cooling fan (belt-driven). The crosses show the location of the temperature sensors where the respective temperature values are represented by T_1, T_2, T_3, T_B and T_{amb}. α represents the thermostat valve opening fraction.

Figure 2-2: Diesel Generator and instrumentation set.

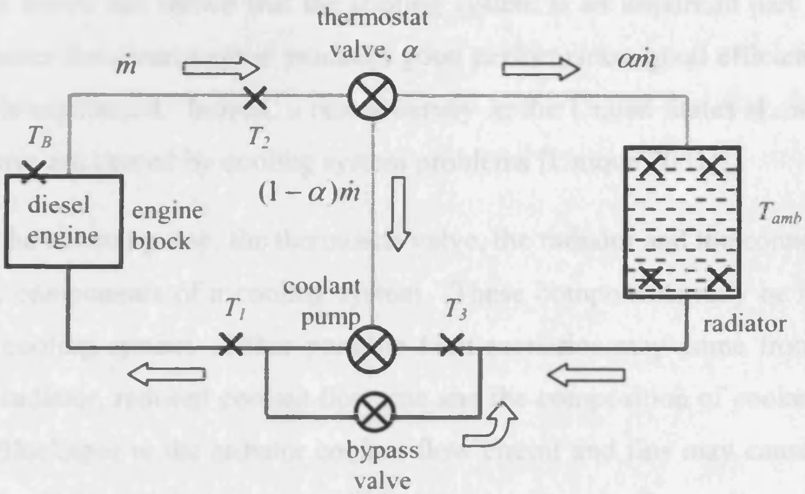


Figure 2-3: Engine cooling system.

2.3.1 The Need for an Effective Cooling System

During the internal combustion process, some energy in the fuel is converted into power which is lost as exhaust and heat energy. A cooling system is important in most internal combustion engines to remove excess heat. The internal combustion temperature can reach as high as 800 Kelvin [Heywood 1988] during the compression stroke. If no cooling system is provided, the engine parts will start to melt from the heat of the burning fuel, and the pistons will expand so much they could not move in the cylinders (called "seize").

Although the exhaust system takes away much of the heat, parts of the engine, such as the cylinder walls, pistons, and cylinder head, absorb large amounts of the heat. If a part of the engine gets too hot, the oil film fails to protect it. This lack of lubrication can ruin the engine.

On the other hand, if an engine runs at too low a temperature, it is inefficient, the oil loses lubrication ability (adding wear and reducing the power produced), deposits form, fuel efficiency is poor and excessive exhaust emissions occur [Autoshop 1998]. Therefore, the cooling system is designed to be inoperative until the engine is warmed up. A thermostat valve ensures this and remains closed until the coolant temperature (or engine block temperature) rises to a desired value. On the research engine, coolant flows through the engine, absorbing heat energy from the engine cylinder wall and block. The hot coolant leaving the engine is cooled by the radiator (via radiator fins) where the hot air is blown away by a radiator fan. This process continues as long as the engine is running, with the coolant absorbing and removing the engine's heat, and the radiator cooling the coolant.

2.3.2 Possible Cooling System Faults

The discussion above has shown that the cooling system is an important part of the diesel engine. It ensures the diesel engine produces good performance, good efficiency and wear of the engine is minimised. Indeed, a recent survey in the United States showed that 50% of engine failures are caused by cooling system problems [Unique 2001].

In Figure 2.1, the coolant pump, the thermostat valve, the radiator and the connecting circuit form the basic components of a cooling system. These components may be the source of failure of the cooling system. Other possible fault scenarios may come from inadequate cooling in the radiator, reduced coolant flow rate and the composition of coolant (water and anti-freeze). Blockages in the radiator coolant flow circuit and fins may cause a reduction in effectiveness of the cooling process. If leakage happens in the cooling system or the coolant pump or thermostat valve malfunction, this will induce a low coolant flow rate in the system. The composition of coolant is also important as it gives protection against rust, scale, and some forms of corrosion. Many of the mentioned faults may induce overheating in the engine. This may pose a threat to the neighbouring engine parts or components and may cause major damage. If monitoring work is not performed consistently, it will become a costly problem. The possible faults in the cooling system are summarised as follows:-

(i) Radiator failure

(ii) Low coolant flow due to pump malfunctioning, or system leakage

(iii) Thermostat valve fails to open

(iv) improper coolant composition

In order to carry out tests under these fault scenarios, various faults are 'injected' into the diesel engine. Further description of the 'simulated' faults is given in Chapter 4.

2.4 An Overview of Fault Diagnostic Techniques

The engine health monitoring process is widely discussed. Fundamentally, the monitoring process involves diagnosing engine parameters which deviate from 'normal' behaviour. Faults can come from several sources. Traditionally, engine faults have only been identified after the engine has broken down. If there is an on board parameter monitoring device to reflect the online engine condition, then diagnosis can be affected earlier. A common parameter indicator (i.e. which can indicate the operating condition of the engine) on automotive engines and electric generators, is temperature. However, this parameter only gives a warning and does not detail the cause. One of the objectives of this research is to design a condition monitoring and fault diagnosis (CMFD) system which can be used to investigate the possible faults at component level for an engine coolant system. The system must be capable of detecting and isolating particular faults on the engine.

There are a number of books that have been published on fault detection for engineering processes [Patton 1989, Chen 1999, Gertler 1998, Chiang 2001] and various fault diagnosis techniques have been reported in the literature. A few detailed survey papers are available [Isermann 1984, Gertler 1988b]. Isermann reported that process faults could be detected when based on the estimation of unmeasurable process parameters and state variables. He also described the suitable choice of using parameter estimation for continuous-time models in fault detection. He then reported the use of process model knowledge [Isermann 1991] and model-based [Isermann 1997] fault diagnosis of technical processes. Gertler pointed out the possibility of using both model-based and model-free methods for fault detection. He asserted that the sensitivity and robustness of models play a role in the selection of a desired model and filtering may be applied to improve sensitivity within a given model framework. Modelling errors affect the fault detection process and may falsify the failure signatures. The following gives a brief overview of some techniques in general terms.

2.4.1 Model-based Technique

The model-based technique is the most common approach to fault detection and isolation. This method utilises an explicit mathematical model of the plant under investigation. The sensor measurements are compared to analytically computed values (i.e. from the model) of the respective variable and the resulting difference is called the residual. The residuals are indicative of the presence of faults in the system. There are a few approaches to residual generation in model based fault diagnosis.

1. Parameter Estimation (Identification)

Parameter estimation is an approach to the detection of parametric faults. A reference model is obtained by identifying the plant in a fault-free condition. The models may be constructed from first principles to relate the model parameters directly to parameters that have physical meaning in the process. Any deviation from the reference plant model serves as a basis for fault detection. This technique may be applied to a non-linear system where the structure is known. For a system without a known structure, the system is treated as a *black box* system where the input and output signals are available for identification. The parameter estimation method is appropriate if the process faults are associated with changes in model parameters (i.e. multiplicative faults), and appropriate mathematical models are available.

2. Observer-based Technique

The observer-based technique can be used for both non-linear and linear plant models. The plant states may provide important information to the plant operating condition. However, some plants states are known to be non-measurable and an observer can be employed to estimate these plant states.

Once the residuals are obtained, evaluation of these residuals may be carried out. Residual evaluation is a decision making stage to classify the particular faults in the plant. Some of the faults gradually build up and may provide a certain fault pattern. This fault pattern can be recognised based on the history of the plant or some theoretical considerations. Once the residuals are obtained, they may be classified into further categories, i.e. good condition, moderate, serious or danger to give the user a clear picture of the plant condition. The users can then react to these decisions as appropriate.

2.4.2 Model-Free Technique

The fault diagnosis techniques which have been reviewed above required a plant model. However, some plants may be too complex to efficiently derive a plant model. A model-free technique is then required. There are several model-free techniques reported by Gertler [Gertler 1998]. Two of the techniques are physical redundancy and spectrum analysis. A brief description of these techniques is presented as follows.

1. Physical Redundancy

This approach can be used to detect sensor faults. To carry out the diagnostic process, multiple sensors are installed to measure the same physical quantity. Any serious discrepancy between the measurements indicates a fault in the plant. If only two sensors are used, fault isolation is not possible. If three sensors are used, a voting scheme can be used to isolate the faulty sensor. However, physical redundancy involves additional hardware and cost.

2. Spectrum Analysis

Spectrum analysis of plant measurements may be used for fault detection. Most plant variables exhibit a typical frequency spectrum under normal operating conditions. Any deviation from this normal condition is an indication of abnormality.

2.5 Diesel Engine Speed Control System

2.5.1 *Engine Speed Control*

The fundamental objective of the speed control system is to keep the system output speed close to the desired speed value. The notion of tracking performance is used to describe the control system's ability to meet this objective during normal operation when the system is subjected to external disturbances such as electrical load. Abrupt changes in the external disturbance generate transients in the system since the physical system cannot react instantaneously. The controller forces these transients produced to decay in a reasonable period of time. Figure 2-4 shows a block diagram of an engine speed control system.

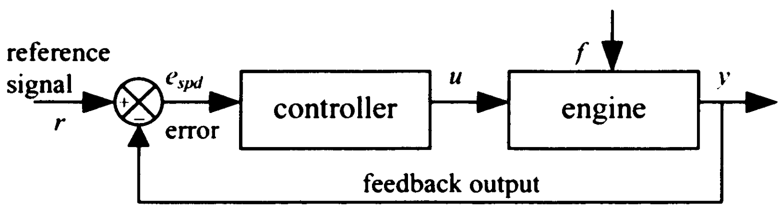


Figure 2-4: A feedback control system for engine speed control.

The control signal u is generated by the controller and applied to the engine to obtain the desired speed value. The measured engine speed, y, is fed back into the controller for comparison with the reference speed (desired speed), r, before it is used in generating the control input, u. The y value is the engine parameter to be controlled, while the reference speed, r, specifies the desired value of the engine speed, y. The external disturbance, f

consists of those components that force the engine to exhibit undesirable behaviour. For a diesel generator, such an external disturbance would be caused by a change of load.

2.5.2 The Need for a Robust Speed Control System

Most diesel engines driving electrical power generators operate at constant speed. It is important to maintain the speed constant in order to maintain the desired electrical frequency. However, due to varying electricity demand, the generator speed will tend to vary due to load changes. Thus, the speed control system must be capable of tracking not only the nominal speed, but also must show capability of rejecting load disturbances (i.e. electrical loading) and be valid across a wide operating envelope.

A good controller may improve the lifetime of an engine. Inconsistency in speed not only generates inconsistent frequency in the power supply but also induces wear and tear on the engine parts due to vibration. Oscillation in speed may occur if an inappropriately tuned controller is employed on the engine. This may induce higher exhaust emissions because of incomplete burn of fuel, as well as reducing fuel efficiency.

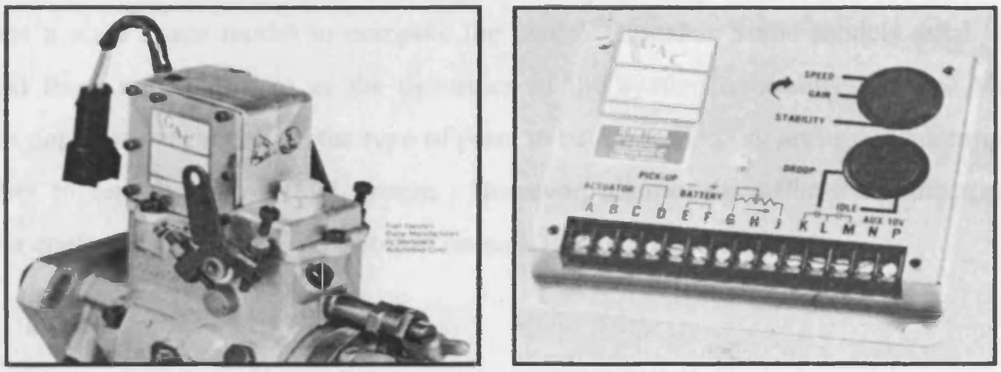
2.5.3 Commercial Speed Control System

The history of early commercial speed control systems can be traced back to James Watt's time when he first applied a 'centrifugal' governor onto an engine [Ogata 1997, Challen 1999]. The Watt governor is purely mechanical. It uses the position of fly weights to set the required fuelling and hence provide a speed control action. Many modern engine speed control systems are electronic and usually programmable or with a tunable control gain. These controllers are normally powered by the engine battery. Common tunable control gains that can be found on the commercial controllers are for speed reference, stability, droop and idle. The engine speed information is obtained from a transducer, i.e. tachometer, shaft encoder or magnetic pickup sensor.

Commercial control systems for diesel engine speed control can be bought off-the-shelf. *Governors America Corporation* (GAC) is one of the manufacturers who supply such control systems. A GAC control system is shown in Figure 2-5. Figure 2-5 (a) shows the actuator (ADC100) and Figure 2-5 (b) the control unit (ESD5100). The actuator is mounted directly on a Stanadyne fuel injection pump. The labels on the control unit indicate that it requires a pick-up signal (engine speed) to feed into the unit, then a command signal is sent out via the label 'actuator' to the actuator. The adjustable parameter gains are shown in Figure 2-5 (b).

2.6 An Overview of Control Techniques

The robustness and effectiveness of any engine speed control system lies in the control algorithm itself and also depends on the plant variations. A good control algorithm can make a substantial difference to both the quality of control and the speed of response. Proportional-Integral-Derivative (PID) controllers have been around for decades and they still provide a major contribution to the area of commercial engine control. Figure 2-6 shows a block diagram of a PID controller with the respective PID gains.



(a) (b)

Figure 2-5: Example of commercial control system for engine speed control.

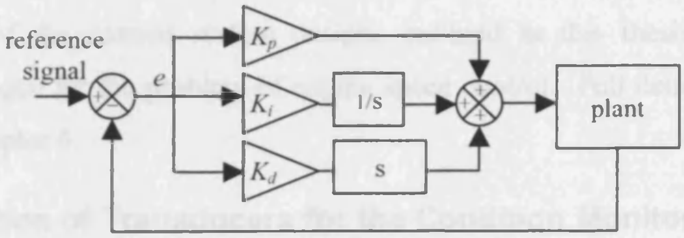


Figure 2-6: PID control setup.

However, there have been a number of alternative algorithms which offer considerable promise. The following brief descriptions highlight the principal techniques that are employed in this thesis.

2.6.1 Variable Structure Control (Sliding Mode Control)

The sliding mode concept is based on a variable structure system (VSS) concept. VSS with a sliding mode comprises a feedback control law and a decision rule. This decision rule is designed to force the system states to reach and remain on a pre-defined surface (i.e. sliding surface) within the state space. The dynamic behaviour of the system when confined to the surface is described as an *ideal sliding motion*. Greater detail of this technique is provided in Chapters 6 and 7 where sliding mode control techniques are utilised and implemented. The detailed theory and properties of sliding mode schemes can be found in Chapter 3.

2.6.2 Model-based Control

A control system which uses a model of the target system to compute control actions is referred to as model-based control. The VSS method is considered a model-based method as it uses a state space model to compute the control signals. Some models need to be identified from real-time data as the dynamics of the system constantly change. Model selection depends very much on the type of plant to be controlled. A prime consideration is the ability to represent the actual system. However, all models suffer from uncertainty. Thus, the control algorithm must be robust enough to reject the uncertainty.

2.6.3 Model Following Control

Model following is a control technique in which the system is controlled to behave like an ideal model. Chan [Chan 1973] described two configurations of model following, known as implicit model following and real model following. In implicit model following, the model is not part of the system but enters only into the design of the control law. Real model following requires the states of the ideal model and the model is a necessary part of the system. One of the control system designs outlined in this thesis presents a model following technique for the problem of engine speed control. Full details of the design are presented in Chapter 6.

2.7 Selection of Transducers for the Condition Monitoring and Control System.

Today, there is an increasing use of measurement and control in industry to improve efficiencies and reduce cost. The transducer is an essential link in such a process. In the dictionary, a transducer is defined as a device that is actuated by power from one system and supplies power, usually in another form, to a second system. The commonly used transducers are actuated by physical variables representing force, pressure, temperature, flow and level, and supply electrical signals to the front end of a measurement and control system.

The diesel engine is formed by a number of different subsystems, for example, the engine block, cooling system, radiator, fuel injection system, speed control system, exhaust gas recirculate (EGR) system etc. Each of these subsystems has their respective functionality within the engine system. Parameters from each of these subsystems may be considered as a measurand for the engine system and may be measured via various available transducers.

In operating the diesel engine, the safety of the operator is a priority. The measured engine parameters must give sufficient indication to warn the operator of possible dangerous situations. For example, engine over-speed and over-heating alarms. These two parameters can be easily measured based on speed and temperature data. Additional transducer information from the engine, such as coolant and lubricant levels, will give the user a clear picture of the engine condition and allow regular maintenance to be scheduled and carried out.

In general, the price of a transducer can vary from £2.00 to a few thousand pounds [Twiddle 2001]. Some transducers may provide a simpler functionality but lack accuracy and robustness. The selection of the parameters which are most appropriate for the design purpose, will be a trade-off between functionality and cost. A list of transducers is shown in Table 2-1 for the respective observed parameters. These are the sensors that were used in the research test.

Transducer type	Parameter
Thermocouples/platinum resistance thermometers (PRT's)	Temperatures
Frequency transducers	Frequency rate
Flow meter/Hall-Effect sensors	Fuel flow rate
Reflective sensor	Flywheel/crank-angle indicator
LED transmitter/detector	Smoke opacity

Table 2-1: List of sensors.

Temperature is a commonly observed variable in engine fault diagnosis. The energy conversion, from diesel fuel energy to electrical energy, produces heat that propagates to the engine block. The heat is carried away by the coolant and is dissipated via the radiator. Engine temperature may determine the state of the engine and efficiency of the cooling system. One of the main objectives of fault diagnostic system design is to minimise the number of sensors. The proposed solutions must be cost effective and give economy for implementation on a commercial engine. Temperature sensors are common in the application of fault diagnosis work and are economically attractive.

In the engine speed control system, engine speed information is required. The speed of the engine indicates the number of revolutions the engine crankshaft rotates per unit time and it is normally measured in revolution per minute (rpm). There are a few transducers capable of providing the speed signal; namely, magnetic, Hall effect sensors, reflective sensors and a shaft encoder. The magnetic sensor is fitted to read pulses of the gear teeth on the crankshaft flywheel. This pulse signal is fed into a frequency-to-voltage instrument to convert the high frequency signal into a readable d.c. voltage level. With a proper sensor gain, the actual engine speed in revolutions per minute (rpm) can be calculated. The Hall effect or reflective sensor can be fitted to read a datum position, i.e. the through-hole on the flywheel. The sensor signal will provide one pulse per revolution. This pulse signal not only provides speed information but can be used for event-based system identification. With knowledge of the datum position, i.e. the TDC of the cylinder, this information can be used to provide event pulses.

Transducers are also capable of measuring other engine parameters like fuel flow (using a Hall effect sensor), air flow and exhaust smoke opacity % (using LED detector/emitter). These parameters may be used for further observation of the engine condition and also for control system performance comparison. Selecting any transducer is always a trade-off between functionality and cost. Factors like robustness, sensitivity, reliability, lifetime and the need for scheduled maintenance must also be taken into consideration.

2.8 Benchmark for the Speed Control System

Benchmarking is a process to measure the performance of a system by selecting performance criteria and comparing them relatively to a so-called base or standard system. This process has been widely used to evaluate how effectively a system performs. In other words, it is a process to set a standard level for comparative purposes.

For a speed control system, a standard specification is set by bodies such as the International Standard Organisation (ISO) and in the United Kingdom, the British Standard Institution (BSI). The speed control system on the diesel engine used in this research is classed as single speed governing. The class of accuracy and the necessary parameters are defined in the British Standard [BS5514-1 1987, BS5514-4 1979]. This standard provides a guideline to justify the performance. More often, this issue is settled by mutual agreement between the manufacturer and the customer (AMC). One must take into consideration that not all kinds of governors are capable of meeting all the requirements of a particular engine and application. In the British Standard, there are five categories (i.e. A_0, A_1, A_2, B_1 and

B_2) to describe the class of governing accuracy. The Class A_1 (i.e. normal requirement of governing accuracy) is the one that applies to the research engine on test [BS5514-4 1979].

Although the above standards may provide a standard specification for speed controller performance, from the practical point of view, the controller performance may depend on the type of engine, its construction and application. AMC is often set to give agreed controller performance. In this thesis, two different control systems are considered and employed in the benchmarking procedure, namely, a proportional-integral (PI) action controller and a Governor America Corp (GAC) control system set. The benchmark results are used to assess the sliding mode controllers developed in the thesis as well as the BSI standards.

2.8.1 Speed Control System Benchmark Criteria

The speed controller is tested against various criteria. Maintaining the engine speed at nominal speed with various types of electrical loading is the first criterion to be assessed since single speed governing is required. Some engines operate at wide operating envelope, such as at different reference speeds. Thus, the control must be able to cope with these changes and at the same time, show robustness against load disturbances. The speed controller is also tested with regard to fuel efficiency and exhaust smoke emission. The following paragraphs will describe the respective criteria in detail.

a) *Engine Speed*

For a diesel generator, the engine speed is fixed at a nominal value or reference speed set point. The control system must be capable of maintaining the speed at this nominal value at the zero loading condition.

b) *Steady State Behaviour - Speed Band*

This test is to determine the width of the envelope of variation of the engine speed under "steady state" conditions. It is expressed as a percentage of the declared speed (i.e. 1500 rpm). It is important to ensure the engine operates at nominal speed to prevent irregular changes in speed and resulting changes in the generated power frequency. Several steady state loading conditions (i.e. electrical load or power) may be applied to the system. The power levels are divided into two categories for testing, $\geq 25\%$ (i.e. $\geq 16.25\,\text{kW}$) and $< 25\%$ (i.e. $< 16.25\,\text{kW}$), as mentioned in the BSI standard. Three data sets for each loading are recorded and the average speed variation values are measured. The data recording can only be started when the speed has reached its steady state. The BSI standard for speed variation is 0.8% (i.e. 0.8% of the 1500 rpm) and 1.0% for $< 25\%$ and $\geq 25\%$ power respectively.

c) *Start of Engine*

The diesel engine must be capable of starting reliably under most conditions. Although tests cannot be carried out in different climates, the speed transient response can be observed. Many diesel generators with nominal speed setting have a speed safety level, above which, the engine fuel will cut off and shutdown the engine automatically. During the start-up phase, the engine controller pumps in maximum fuel, trying to bring the engine speed from 0 rpm to its nominal speed. At this moment, the engine speed tends to generate maximum overshoot from its nominal speed. The exhaust smoke emission also tends to be maximum (i.e. 100% opacity) at this instant because of changes to the air to fuel ratio. The controller must control this speed overshoot so that the safety level is not exceeded and minimise the transient time.

Two different engine parameters will be investigated during the starting of the engine, i.e. engine speed transient response and exhaust emission. The transient response covers speed overshoot, rise time, settling time and delay. The speed overshoot during the start-up phase is a measure of the speed above the nominal speed and it is measured as a percentage of the nominal speed. The rise time and delay time measure the duration of the speed rising from 0 rpm to 90% (i.e.1350 rpm) and to 50% (i.e. 750 rpm) of the nominal speed respectively. The settling time records the time for the speed to rise from 0 rpm to the stage where the transient response decreases and stays within an allowable speed band. The delay parameter measures the time taken for the speed to rise from 0 rpm to 50% of the nominal speed. The next parameter is a measure of the duration of the exhaust emissions, which reach 100% opacity during the test. The settling time of the exhaust emissions measures the duration from the rising of the opacity reading to it settling down to a steady state value.

The starting characteristics depend on several factors, for example air temperature, temperature of the reciprocating internal combustion (RIC) engine, starting air pressure, condition of starter battery, viscosity of oil, total inertia of the generating set and quality of the fuel and the state of the starting equipment.

d) *Engine Speed Response to Large Step-load Change*

The dynamic behaviour of interest is the speed change with load. The transient response is defined as the maximum deviation of speed after a sudden load change from a previous speed to steady state level and the speed change is expressed as a percentage of the operating speed (i.e. 1500 rpm). These tests involve applying a large electrical load (i.e. maximum load) to the system during steady state conditions and observing the speed transient responses.

The first objective of the test is to assess the controllers in the presence of a large step change in load. This corresponds to a sudden high power demand by the consumer. In the BSI standard, it states that the test should use the maximum possible load. However, for reasons of safety, the applied load is limited to 92.3% (i.e. 60 kW) of the maximum load (i.e. 65 kW). In the BSI standard, only a 10 % nominal speed drop is allowed and the speed must recover within 8 seconds.

e) *Engine Speed Response Small Varying Step-load change*

This test criterion is similar to the previous one. The objective is to simulate small varying electricity demands by the consumer. A series of electrical load changes may be applied to the controller, i.e. step-loads of 10 kW or 20 kW, gradually increasing the load to 60 kW and then reducing it back to zero load. The engine speed response is investigated and the transient response is analysed.

f) *Robustness/Wider Operating Range Test*

The objective of this test is to show the robustness of the controller at different operating conditions, i.e. at different reference speed settings. The reference speed of the diesel engine may be set to 1350 rpm or 1200 rpm. A large load of 60 kW may be applied to the engine and the response studied. A robust controller should be capable of maintaining the engine speed at different speed set points (three) and also be able to cope with load changes.

g) *Fuel Consumption*

The fuel consumption is the quantity of fuel consumed by the genset per unit of time at a stated power. A control system that is not tuned appropriately may cause irregularity in speed. This undesirable condition may affect the fuel efficiency. Thus, fuel consumption is used to assess the controller performance.

Fuel consumption calculations may be presented in *volume per second* (10^{-9} m^3/s) and *volume per unit of power* (10^{-9} m^3/kW). The test was designed to run with the genset operating at different loading conditions, i.e. from 0 kW up to 60 kW with each step-load of 10 kW. Upon each load change, the measurements are only made when the speed reaches steady state.

The fuel that flows into the engine is measured by an electronic fuel system (Enviro Systems Ltd.) using a Hall effect sensor. The pulse series (are obtained from the sensor which are induced due to the rotating magnetic rod in the system) is then converted to a d.c. signal to give the fuel flow value. The fuel system was previously calibrated using an optical fuel meter and a Woodward Governor Corp (WGC) control system (i.e. Woodward ST-125 control and actuator set). A function of '*rate of pulse*' is obtained from the calibration to represent the actual fuel consumption.

h) *Engine Exhaust Emission*

Combustion is a chemical reaction that converts the energy contained in the fuel to the internal energy of product gases. The internal combustion engine serves as a mechanism to convert this internal energy into useful work. The given stoichiometric reaction for a particular diesel fuel type is as follows:

$$C_{13.2}H_{24.4} + 91.9(0.21O_2 + 0.79N_2) \Rightarrow 13.2CO_2 + 12.2H_2O + 72.60N_2$$

The fuel-air ratio can be obtained using the molar air-fuel ratio. The *equivalent ratio* is defined as the actual fuel-air ratio divided by the stoichiometric fuel-air ratio. If this equivalent ratio is above 0.7, excessive smoke emissions occur.

In reality, combustion is always incomplete and leaves behind unburned hydrocarbons from the fuel and partially oxidised products like carbon monoxide, as well as visible smoke. The objective of this test is to measure the quality of the exhaust emissions (i.e. visible smoke) from the genset. A device called an Opacity meter is used to measure the relative light absorption of the smoke discharged from the genset exhaust. The relative light energy loss is translated into opacity (%). A definition of opacity, N, as a percentage, is given below.

$$N = \left(1 - \frac{I}{I_o}\right) \text{x } 100$$

where I and I_0 are the measurement of the light intensity reaching the photodiode with smoke source and without smoke source respectively. The opacity readings are recorded for load changes from 0 kW to 60 kW at each step change of 10 kW. The exhaust during the start up phase of the engine is also observed.

For each of the above test parameters, three sets of data are recorded to ensure repeatable performance. For tests relating to fuel flow, smoke and steady state speed variation, a steady condition must be achieved before data is acquired. The average value of the three sets of data for the respective parameter is then considered.

The ambient condition around the engine may contribute to the engine performance. Atmospheric pressure, air temperature and relative humidity are the contributors. The ambient conditions stated in BSI are air temperature 25°C, atmosphere pressure of 1.0×10^5 Pa and relative humidity of 30%. It is assumed that these conditions remain constant throughout the tests.

2.9 Requirements For the Diesel Engine Fault Diagnostic and Control System

The preceding sections have addressed the problem of fault diagnosis and engine speed control. Several methods have been reviewed and selection of appropriate sensors considered. This section addresses several aspects that need to be considered prior to the design and implementation of an on board fault diagnostic system and speed control system.

2.9.1 Instrumentation

Once the measurands for both the fault diagnostic and control systems have been decided, the next step is to process the raw signal. The measured signals from ADC tend to have high frequency components from, for example, electrical noise. This undesired high frequency noise is filtered out through the use of a 5th order Butterworth filter. The 5th order provides a steep cut-off rate. The selection of sampling frequency depends on the detail of data that is required in the test. According to the Nyquist-Shannon rule, the sampling frequency must be at least twice as large as the signal frequency. An improper sampling rate may cause possible loss of signal information or aliasing effects.

2.9.2 Data Acquisition System

The data acquisition and control hardware systems that were used in the research consist of a dSPACE^{TM} DS1103 real-time processor board and a computer set. The real-time processor board comprises a Texas Instrument TMS 320F240 DSP microcontroller and the I/O is expanded through a number of different socket connections, i.e. sixteen channels with 16-bits A2D, eight channels with 16-bits D2A, a Slave DSP Bit I/O Unit. dSPACE™ integrated software, the ControlDesk, allows real time implementation, parameter setting, displaying I/O in graphical format, on-board control and data acquisition.

Matlab™, Simulink™ and RealTime Toolbox software are employed to design and implement both the data acquisition system and control system on the dedicated dSPACE™ controller board. The software also allows the design to be simulated within the Matlab/Simulink environment prior to implementation. The compatibility of dSPACE™ with Matlab/Simulink means that the Maltab/Simulink model can be directly implemented on the hardware. This feature offers a speedy implementation for real-time application on the diesel engine test-bed.

The final design of Matlab/Simulink model is converted into C code using a specific function in the Matlab/RealTime Toolbox. This C code is further cross-compiled into an executable file to be run on the Texas Instrument microcontroller. This microcontroller is linked to the dSPACE hardware interface which is connected to the transducers on the diesel engine. A full description of the data acquisition procedure can be found in Appendix B.5.

Other sub-systems that are used for acquiring data are sensors, frequency-to-voltage converter instruments, oscilloscopes, exhaust gas meter, electronic fuel flow meter and air-flow meter.

2.9.3 Equipment Calibration

Prior to the benchmarking process, equipment calibrations were carried out for the fuel flow meter, exhaust gas meter and air flow meter. These devices measure extra engine parameters for observation and controller performance comparison such as the consumption of diesel, the opacity of the exhaust gas and the air drawn into the engine. Details of the equipment setup for calibration are given in Appendix C.

2.10 Diagnostic Technique and Control System Technique

Sections 2.4 and 2.6 have described various techniques for the problem of fault diagnosis and speed control respectively. This section gives a brief introduction to the theoretical and methods used in the research

2.10.1 Engine Cooling System Fault Diagnostic Technique

For the problem of engine cooling system fault diagnosis, a model based approach is considered. Since the cooling system is non-linear in nature, a non-linear model is considered for the design. The diagnostic technique employs a sliding mode observer to estimate the system states and re-construct the parameters which have physical meaning. The residuals are generated from estimated parameter values and output from a non-linear fuzzy based cooling system model.

2.10.2 Control System Technique

For the sliding mode control system design, two approaches are proposed, namely, a model-based approach and model-free approach. The nature of the '*classical*' sliding mode control design is model-based. The design process used the model parameters to compute the control algorithm. Two types of controllers are considered in model-based control design: model following control and integral action control. Both of the controller approaches give speed tracking capability.

In model-free controller design, another class of sliding mode controller is considered. It is called a higher order sliding mode (HOSM) controller.

2.10.3 Model Identification

For the model-based controller design, an engine model for speed control is required. An event-based closed-loop dynamic identification technique is proposed to obtain the model. In order to carry out this identification process, there is a need to break into the speed control loop. Since it is impossible to break into the commercial speed control loop, a proportional and integral (PI) action controller was designed to replace the existing commercial controller. The setup of the PI controller is described in the following section.

2.11 PI Controller, Setup and Implementation

The classical Proportional-Integral-Derivative (PID) controller is the most popular and most widely used industrial controller. Its general properties with regard to effectiveness and simplicity are well recognised [De Santis 1994]. Although the adoption of PID control has spanned more than five decades, it is still extremely popular. More than half of the industrial controllers in use today utilise PID or modified PID control schemes [Ogata 1997]. This control scheme is used in the dynamic identification process and also for the designed controllers comparison. In this research the control scheme only involves the proportional and integral component - PI control. The derivative gain, D, is not considered in the test as the plant noise is undesirably amplified at high frequency causing engine speed fluctuations. This observation was also reported by Clarke [Clarke 1984].

2.11.1 PI Setup and Gain Tuning

The objective of the PI configuration in the closed-loop is to control the speed of the diesel engine to provide a stable condition for the dynamic identification process and the controller benchmarking. The block diagram in Figure 2-7 shows the PI setup on the diesel engine. The identification setup consists of a PI loop in order to control the speed of the engine during the identification process. In order to run the engine in a reasonably stable condition, the controller gains are required to be tuned to a suitable and acceptable level.

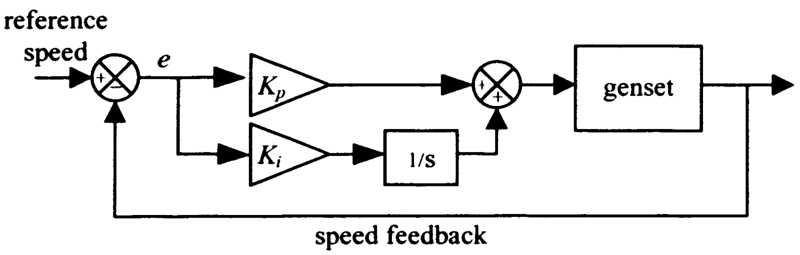


Figure 2-7: PI controller setup

The initial tuning of the PI gain was based on the Ziegler-Nichols method [Ziegler 1942, Ogata 1997]. Ziegler and Nichols suggested two methods of tuning PID controllers. The first method is based on the step response and the second method is based on the value of K_p that results in marginal stability when only the proportional control action is used. The second method was employed in the gain tuning process. By using the second method, the K_p was increased from 0 to a critical value K_{cr} where the system output (engine speed) first exhibited sustained oscillations. The critical gain K_{cr} and the corresponding period P_{cr} were experimentally obtained. P_{cr} was calculated from the engine speed oscillations. Both of these values were used to calculate the controller parameter values, K_p and K_i based on the following formulae.

$$K_p = 0.45K_{cr}, \; K_i = \frac{K_p}{T_i} \quad \text{where } T_i = \frac{1}{1.2}P_{cr}$$

Equation 2-1

However, the obtained K_p and K_i gains did not maintain the engine speed under specific loading conditions. Both of the gains were then fine-tuned to obtain a satisfactory transient response under varying electrical load. A small step load (step-up load) was applied to the generator and the transient response observed. The gains were iteratively tuned.

2.11.2 Integral Anti-windup Considerations

In the PI/PID control implementation, the integral term often leads to a phenomenon known as integral windup which causes long periods of overshoot in the controlled response. Integral windup occurs when the control variable is limited either at its low or high limit effectively breaking the feedback loop [Airikka 2004]. This will cause a continuous integration of the control error forcing the control signal to remain at the limit. Integral anti-windup is used to modify the integral part of the control when the control reaches its limits to prevent integration of the error. For design simplicity, the integrator Simulink block limit is constrained to some fixed value.

2.11.3 Signal Conditioning

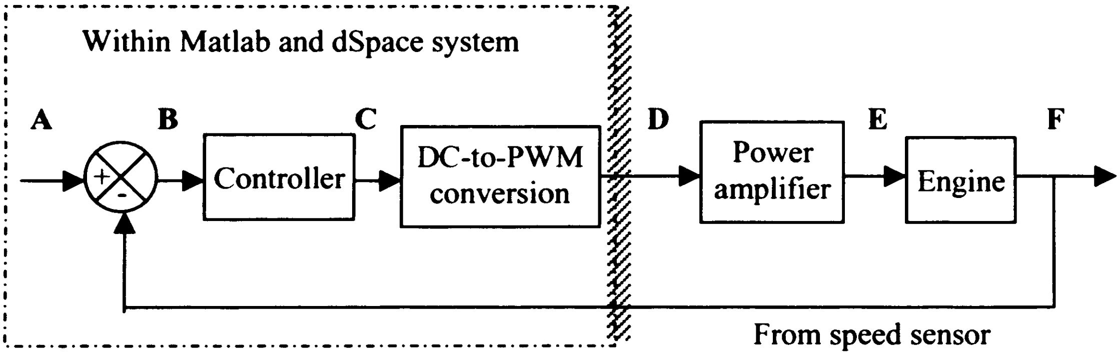
In many speed control systems, actuators are the final-stage devices to perform the control processes. More often than not, the actuators require larger control signal amplitude specification [GAC 1999a 1999b] to drive the actuator. For example, the GAC actuator (ADC100) requires high current to drive an electromagnetic servo. Not many of the PC-based signal-processing cards (i.e. dSPACE DSP board) can supply an adequate level of signal for the actuators. The actuator in use in the research test requires the electrical specification as shown in Table 2-2.

Additional hardware has been designed to amplify both the current and voltage to the demand level. Full details of the voltage and current amplifier design are described in Appendix D. The signal type required by the actuator is PWM. The movement of the actuator shaft depends on the duty cycle (i.e. proportion of time/cycle/period the device is operated) of the PWM signal. The control signal is thus constrained to the limit from 0 to 1 (i.e. 0% to 100% duty cycle). The sum of the d.c. control signal, from P and I components, is converted into a PWM signal via the Simulink Real-Time Toolbox block. The operating frequency is set to 500Hz and the amplitude is at the maximum of the DSP board supply which is 5V. This PWM signal is then fed to the voltage and current amplifier.

Specification	Value
Amplitude	~12 volts (i.e. engine battery)
Frequency	From 320 Hz to 430 Hz.
Current required	1.5 A to 2.7 A (max)
Input signal type	Pulse-width modulation (PWM)

Table 2-2: GAC actuator electrical specification.

Practical actuators have limits on the allowable range of values and often the corresponding variables are scaled in the range 0-100%. The control signal range is 0 to 1 from the Matlab Simulink model. Figure 2-8 shows the control signal flow in the system. The actuator shaft movement is linear with the signal [GAC 1999a]. It is assumed that the fuel pump throttle has a similar linear response to the shaft movement.



A: reference speed (i.e. 1200 rpm, 1350 rpm and 1500 rpm)
B: Speed error (i.e. 0 to 1500 rpm)
C: Control variable value (i.e. from 0 to 1)
D: PWM signal (with magnitude of 5V, 500Hz)
The duty cycle of PWM signal (0-100%) is proportional to the control variable value (0 to 1)
E: PWM signal (with magnitude of 12V, 500Hz)
The duty cycle of PWM signal (0-100%) is proportional to the control variable value (0 to 1)
F: measure speed (i.e. 0 to 1500 rpm)

Figure 2-8: Control signal flow.

2.11.4 Starting of the Engine

A 'gain-scheduling' method is applied for the engine startup phase and running phase. There are two sets of PI gain used to run the test.

During the startup phase, the speed error tends to be the largest. In order to bring the engine speed to its speed set point in reasonably fast time with no excessive overshoot, the proportional gain is normally set to a larger value than the value running at steady state situation. However, the integral component integrates the large speed error of 1500 rpm and generates a large integral control value. This integral control value has a large impact on the engine speed when the speed almost reaches the set point value. The engine speed increases at a fast rate and eventually overshoots by a large amount. On many occasions, this overshoot hits the engine overshoot safety limit and causes the engine to stall. Thus, the integral gain is set to zero during engine startup.

When the engine speed reaches the set point, the proportional gain is adjusted to a lower value to avoid speed oscillation effects. The integral component is included during the normal operating condition to counter the steady state error. Thus, a gain scheduling type of PI control setup is employed in the setup. The gain scheduling method is not new and can be found in different applications [Jimenez-Garcia 2000, Miao 2002]. The method enables a controller to respond rapidly to changing operating conditions. It is important that the selected scheduling variables (in this case the engine speed) reflect changes in plant dynamics as operating conditions change (i.e. from the startup of the engine to the normal operating point).

2.12 Conclusion

This chapter has introduced the diesel engine and its cooling system. Several cooling system faults have been discussed. The fault diagnosis techniques to be used for the cooling system problem have been clearly stated and selection of the sensors has been discussed.

For the diesel engine control system design, the control techniques to be applied have been proposed:

Model-based : (a) Model-following control with speed tracking

(b) Integral action control with speed tracking

Model-free : Higher order sliding mode control

A list of control criteria has been presented for use in the controller performance comparison. Equipment calibration has been described and the detail of the work can be found in Appendix C. The setup and tuning of the PI controller has been described. Signal conditioning on the control signal has been presented. The next Chapter will present the sliding mode concept and its properties. This technique is central to the monitoring and control strategies to be used in the thesis.

Chapter 3

3 Sliding Mode System (Variable Structure System)

This chapter concentrates on the sliding mode concept and technique. This method is used here because it is robust against system uncertainty. A brief history of the sliding mode concept is given. Sliding modes have several attractive properties and these properties are described in detail and demonstrated using a simple model. The stability issues of sliding mode control are also discussed.

The organisation of this Chapter is as follows: Section 3.1 presents the historical background of VSS and higher order sliding mode (HOSM) techniques. The VSS concept is described in detail in Section 3.2 with examples for illustration. The reachability condition required to maintain the sliding motion is discussed in Section 3.3. Section 3.4 highlights the properties of the sliding motion. The possible solution to chattering in sliding mode design is discussed in Section 3.5. Section 3.6 discusses the equivalent control concept of the sliding mode. Section 3.7 shows the application of state space techniques to sliding mode design and Section 3.8 describes the design approach. Section 3.9 highlights the HOSM concept. The need for an observer for state estimation is motivated and sliding mode observer design is discussed in Section 3.10.

3.1 History

The sliding mode concept is based on the notion of a variable structure system (VSS) which emerged in the late 1950s in Russia. Much of the early literature is in Russian. The concept became known to the wider world when some of the literature was translated into English [Utkin 1977, Utkin 1978]. The 1977 survey paper entitled '*Variable structure systems with sliding modes*' is one of the earliest papers published in English. One of the earliest books on VSS in English was published in the following year.

Since this wider publication, the concept has received a great deal of attention across the world. [Zinober 1990, Zinober 1994, Spurgeon 1993, Edwards 1999, Bartolini 1995, DeCarlo 1988]. The literature has shown that the application of the sliding mode concept covers both control system design and fault diagnosis. The advantageous properties of sliding modes have been the motivation and driving force in the research.

In the 1980s, the *higher order sliding mode* (HOSM) concept emerged as a more general class of sliding mode systems. The standard sliding mode control is seen to be a sub-class of HOSM controllers. The HOSM technique has brought a solution to the chattering problem which is inherent in standard sliding mode design. Research work in this area is presented in [Levant 1993 1997 2001], [Bartolini 2000] and [Emel'yanov 1996].

3.2 Variable Structure Systems

The variable structure systems design is different from the linear state controller design. Consider a linear model as follows:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

Equation 3-1

In the linear state feedback controller design, the structure of the state feedback is fixed as

$$u(t) = -Kx(t)$$

Equation 3-2

where the constant parameters are chosen according to a given design procedure, such as eigenvalue placement. In VSS, the control is allowed to change its structure and switch at any instant from one to another set of usually continuous functions of the state. With proper selection of the parameters for each control structure and an appropriately defined switching logic, an asymptotically stable system may be obtained. Two examples will be described to demonstrate the advantages of changing control structure during the control phase.

3.2.1 First Example of a Variable Structure System

In the first example, consider a second-order system

$$\ddot{x}(t) = u(t)$$

Equation 3-3

Suppose the control law $u(t)$ is chosen as

$$u = \begin{cases} -\alpha_{vss} x(t) \\ \alpha_{vss} x(t) \end{cases}$$

Equation 3-4

where $\alpha_{vss} > 0$. The trajectories of the first composite system are ellipses and in the second system, they form a saddle. Figure 3-1 (a) and (b) show the respective phase plane plots. From Figure 3-1 (b), there is a unique stable subspace defined by the equation

$$\dot{x}(t) + \sqrt{\alpha_{vss}}\, x(t) = 0$$

Divide the phase plane into four regions as shown in Figure 3-1 (c). The control is selected as,

$$u(t) = \begin{cases} -\alpha_{vss} x(t) & \text{in regions II} \\ \alpha_{vss} \mathrm{x}(t) & \text{in regions I} \end{cases}$$

The trajectories take the form shown in Figure 3-1 (c) and the system is globally asymptotically stable. It is possible to choose for the system to switch on a line other than $\dot{x}(t) + \sqrt{\alpha_{vss}}\, x(t) = 0$, i.e.

$$\dot{x}(t) + cx(t) = 0$$

Equation 3-5

where $0 < c < \sqrt{\alpha_{vss}}$.

The corresponding trajectories are shown in Figure 3-2 (a). To understand the behaviour of the system on the switching line, two switching lines are introduced as shown in Figure 3-2 (b). The system with elliptical trajectories switches on the line l_1 with $c = c_1$ and the 'saddle' system switches on the line l_2 with $c = c_2 (c_1 > c_2)$.

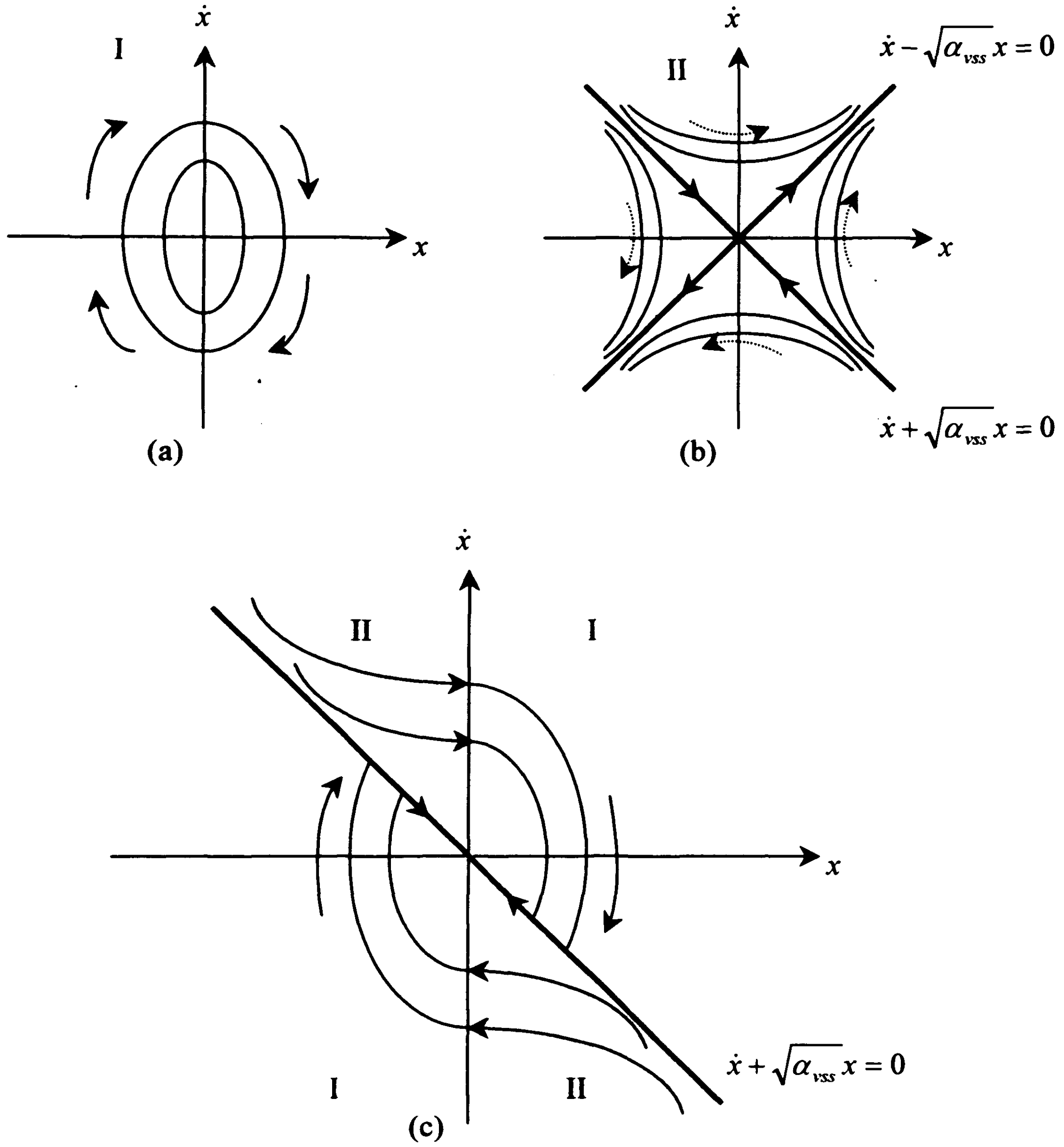


Figure 3-1: Asymptotically stable variable structure system consisting of two unstable structures.

The system trajectory is an ellipse until it hits the line l_1 at which point it switches to the saddle trajectory and is directed to the line l_2. Another switch occurs on l_2 and the trajectory again becomes elliptical. The result of the switching is that the trajectory will remain between the lines l_1 and l_2 and zigzag to the origin. If $c_1 = c_2 = c$, then the trajectory switches with infinite frequency on the line $\dot{x}(t) + cx(t) = 0$. The trajectories are forced to remain on the switching line. Thus, the system moves stably down the switching line to the origin and this is called the sliding mode or sliding motion.

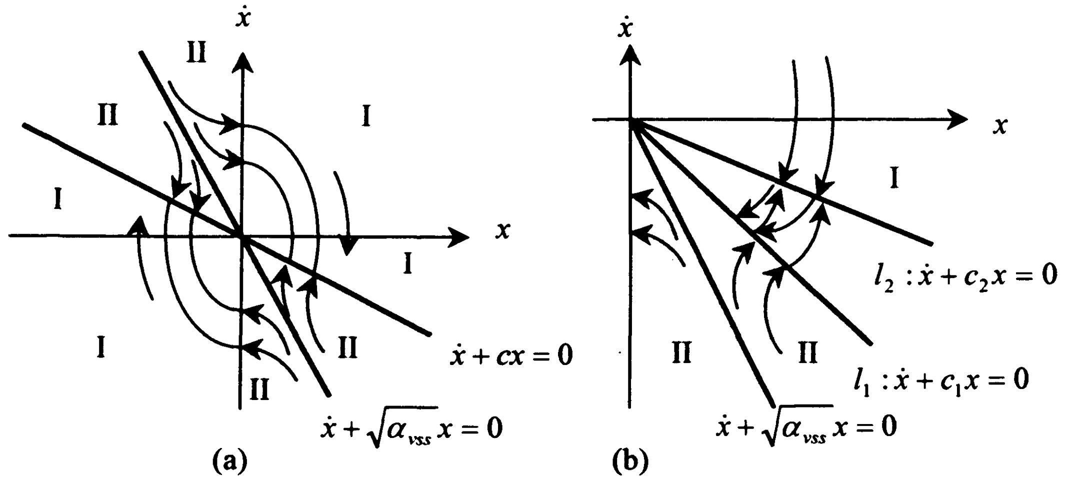


Figure 3-2: Sliding motion.

Equation 3-6

From the above example, the following summary is made. Both the structures in equation 3-4 are unstable as shown in Figure 3-1 (a) and (b). However, with a proper switching logic and appropriate switching line, asymptotic stability can be achieved.

3.2.2 Second Example of a Variable Structure Systems

In the second example, a similar second order system is used

$$\ddot{x}(t)=u(t)$$
$$u(t)=-\varpi x(t)$$

Equation 3-7

where $\varpi=\alpha_1^2$ and $\varpi=\alpha_2^2$, $\alpha_1^2>1>\alpha_2^2$. The phase portraits of both structures are shown in Figure 3-3 (a) and (b). Asymptotic stability is achieved if the following switching structure is applied.

$$\varpi=\begin{cases}\alpha_1^2 & \text{if } x\dot{x}>0\\ \alpha_2^2 & \text{if } x\dot{x}<0\end{cases}$$

Equation 3-8

The resulting trajectory is shown in Figure 3-3 (c). The structure of the system is seen changing on the coordinate axis.

Each time switching occurs on the axis, the distance from the origin to the trajectory decreases. Eventually, the system trajectory will spiral in toward the origin and result in an asymptotically stable motion.

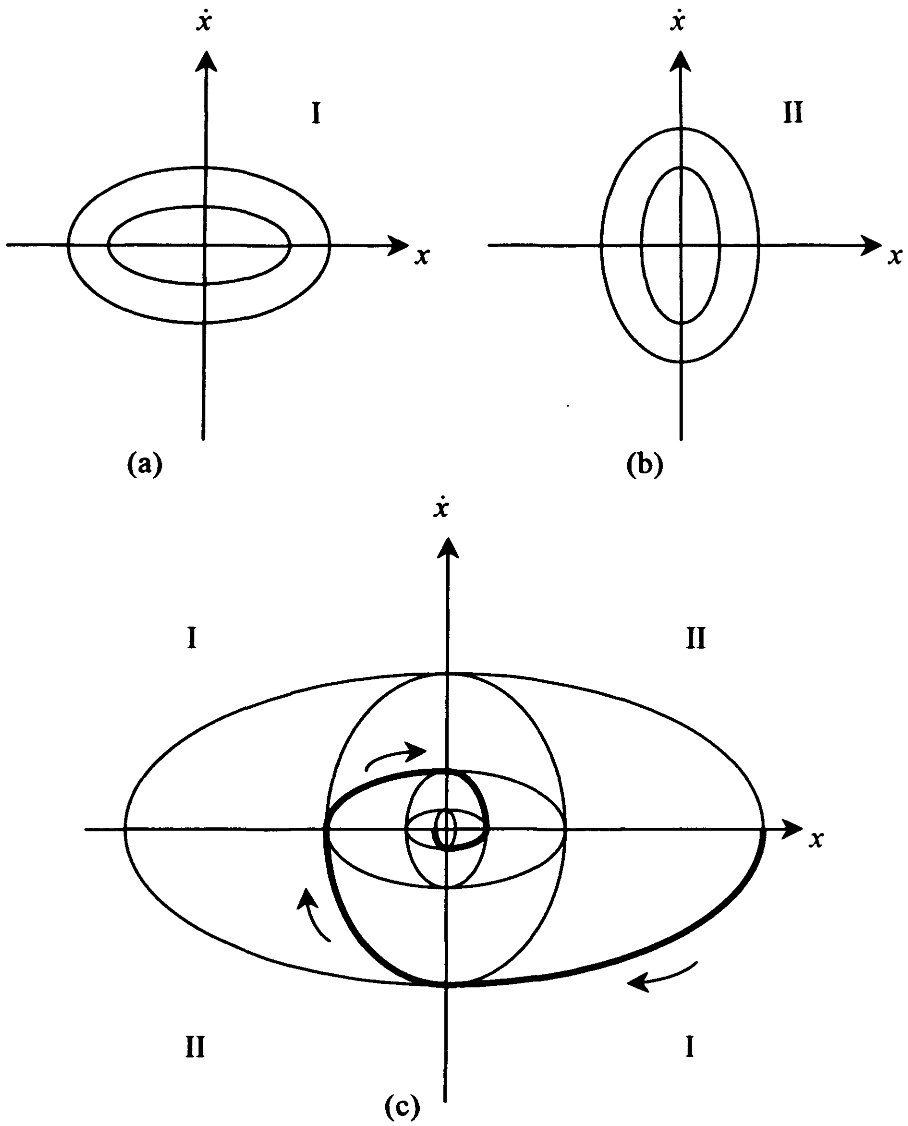


Figure 3-3: Asymptotically stable variable structure system consisting of two unstable structures.

From the above two examples, one can summarise that the VSC comprises a switching function, a feedback control law and a decision rule. This decision rule is designed to force the system states to reach and remain on the switching surface within the state space. The dynamic behaviour of the system when confined to the surface is described as an *ideal sliding motion* [Edwards 1998]. The concept is a non-linear control technique as the control is non-linear in nature.

3.2.3 Behaviour of Variable Structure Systems

To demonstrate the sliding mode behaviour, consider the system in equation 3-3 with the control law as follows:-

$$u(t) = \begin{cases} -1 & \text{if } s(x(t), \dot{x}(t)) > 0 \\ 1 & \text{if } s(x(t), \dot{x}(t)) < 0 \end{cases} \quad \text{or} \qquad u(t) = -\operatorname{sgn}(s)$$

Equation 3-9

where sgn is the signum function which has the property of $\text{sgn}(s)s = |s|$. s is normally chosen as a function of the system states, i.e.

$$s(x(t)) = Sx(t)$$

where s defines a manifold (the so called sliding surface) of the system. There are two phases in controller design, namely design of the switching function and choice of control to ensure the system trajectories are attracted to and remain on the switching surface (i.e. reachability). When a sliding motion occurs, $Sx(t) = 0$. The choice of sliding manifold depends on S which will determine the system dynamics in the sliding mode. To illustrate a specific behaviour, consider $s(x(t), \dot{x}(t)) = cx(t) + \dot{x}(t)$, where c is a positive design scalar. The phase trajectory of this system is shown in Figure 3-4. It is seen that the choice of switching function gives a first order dynamic in the sliding mode that is entirely defined by the choice of switching function.

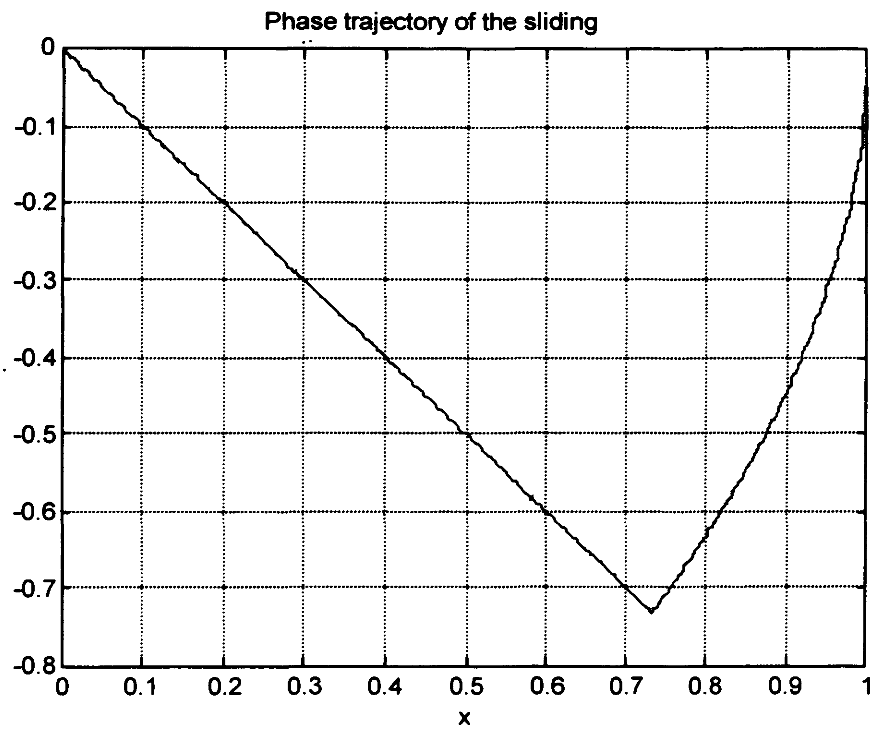


Figure 3-4: Phase trajectory of the system.

The condition to ensure the sliding mode is attained will now be explored. Consider

$$s\dot{s} = s(c\dot{x}(t) + \ddot{x}(t)) = s(c\dot{x}(t) - \text{sgn}(s)) < |s|(c|\dot{x}(t)| - 1) < 0$$

when $c|\dot{x}(t)| < 1$, so that $s\dot{s} < 0$. The trajectory will switch along the line $s = 0$ at high frequency, moving toward the origin and sliding motion will occur. One important topic in sliding mode control design is the *reachability condition* which is the condition to keep the system in a sliding motion. The following section will describe the condition in more detail.

3.3 Reachability Condition

One of the prime tasks of designing the sliding mode controller is to keep the sliding variable on the sliding surface and hence maintain a sliding motion. This is achieved by setting a condition on the control law under which the system will find the sliding surface attractive. This condition is called the *reachability condition.* The usual approach is to perform quadratic stability analysis using a Lyapunov function as follows:

$$V(s) = \frac{1}{2}s^2$$

Equation 3-10

where s is the sliding variable. It follows that

$$\dot{V}(s) = \dot{s}s$$

The sliding variable, s, will only decay to zero under the following condition:

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

Equation 3-11

The reachability condition in equation 3-11 can be expressed by a non-linear (i.e. discontinuous) and a linear terms as follows:

$$\dot{s} = -\rho \operatorname{sgn}(s) \text{ and } \dot{s} = -h_L s$$

Equation 3-12

The discontinuous term will introduce chattering in the control action. These reachability conditions will make sure that $s\dot{s} < 0$ and attain a sliding motion at least asymtotically. The rate of decay is controlled by the choice of h_L in the second case. For the purpose of illustration, consider the second order system in equation 3-3 and re-write the system as follows:

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = u(t)$$

Equation 3-13

where $u(t)$ is the system input. The sliding surface is defined as:

$$s(t) = cx_1(t) + x_2(t)$$

Equation 3-14

where c is a positive scalar which determines the dynamics during sliding motion. The reachability condition is selected as:

$$\dot{s} = -\rho \operatorname{sgn}(s(t)) - h_L s(t)$$

Equation 3-15

where ρ and h_L are positive scalars. By differentiating equation 3-14

$$\dot{s}(t) = cx_2(t) + u(t)$$

Equation 3-16

From equation 3-15 and equation 3-16, the control law, $u(t)$ is thus given as:

$$u(t) = -cx_2(t) - \rho \operatorname{sgn}(s(t)) - h_L s(t)$$

Equation 3-17

Where the reachability condition yields,

$$\begin{aligned} s\dot{s} &= s[-\rho \operatorname{sgn}(s(t)) - h_L s(t)] \\ &= -\rho|s(t)| - h_L[s(t)]^2 \\ &\leq -|s(t)|[\rho + h_L|s(t)|] \end{aligned}$$

Equation 3-18

The inequality in equation 3-11 is true if ρ and h_L are strictly positive. Figure 3-5 shows the phase portrait. The parameters used are

$c = 1$, $\rho = 1$ and $h_L = 1$

The initial conditions of the system are given by $x_1(0) = 1, x_2(0) = 0$ and the system satisfies the inequality of equation 3-11. Figure 3-5 clearly shows the two stage nature of the dynamics of the sliding motion. During the first stage, the controller drives the system onto the sliding surface (i.e. the dotted line). This is known as the reaching phase. The system hits the sliding surface at the point S and the sliding phase take place. After reaching the sliding surface, the system stays there and moves towards the origin.

However, when the reaching condition in equation 3-18 is violated, with $c = 1$, $\rho = 1, h_L = -1.001$, the sliding surface is no longer attractive to the sliding variable. The system does not satisfy equation 3-11 and will not stay on the sliding surface. By keeping the same initial conditions, Figure 3-6 shows the system moving away from the sliding surface.

3.4 Properties of the sliding motion

The variable structure control induces a sliding motion when the system trajectories hit the pre-defined sliding surface if an appropriate reachability problem is solved. The resulting sliding motion has several advantages. The next sub-sections will highlight these properties.

3.4.1 Reduced Order System

Firstly, there is a reduction in system order during sliding. By referring back to the second order system in equation 3-3 and the switching function as $s(x(t),\dot{x}(t)) = cx(t) + \dot{x}(t)$, it is seen that when the system trajectories hit and remain on the sliding surface, the switching function becomes $s(x) = 0$. The dynamics become

$$cx(t) + \dot{x}(t) = 0$$
$$\dot{x}(t) = -cx(t)$$

Equation 3-19

The system now is of a lower order than the original system and the motion is independent of the control when sliding motion occurs.

3.4.2 Uncertainty Rejection

Another important property of sliding motion is the rejection of a class of uncertainties and disturbances. Consider the system in equation 3-3 with the additional term $-a\sin x(t)$ where a is a positive scalar:

$$\ddot{x}(t) = -a\sin x(t) + u(t)$$

Equation 3-20

The same switching function is used for this demonstration. By setting $a = 0.2$ and with initial conditions $\dot{x}(0) = 0,\ x(0) = 1$, the phase plane plot (i.e. the dotted line) is obtained as in Figure 3-7. A comparison plot (inner curve line) is obtained when $a = 0$. The plot shows that, once sliding motion occurs, this new system behaves as,

$$\dot{x}(t) = -cx(t)$$

The additional term is a disturbance or uncertainty in the original system. When sliding motion takes place, this term is completely rejected. The system is said to be robust. It is insensitive to some mismatches between the model used for control law design and the plant on which it will be implemented.

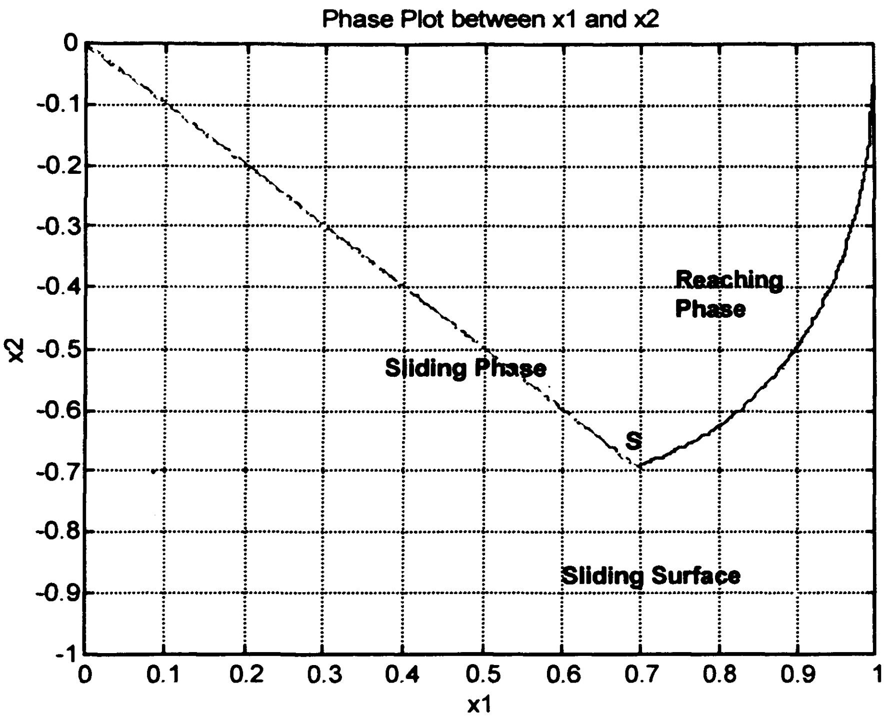


Figure 3-5: Phase portrait of a sliding motion with $\rho = 1$ and $h_L = 1$

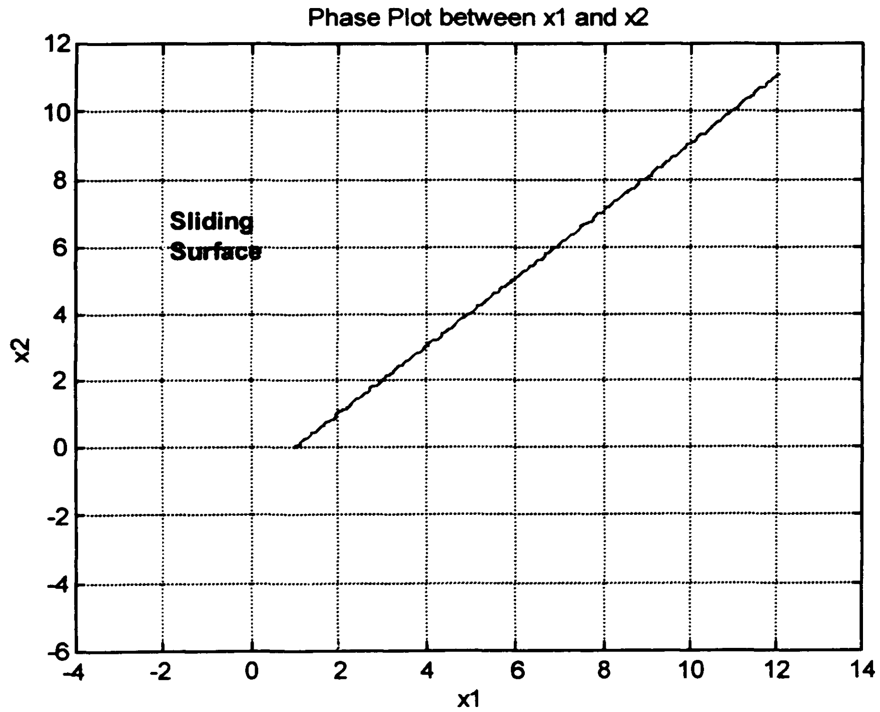


Figure 3-6: Phase portrait of a sliding motion with $\rho = 1$ and $h_L = -1.001$

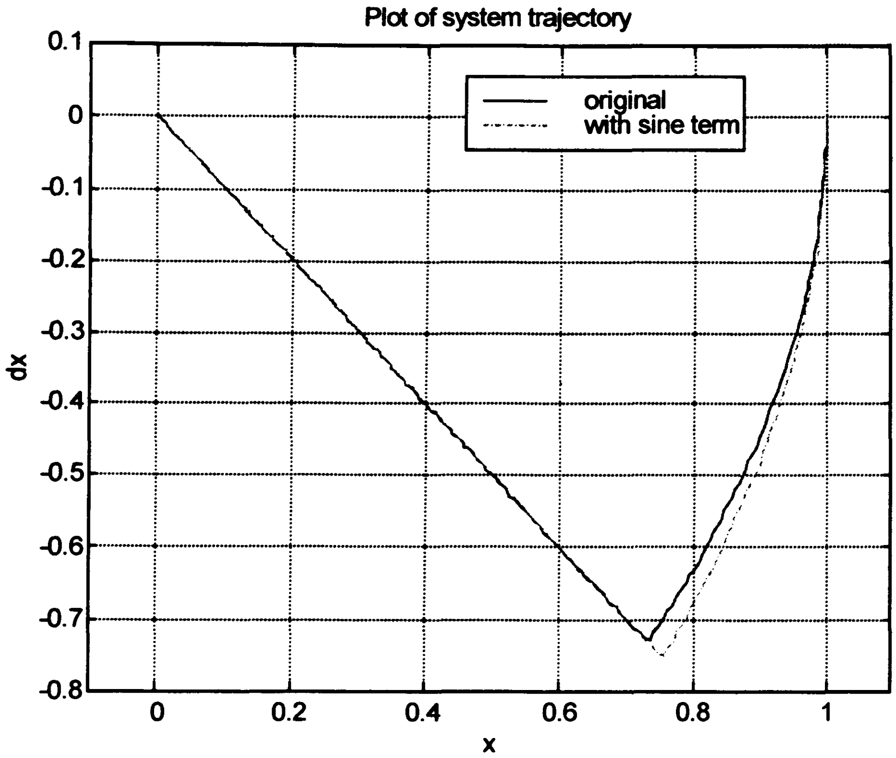


Figure 3-7: System trajectories.

The reason for this uncertainty rejection can be explained by the equivalent control concept which will now be introduced. Once a sliding mode has been attained, $s(t) = 0$ and $\dot{s}(t) = 0$ for all subsequent time. The switching function, $s(x(t), \dot{x}(t)) = cx(t) + \dot{x}(t)$ yields

$$\dot{s}(t) = c\dot{x}(t) - a \sin x(t) + u(t)$$

Equation 3-21

The equivalent control is obtained by equating the expression in equation 3-21 to zero which yields

$$u_{eq}(t) = -c\dot{x}(t) + a \sin x(t)$$

Equation 3-22

With the effective control defined in equation 3-22, the uncertainty term ($a \sin x(t)$) in the system is cancelled.

3.5 Chattering in Sliding Modes

The sliding mode occurs because the control structure switches from one value to another infinitely fast. The plot of the control signal for the system with the control structure in equation 3-9 is shown in the upper subplot of Figure 3-8. The sliding motion starts after 0.73 seconds when the high frequency switching occurs. The respective switching function graph is shown in the lower subplot of Figure 3-8 and it can be seen that $s(x,\dot{x}) = 0$ after 0.73 seconds.

In practical systems, it is often impossible or impractical to achieve the high frequency switching that is necessary for such VSC designs. There are several reasons for this. One cause is the presence of finite time delays for control computation. The second cause is the limitation of physical actuators. Since it is impossible to switch the control at an infinite rate, chattering always occurs in the sliding motion. Since chattering is often undesirable, there is a need to eliminate or reduce the effects. Several methods are described below.

3.5.1 Pseudo Sliding Method

A natural solution to the chattering effect is to smooth the discontinuous control action. To obtain the smooth action, consider equation 3-9. The property ($\mathrm{sgn}(s)s = |s|$) yields

$$u(t) = -\mathrm{sgn}(s(t)) = -\frac{s}{|s|}$$

Equation 3-23

The control law is seen to be relay-like in nature. The ideal relay characteristic is practically impossible to implement. In order to smooth the control action, a small positive scalar, δ is added to the denominator of the control law to yield

$$u(t) = -\frac{s}{|s| + \delta}$$

Equation 3-24

The variable, δ is used to trade off ideal sliding mode performance against a smooth control action. To illustrate the smooth control action, consider the system in equation 3-3 again with the control action in equation 3-24. Set the initial condition to be $x(0) = 1, \dot{x}(0) = 0$ and $\delta = 0.01$. The resulting smooth control signal is shown in the upper subplot of Figure 3-9. With this smooth control action, the ideal sliding mode no longer occurs and after 0.73 seconds, the trajectories only stay close to $s = 0$.

3.5.2 Low Pass Filter Method

Another reported method to smooth the discontinuous control signal in the literature is by employing a low pass filter to the control signal to filter out the high frequency component and retain the low frequency component. For example,.

$$\beta \dot{u}_{low}(t) + u_{low}(t) = u(t)$$

Equation 3-25

where $u_{low}(t)$ is a low frequency component in the control signal. For simplicity, a low pass transfer function can be obtained by applying the Laplace transform to equation 3-25 to yield.

$$u_{low}(t) = \frac{1}{1+s\beta}u(t)$$

Equation 3-26

The resulting effect on the control signal is shown in lower subplot of Figure 3-9.

3.5.3 Tuning Reaching Law Method

The chattering effect can also be reduced by tuning the reaching law parameters. i.e. parameters ρ and h_L in equation 3-15. Consider the system of equations 3-13, 3-15 and 3-17.

From the original setting of $\rho = 1$ and $h_L = 1$, the control signal exhibits chattering behaviour as shown in the upper subplot of Figure 3-10. When these parameters are tuned to $\rho = 0.001$ and $h_L = 2$, a reduced chattering effect is obtained. The lower subplot of Figure 3-10 shows this reduction effect. By selecting the gain ρ small, the amplitude of the chatter will be reduced. However, ρ cannot be chosen equal to zero because the reaching time would become infinite. A large value for h_L increases the reaching rate when the state is not near the switching surface. By this method, chattering can practically be suppressed.

In summary, chattering is a hurdle in the use of VSC in practical control systems. Elimination or reduction of chattering remains an important problem.

3.6 Equivalent Control, $u_{eq}(t)$

The equivalent control is effectively the average value of $u(t)$ which maintains the state on the switching surfaces $s(x) = 0$. The actual control $u(t)$ consists of a low-frequency (average) component, $u_{eq}(t)$ or $u_{low}(t)$ as mentioned above, and a high frequency (chatter) component.

To demonstrate this equivalent control signal, consider the system in equation 3-3 and switching function, $s(x(t),\dot{x}(t)) = cx(t) + \dot{x}(t)$.

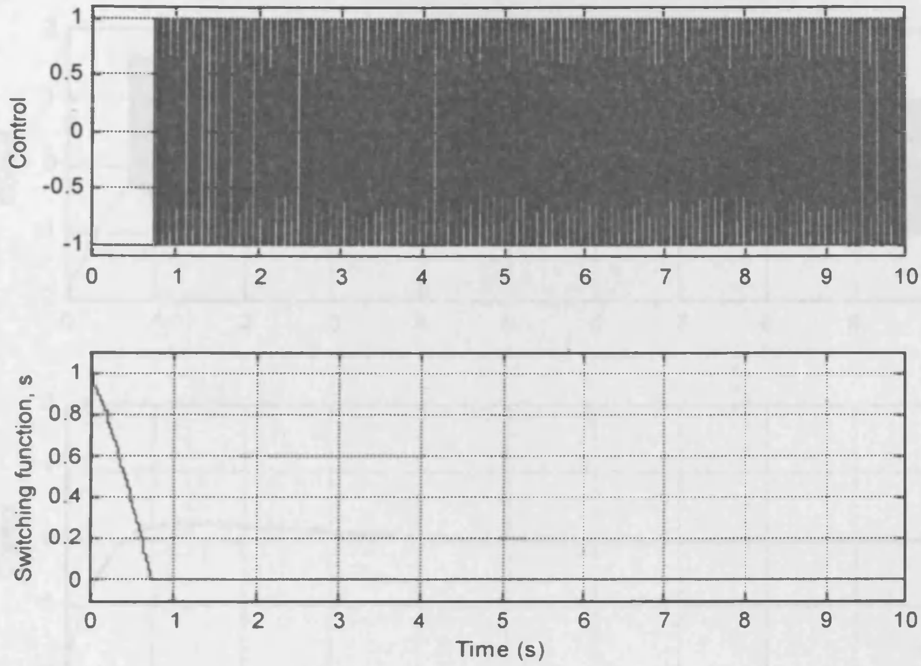


Figure 3-8: Discontinuous control action and switching function.

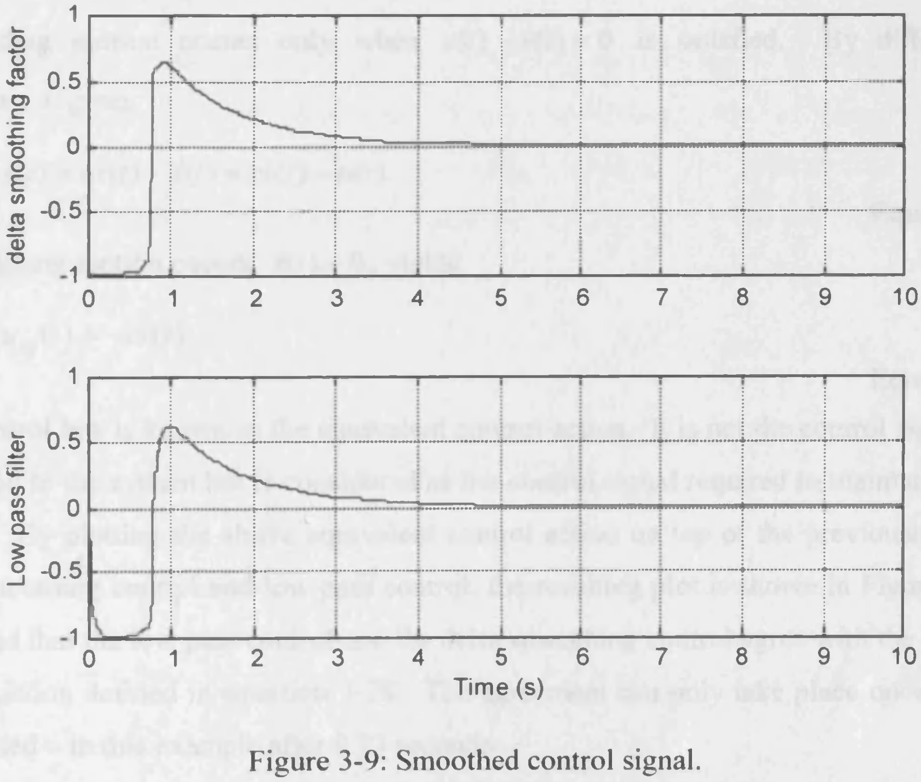


Figure 3-9: Smoothed control signal.

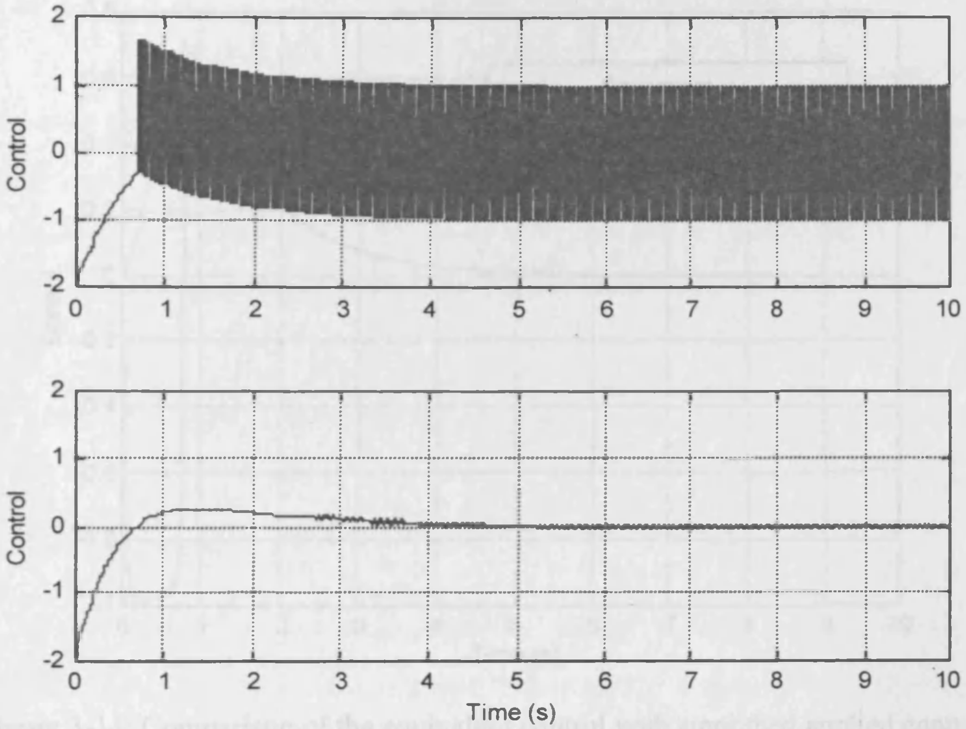


Figure 3-10: Smoothed control signal using 'tuning reaching law approach'.

The sliding motion occurs only when $s(t) = \dot{s}(t) = 0$ is satisfied. By differentiating $s(t) = cx + \dot{x}$ gives

$$\dot{s}(t) = c\dot{x}(t) + \ddot{x}(t) = c\dot{x}(t) + u(t)$$

Equation 3-27

When sliding motion occurs, $\dot{s}(t) = 0$, yields

$$u_{eq}(t) = -c\dot{x}(t)$$

Equation 3-28

This control law is known as the equivalent control action. It is not the control signal which is applied to the system but is considered as the control signal required to maintain a sliding motion. By plotting the above equivalent control action on top of the previously obtained delta smoothing control and low pass control, the resulting plot is shown in Figure 3-11. It is noticed that the low pass control and the delta smoothing control agree with the equivalent control action defined in equation 3-28. The agreement can only take place once sliding is established – in this example after 0.73 seconds.

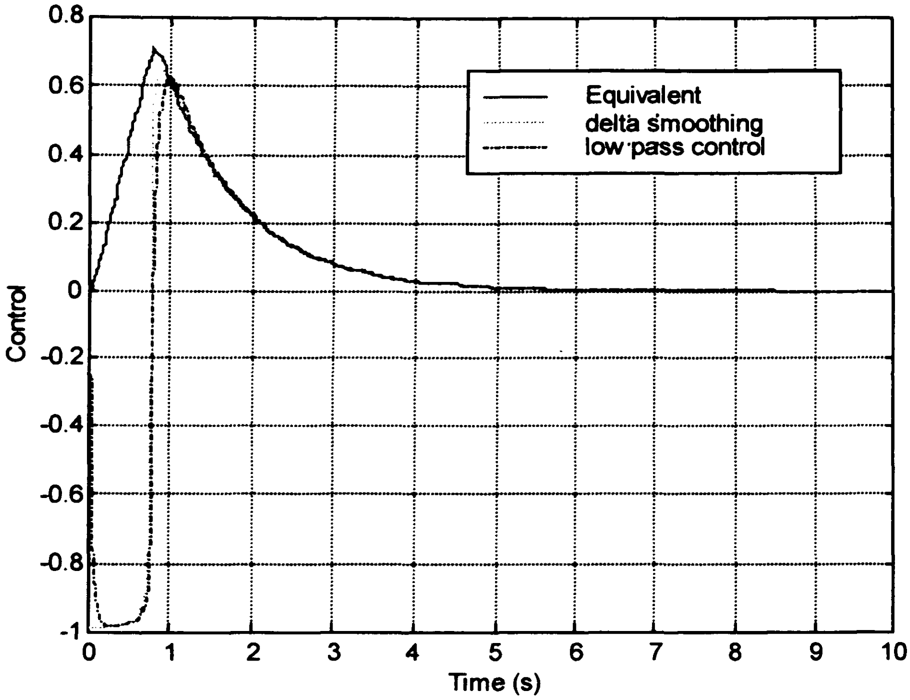


Figure 3-11: Comparison of the equivalent control with smoothed applied control.

3.7 A State Space Approach

For the examples in the preceding sections, the analysis is based on second order systems and phase portraits are used. However, for a multivariable system of high order, a more general framework must be obtained for analysis purposes. The state space approach is a solution to this. Consider the system in equations 3-3 and 3-13. The following state vector is introduced.

$$x(t) = \begin{bmatrix} x_1(t) \\ \dot{x}_1(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Equation 3-29

The equation 3-29 can be written in state space form as

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t)$$

Equation 3-30

The switching function, $s(x(t), \dot{x}(t)) = cx(t) + \dot{x}(t)$, can be expressed in matrix form as

$$s(x_1(t), x_2(t)) = Sx(t)$$

Equation 3-31

where

$$S = \begin{bmatrix} c & 1 \end{bmatrix}$$

Equation 3-32

For the system in equation 3-20, the state space representation is as follows:

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) - \begin{bmatrix} 0 \\ 1 \end{bmatrix} a \sin x_1(t)$$

Equation 3-33

It can be seen that the uncertainty, $a \sin x_1(t)$ acts in the input channel. (i.e. the second differential equation). The uncertainty which is acting in the input channels is called matched uncertainty. Uncertainty which does not act in the input channels is referred to as *unmatched uncertainty*. It will be seen that sliding systems are totally insensitive to matched uncertainty.

For the rest of the thesis, the following uncertain linear time invariant system will be considered:

$$\dot{x}(t) = Ax(t) + Bu(t) + f(t, x(t), u(t))$$

Equation 3-34

where $A \in \Re^{nxn}$, $B \in \Re^{nxm}$, $f(\cdot)$ is an unknown bounded function and B is full rank. The system is assume to have m inputs in general. The specific system involved in this research is a single input and single output (SISO) system; the value of m is 1. The general form for the switching function is

$$\begin{aligned} &s(x(t)) = Sx(t) \\ &\ell = \left\{ (x(t)) \in \Re^n : \; s(x(t)) = 0 \right\} \end{aligned}$$

Equation 3-35

where $S \in \Re^{mxn}$.

3.8 Sliding Mode Design Approach

To illustrate sliding mode design in a general framework, consider the system in equation 3-34. Without loss of generality, the system can be transformed into a suitable canonical form via 'QR' decomposition on the input distribution matrix B. There exists an orthogonal matrix T_r which satisfies the following equation.

$$T_r B = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}$$

Equation 3-36

where $B_2 \in \Re^{mxm}$. From the computational point of view, it is more convenient to deal with orthogonal matrices, since the inverse can be obtained by straight forward transposition. If the states are partitioned as

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Equation 3-37

where $x_1 \in \Re^{n-m}$ and $x_2 \in \Re^m$, the nominal linear system becomes

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t)$$

Equation 3-38

$$\dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t)$$

Equation 3-39

With this transformation, the system is decomposed into two connected subsystems, one acting in the null space, $N(S)$ and the other acting in the range space, $R(B)$. The control signal is now acting in the lower subsystem. When a sliding mode takes place, the model order will be reduced and this lower part of the model equation will have no effect on the performance. The representation of equation 3-38 and equation 3-39 is referred to as *regular form*. The resulting switching function matrix and switching function becomes

$$S = [S_1 \quad S_2]$$
$$s(x_1(t), x_2(t)) = S_1x_1(t) + S_2x_2(t)$$

Equation 3-40

where $S_1 \in \Re^{mx(n-m)}$ and $S_2 \in \Re^{mxm}$ is non-singular. Then,

$$\det(SB) = \det(S_2B_2) = \det(S_2)\det(B_2)$$

It follows that SB is full rank. During the ideal sliding mode, the switching function becomes

$$S_1x_1(t) + S_2x_2(t) = 0$$
$$x_2(t) = -Mx_1(t)$$

Equation 3-41

where $M \in \Re^{mx(n-m)}$ is defined to be

$$M = S_2^{-1}S_1$$

Equation 3-42

Substituting equation 3-41 into equation 3-38 yields

$$\dot{x}_1(t) = (A_{11} - A_{12}M)x_1(t)$$

Equation 3-43

It can be seen that the matrix S_2 has no direct effect on the dynamics of the sliding motion and acts only as a scaling factor for the switching function. The matrix $\overline{A}_{11} = (A_{11} - A_{12}M)$ must have stable eigenvalues. The hyperplane design problem can be solved by choosing a state feedback matrix M to fulfill the required performance of the reduced order system (A_{11}, A_{12}). This can be achieved if the equivalent system (A_{11}, A_{12}) is controllable and this equivalent system is controllable if and only if the pair (A, B) is controllable [Edwards 1998].

From equation 3-42, the switching function matrix can be written as

$$S_2 M = S_1$$

Equation 3-44

In order to determine the hyperplane matrix S from M, the second element of S is set to an identity matrix (i.e. $S_2 = I_m$). The resulting hyperplane matrix becomes

$$ST_r^{-1} = [M \quad I_m]$$

Equation 3-45

This solution to S minimises the calculation in proceeding from M to S. There are various methods to solve for the state feedback matrix M, namely robust pole assignment, linear quadratic approach etc [Edwards 1998]. Robust pole assignment is employed in the controller design in the later chapter.

3.9 Higher Order Sliding Modes (HOSM)

The above sliding mode control design considers the so-called first order sliding mode. This method is now referred to as '*classical*' sliding mode control for the rest of this thesis. There is a wider class of sliding mode control called *higher order sliding mode* control. This technique has received a great deal of attention recently in the control community [Levant 1993 1997 2001, Bartolini 2000, Emel'yanov 1996].

HOSM control is said to be a generalisation of classical sliding mode control. An r-th order sliding (also know as r-sliding) is determined by the following equalities,

$$s = \dot{s} = \ddot{s} = \ldots = s^{(r-1)} = 0$$

Equation 3-46

This forms an r-dimensional condition on the state of the dynamic system. It can be seen that the discontinuous control of the HOSM is acting on the higher order time derivative $s^{(r)} = 0$. The classical sliding mode control is considered to be a special case which is the subset of HOSM control when $r = 1$. While classical sliding modes feature finite-time convergence, convergence to a HOSM may be asymptotic [Levant 2001]. Levant and Bartolini have given examples of 2-sliding modes attracting and converging in finite time [Levant 1993, Bartolini 1998a].

The HOSM control has the potential to provide greater accuracy and to eliminate the chattering problems inherent in many classical sliding mode control configurations as previously mentioned. It is known that in practice, the classical sliding mode precision is proportional to the time interval between the measurements or to the switching delay. An r-sliding mode realization may provide up to r-th order of sliding precision with respect to the measurement interval. Properly used, HOSM control also totally removes the chattering effect [Levant 1993 2001]. The 2-sliding mode control can be used for smoothing out the control signal. The sliding order is a measure of the degree of smoothness of the sliding variable in the vicinity of the sliding mode. The sliding order is defined as the number of continuous total time derivatives of the sliding variable [Levant 2001].

The robustness described in the preceding sections is retained in the HOSM control configuration. One of the objectives of this thesis is to design HOSM controllers for idle speed control. The design and implementation work is presented in Chapter 7.

3.10 Sliding Mode Observer Design

Sliding mode control design based on a linear state space system requires the system states to be available. In practice, it may not be feasible to measure all the state variables as only a subset of the state information may be available. One method to overcome this problem is via the use of an *observer*. One of the early observers is the Luenburger observer [Luenberger 1971]. This observer considers the following system equations.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

Equation 3-47

The observer structure is defined as

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) - K_{obs}(y(t) - C\hat{x}(t)) \\ &= (A + K_{obs}C)\hat{x}(t) + Bu(t) - K_{obs}y(t) \end{aligned}$$

Equation 3-48

where the design matrix $K_{obs} \in \Re^{nxp}$ is chosen to ensure the eigenvalues of $(A + K_{obs}C)$ have negative real parts. The observer acts like a model of the plant which is driven by the mismatch between the plant and observer outputs. The state error is defined as

$$e(t) = x(t) - \hat{x}(t)$$

Equation 3-49

The error dynamics are thus given by

$$\dot{e}(t) = (A + K_{obs}C)e(t)$$

Equation 3-50

If $(A+K_{obs}C)$ is a stable matrix, the output error will converge to zero from any initial condition and $\hat{x}(t)$ will converge to $x(t)$. The above has illustrated the use of an observer to estimate the system state using a Luenburger observer. Note that the analysis above requires the plant model to be known accurately. The next section will outline the design of a sliding mode observer.

3.10.1 Observer Design From Output Information

A general framework for sliding mode observer design [Edwards 1998] for a linear system is illustrated here. Consider the system in equation 3-47. It is assumed that the states of the system are unknown and only the signals $u(t)$ and $y(t)$ are available. The aim is to generate a state estimate $\hat{x}(t)$. The output error is,

$$e_y(t)=\hat{y}(t)-y(t)$$

Equation 3-51

The output error, $e_y(t)$ is forced to zero in finite time by forcing a sliding motion on $e_y(t)$ so that $e_y(t)=0$. The observer has the following structure

$$\dot{\hat{x}}(t)=A\hat{x}(t)+Bu(t)-G_l e_y(t)+Bv_{obs}$$

Equation 3-52

where G_l is the linear observer gain and

$$v_{obs}=\begin{cases}-\rho_{obs}\dfrac{Fe_y}{\|Fe_y\|} & \text{if } e_y\neq 0\\ 0 & \text{otherwise}\end{cases}$$

Equation 3-53

The block diagram in Figure 3-12 shows the representation of the observer design in the Matlab Simulink environment.

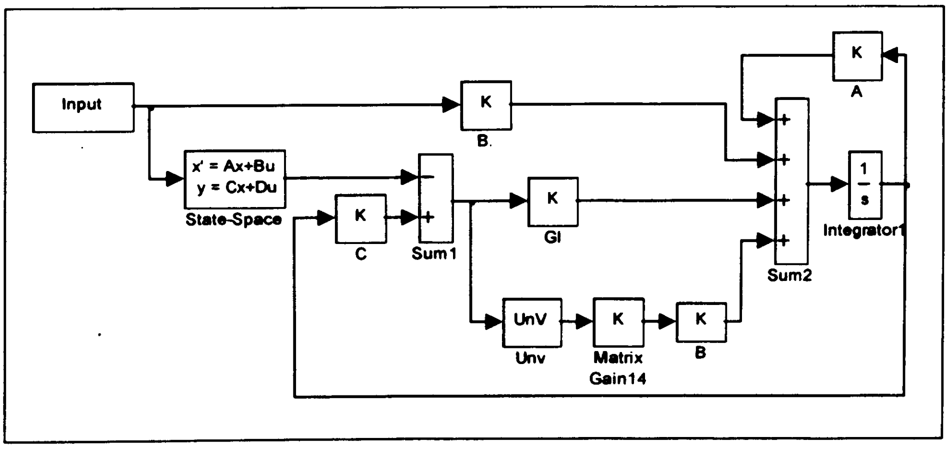


Figure 3-12: The observer block diagram.

The observer error dynamic becomes

$$\dot{e}_y(t) = (A - G_l C)e_y(t) + Bv_{obs}$$

Equation 3-54

The gain G_l is chosen such that the closed loop observer matrix, $A_o = (A - G_l C)$ has stable eigenvalues and satisfies the Lyapunov equation:

$$PA_o + A_o^T P^T = -Q$$

Equation 3-55

where Q is some positive definite matrix and P is a Lyapunov matrix which satisfies the structural constraint

$$C^T F^T = PB$$

Equation 3-56

where $F \in \Re^{mxm}$ is a non singular design matrix [Edwards 1996].

To prove that the observer guarantees quadratic stability of the error system, the following candidate Lyapunov function is considered.

$$V(e) = e_y^T P e_y$$

Equation 3-57

Differentiating the equation and substituting the error dynamic equations yields,

$$\begin{aligned}\dot{V}(e) &= \dot{e}_y^T P e_y + e^T P \dot{e}_y \\ &= \left(e_y^T A_o^T + Bv_{obs}\right) P e_y + e_y^T P\left(A_o^T e_y + Bv_{obs}\right) \\ &= e_y^T \left(PA_o + A_o^T P^T\right) e_y + 2e^T PBv_{obs} \\ &\leq - e_y^T Q e_y - 2\rho \left\| FCe_y \right\|\end{aligned}$$

Equation 3-58

If ρ is selected so that $\rho > 0$, then the convergence of the observer output error is guaranteed and the system is quadratically stable.

As well as being used for state estimation, the sliding mode observer technique will also be used for fault monitoring. The CMFD system is used to generate an alarm when an abnormal condition develops in the process being monitored. This observer based approach can be used to re-construct the fault signals. An analysis of the discontinuous signals required to maintain a sliding motion can be used to determine abnormalities in the process. The designed observer is thus used to estimate the system states and also re-construct system parameters. A nonlinear sliding mode observer is considered in Chapter 4 and applied to a particular diesel engine cooling system.

3.11 Conclusion

This chapter has introduced the sliding mode technique which is based on the variable structure concept. The chapter has given the description and illustration of the VSS. Then the important reachability condition required to maintain a sliding motion was developed. Two advantageous properties of sliding motion have been highlighted, namely:

(a) Reduction in system order
(b) Uncertainty rejection

The equivalent control, also know as the *average* control required to maintain a sliding motion, has been discussed. This chapter discussed the chattering problem and possible solutions have been discussed. Two methods of sliding surface design have been presented The HOSM control concept was discussed briefly. It has the potential to eliminate chattering. Lastly, this chapter has presented a general design framework for the sliding mode controller and observer design.

Chapter 4

4 Robust Fault Diagnosis of the Diesel Engine

Chapter 2 has described several fault diagnosis issues including fault diagnostic techniques, transducer selection, instrumentation requirements and the data acquisition system. In Chapter 3, the robust sliding mode concept has been described. This chapter will concentrate on the design of an on-board fault diagnosis system using the sliding mode concept. The designed technique is then implemented onto a particular diesel engine cooling system.

This chapter starts with an introduction to the cooling system fault diagnosis problem and a brief review of the applicability of the sliding mode technique to fault diagnosis. Section 4.2 raises some diagnosis issues particularly relevant to the engine. The approach to the diagnosis problem is mentioned in Section 4.3. Section 4.4 describes the development of the cooling system model and the modifications made to the model to represent fault scenarios. Sliding mode observer design is described in Section 4.5 and a discussion of system stability is highlighted in Section 4.6. Fault determination via system parameters is discussed in Section 4.7. Section 4.8 addresses the hardware modifications needed to represent fault scenarios as well as the hardware setup and data acquisition for the diagnostic system. Finally, implementation results are discussed in Section 4.9.

4.1 Introduction

A fault is caused by any kind of malfunction in the actual dynamic system that may lead to an unacceptable overall system performance. Faults will often be harmful to an engineering process if early detection is not made. Prompt detection helps to minimise the maintenance and repair costs of the system and contributes towards increased system reliability. In terms of the coolant system under consideration here, problems with the radiator or coolant pump, for example, would have the effect of causing the engine to overheat. Appropriate sensors to monitor the condition of such components directly may be expensive or the variables may be difficult to measure. In this chapter a key issue will be the desire to minimize the number of sensors to make usage cost effective. It will be seen that an appropriately designed sliding mode observer which uses straightforward temperature measurements can be used to monitor sophisticated system parameters.

A number of authors have considered the application of sliding mode approaches to the model-based fault detection problem. A design approach is adopted which seeks to make the occurrence of a fault synonymous with a break in the sliding motion [Hermans 1996]. However, such a break can be difficult to detect in practice. Adaptive sliding observers have been considered [Yang 1995]. A design framework which uses a nominal linear system representation for design of the sliding mode observer and then applies the equivalent input injection principle for fault reconstruction can be found in [Edwards 1999]. Here the observer design is performed for a cooling system description and the equivalent input injection principle is applied for parameter estimation. The state and parameter estimates can then be used to generate residual signals which act as an indicator of fault conditions.

4.2 Engine Fault Diagnosis Issues

Chapter 2 has discussed general fault diagnosis issues. In this section there is a brief discussion of issues regarding engine fault diagnosis. One of the readily observed parameters of the commercial engine is temperature. This temperature information only provides the overall performance condition of the engine. It does not give information about faults down to the component level. At the component level, different fault modes may be observed. Two general fault modes are:

(a) Abrupt faults, i.e. step-like changes.

(b) Incipient faults, i.e. bias or drift.

The abrupt faults occur suddenly and without any warning, whereas the incipient faults slowly develop in the system. Coolant pump failure is often an example of an abrupt fault and radiator (of the cooling system) blockage is usually as an incipient fault. Both of these faults may lead to a rise in temperature.

4.2.1 Modern Engine

Modern engines today are complex in construction and monitoring work has to cope with this complexity. Many diesel engines are placed in remote areas, where accessibility is impossible at times. Many of these engines are controlled remotely. In order to monitor such engines effectively, the monitoring method must be accurate and provide high confidence in finding out a particular engine fault on a particular sub-system. Constantly monitoring the engine condition will help to pinpoint the possible sub-system fault on the engine and allow maintenance work to be carried out at the earliest possible time.

4.2.2 Limitations in Monitoring

The commonly observed parameters on a diesel engine are speed, temperature, cylinder pressure and load. These parameters provide instant condition information on the engine. However, there are limitations in this monitoring work. Some quantities are complex to derive, such as fuel consumption rate, coolant flow and heat transfer coefficient. There are also limitations due to human reactions - the alertness of a watchkeeper when an engine fault occurs, for example. The remedy and any warnings issued depend on the response of the watchkeeper. Another possible limitation arises due to long term trends. This unforeseen circumstance often goes unnoticed.

4.2.3 Parameter Measurement Methods

According to Challen [Challen 1999], in practice, parameter measurement methods can be classified as direct and indirect methods as follows:

(a) Direct Method

The direct method is concerned with the actual performance of the engine. Parameters such as speed, temperature, cylinder pressure and load are considered in the direct method. These parameters represent the condition of the engine directly. In practice, most of the instrumentation and transducers are employed to capture direct monitoring data.

(b) Indirect Method

The indirect method is concerned with the state of parameters relating to lubricants, coolant water and diesel fuel. For example, the lubricant can be subject to adverse influences such as dilution by fuel, changes in viscosity and contamination by particles resulting from wear. For the coolant water, the heat transfer coefficient can determine the effectiveness in absorbing the engine heat and dissipation of the heat via the radiator. The heat transfer coefficient can not be measured directly but using special techniques, the value can be reconstructed based on only engine temperature information, by means of an observer.

4.3 Fault Diagnosis Approach

This chapter focuses on the specific application of sliding mode observers to detect possible faults in a diesel engine cooling system. The emphasis here is on the rig trials, which demonstrate the proof of concept. However, it should be noted that the methodology is equally relevant for a range of fault diagnostic problems. The work considers the estimation of system states as well as pertinent diesel engine coolant system parameters. The observer system is required to remain in the sliding mode under all system conditions in order to effectively estimate parameters. A fully nonlinear model [Twiddle 2001] is used to verify that the parameter estimates are reasonable. For the cooling system, malfunctions may occur in the following three main components: the thermostat bypass valve, the coolant pump and the radiator. Possible malfunctions in any of these sub-systems will cause the engine to overheat and problems may propagate to other engine sub-systems. A range of fault scenarios has been created to demonstrate the approach.

4.4 Model Development for the Diesel Engine Cooling System

The diagram shown in Figure 2-3 represents a diesel engine coolant system. The engine block represents a heat source. The thermostat valve divides coolant flow according to its opening level, α. The radiator acts as a heat sink to the atmosphere. Arrows dictate the direction of coolant flow. While the thermostat valve is closed ($\alpha = 0$), no coolant can flow through the radiator and coolant circulates through the left circuit. The coolant will only flow through the radiator when the thermostat valve is open. The bypass valve is used to bypass part of the coolant mixture and is merely a method used to simulate fault conditions on the rig and is not an inherent component of a typical diesel engine generator.

A model of the coolant system developed in [Chiang 1982] and the fault diagnostic technique proposed by Bhatti [Bhatti 1999b] for this system are employed here. Bhatti *et al* has shown the potential of the proposed diagnostic technique in the simulation work. The model equations are based on the balance of thermal energy between the engine block and the coolant, and also between the coolant and the radiator where heat dissipates to the atmosphere. Heat transfer from the engine block is balanced by heat absorbed by the coolant and heat loss through convection. In addition to the heat balance, some of the heat may dissipate through the engine surface by convection and radiation processes. Heat is also lost through conduction when it propagates to another sub-system or to the atmosphere.

4.4.1 Cooling System Components

The cooling system under consideration is built up of several sub-components, namely, thermostat bypass valve, coolant pump and radiator. Failure in any component may cause the engine to overheat. The particular sub-components are described in the following sub-sections, as well as the potential faults that may occur.

(a) Thermostat bypass valve

The thermostat bypass valve is used to control the coolant flow to the radiator. It is usually closed $(\alpha = 0)$ to give a warm start during start up of the engine. It allows the engine block to warm up to a suitable operating temperature. For the Perkins engine, it is set to open at about 74°C (347K). If the valve behaves abnormally, where it may totally close or open at a later stage, then the engine may overheat. Possible problems with the thermostat valve are that it occasionally gets 'stuck' and the valve fails to open [Autoshop 1998]. As a consequence, the engine gets overheated and excessive wear is caused.

(b) Coolant pump

Another important part of the cooling system is the coolant pump which is used to maintain a constant flow rate ($\dot{m}$) of coolant while the engine is running. Although the coolant flow rate depends on the engine speed, with higher speeds yielding higher flow rates, the engine used here is running at constant speed and a constant flow of coolant is assumed. Malfunctioning of the pump will critically affect the dissipation of heat in the engine block. The capacity of the coolant flow in the system must be closely monitored before any faults occur.

(c) Radiator

The radiator is an essential component of the cooling system. It is responsible for the dissipation of engine heat to the atmosphere and for maintaining the engine block temperature within the appropriate operating range. It is constructed in a way in which heat can dissipate at a rather faster rate through a large surface area of the 'zig-zag' fins and circuits. The heat from the hot coolant is transferred to the surface of the fins and blown away by a radiator fan. There are ways by which radiator blockage and failure of the radiator fan can occur. An example is when the radiator circuit becomes clogged due to internal corrosion which causes low coolant flow. This reduces the area available for heat transfer and causes the engine to overheat. In this case, the heat transfer coefficient will determine the efficiency of the radiator system.

4.4.2 Cooling System Model Formulation

Let T_1 be the engine block inlet temperature, T_2 be the engine outlet temperature and radiator inlet temperature, T_3 be the radiator outlet temperature and T_B be the engine block temperature. A balance of the thermal energy with respect to the engine block and coolant flow is shown in the equations below.

An expression for the heat transfer for coolant in the engine block is

$$(mc_{block\text{-}to\text{-}coolant})\frac{dT_2}{dt} = \dot{q}_c - \dot{q}_h$$

Equation 4-1

where, $\dot{q}_c = (h_{bc}A_{block\text{-}to\text{-}coolant})(T_B - T_2)$ is the engine block to coolant heat transfer, $\dot{q}_h = (\dot{m}c)(T_2 - T_1)$ the heating of coolant flow/coolant heat gain, h_{bc} the coolant side heat transfer coefficient, $A_{block\text{-}to\text{-}coolant}$ the heat transfer area, m the mass of coolant, $c_{block\text{-}to\text{-}coolant}$ the specific heat of coolant in block, c the specific heat capacity and $\dot{m}$ the mass flow rate of coolant through block.

The radiator dissipates heat from the engine block through its fin area to keep the temperature of the engine block within the operating temperature range. The inlet temperature is assumed to be the same as the output temperature of the engine block. The heat dissipation to ambient via the radiator is then given by

$$(mc_{rad})\frac{dT_3}{dt} = \dot{q}_{cL} - \dot{q}_{L\_rad}$$

Equation 4-2

where $\dot{q}_{cL} = \dot{m}c(T_3 - T_2)$, $\dot{q}_{L\_rad} = \dot{q}_{conv\_rad} + \dot{q}_{rad\_rad}$, the radiation losses, $\dot{q}_{rad\_rad}$ are assumed to be small (≈0), $\dot{q}_{conv\_rad} = (h_{rad}A_{rad})(T_3 - T_{amb})$, m the mass of coolant, c_{rad} the specific heat capacity of coolant, h_{rad} the radiator heat transfer coefficient and A_{rad} the total area of radiator.

From Figure 2.3, consider an energy balance with the system boundary defined at the inlet of the coolant pump where the coolant flows from the bypass and the radiator are conjoined. Assuming there is no heat transfer from the system pipe-work between thermostatic valve, radiator outlet and the inlet to the pump, then [Twiddle 2001]

$$\dot{m}_B c T_1 = (1-\alpha)\dot{m}_B T_2 + \alpha\dot{m}_B c T_3$$

In terms of volume flow rate, the above equation becomes

$$\rho\dot{V}_B c T_1 = (1-\alpha)\rho\dot{V}_B c T_2 + \alpha\rho\dot{V}_B c T_3$$

where ρ Coolant density

$\dot{m}_B$ Mass rate flow through engine block

$\rho\dot{V}_B = \dot{m}$ Coolant mass flow rate at pump outlet.

$\alpha\rho\dot{V}_B = \rho\dot{V}_{rad} = \dot{m}_{rad}$ Coolant mass flow rate at the radiator inlet.

$\dot{V}_B$ Coolant volume flow-rate through engine block

$\dot{V}_{rad}$ Coolant volume flow-rate at radiator inlet

$\dot{m}_{rad}$ Mass rate flow through radiator

Further to this, assume that ρ and c are constant, over the temperature range observed in the system. Therefore, the relationship between T_1, T_2, T_3 and α is shown in the following equation:

$$T_1 = \alpha T_3 + (1-\alpha)T_2$$

Equation 4-3

By re-arranging and substituting the appropriate parameters into the equations yields,

$$\dot{T}_2 = [-k_1 - \dot{m}k_2]T_2 + \dot{m}k_2T_3 + k_1T_B$$

Equation 4-4

$$\dot{T}_3 = -\dot{m}k_3T_2 + (\dot{m}k_3 - h_{rad}k_4)T_3 + h_{rad}k_4T_{amb}$$

Equation 4-5

A third equation, denoted as T_{2a}, is derived from the equation 4-4 for thermostat valve parameter estimation. The dynamic characteristics of T_2 and T_{2a} are exactly the same. However, the differential equation for T_2 is parameterised in terms of variable coolant mass flow rate and T_{2a} is parameterised in terms of variable thermostat valve opening. The equation T_{2a} can be achieved by rearranging the k_2 term in order to include the thermostat valve term in the equation. This yields the following final model equations.

$$\dot{T}_2 = (-k_1 - \dot{m}k_2)T_2 + \dot{m}k_2T_3 + k_1T_B$$

Equation 4-6

$$\dot{T}_{2a} = (-k_1 - \alpha k_{2a})T_{2a} + \alpha k_{2a}T_3 + k_1T_B$$

Equation 4-7

$$\dot{T}_3 = -\dot{m}k_3T_2 + (\dot{m}k_3 - h_{rad}k_4)T_3 + h_{rad}k_4T_{amb}$$

Equation 4-8

where $\dot{m}$, h_{rad} and α are the coolant mass flow rate, the radiator heat transfer coefficient and the opening level of the thermostat valve, respectively. These represent the parameters that could be useful to monitor the health of the engine but are not straightforward to measure. Bhatti derived the equation 4-6 based on the equation 4-4 for the coolant flow estimation [Bhatti 1999b]. The equation 4-7 is essentially derived from equation 4-4 for the particular thermostat valve estimation. k_1, k_2, k_3, k_4 and k_{2a} are given by

$$k_1 = \frac{h_{bc}A_{block-to-coolant}}{mc_{block-to-coolant}},\ k_2 = \frac{\alpha c}{mc_{block-to-coolant}},\ k_{2a} = \frac{\dot{m}c}{mc_{block-to-coolant}},\ k_3 = \frac{\alpha c}{mc_{rad}},\ k_4 = \frac{A_{rad}}{mc_{rad}}$$

4.4.3 Modifications Made to the Model used for Fault Detection

During fault scenarios, for example, a pump fault may cause a variation in the coolant mass flow rate, $\dot{m}_2$. Let $\hat{\dot{m}} = \dot{m} + \Delta\dot{m}$, $\hat{h}_{rad} = h_{rad} + \Delta h_{rad}$ and $\hat{\alpha} = \alpha + \Delta\alpha$ represent deviations $\Delta\dot{m}$, Δh_{rad} and $\Delta\alpha$ from nominal conditions $\hat{\dot{m}}$, $\hat{h}_{rad}$ and $\hat{\alpha}$. Assume the constants k_1, k_2, k_3, k_4 and k_{2a} are evaluated at the nominal operating point determined by $\dot{m}$, h_{rad} and α. Then

$$\begin{aligned}\dot{T}_2 &= [-k_1 - (\dot{m} + \Delta\dot{m})k_2]T_2 + (\dot{m} + \Delta\dot{m})k_2T_3 + k_1T_B \\ &= (-k_1 - \dot{m}k_2)T_2 + \dot{m}k_2T_3 + k_1T_B - \Delta\dot{m}k_2T_2 + \Delta\dot{m}k_2T_3\end{aligned}$$

Equation 4-9

$$\begin{aligned}\dot{T}_{2a} &= [-k_1 - (\alpha + \Delta\alpha)k_{2a}]T_{2a} + (\alpha + \Delta\alpha)k_{2a}T_3 + k_1T_B \\ &= (-k_1 - \alpha k_{2a})T_{2a} + \alpha k_{2a}T_3 + k_1T_B - \Delta\alpha k_{2a}T_{2a} + \Delta\alpha k_{2a}T_3\end{aligned}$$

Equation 4-10

$$\begin{aligned}\dot{T}_3 &= -(\dot{m} + \Delta\dot{m})k_3T_2 + [(\dot{m} + \Delta\dot{m})k_3 - (h_{rad} + \Delta h_{rad})k_4]T_3 + (h_{rad} + \Delta h_{rad})k_4T_{amb} \\ &= -\dot{m}k_3T_2 + (\dot{m}k_3 - h_{rad}k_4)T_3 + h_{rad}k_4T_{amb} - \Delta\dot{m}k_3T_2 + (\Delta\dot{m}_2 - \Delta h_{rad})k_4T_3 \\ &\quad + \Delta h_{rad}k_4T_{amb}\end{aligned}$$

Equation 4-11

Equations 4-9, 4-10 and 4-11 represent the assumed behaviour of the uncertain cooling system dynamics. The nominal dynamics is obtained by setting $\Delta\dot{m} = \Delta h_{rad} = \Delta\alpha = 0$.

4.5 Sliding Mode Observer Design

In Chapter 3, the switching function is chosen to ensure desirable performance is exhibited by the system of interest. For an observer problem, the switching function may most appropriately be defined as the output error. By forcing this switching function to zero, the observer output is forced to equal the plant output and a set of estimated states that yield the measured system output are obtained. The system is then said to have attained a sliding mode.

Having defined an appropriate switching function, it is necessary to choose an injection signal that will force the sliding mode condition to be attained and maintained. This is termed the reachability problem and is typically ensured by choosing the injection signal so that $s(t)$ and $\dot{s}(t)$ are forced to have opposite sign. Essentially, discontinuous injection signals are used to maintain some appropriately chosen switching function at zero. During the sliding motion, the system exhibits total robustness (as mentioned in Chapter 3) to a class of disturbance signals and any uncertainty that is implicit in the channels where the discontinuous injection signal is applied.

For the fault diagnostic situation considered here, the switching function will be the observer error. Further, it will be seen that the discontinuous signal employed to maintain the observer error at zero provides effective information for parameter reconstruction and fault diagnosis. A proposed structure of the sliding mode observer for the cooling system is as follows:

$$\dot{\hat{T}}_2 = (-k_1 - \dot{m}k_2)\hat{T}_2 + \dot{m}k_2T_3 + k_1T_B + v_2$$

Equation 4-12

$$\dot{\hat{T}}_{2a} = (-k_1 - \alpha k_{2a})\hat{T}_{2a} + \alpha k_{2a}T_3 + k_1T_B + v_{2a}$$

Equation 4-13

$$\dot{\hat{T}}_3 = -\dot{m}k_3T_2 + (\dot{m}k_3 - h_{rad}k_4)\hat{T}_3 + h_{rad}k_4T_{amb} - \beta e_3 + v_3$$

Equation 4-14

where,

$$v_i = \begin{cases} -K_i\dfrac{e_i}{\|e_i\|} & \text{if } e_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Equation 4-15

where,

$$e_2 = T_2 - \hat{T}_2 \ , \quad e_{2a} = T_{2a} - \hat{T}_{2a} \ \text{ and } \ e_3 = T_3 - \hat{T}_3$$

Equation 4-16

where $i = 2, 2_a, 3$ and K_2, K_{2a} and K_3 are the gains given to the discontinuous signals v_2, v_{2a} and v_3 which must be chosen to satisfy the reachability problem (as demonstrated in Chapter 2) and ensure sliding motion occurs even in the presence of parameter variations and faults. It should be noted that, although k_1, k_2, k_3, k_4 and k_{2a} are assumed to be constant although they are dependent on the variations $\dot{m}$, h_{rad} and α. However, with the proposed observer structure a sliding motion will still be maintained and this parameter uncertainty will be rejected, as it appears in channels where the v_i are injected. Maintenance of the sliding condition is essential for parameter estimation and fault diagnosis in this scheme.

β acts as a gain parameter to tune asymptotic error decay. This parameter can increase the rate of error decay in the absence of uncertainties. The system stability test will be conducted in next sub-section. The observer error dynamics can be written as follows.

$$\dot{e}_2 = (-k_1 - \dot{m}k_2)e_2 + \Delta\dot{m}k_2(T_3 - T_2) - v_2$$

Equation 4-17

$$\dot{e}_{2a} = (-k_1 - \alpha k_{2a})e_2 + \Delta\alpha k_{2a}(T_3 - T_{2a}) - v_{2a}$$

Equation 4-18

$$\dot{e}_3 = (\dot{m}k_3 - h_{rad}k_4 - \beta)e_3 + \Delta h_{rad}k_4(T_{amb} - T_3) + \Delta\dot{m}k_3(T_3 - T_2) - v_3$$

Equation 4-19

4.6 System Stability

To determine the stability of the observer system, consider the candidate Lyapunov function as follows,

$$V(e) = \frac{1}{2}(e_2^2 + e_{2a}^2 + e_3^2)$$

Equation 4-20

Differentiating the equation and substituting the error dynamic equations yield,

$$\dot{V}(e) = e_2\dot{e}_2 + e_{2a}\dot{e}_{2a} + e_3\dot{e}_3$$

Equation 4-21

$$\begin{aligned}\dot{V}(e) = e_2[(-k_1 - \dot{m}k_2)e_2 + \Delta\dot{m}k_2(T_3 - T_2) - v_2] + e_{2a}[(-k_1 - \alpha k_{2a})e_{2a} + \\ \Delta\alpha k_{2a}(T_3 - T_{2a}) - v_{2a}] + e_3[(\dot{m}k_3 - h_{rad}k_4 - \beta)e_3 + \Delta h_{rad}k_4(T_{amb} - T_3) + \\ \Delta\dot{m}k_3(T_3 - T_2) - v_3]\end{aligned}$$

$$\begin{aligned}\dot{V}(e) = & -k_1 e_2^2 - \dot{m}k_2 e_2^2 - k_1 e_{2a}^2 - \alpha k_{2a} e_{2a}^2 - (h_{rad}k_4 + \beta - \dot{m}k_3)e_3^2 - \\ & e_2 \Delta \dot{m} k_2 (T_2 - T_3) - K_2 \frac{\|e_2\|^2}{\|e_2\|} - e_{2a}\Delta \alpha k_2 a (T_{2a} - T_3) - K_{2a}\frac{\|e_{2a}\|^2}{\|e_{2a}\|} - \\ & e_3 \Delta h_{rad} k_4 (T_3 - T_{amb}) - e_3 \Delta \dot{m} k_3 (T_2 - T_3) - K_3 \frac{\|e_3\|^2}{\|e_3\|}\end{aligned}$$

Equation 4-22

as $e_i^2 = \|e_i\|^2$

$$\begin{aligned}\dot{V}(e) = & -k_1 e_2^2 - \dot{m}k_2 e_2^2 - k_1 e_{2a}^2 - \alpha k_{2a} e_{2a}^2 - (h_{rad}k_4 + \beta - \dot{m}k_3)e_3^2 - \\ & e_2 \Delta \dot{m} k_2 (T_2 - T_3) - K_2 \|e_2\| - e_{2a}\Delta \alpha k_{2a} (T_{2a} - T_3) - K_{2a}\|e_{2a}\| - \\ & e_3 \Delta h_{rad} k_4 (T_3 - T_{amb}) - e_3 \Delta \dot{m} k_3 (T_2 - T_3) - K_3 \|e_3\| \\ \le & -k_1 e_2^2 - \dot{m}k_2 e_2^2 - k_1 e_{2a}^2 - \alpha k_{2a} e_{2a}^2 - (h_{rad}k_4 + \beta - \dot{m}k_3)e_3^2 - \\ & \|e_2\|(K_2 - \|T_2 - T_3\|\Delta \dot{m} k_2) - \|e_{2a}\|(K_{2a} - \|T_{2a} - T_3\|\Delta \alpha k_{2a}) - \\ & \|e_3\|(K_3 - \|\Delta h_{rad} k_4 (T_3 - T_{amb})\| - \|\Delta \dot{m} k_3 (T_2 - T_3)\|\end{aligned}$$

Equation 4-23

If the uncertainty is bounded as

$$\begin{aligned} & K_2 \ge \|\Delta \dot{m} k_2 (T_2 - T_3)\| \\ & K_{2a} \ge \|\Delta \alpha k_{2a} (T_{2a} - T_3)\| \\ & K_3 \ge \|\Delta h_{rad} k_4 (T_3 - T_{amb}) + \Delta \dot{m} k_3 (T_2 - T_3)\|\end{aligned}$$

and β is selected so that $\beta > \dot{m}k_3$, convergence of the observer error is proved since all the nominal constants involved in the above equation are positive, i.e.

$$k_1 > 0,\ k_2 > 0,\ k_3 > 0,\ k_4 > 0,\ \dot{m} > 0,\ k_{2a} > 0,\ h_{rad} > 0$$

4.7 Fault Determination

The gains of the discontinuous terms, K_2, K_{2a} and K_3, are designed so that the observer remains in a sliding mode and the error is zero. Thus, the observer errors (e_2, e_{2a} and e_3) and their derivatives will converge to zero. The perturbations due to faults are effectively compensated for by the discontinuous signals v_2, v_{2a} and v_3 respectively. In the sliding mode, equations 4-17, 4-18 and 4-19 become :-

$$0 = \Delta \dot{m} k_2 (T_3 - T_2) - v_2$$

Equation 4-24

$$0 = \Delta \alpha k_{2a} (T_3 - T_{2a}) - v_{2a}$$

Equation 4-25

$$0 = \Delta h_{rad} k_4 (T_{amb} - T_3) + \Delta \dot{m} k_3 (T_3 - T_2) - v_3$$

Equation 4-26

and re-arranging these equations gives,

$$\Delta \dot{m} = \frac{v_2}{k_2(T_3 - T_2)}$$

Equation 4-27

$$\Delta \alpha = \frac{v_{2a}}{k_{2a}(T_3 - T_{2a})}$$

Equation 4-28

$$\Delta h_{rad} = \frac{v_3 - \Delta \dot{m} k_3 (T_3 - T_2)}{k_4 (T_{amb} - T_3)}$$

Equation 4-29

It is seen that v_2 v_{2a} and v_3 effectively provide a means to estimate the change in mass flow rate, thermostat valve opening and radiator heat transfer coefficient. The signals v_2 v_{2a} and v_3 are discontinuous injection signals and to obtain parameter values, the average value of v_2 v_{2a} and v_3 must be used. It is thus necessary to 'filter' v_2 v_{2a} and v_3 to construct estimates of the change in mass flow rate, the thermostat valve opening and the radiator coefficient. In order to obtain the average value of v_2 v_{2a} and v_3, the smoothing method described in Section 3.5.1 in Chapter 3, is used. Thus the equation 4-15 can be re-written as follows:

$$v_i = \begin{cases} -K_i \dfrac{e_i}{\|e_i\| + \delta} & \text{if } e_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

where δ is the positive smoothing factor. Hence, the parameter information can then be used to form a fault detection scheme for the coolant system.

4.8 Experimental Setup

The sliding mode observer developed in the previous section is implemented on a diesel power generator. Section 4.4.1 has mentioned the sub-components of the cooling system and the possible faults which may occur. In order to simulate the mentioned component faults, various faults are introduced in the engine in order to test the efficiency of the proposed design approach.

4.8.1 Hardware Modification for Fault Scenarios

(a) Thermostat Valve Faults

The valve is located between the engine and the radiator. Although the valve is easy to access externally, it is rather difficult to control the opening of the valve on line. As a result, the fault was simulated using a different set of valves set to open at the higher temperature of 82° C (355K). This was done by changing the spring in the valve that controls the opening. As the engine coolant warms up, the wax in the valve expands, forcing the valve to open. Estimating the valve opening α will help in the fault monitoring process. This can be constructed from equation 4-28 and the nominal thermostat valve opening level.

(b) Low Coolant Flow

Low coolant flow occurs if there is coolant leakage in the cooling system, for example in the radiator or connecting hose. Malfunction in the coolant pump can contribute to low coolant flow. Such faults may result in the engine block getting overheated. To simulate this situation on the rig is rather difficult, as it is important not to damage the engine. Therefore, turning off the coolant pump is not an option. As a result, a by-pass passage around the coolant pump was considered to be a feasible solution to mimic the fault condition for data acquisition purposes. The modification is shown in Figure 2-3. With this passage, a small volume of the coolant, approximately 5% of the total coolant, is circulated around the coolant pump without letting it pass through the engine block. The flow was thus reduced. This condition also allows the sensitivity of the diagnostic system to be evaluated for small changes in nominal conditions. Estimating the coolant flow rate, $\dot{m}$, can provide coolant flow information. This can be readily constructed from equation 4-27 and the nominal mass flow rate.

(c) Radiator Blockage

In order to simulate a radiator blockage on the research engine, the radiator was covered, by means of covering the radiator fins. Two cases were considered: approximately 10% and 25% covered. The engine was run for only a short time during data acquisition to prevent excessive overheating. The heat transfer coefficient can then be constructed from equation 4-29 and the nominal radiator coefficient.

Hence, the sliding mode observer is not only used here to estimate the system states but also to predict three important parameters, the thermostat valve opening, the coolant flow rate, and the radiator heat transfer coefficient.

4.8.2 Hardware Setup

The hardware interface for the cooling system fault diagnosis comprises of a dSPACE data acquisition system and temperature sensors. The sensors are fitted at the locations shown in Figure 2-3. The raw sensor signals are small and a signal amplifier is needed to amplify the signal before feeding it into the dSPACE data acquisition system. With an appropriate gain factor, these signals are then converted to a Kelvin unit representation.

A block diagram in Figure 4-1 shows the signal flow of the system starting from the transducer, as well as the signal from the sliding mode observer. The sliding mode observer design is written in Matlab C-code C-mex S-function format. Once the observer design and the interface in the Matlab Simulink environment is completed, the Simulink model is then converted into C-code format in order to be downloaded onto a DSP chip. The procedure for C-code generation via the Real Time Workshop is described in Appendix B.

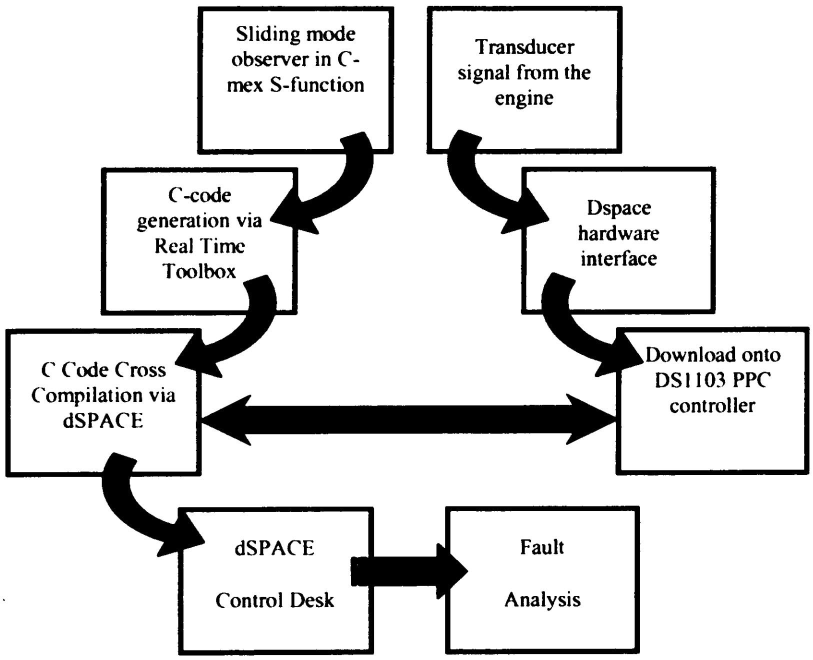


Figure 4-1: Block diagram showing the system interface and signal flow.

4.8.3 Data Acquisition

A 5th order Butterworth filter was employed to remove the high frequency components of the signal noise. The filter had a sampling frequency of 1000 Hz with a cut-off frequency of 1 Hz. Raw data, initially sampled at 1000 Hz, was down sampled to 100 Hz before saving due to the memory constraints of the online computer. The reason for this is to make sure signal properties are captured with no loss of information and no aliasing. Further details regarding data sampling can be found in Appendix A.

4.8.4 System Parameter Setting

The design parameter settings are $\alpha = 1$, $\delta = 1$ and $\beta = 100$. The estimated values from the genset are $A_{rad} = 12\, m^2$, $A_{block\text{-}to\text{-}coolant} = 1.5\, m^2$, $m = 15\, kg$ and $m_{rad} = 16.5\, kg$. The control gain design values are $K_2 = 800$, $K_{2a} = 800$ and $K_3 = 800$ [Bhatti 1999b]. The standard values from data book are $h_{bc} = 0.15\, \mathrm{kW/m^2}K$ and $c_{block\text{-}to\text{-}coolant} = c_{rad} = 4.2\, kJkg^{-1}K^{-1}$

4.9 Experimental Results

The following figures were generated from tests on the diesel engine. The first set of results corresponds to normal operating conditions with a step load applied to the engine from 0kW to 65 kW, followed by a step down load. The process was repeated once.

The estimates of the system states, T_2 and T_3 are shown in Figure 4-2 and Figure 4-3 respectively. This value is then compared with the value which was measured online while the experiment was carried out. The error between the measured and estimated values is of the order of 0 ~ 0.06K, which is about 0.02%. It is believed the small error is due to the property of sliding mode observer which is insensitive to parameter variations. The injection signal, which is non-linear in nature, is employed in the observer equations for this purpose. Thus, the observer has been validated under normal operating conditions. A non-linear model described in [Twiddle 2001] is employed here to provide a comparison between parameter estimates developed from the nonlinear model, and those estimated by the sliding mode observer. It should be noted that this is not the non-linear model used to design the observer but is a model developed using black box techniques and a large amount of empirical engine data. Further, it is the change in system parameters, as monitored via the equivalent injection analysis, which is to be used to indicate the presence of faults. The results are shown in Figure 4-4, Figure 4-5 and Figure 4-8.

During the transition between the thermostat valve opening and closing (which happens under normal conditions), the error between the response of the non-linear model and the measured response is relatively large, approximately 10%. When the thermostat valve is opened, coolant flowing to the radiator and the system undergoes a transition in thermodynamic properties. Due to this imbalance, the system parameters change accordingly in practice but this process is not modelled accurately. The hysteresis effect of the thermostat valve itself is also thought to contribute to the mismatch.

A fault is now introduced in to the system by opening the bypass valve to reduce the coolant flow rate through the system. Figure 4-6 shows the rate of change of coolant flow. The bypass valve was opened for the period of time between 60s and 125s. It is obvious that the response of the system does not change instantaneously when the valve was opened and closed. A likely explanation for this is that the coolant still in the bypass valve circuit has a lower temperature as compared to the coolant circulating in the system. When the bypass valve opens, these two sets of coolant which are at different temperatures mix together and after sometime, the coolant in the bypass valve eventually rises to the circulating coolant temperature. As a result the system takes a longer time to reach its steady state.

Next the thermostat valve is set to open at a later stage when compared to normal conditions. Figure 4-7 shows the plot of the thermostat valve opening during normal conditions and in the presence of a fault. It is proved that when the valve opens late, or indeed does not open at all, the system can accurately detect the fault.

For the radiator system, 10% and 25% of the radiator is covered. Figure 4-8 shows a plot of the heat transfer coefficient against ambient temperature. During normal operating conditions, the data forms a relatively concentrated cloud. This shows that the engine heat dissipates at a constant rate and the engine block maintains its temperature. As the engine becomes hotter, variation from ambient temperature occurs, but the variation is relatively small (about 4K). It is seen that the heat transfer coefficient reduces when the radiator is covered. Also the clouds of data points become more elongated. This is because the cooling effect of the radiator is less efficient and thus the observed temperature drifts are larger. The ambient temperature is lower when the radiator is 25% covered as compared to when it is 10% covered due to the correspondingly larger reduction in the efficiency of the radiator. This is because less heat is transferred to the ambient. The heat transfer coefficient when the radiator is 10% covered is marginally larger than when the radiator is 25% covered. For compact automotive radiators, typical values for h_{rad} are given as $0.08-0.13$ according to the book of Lilly [Lilly 1984]. The obtained heat transfer coefficient for the particular test diesel engine during normal operating condition is approximately 0.065 to 0.075 kW/m^2.K. During fault conditions, the values reduce to the range between 0.045 to 0.055 kW/m^2.K.

4.10 Conclusion

This Chapter explored the application of sliding mode concepts for fault detection and isolation. The experimental results have shown that perturbations in component parameters can be recovered by manipulating the equivalent input injection signal to the observer. Particular cooling system subsystems that are considered here are:

(a) Thermostat valve

(b) Coolant pump

(c) Radiator

It has been demonstrated that faults present in diesel engine components can be detected. It is seen that the behaviour of relatively sophisticated components can be inferred from straightforward temperature measurements. Although a specific system has been considered here, the core results have potentially wide applicability to a range of engineering systems. This on-board parameter monitoring technique has been shown to be efficient and most importantly, it is economically applicable as only temperature sensors are involved in the monitoring work.

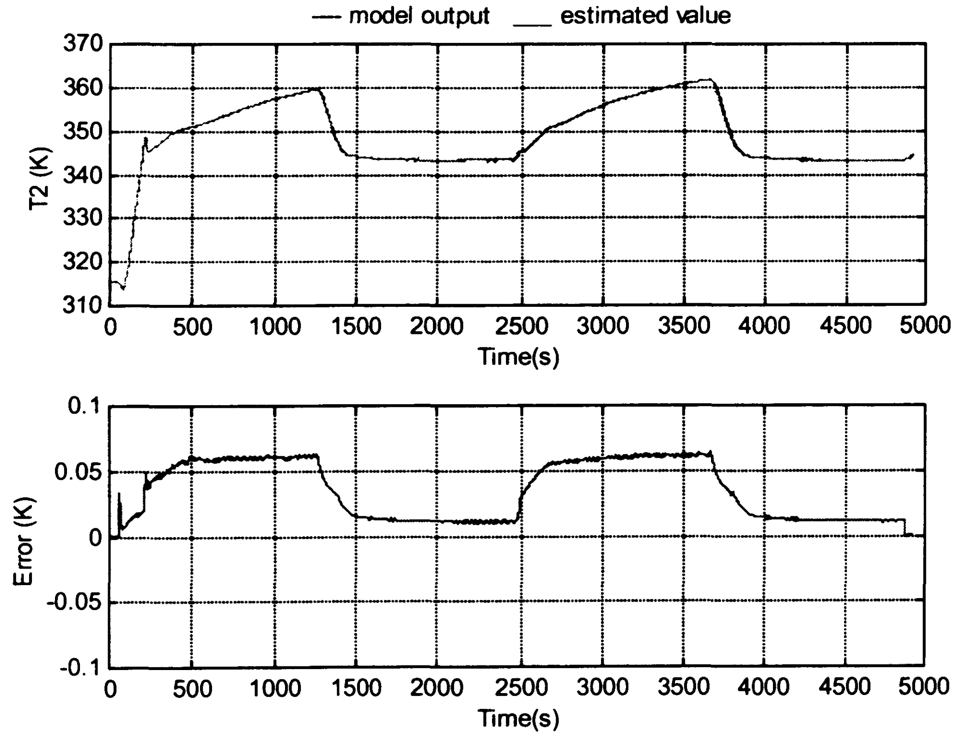


Figure 4-2: State estimate, T_2 and the error between the measured value and the estimated value.

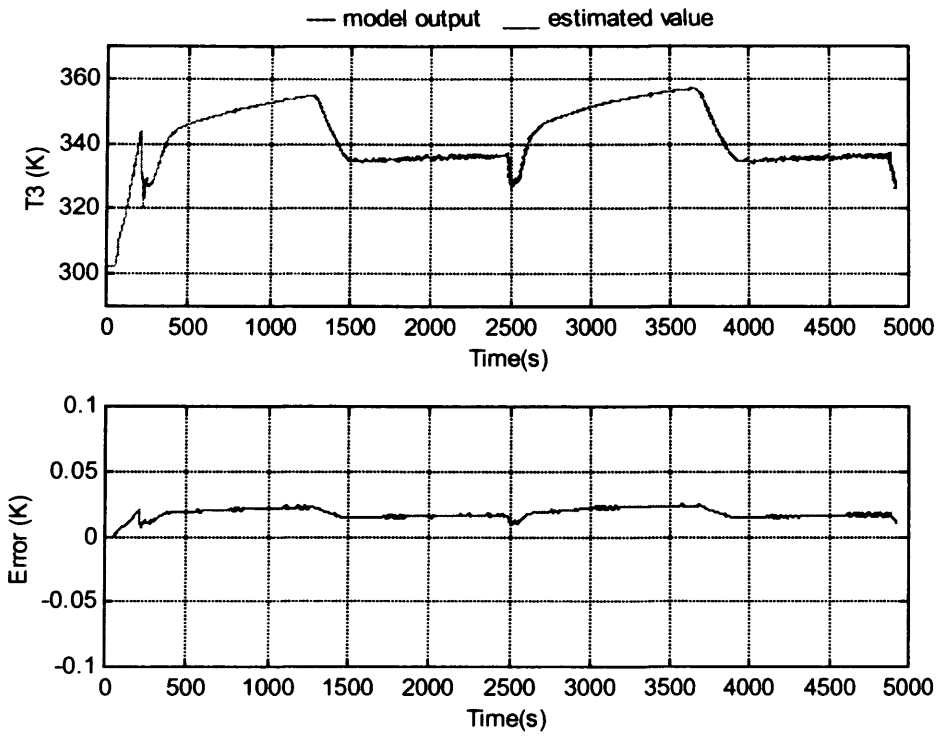


Figure 4-3: State estimate, T_3 and the error between the measured value and the estimated value.

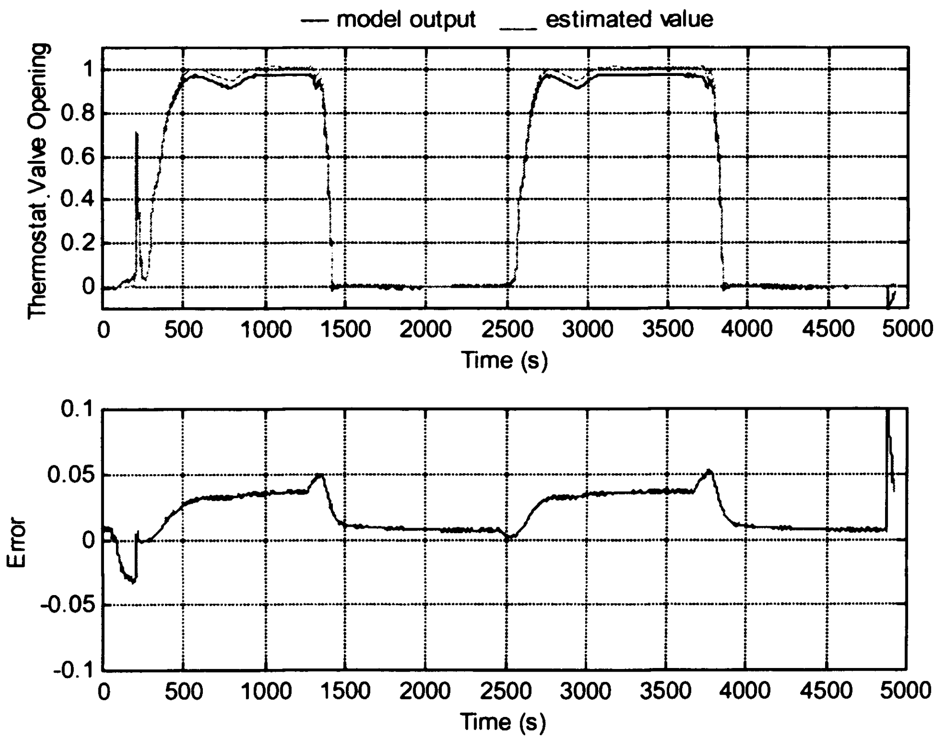


Figure 4-4: Component parameter estimation, α and the error between the measured value and the estimated value.

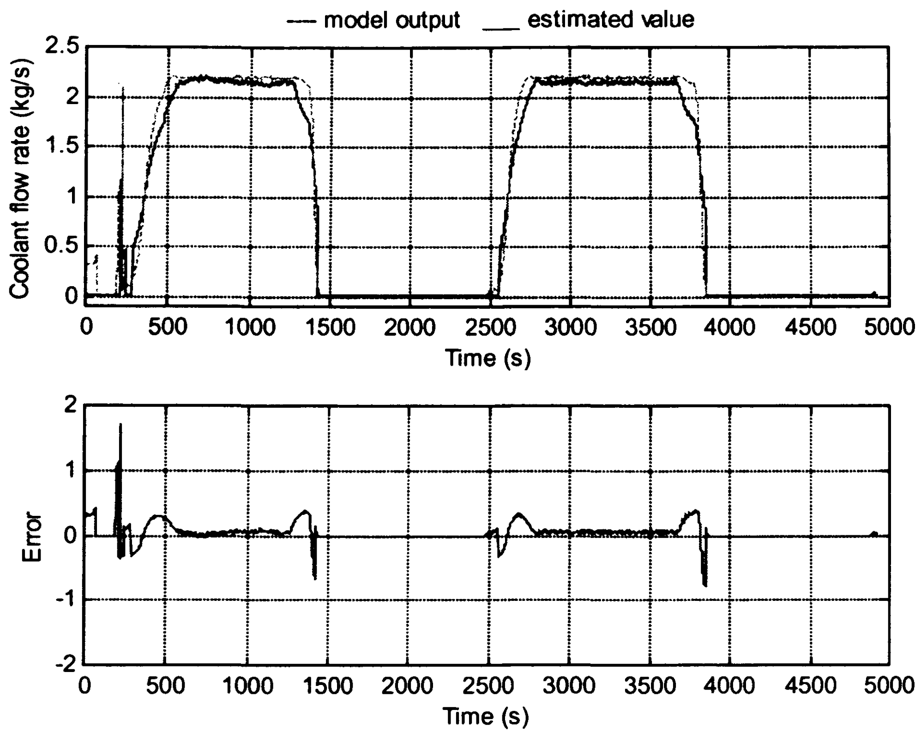


Figure 4-5: Component parameter estimation, $\dot{m}$ and the error between the measured value and the estimated value.

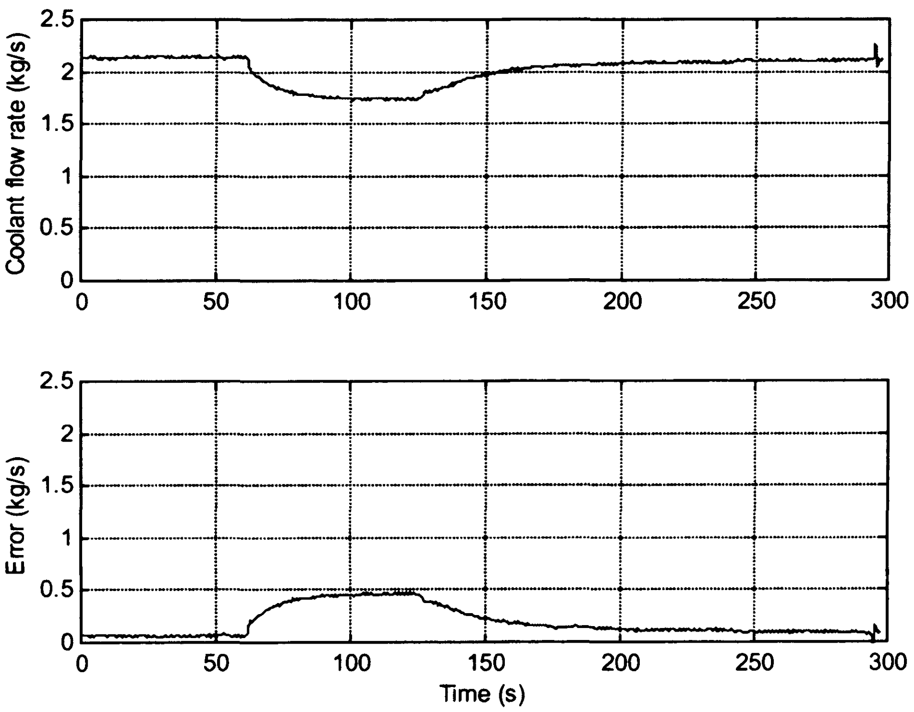


Figure 4-6: Opening of the bypass valve.

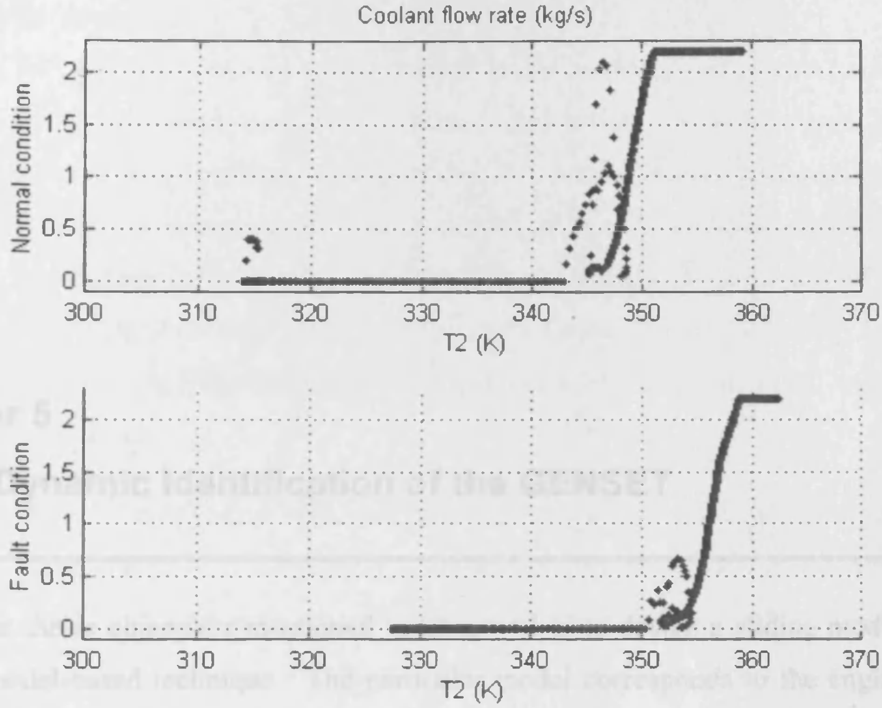


Figure 4-7: Thermostat valve open later than normal.

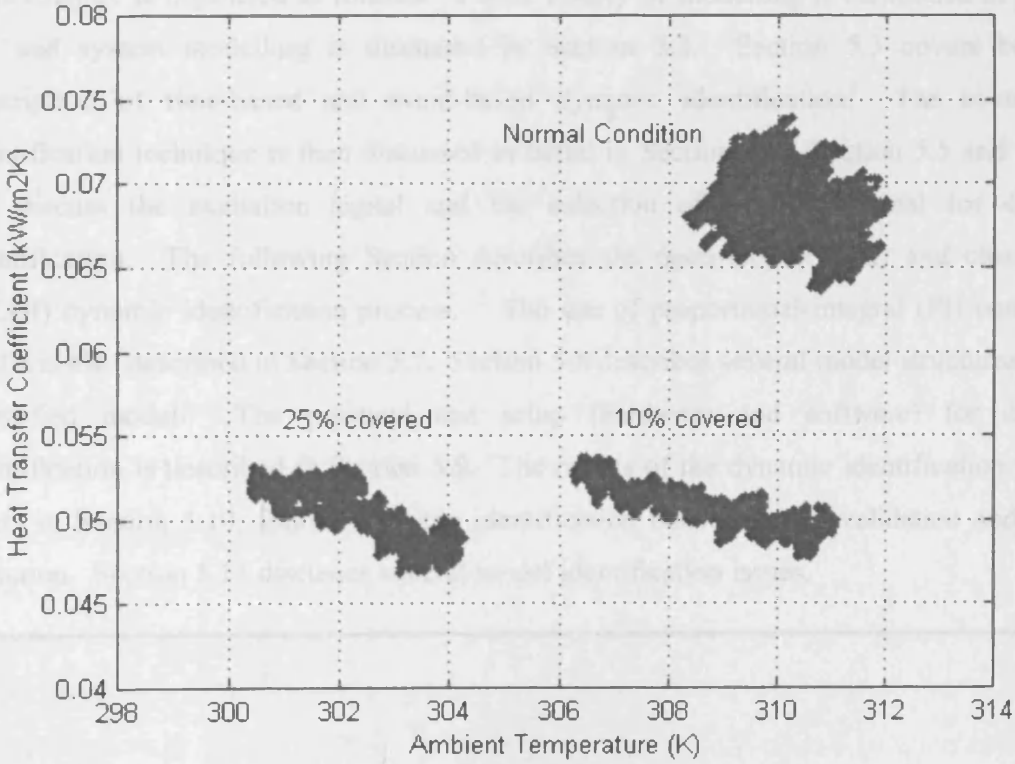


Figure 4-8: Plot of estimated heat transfer coefficient at the radiator for normal and faulty conditions.

Chapter 5

5 Dynamic Identification of the GENSET

One of the thesis objectives mentioned in Chapter 1 is to design a sliding mode controller using a model-based technique. The particular model corresponds to the engine speed at 1500 rpm with no load on the generator. In this chapter, the specific system will be modelled using closed-loop event-based identification techniques.

This Chapter is organised as follows: A brief history of modelling is mentioned in Section 5.1 and system modelling is discussed in Section 5.2. Section 5.3 covers both the description of time-based and event-based dynamic identification. The event-based identification technique is then discussed in detail in Section 5.4. Section 5.5 and Section 5.6 discuss the excitation signal and the selection of a PRBS signal for dynamic identification. The following Section describes the open-loop (OLDI) and closed-loop (CLDI) dynamic identification process. The use of proportional-integral (PI) control for CLDI is also described in Section 5.7. Section 5.8 describes several model structures for the identified model. The practical test setup (hardware and software) for dynamic identification is described in Section 5.9. The results of the dynamic identification process starts in Section 5.10, followed by the identification results, model validation and model selection. Section 5.11 discusses several model identification issues.

5.1 Introduction

Modelling is a process to obtain a relationship (in the mathematical sense), between system variables, which represents an actual system. This process is widely used in areas like engineering, economics, biology, medicine, the life sciences, speech recognition, astronomy and agriculture. Application areas may be found as in the modelling of the human body (knee and leg movement) [Jones 2000, Lim 2002], modelling of the engine for engine fault diagnosis [Twiddle 2001, Goh 2002] and for control system design [Bhatti 1999]. Today, many engineering applications require system models which normally represent some engineering processes.

There are many identification methods and model structures available to estimate the dynamics of a system. This chapter concentrates on both the Auto-Regressive with eXogenous input (ARX) and Auto-Regressive Moving Average with eXogenous input (ARMAX) identification method to identify the diesel engine model for idle speed control. During the identification process, several areas are taken into consideration such as special techniques for data acquisition, hardware design, data processing, the modelling algorithm and structure and assumptions made.

This chapter presents an introduction to system modelling and the general procedures used. It focuses on the system dynamic identification method. The chapter also presents the application of advanced techniques, specifically a crank angle event-based technique, to identify a dynamic model of the speed (at 1500rpm) control system of a diesel engine. During the identification test, a proportional-integral (PI) control system is employed to run the engine in closed-loop in a stable condition. The detail of the PI control system design is described in this chapter. A linear model is derived using the ARX method where a pseudo-random binary sequence (PRBS) signal is injected into the control signal. Lastly, experimental validation is performed using data from the engine.

5.2 System Modelling

The modelling process deals with the problem of building a mathematical model (i.e. a set of equations) to represent the dynamics of an actual system. The process requires two kinds of knowledge. One is the actual knowledge and insights of the process's way of functioning and its properties. The other is the knowledge of how these facts can be transferred into a useful model [Ljung1994]. In general, there are two methods to model the actual system.

1. *Dynamic analysis model* (system analysis) - the model is obtained by analysis on the physical system at a fundamental level (i.e. using first principles) and involving

sufficient approximation to simplify the model to a differential equation form. This is an analytic approach. This method was used to model the cooling system in the previous chapter.

2. *Dynamic identification model* (system identification) - the model is obtained by inference from the observed behaviour of the physical system (i.e. input and output signals of the system). This is an experimental approach. Some experiments are performed on the system. A model is then fitted to the recorded data by assigning suitable numerical values to its model.

Dynamic analysis modelling provides many details of the physical system. It is most desirable to derive the physical model using first principles. However, this may not always be feasible. The generation of models from first principles requires considerable effort, accuracy and detailed knowledge of the engine. Representative work using dynamic analysis models (a physical engine model) may be found in the references [Cartona 2001, Gertler 1995, Dalton 1998, Chin 1986] and great understanding of the physical system is required. Hence this identification method is often impractical for model based controller (MBC) design. When this is the case, the second method is considered. It is also known as a 'black-box' modelling method [Soderstrom 1989]. This method is employed for the diesel engine model identification. The model is obtained using system I/O signal information and the identification process is iterative. The schematic in Figure 5-1 shows the dynamic identification model process. The model that fulfils the I/O relations the best is chosen.

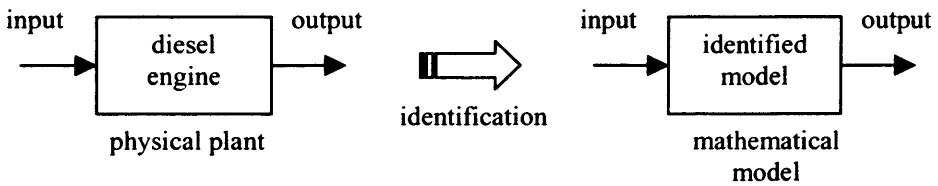


Figure 5-1: Dynamic identification models.

5.3 Time-based versus Event-based Dynamic Identification (TBDI versus EBDI)

In dynamic identification, the choice of data acquisition method is important. The most common method is a time-based method (TBM). The generated data is in a time series. Another, more advanced, method is the event-based method (EBM). The EBM generates data which depends on certain events that occur in the system. This section describes both the TBM and EBM in detail. It also discusses the advantages and disadvantages of both methods in relation to the diesel engine.

5.3.1 Time-based Method for Dynamic Identification

In many scientific experiments, the observed variables are commonly recorded and plotted on a graph against time. The recorded signals show the processes change via a time series. Equations 5-1 and 5-2 show examples of time-series data with a total number of n data points.

$$\text{System input, } u(t) = [u(1), u(2), u(3), ..., u(n)]$$

Equation 5-1

$$\text{System output, } y(t) = [y(1), y(2), y(3), ..., y(n)]$$

Equation 5-2

This method is sufficient to monitor the system dynamic behaviour throughout the data-recording phase. The details of the system dynamics, assuming the sampling rate is sufficiently high, are recorded at every sampling interval without losing important information about the system. It is the most commonly used method for dynamic identification.

5.3.2 Event-based Identification

The event-based identification method is independent of time but synchronised with the events which occur during the process. This method is normally applied to more complex systems such as the diesel engine, which has complex dynamics. Studies of event-based identification of engines can be found in [Chin 1986, Jones 1995, Vantine 2001]. The dynamics of an internal combustion engine are intrinsically linked to the combustion events [Chin 1986] and hence it is believed that the system dynamics are linked to these events. It only picks up the specific system dynamics which are captured along with the series of events. The EBM sample the data at the point where certain events occur and not the whole cycle of the combustion system. Part of the undesired system dynamic can be filtered out in the process. A lower sampling rate is required for EBM compared to TBM. Dynamics outside the events are not of interest and not recorded. This also frees up computing processor resources.

One of the disadvantages of this method is that a special hardware setup is needed to capture the event. This also depends on the accessibility of the events. Fitting sensors on high-speed moving parts of an engine and in the presence of high internal cylinder pressure is expensive and possibly dangerous. Besides, a special software technique is required to form a triggering system for data acquisition at every single event.

5.4 Advanced Techniques for Event-based Identification

The previous section compared and described both the TBM and EBM. An advanced EBM technique is employed here using a synchronous event-based crankshaft angle method. This method was reported in use on a diesel engine for dynamic identification as early as the 1970s [Flower 1977-a, Flower 1977-b, Vantine 2001]. Flower's investigation involved data sampling via timing-pulses, which were obtained from a pick-up sensing magnet on the flywheel. Vantine analysed the event-based diesel engine model for control purposes and compared it to the more traditional time-based model.

5.4.1 Engine Processes/Events Consideration

A four-stroke diesel engine works on the basis of four cylinder processes (i.e. stroke/cycle): intake, compression, ignition (or expansion) and exhaust. These events each occupy approximately 180° of crankshaft angle. It is assumed that each cycle (i.e. 180° of crankshaft angle) generates a unique engine dynamic. Thus, data that simulates at a fixed crankshaft angle gives the unique overall engine dynamic response. This method has seen application in some conventional engine controllers and it is said to be an accepted practice to sample the engine variables synchronously with these events [Hendricks 1994].

The technique used in this research study is to configure the data acquisition system to sample data at the crankshaft angle where the cylinder 1 (nearest to the radiator) is at its top-dead-centre (TDC) and bottom-dead centre (BDC). This method gives a better estimation of the diesel engine model, as it is believed that the dynamics of an internal combustion engine are highly linked to the combustion events and hence the engine speed.

5.4.2 Hardware Design and Setup

The crankshaft angle identification technique involves special hardware and software design. The author designed both the hardware and software. The workshop assistant in the department fabricated the hardware. Full details of the design work are described in Appendix E. The hardware construction consists of an aluminium disc, a surface reflecting sensor, TDC indicator and appropriate wiring to the dSPACE data acquisition system. The disc is mounted on one end of the crankshaft rotating at equal speed. On the aluminium disc, the designed two through-holes, each 180-degree apart, are used to indicate the TDC and BDC when they pass through the sensor. This sensor generates a pulse when each hole passes through it. Graphs in Appendix E show how the time-based data is captured, processed and finally produces the event-based data set for dynamic identification. The schematic in Figure 5-2 shows the process flow of the event-based data acquisition. The triggering system is built in the Matlab/Simulink environment by using the event information of the engine. The system produces a series of triggering/timing pulses for data acquisition. The event-pulses are also used to control the injection of the engine input signal (i.e. control signal) and the PRBS signal.

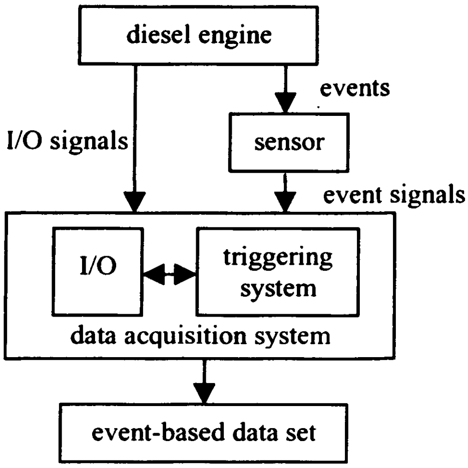


Figure 5-2: Process flow of event-based data acquisition.

5.4.3 Measurement of Engine Speed

Engine speed is sampled synchronously with the mentioned event-pulses (from the two reflective sensors on the aluminium disc). The sampling interval of the speed signal was established by the duration between two event pulses. An example of engine speed measurement using the event-based method is shown in Figure 5-3. The speed reference signal has a pseudo random binary sequence signal injected into it and the resulting speed signal can be seen to fluctuate around 1500 rpm. The first graph in Figure 5-3 is the continuous speed signal in the time base. The second graph shows the event pulses occurring during one second and the third graph shows the event-based speed signal.

5.5 Excitation Signal

An important concept in dynamic identification is the persistence of excitation. If operating loads[1] (during steady state operation) are insufficient for dynamic tests of the system under investigation, the system has to be excited by a suitably chosen excitation signal. The excitation signal is sometimes known as a perturbation signal.

[1] The operating loads refer to the unbalanced forces of the rotating engine (diesel generator) at zero kW of electrical load.

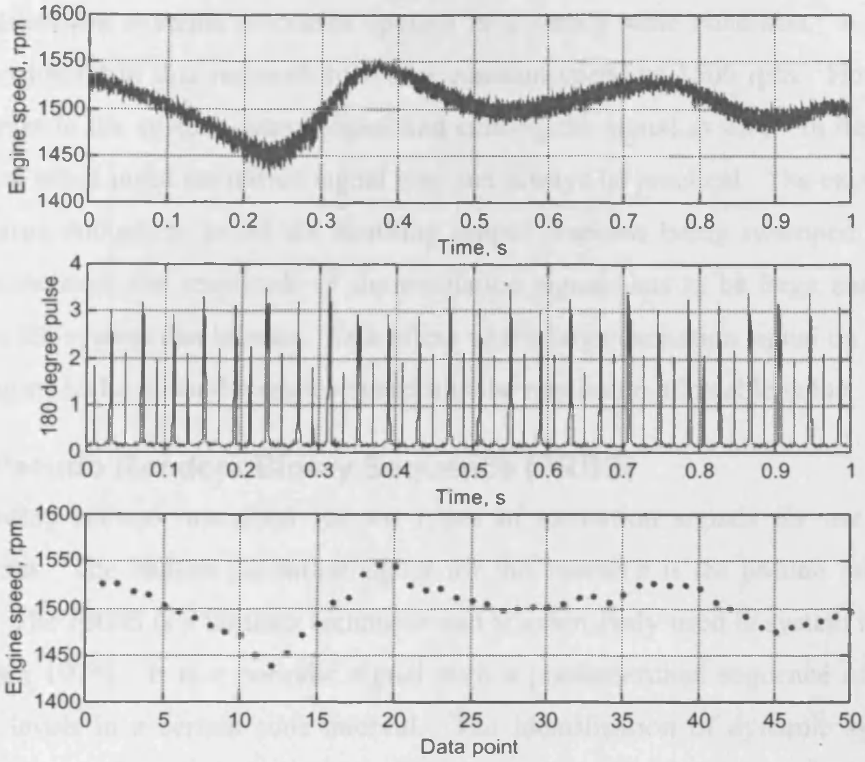


Figure 5-3: Speed signal recorded using different techniques.

It is normally smaller in amplitude compared to the input signal of the system. With the injection of the excitation signal, ideally, all the modes of the process are excited during the identification process. In experimental design, the selection of the test signal is a topic for off-line identification. The choice of excitation signal depends on the application, available equipment, the presence of measurement noise and time available in which to execute the test [Brown 1977]. Traditional techniques involve subjecting the system to step, ramp, pulse or sinusoidal input signals, and then carrying out relatively simple analysis of the output response curves [Schwarzenbach 1992]. A detailed survey on the excitation signal can be found in [Brown 1977, Schoukens 1988]. The characteristics of the signals were introduced as well as the advantages and disadvantages of the different signals over one another. The following section describes various excitation signals, their properties and the selection of the signal for dynamic identification. The examples of excitation signal are sinusoidal, pseudo-random, periodic chirp, periodic random and transient signals [Brown 1977, Natke 1988].

5.5.1 Limitations of the Excitation Signal

Many engineering systems/processes operate in a steady state condition. e.g. the diesel engine considered in this research runs at a constant speed of 1500 rpm. However, noise often appears in the system output signal and causing the signal to vary. In the presence of this noise, a small input excitation signal may not always be practical. The excitation signal must be large enough to avoid the resulting output response being swamped by the noise signal. Sometimes the amplitude of the excitation signals has to be large and more often larger than the system can tolerate. One effect of the large excitation signal on the engine is that the engine stalls when the engine speed hits the maximum allowable value.

5.6 Pseudo Random Binary Sequence (PRBS)

The preceding section described various types of excitation signals for use in dynamic identification. The chosen excitation signal for this research is the pseudo random binary sequence. The PRBS is a familiar technique and is extensively used in system identification [Verbruggen 1975]. It is a periodic signal with a predetermined sequence of $+A$ and $-A$ amplitude levels in a certain time interval. The identification of dynamic systems using PRBS test input sequences has received great attention since the 1960s [Briggs 1966, Funahashi 1974, Flower 1976a, Tzafestas 1977, Schoukens 1988].

5.6.1 Description of the PRBS

The periodic PRBS signal can be generated by a number of sequences [Soderstrom 1989] of ones and zeros, with sequence length, $N_{data} = 2^n - 1$ with n positive real integer (i.e. $N_{data} = 15, 31, 63, 127, 255...$). A signal generated by this sequence is produced with a $+A$ level over discrete intervals of time, $t = k\Delta t$ (where $k = 0,1,2,...$), as the sequence produces a one and a $-A$ level as the sequence produces a zero. A possible change in signal level at time instants t is controlled by a clock pulse. The signal has period of $T_{per} = N\Delta t$ second. An example of a PRBS signal is shown in Figure 5-4.

5.6.2 A Special Technique For PRBS Injection

A special technique is employed to inject the PRBS signal. In the conventional method, the time interval of the PRBS signal points is set in the signal generator or software. In this research test, a different technique is used. Instead of fixing the time interval, the time interval is controlled by the event-pulses. Each event pulse triggers the PRBS signal point. This means that when the PRBS changes its signal level, the cylinder 1 is at its TDC and BDC. By applying this technique, the PRBS signal is forced to change the state according to the system event. Thus, the PRBS signal injection is synchronised with the data acquisition system.

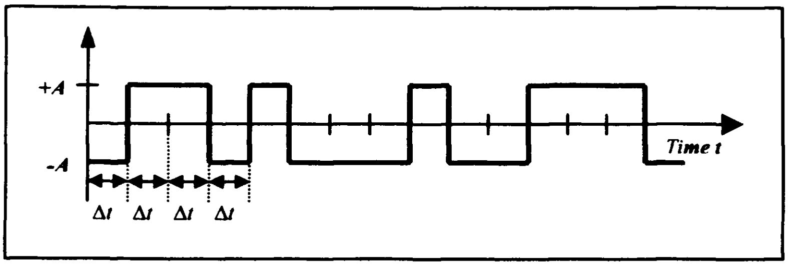


Figure 5-4: Pseudo random binary sequence.

Figure 5-5 shows how this technique is applied. A PRBS signal indexing (i.e. shown in the third graph) is designed using the discrete time integrator Simulink block which resets itself at the length of the PRBS. The increment of the index depends on the event-pulses. This index is then referred to a look-up table which contains the PRBS signal sequence. When the index swings from zero to the length of PRBS, a full period of PRBS is generated at the output. A close up on the signal is shown in Figure 5-6 (not to scale). The vertical line and the stair-like graph represent the event-pulses and the indexing respectively. The PRBS (square graph) is shifted by 2ms to show the signal changing state more clearly. This shows that the PRBS signal and PRBS index value change when the event-pulses trigger.

5.6.3 Advantages of PRBS

One of the advantages of PRBS is that it is easy to generate and introduce into a system using the available software and hardware, i.e. the Matlab software and dSPACE data acquisition hardware. The PRBS signal is random within a period and signal frequencies vary according to the binary pulse width. This random behaviour forms a wide frequency range in the signal. This allows the system dynamic to be excited over a wide frequency spectrum. A periodic PRBS signal and averaging many cycles of PRBS are reported to reduce random measurement noise [Schwarzenbach 1992].

5.7 Identification approach

In general, there are two broad classes of control system, namely, open-loop control and closed loop control. Figure 5-7 shows the variables involved in both of the systems. The dynamic identification of these different systems requires some insight and practical considerations. It is widely thought experiments carried out in open loop are more straightforward. The following sections describe both the open-loop and closed-loop issues

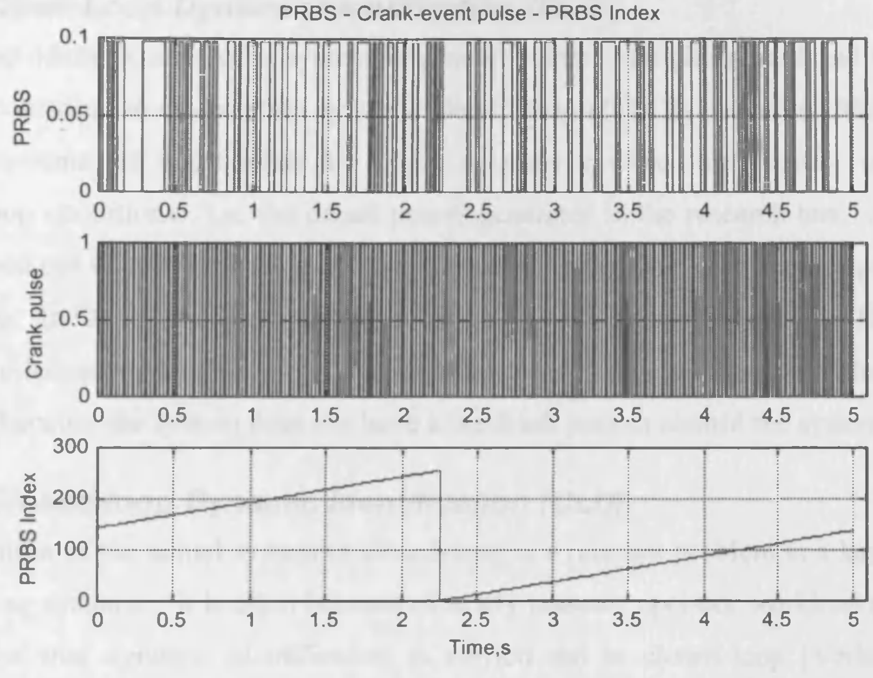


Figure 5-5: Event-based triggering of PRBS states.

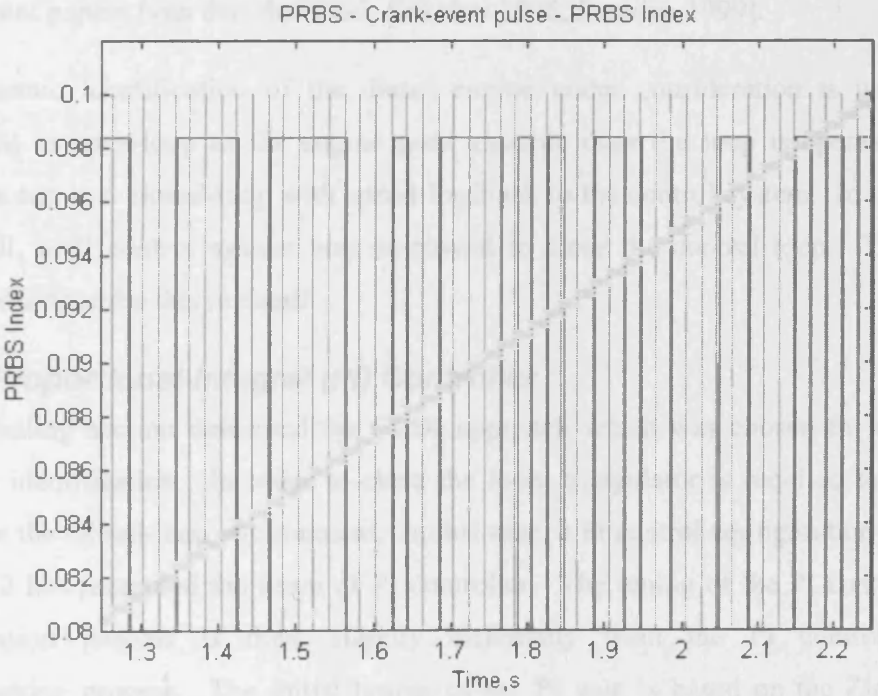


Figure 5-6: The triggering pulse

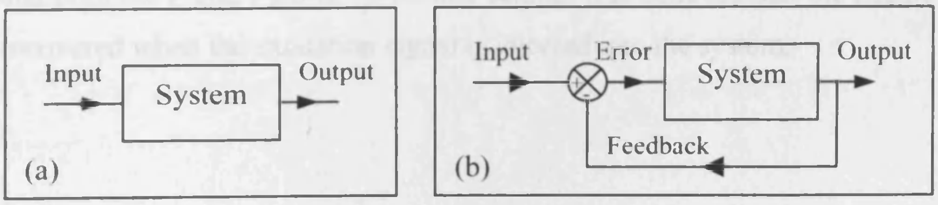


Figure 5-7: Open-loop (a) and closed-loop (b) control system.

5.7.1 Open-Loop Dynamic Identification (OLDI)

Open-loop identification of a system is generally more straightforward and the preferred method compared to closed-loop dynamic identification (CLDI) [Pasadyn 1999]. However, not all systems are appropriate for OLDI as some systems must strictly operate under closed-loop conditions. i.e. the diesel power generator in the research test. A simple test was carried out with a fixed control value applied to the engine. The engine speed drifted in open-loop. In OLDI, the excitation signal magnitude is an issue of concern. The magnitude of the excitation signal is set in such a way that the system must operate without becoming unstable because the system does not have a feedback path to control the system.

5.7.2 Closed-loop Dynamic Identification (CLDI)

Identification of the actual system in closed-loop is a relevant problem in a large number of engineering systems. It is often because of safety reasons, operator workload or production restrictions that dynamic identification is carried out in closed-loop [Verhaegen 1993]. Theoretical study of CLDI is documented in an early contribution [Soderstrom 1975] and in more recent papers [van den Hof 1995, Sokolov 1996, Pasadyn 1999].

The dynamic identification of the diesel engine under consideration is unlikely to be performed in open-loop as the engine goes unstable once the loop is opened. Thus, the system is run in a closed-loop with speed feedback to the control system. In order to carry out CLDI, a PI control system was employed to close the control loop. The following section will describe this in detail.

5.7.3 Proportional-Integral (PI) Controller

The preceding section described the CLDI approach which was chosen for diesel engine dynamic identification. In order to close the loop, a regulator is required to regulate the system in the closed-loop environment. In this case, a PI control configuration is employed. Chapter 2 has presented the setup of PI controller. The tuning of the PI controller for the identification process is done slightly differently from the PI controller for the benchmarking process. The initial tuning of the PI gain is based on the Ziegler-Nichols method [Ziegler 1942, Ogata 1997]. The secondary tuning is to loosen the control action by decreasing both the P and I gains. With this setting, it is believed that the engine dynamics can be recovered when the excitation signal is injected into the system.

5.8 Dynamic Identification Methods

In dynamic identification, when a model structure is decided, the next stage is to consider the type of identification algorithms most suitable to produce the selected model structure. Both the ARX and ARMAX are considered here.

(a) ARX Structure

The ARX model structure is represented by a simple linear difference equation as follows:

$$y(t)+a_1y(t-1)+\ldots+a_{an}y(t-na)=b_1u(t-nk)+\ldots+b_{nb}u(t-nk-nb+1)+e(t)$$

Equation 5-3

which relates the current output $y(t)$ to a finite number of past outputs $y(t-k)$ and inputs $u(t-k)$. This structure is defined by the three integers *na*, *nb* and *nk*. *na* is equal to the number of poles, *nb-1* is the number of zeros and *nk* is the pure time-delay in the system. The parameter/coefficient vector is taken as

$$\theta=(a_1 \ldots a_{na} \;\; b_1 \ldots b_{nb})^T$$

Equation 5-4

A least squares method is used to estimate the coefficients *a* and *b* in the ARX model structure. This method minimises the sum of squares of the right-hand side minus the left-hand side of the expression above, with respect to *a* and *b*.

(b) ARMAX Structure

Let $y(t)$ and $u(t)$ be scalar signals (i.e. output and input signals) and consider the ARMAX model structure

$$\begin{aligned}&y(t)+a_1y(t-1)+\ldots+a_{an}y(t-na)=b_1u(t-nk)+\ldots+b_{nb}u(t-nk-nb+1)\\&+e(t)+c_1e(t-1)+\ldots+c_{nc}e(t-nc)\end{aligned}$$

Equation 5-5

where $e(t)$ is an uncorrelated sequence of noise with zero mean and variance. The parameter vector is taken as

$$\theta=(a_1 \ldots a_{na} \;\; b_1 \ldots b_{nb} \;\; c_1 \ldots c_{nc})^T$$

Equation 5-6

ARMAX is a combination of the auto-regressive (ARX) and moving average (MA) processes.

For the sliding mode application, many of the design processes employ a linear state-space model [Utkin 1978]. It is logical to describe a sliding mode controller design technique based on linear models to tackle an industrial control problem. Hence, the final model structure is in state equation form.

5.9 Dynamic Identification Setup Procedure

5.9.1 Experimental Hardware Setup

The full hardware setup for the dynamic identification data acquisition test is shown in Figure 5-8. The magnetic-pickup pulses are obtained from flywheel teeth (i.e. the starter ring). These pulses are fed to a frequency-to-voltage converter before being sent to the dSPACE system.

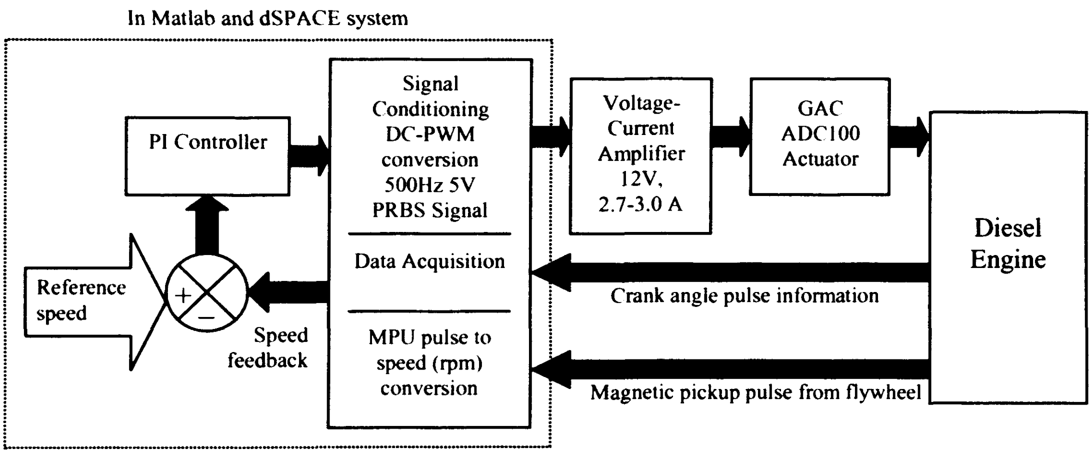


Figure 5-8: Overall setup for dynamic identification

5.9.2 Practical generation of the PRBS Excitation Signal

The PRBS generation procedure in [Ljung 1999] was used. Different lengths of PRBS were designed. The injection of the PRBS (the change of state) is synchronised with the crank-angle of the engine and this will excite the engine dynamic at a particular combustion cycle and hence give a prominent event excitation.

5.9.3 dSPACE-Matlab-Simulink Setup

The Matlab Simulink setup is shown in Figure 5-9. The signal-conditioning block contains sensor gain factors and it converts the measured signals their real value. It is used to produce a series of triggering pulses depending on the event pulses. The PRBS injection, control signal generation and data recording are synchronised with the event-pulses.

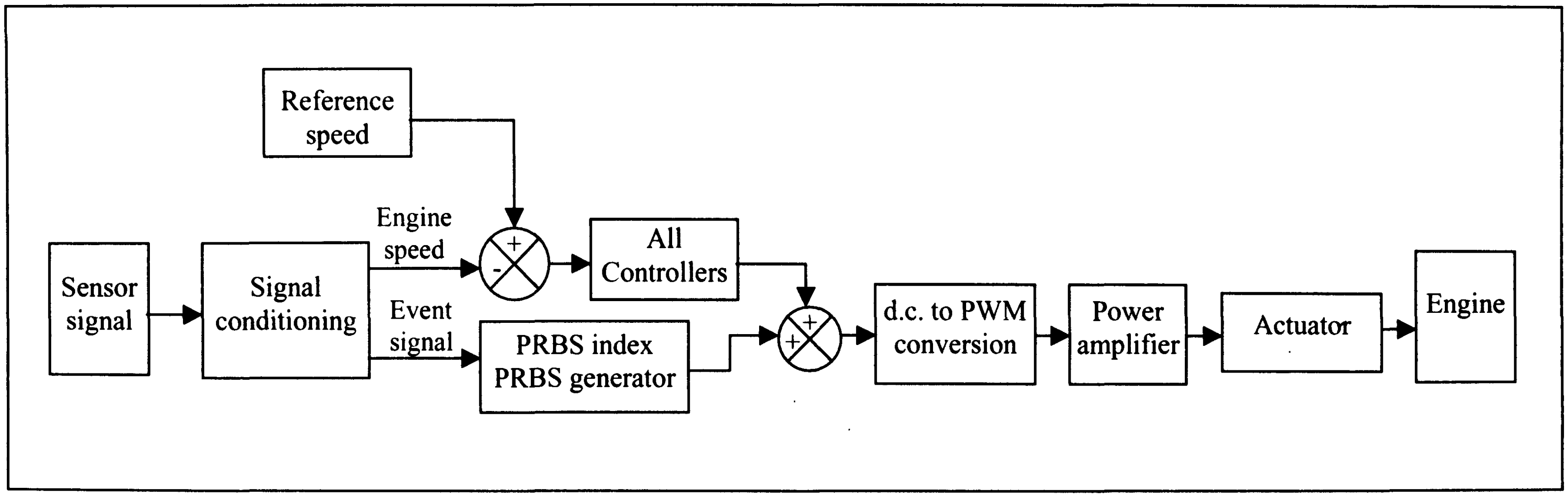


Figure 5-9: Matlab/Simulink setup for the identification test.

The PRBS sequence is produced using a Matlab command called '*idinput*'. The amplitude of the PRBS is set between –0.1 and 0.1 which the unit is equivalent to the ±10% duty cycle of the control signal (in term of PWM signal). This value is also equivalent to 20% of the input control signal range which is large enough to excite the engine without tripping the generator set speed limiting switch. The engine speed was perturbed within the 1400rpm and 1600rpm range. The test is designed in a way that the PRBS amplitude can be adjusted online for different amplitude level test and data collection. Two different lengths of PRBS sequence are considered in the test, for example 255 and 1023. An example of a PRBS signal of length 255, at amplitude –0.1 and 0.1 is shown in Figure 5-10.

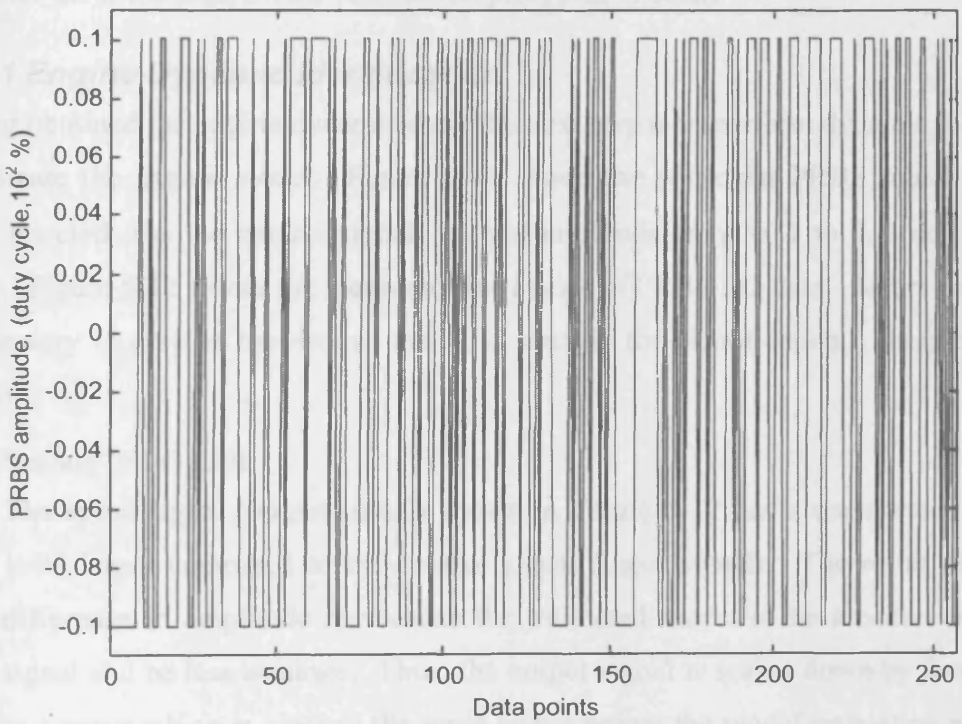


Figure 5-10: PRBS signal of length 255.

5.9.4 Sampling Rate

The sampling rate is determined by considering the highest frequency parameter in the test, which is 50 Hz from the event pulse sensors. According to the Nyquist-Shannon result in Appendix B, the sampling frequency must be greater than twice the highest frequency of this signal in order to be able to reconstruct the original signal perfectly from the sampled version. Thus, the sampling frequency is set to 10,000 Hz. However, due to hardware limitations, the identification data is only recorded at lower rates of 100 Hz and 50 Hz.

5.10 Dynamic Identification: Results, Model Analysis and Verification

The goal of the identification procedure is to obtain a good and reliable model. To do this, a few tests are introduced. First the *a priori* assumptions of the system, such as linearity are tested. The experiment is repeated with a few different PRBS amplitudes in order to verify for what operating range a linear model is adequate. Another aspect is to test the time invariance of the system. A convenient way to conduct this test is to divide the recorded data into two parts. The model is identified using the first data set. Then the model output is computed for the second set of data using the identified model. For a time invariant process, the model should produce the output equally well for both data sets. This procedure is sometimes called *cross-checking* or *cross-validation* [Soderstrom 1989]. The following describes the mentioned model verification processes in detail.

5.10.1 Engine Dynamic Identification

Having obtained the engine dynamic data, the next step is to perform dynamic identification to estimate the engine model. Figure 5-11 shows the particular PRBS signal which was being injected into the control signal. It has amplitude from –0.1 to 0.1 (i.e. ±10% duty cycle). Figure 5-12 shows the measured five cycles of PRBS I/O data. Before estimation, it is necessary to process the data so that it is suitable for identification. The preprocessing involves:

(a) Scaling of I/O Data

The speed signal (output signal) shown in Figure 5-12 has a scale which is a factor 1000 larger compared to the control signal (input signal). The result of this large difference in amplitude may cause the estimated model to be sensitive to the input signal and be less accurate. Thus, the output signal is scaled down by factor of 1000 to a range which is close to the input signal before the model estimation process. At the end of the estimation process, the output signal of the model is then re-scaled back by a factor of 1000. For a state-space model, the output matrix C has to be scaled.

(b) Trend correction

In an industrial process, the system output may tend to drift with time. This variation is called trend. Trend can be compensated by a controller but its presence in the estimation data affects the accuracy of the identified model. Removing this effect is called detrending. Detrending can be done using the Matlab command '*dtrend*'. This command can remove the mean values or linear trends from the data. The detrended measured engine I/O signals are shown in Figure 5-13.

(c) Selecting Data Ranges

It is often the case that the whole estimation data is not suitable for identification, due to various undesired features, such as, missing or 'bad' data, outbursts of disturbances, level changes etc. Only a portion of the measured data can be used for estimation purposes and another portion for validation purposes.

(d) Prefiltering

The measured signal usually contains 'noisy' high frequency components which are undesirable for model estimation. This high frequency component can be filtered out through a linear filter and focus the model's fit to the system to a specific frequency range. This high frequency component filtering does not apply to the data process as every signal point of the event-based data is needed for model estimation.

5.10.2 Dynamic Identification Results

During the data acquisition stage, the data are recorded at different loading condition (i.e. 0 kW, 10kW, 20kW, 30kW, 40kW, 50kW and 60kW), at 1500rpm and different levels of PRBS amplitude (i.e. 0.1, 0.2 and 0.25). The experimental I/O data are recorded for ten cycles of PRBS signal. These data are divided into two equal size data sets. The first five cycles of the signal are used for dynamic identification and the other five cycles are used to validate the identified model. The estimated model is returned in *theta* format. This is a matrix containing information about model structure, estimated parameters and their estimated accuracy (refer to [Ljung 1999] for detail of the theta structure).

5.10.3 Validation of Identified Models

Model validation is the heart of the identification problem because there is no absolute procedure for approaching it. A variety of different tools are available to evaluate the model qualities. A few of these tools are addressed here.

(a) Data Correlation

The first stage of the validation process involves comparing the output of the second half of the five cycle data at the identified model output. A Matlab command called '*compare*' is used. The input signal (i.e. the control signal) is fed into the model and the output is obtained. A graph of both output signals is plotted and a *mean square fit* value (MSFV) is generated. The smaller this value, the better the fit of the identified model to the actual system. The best-fit model is selected.

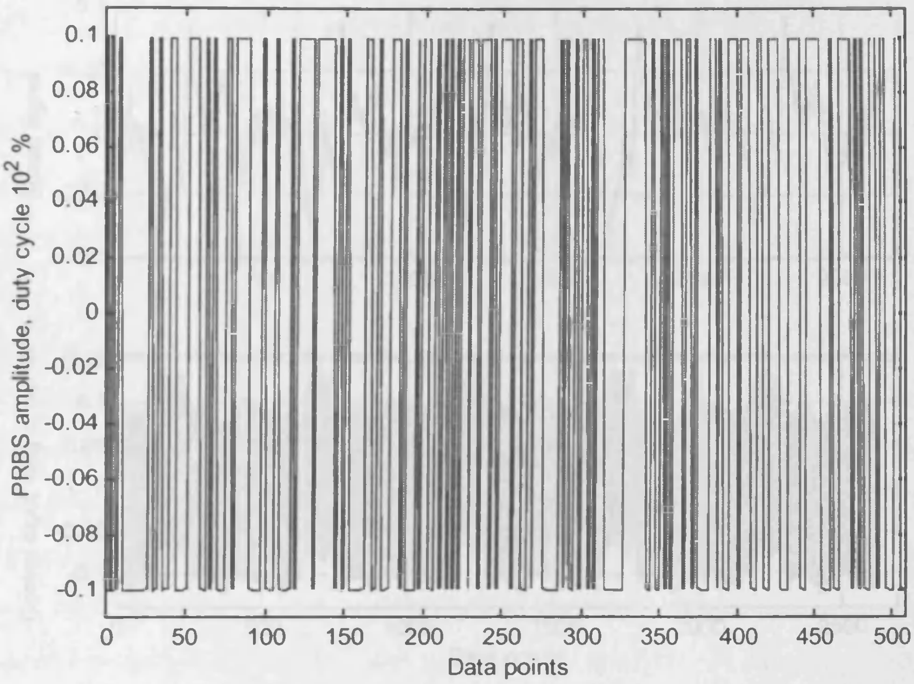


Figure 5-11: Injected PRBS signal. (zoom at first 500 data points)

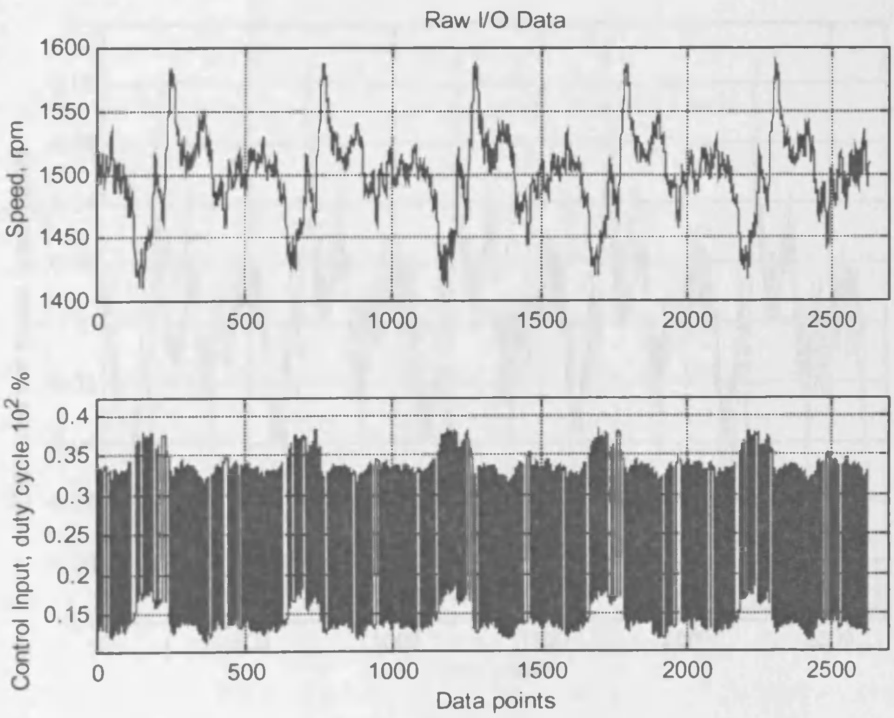


Figure 5-12: Raw data: Speed and control signal.

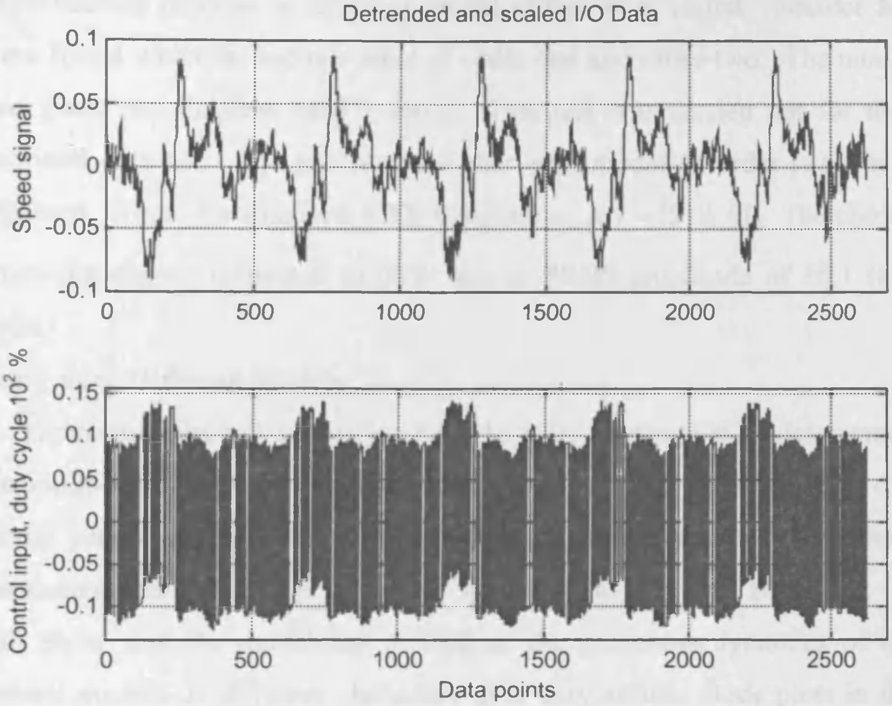


Figure 5-13: Detrended data: Speed and control signal.

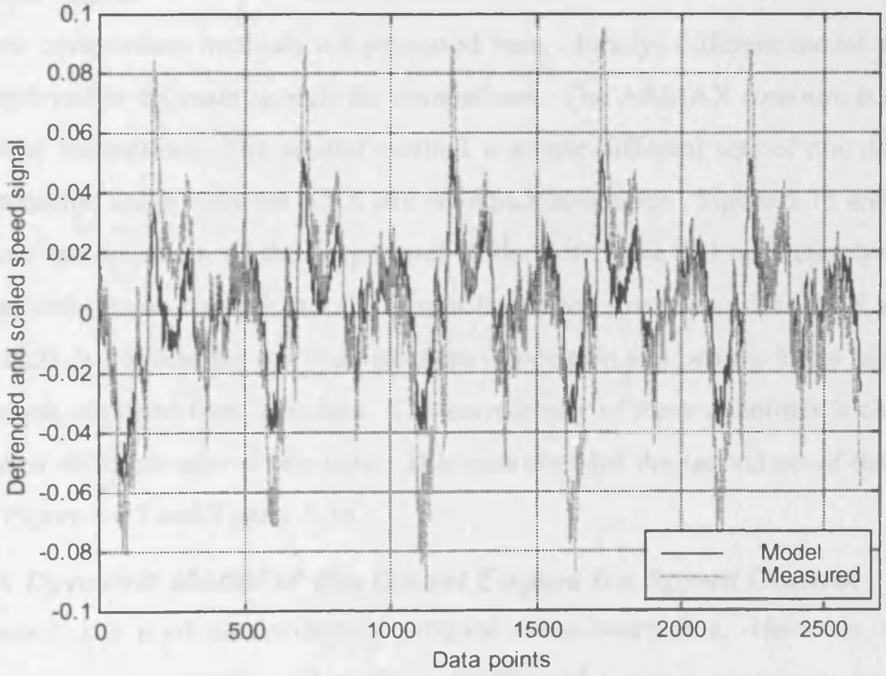


Figure 5-14: Model validation using second half of the 5 PRBS cycle data.

(b) Mean Square Fit Value (MSFV)

The MSFV is a value which is calculated from the measured data to estimate the quality of the obtained model. It is used to evaluate how well the set of model data fits the measured data set. The smaller the MSFV value is, the better. In the identification process, a series of model orders were tested. Smaller MSFV values were found when the models were of order one and order two. The model with order two gives the smallest MSFV error. This test was carried out for the rest of the obtained data set. The test revealed that with model of order two, the best fit was obtained. Thus, the finalised ARX structure is $NN = [2\ 2\ 0]$. The chosen data set is when the engine subjected to 0kW and at PRBS amplitude of ±0.1 (i.e. 20% duty cycle).

(c) Comparing Different Models

A very conveneint way to analyse a model is by plotting the model parameters such as the input-output response as previously demonstrated. Alternatively, displaying the model properties in terms of quantities will have more physical meaning than the parameters themselves, i.e. frequency spectrum analysis on a Bode plot. This analysis will show that the model has picked up the properties/dynamics of the plant. If several models of different characters give very similar Bode plots in the frequency range of interest, one can be fairly confident that these must reflect true features of the diesel engine.

Two comparison methods are proposed here. Firstly, different model structures are employed to estimate models for comparison. The ARMAX structure is used here for model estimation. The second method is to use different sets of raw data for model estimation using both the ARX and ARMAX structures. Figure 5-15 and Figure 5-16 show the spectrum of the output and Bode plot of the I/O respectively (from the 1st data set). It can be seen that the output frequency spectrum of the ARX and ARMAX models fall within the raw data frequency spectrum and behave in the same fashion as the one obtained from raw data. The consistency of these quantities is checked by the use of different sets of raw data. The quantities for the second set of data are plotted in Figure 5-17 and Figure 5-18.

5.10.4 A Dynamic Model of the Diesel Engine for Speed Control

The obtained data is of non-uniformly sampled event-based data. However, it is assumed that the variations in sampling times are negligible and hence, a continuous time model can be obtained using the available identification toolbox from Maltab™. The final model in theta format is then converted to a state space model format (i.e. the triple (A, B, C)) conveniently by using the Matlab command. The output matrix of the state space equation, as it has been previously scaled down, is now scaled back to its actual value. Thus, the second order state-space model for engine idle speed control is as follows:

$$\dot{x}(t) = \begin{bmatrix} -0.7225 & -0.0039 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 23.2177 & 17.6746 \end{bmatrix} x(t)$$

Equation 5-7

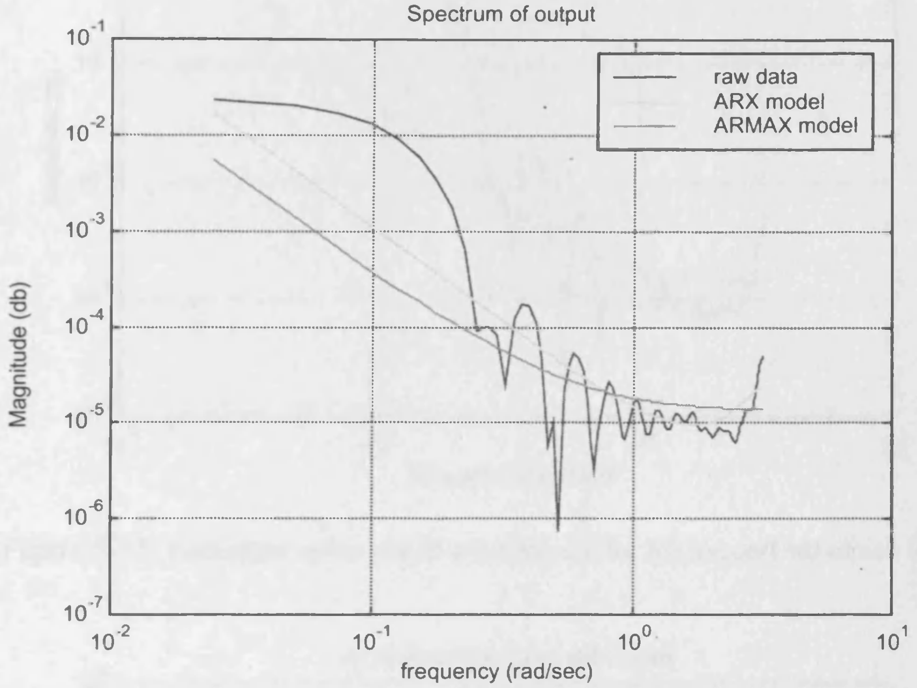


Figure 5-15: Estimated noise spectrum analysis for the first set of raw data.

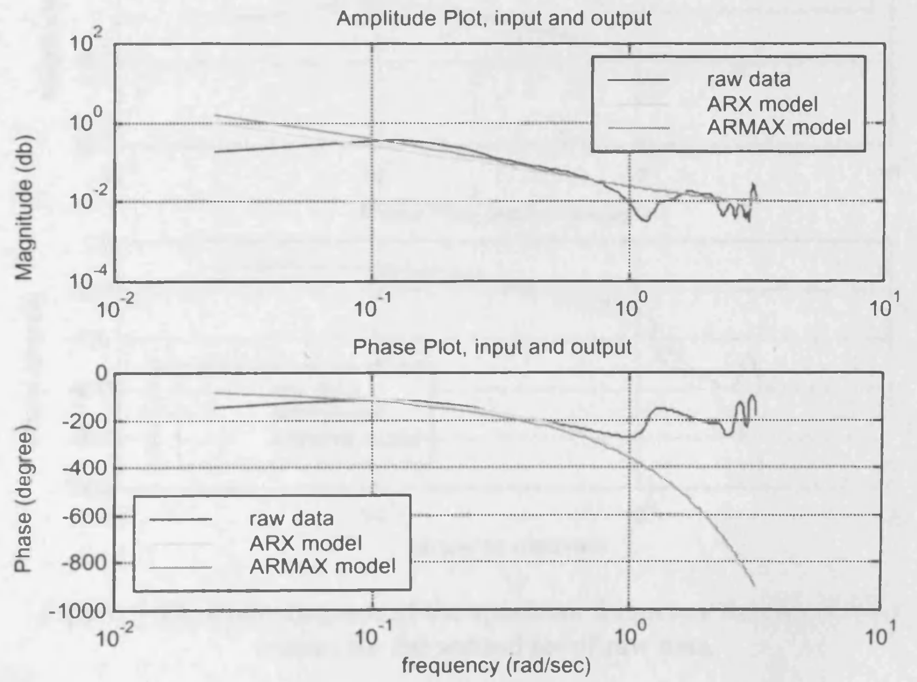


Figure 5-16: Bode diagram of the spectrum frequency function of the input and output for the first set of raw data.

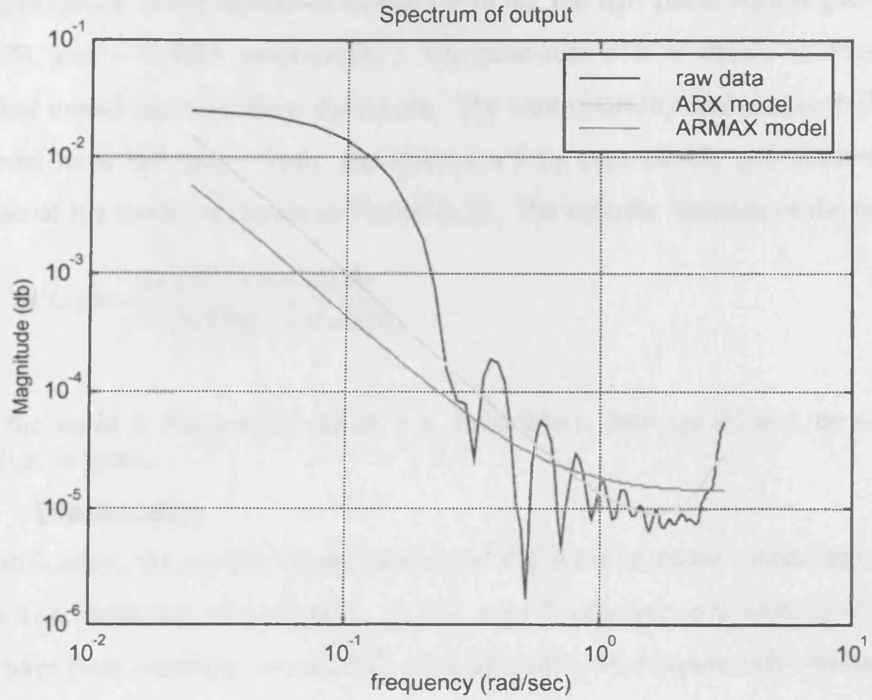


Figure 5-17: Estimated noise spectrum analysis for the second set of raw data.

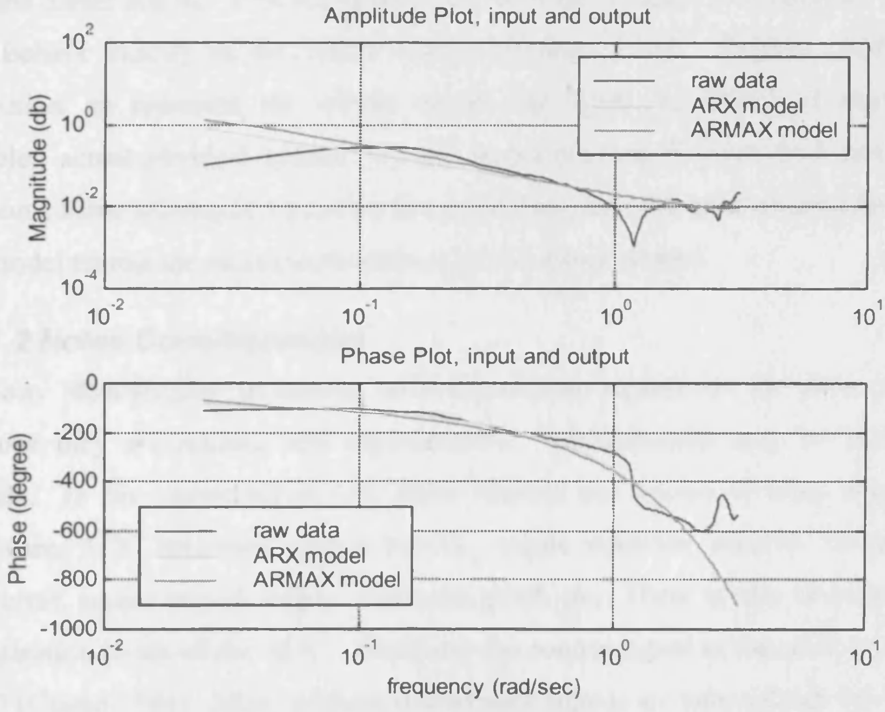


Figure 5-18: Bode diagram of the spectrum frequency function of the input and output for the second set of raw data.

The eigenvalues of the identified model are in the left half plane with negative real parts of -0.7170 and -0.0055 respectively. The pole-zero plot is shown in Figure 5-19. The identified model has very slow dynamics. The controllability and observability matrices for the model have full rank. Thus, the system is fully controllable and observable. The step response of the model is shown in Figure 5-20. The transfer function of the model is

$$W(s) = \frac{23.2177s + 17.6746}{s^2 + 0.7225s + 0.0039}$$

Equation 5-8

where the input is the control signal (i.e. from 0 to1, duty cycle) and the output is engine speed (i.e. in rpm).

5.11 Discussion

In identification, the model ideally represents the actual process. Although there are some aspects (i.e. selection of excitation signal, model structure and identification algorithms) which have been carefully considered, there are some other issues to be discussed.

5.11.1 Transition of the actual system to model

The transition of the actual system to a model is not perfect, especially for a complex system like the diesel engine. It is not possible to construct a model of a complex real system that will behave exactly as the actual system [Popken 1999]. Popken used the term 'low resolution' to represent the system model and 'high resolution' to represent the more complex actual physical system. In the transformation between high and low-resolution versions, some information must be lost (by definition). An information loss is acceptable if the model retains the main characteristics of the actual system.

5.11.2 Noise Considerations

In many identification processes, noise/disturbance signals are the main obstacles simply because they are random and unpredictable. Measurement may be corrupted by such signals. In the identification test, noise sources can appear at many stages. e.g. DAC hardware, ADC hardware, engine battery, engine vibration, sensors, frequency-to-voltage converter, sensor power supply, electrical glitch etc. There is also noise generated by the quantisation levels of the ADC. Similarly the control signal to the plant is quantised by the DAC [Clarke 1984]. Many of these disturbance signals are unavoidable but their effects can possibly be minimised by repeating the test and averaging the recorded signal (here data were collected for 10 cycles of PRBS).

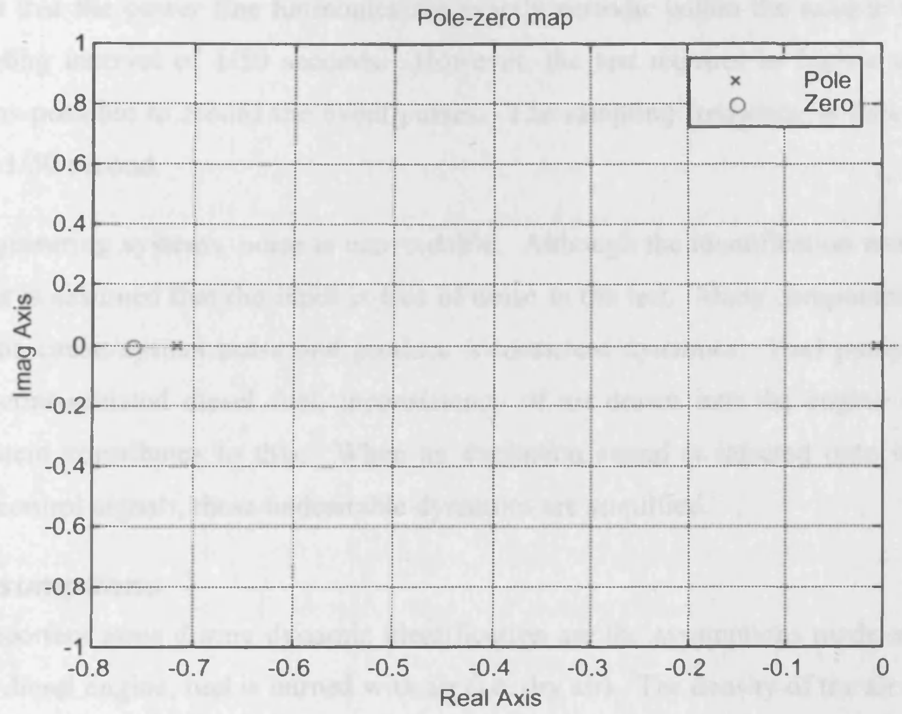


Figure 5-19: Poles and zeros of the identified model.

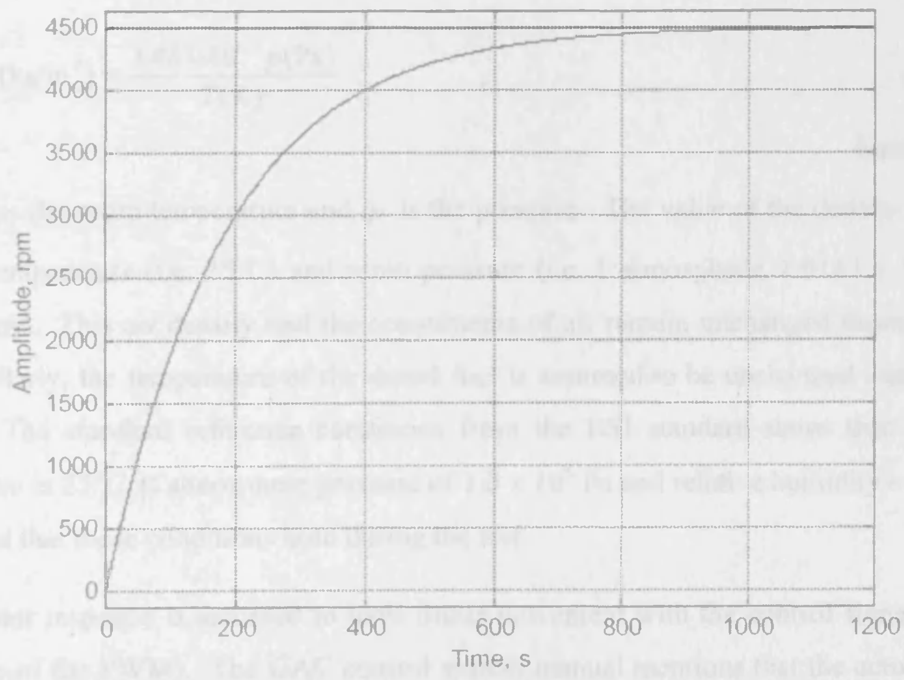


Figure 5-20: Step response of the identified model.

Having said that, some of the noise appearing in the system is consistent with periodic behaviour. One common source of periodic noise comes from electrical noise occurring at harmonics of the power line frequency. This noise can be minimised by choosing a sample period such that the power line harmonics are exactly periodic within the sample window, i.e. a sampling interval of 1/50 seconds. However, the test requires as high a sampling frequency as possible to record the event pulses. The sampling frequency is thus set to a multiple of 1/50 second.

In most engineering systems, noise is unavoidable. Although the identification results may be biased, it is assumed that the input is free of noise in the test. Many components of the diesel engine cause system noise and produce inconsistent dynamics. Fuel pump throttle responses, contaminated diesel fuel, inconsistency of air drawn into the engine or turbo charger system contributes to this. When an excitation signal is injected onto the input signal (i.e. control signal), these undesirable dynamics are amplified.

5.11.3 Assumptions

Another important issue during dynamic identification are the assumptions made about the test. In the diesel engine, fuel is burned with air (i.e. dry air). The density of the air depends on temperature (i.e. ambient temperature) and altitude. In the test, the air temperature is taken into consideration. The density of the air has the following relationship [Heywood 1988].

$$\rho(\text{kg/m}^3) = \frac{3.483x10^{-3}\, p(\text{Pa})}{T(\text{K})}$$

Equation 5-9

where T is the room temperature and p is the pressure. The value of the density of the air at room temperature (i.e. 25°C) and room pressure (i.e. 1 atmosphere, 1.0133 x 10^5 Pa) is 1.184 kg/m^3. This air density and the constituents of air remain unchanged throughout the test. Similarly, the temperature of the diesel fuel is assumed to be unchanged and so is the density. The standard reference conditions from the BSI standard states that the room temperature is 25°C at atmosphere pressure of 1.0 x 10^5 Pa and relative humidity of 30%. It is assumed that these conditions hold during the test.

The actuator response is assumed to have linear movement with the control signal (i.e. the duty cycle of the PWM). The GAC control system manual mentions that the actuator shaft rotation is proportional to the actuator current, hence the control signal.

5.11.4 Uncertainty

The uncertainty of a system is a set-valued model describing the variability of the plant dynamics and/or incomplete knowledge of the plant dynamics. Two important points were raised by [Van den Hof 1995]. Firstly, it is generally not possible to bound the system uncertainty in identified models and secondly, it is not clear what kind of models are best suited for model-based control design. Thus, the identification methods deliver a nominal model which is just an approximation to the actual system. Based on this nominal model, a controller has to be designed, assuming a certain level of accuracy of the nominal model.

5.11.5 Model-based Control Related Issues

Model-based controller design methods such as sliding mode control are highly dependent on the properties of the identified model. There is no specific model for the controller design as the possible suitable models for the controller design are large. There can be a large set of relevant models (i.e. nominal models) for a given actual system and a given controller design method. It is still not obvious how to obtain such models in a systematic manner using the methods of identification for nominal models. As a result, the identification is an iterative process. The whole process of identification is shown in Figure 5-21.

The controller is designed with reference to a mathematical model of the actual system. In general, this model only approximates the behaviour of the actual physical plant. In addition, the designed controller is never exactly equal to the desired design. The control system should, therefore, be insensitive to errors in the mathematical models of the actual system and the controller.

5.12 Conclusion

This chapter presented the dynamic identification of the diesel engine for idle speed control system design. Techniques such as hardware event-based techniques and the PRBS injection have been described and further details can be found in the Appendix E. A system model of 2nd order has been identified and is written as a state space equation as follows.

$$\dot{x}(t) = \begin{bmatrix} -0.7225 & -0.0039 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 23.2177 & 17.6746 \end{bmatrix} x(t)$$

The validation of the model has been carried out and cross validation has been considered using different model structures, orders and different data sets. Some identification issues and assumptions have been taken into account.

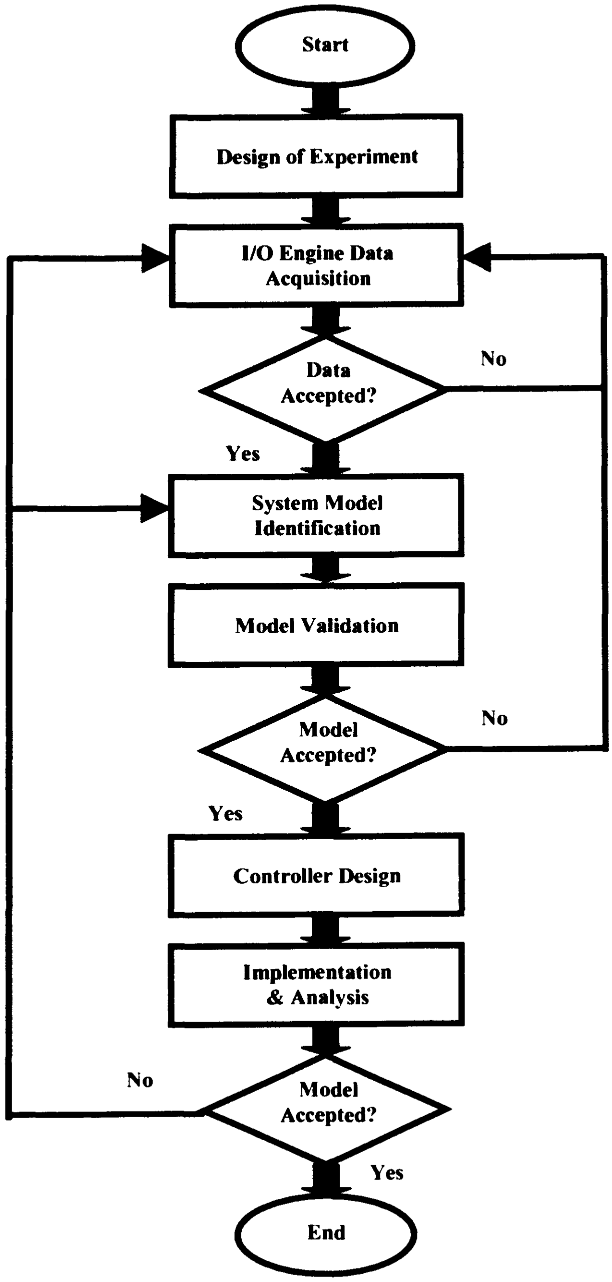


Figure 5-21: Dynamic identification process chart.

Chapter 6

6 Classical Sliding Mode Control

In the previous two chapters, the sliding mode concepts have been described and the engine model for speed control has been identified. This chapter describes the application of both the mentioned concepts and model on to a commercial diesel generator. The classical method of sliding mode control is considered in this chapter. Simulation of the model is first studied, followed by a number of controller designs. The designed controllers are then implemented on to the diesel engine and the performance is studied and discussed.

This chapter is organised as follows: Section 6.1 gives an overview of early techniques used for the speed control (i.e. at 1500rpm) problem. The speed control problem and controller design strategy are mentioned in Section 6.2 and Section 6.3 respectively. Section 6.4 describes the design of an integral action based sliding mode controller and Section 6.5 discusses several implementation issues. Section 6.6 presents implementation results. The design of a model following sliding mode controller is started in Section 6.7 with practical issues discussed in Section 6.8. Implementation results are described in Section 6.9. Several practical issues emerged during implementation and these issues are discussed in Section 6.10.

6.1 Overview

The engine speed control problem is no longer a new topic in the literature. One of the oldest methods is via the use of a centrifugal governor which is fully mechanically governed. This method is getting less popular now because advanced electronics based methods have emerged. Chapter 2 and Chapter 5 have mentioned the popularity of PID controllers in industry.

This Chapter will introduce some controller design frameworks to the diesel engine idle speed control problem. The frameworks will take account of all stages of the controller design process, from the obtained model (i.e. from Chapter 5) to the implementation on a diesel engine. Most importantly the controller design frameworks must allow non-experts to maintain and be able to tune the resulting controller conveniently. Thus, a simple and straightforward controller-tuning algorithm is required.

6.2 Diesel Engine Speed Control Problem

The main concern in generator speed control is to counter the possible disturbances in the nature of usage. The presence of a variety of disturbances affects the engine performance (i.e. cause the engine speed to vary). The engagement and disengagement of various electrical appliances and sudden demands of electricity during peak hour aggravate the problem further.

A disturbance rejection strategy is essential for proper speed control. While tracking the engine speed, the controller should also show robustness to cope with the non-linearities present in the system and at the same time, the controller should be capable of rejecting load disturbances. The common non-linearities present are friction, i.e. Coulomb friction and Stribeck friction [Armstrong-Helouvry 1994]. Chapter 2 has described several requirements for a speed controller. On the diesel generator, maintaining the engine speed at the reference speed setting means that the output frequency of the electricity will be constant.

The objective of the controller is to keep the engine speed at the reference speed, i.e. 1500 rpm and thus keep the output frequency at 50 Hz. This practical engine speed control problem has been approached in the past using different controller techniques [Banisoleiman 1990, Takahashi 1985, Bhatti 1999]. Banisoleiman *et al* proposed a strategy controlling the inlet manifold of the engine and increasing the air mass flow rate of an engine. At the same time, the fuel pump is set to produce full load Brake Mean Effective Pressure (BMEP) where BMEP denotes the average effective pressure of all stroke cycles. This system can improve the transient response of the engine speed. However, the control strategy involves at least two control variables (i.e. inlet manifold and fuel pump) and a complicated process. Takahashi *et al* applied a linear quadratic and integral (LQI) control strategy to a simple simulation engine model. The control process required three control variables. Bhatti *et al* proposed a sliding mode controller, which involved designing an observer to reconstruct the system states for use by the controller. This design was applied on to a car engine.

The idle speed control problem here is formulated as a speed tracking and disturbance (i.e. electrical load) rejection problem. The tracking requirement is to ensure that the engine speed follows a reference speed set point. The disturbance rejection requirement is to ensure that the engine speed does not deviate too much from the set point in the presence of electrical load disturbance and engine speed recovers back to reference speed as soon as possible so that the required frequency of the generator electricity is maintained.

6.3 Controller Design Strategy

This Chapter concentrates on controller design using model-based methods. The identified linear model (in state space form) from the previous Chapter will be used. Two model-based sliding mode controller design schemes are proposed; one with integral action and another employing a model-following method. Both the controller designs are described in the following two sections.

6.4 Integral Action Sliding Mode (IASM) Control

The IASM control technique introduces an additional state (an integral error state) into the system. The controller synthesis is then to minimise the tracking error. The control law requires all the internal states of the system to be available. Thus, a sliding mode observer is designed to estimate these states for the controller. The engine speed system is first considered.

6.4.1 Diesel Engine Speed Representation

The diesel engine speed system is identified in a linear state equation set. The engine speed system is a single input single output (SISO) system and has the following state equation representation.

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + f(t,x,u) \\ y(t) &= Cx(t) \end{aligned}$$

Equation 6-1

where $A \in \Re^{nxn}$, $B \in \Re^{nxm}$, $C \in \Re^{pxn}$ and $m = p$. i.e. the system is square. The matrices A, B and C are the system, input distribution and output distribution matrices respectively. The variables $u(t)$ and $y(t)$ are referred to as the input and output respectively. The function $f(t,x,u)$ represents the system uncertainties plus any model uncertainties in the system. The function $f(t,x,u)$ is assumed unknown but bounded. The nominal system is determined by setting $f(t,x,u) = 0$.

The sliding mode integral tracking control algorithms that will be considered here are modified from [Edwards 1996]. The necessary assumptions on the linear system (A, B, C) for the design process are:

(a) the pair (A, B) is controllable
(b) $\det(CB) \neq 0$
(c) the invariant zeros of (A, B, C) are stable (i.e. in the open left half complex plane)

The invariant zeros of the linear system (A, B, C) are defined from Rosenbrock's System Matrix, $R(s)$ as follows:

$$R(s) = \begin{bmatrix} sI - A & B \\ -C & 0 \end{bmatrix}$$

The invariant zeros are then given by the following set:

$$\{ s \in C : \det R(s) = 0 \}$$

Equation 6-2

where s is a set of complex frequencies (Edwards 1998).

6.4.2 Controller Formulation

For the application of sliding mode controller/observer design, the linear system has to be in *regular form*. The regular form can be obtained via coordinate transformation $x \mapsto Tx$ and the new system triple $A_{reg} = TAT^{-1}$, $B_{reg} = TB$ and $C_{reg} = CT^{-1}$. The transformed matrices are partitioned as follows

$$A_{reg} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B_{reg} = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}, \quad C_{reg} = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

Equation 6-3

and the system states are partitioned as

$$x_{reg} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Equation 6-4

where $A_{11} \in \Re^{(n-m)x(n-m)}$, $B_2 \in \Re^{mxm}$ and $C_2 \in \Re^{pxp}$. The square matrix B_2 is non-singular because the input distribution matrix is assumed to be of full rank. The square matrix C_2 is also non-singular because $C_2 B_2 = CB$ which is non-singular by assumption and B_2 is non-singular by construction. It is obvious from the structure of the regular form that matrix B transform into $[0 \ B_2]^T$, that the input channels only affect m states and thus the regular form isolate the inputs from the remaining $(n-m)$ states. The transformation matrix, T, is orthogonal and the inverse of the transformation matrix is its transpose. The transformed nominal system can be written as

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t)$$

Equation 6-5

$$\dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) + B_2 u(t)$$

Equation 6-6

The tracking control technique uses an integral action methodology. Consider the introduction of an additional state, $x_r \in \Re^p$, given as:

$$\dot{x}_r(t) = r(t) - y(t)$$

Equation 6-7

where the differentiable signal $r(t)$ satisfies

$$\dot{r}(t) = \Gamma(r(t) - R)$$

Equation 6-8

where $R \in \Re^p$ is a reference demand vector and $\Gamma \in^{pxp}$ is a stable design matrix. Equation 6-8 represents an ideal reference model where $r(t)$ defines a dynamic profile which ultimately converges to the demand vector, R. The matrix Γ will determine the dynamic response of the system output. The selection of $\Gamma \in^{pxp}$ will be discussed in the sliding mode design algorithm section later in this chapter. The integral state x_r is introduced to provide a framework for engine speed tracking. Combine both the integral action state and the previous system state to yield the new state vector as follows:

$$x_{tol} = \begin{bmatrix} x_r \\ x \end{bmatrix}$$

Equation 6-9

The new state vector has dimension $(n+p)$. The first p states correspond to integrals of the output error and the remaining n states are the system states. Similarly to the previous procedure, assume a partition of the state vector, x_{tol}, to isolate the input channel and the new partitioned state is defined as follows:

$$x_{tol} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Equation 6-10

where $z_1 \in \Re^n$ and $z_2 \in \Re^m$. z_1 now contains the first n states of x_{tol} and z_2 has the last m states of x_{tol}. The system input now only affects the states in z_2. The new augmented system can be re-partitioned in the form:

$$\begin{aligned} \dot{z}_1(t) &= \tilde{A}_{11}z_1(t) + \tilde{A}_{12}z_2(t) + B_r r(t) \\ \dot{z}_2(t) &= \tilde{A}_{21}z_1(t) + \tilde{A}_{22}z_2(t) + B_2 u(t) \end{aligned}$$

Equation 6-11

where B_r is the input distribution matrix for the demand signal $r(t)$:

$$B_r = \begin{bmatrix} I_p \\ 0 \end{bmatrix}$$

The newly partitioned state matrix is given as:

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} = \left[\begin{array}{cc|c} 0 & -C_1 & -C_2 \\ 0 & A_{11} & A_{12} \\ \hline 0 & A_{21} & A_{22} \end{array}\right]$$

Equation 6-12

The sliding mode controller seeks to induce a sliding motion on a sliding surface. A possible sliding surface can be proposed as:

$$\begin{aligned} \ell_{IASMC} &= \left\{ (z_1, z_2, r) \in \Re^{n+p} : \; s(z_1, z_2, r) = 0 \right\} \\ s(z_1, z_2, r) &= S_1 z_1 + S_2 z_2 - S_r r \end{aligned}$$

Equation 6-13

where $S_1 \in \Re^{m \times n}$, $S_2 \in \Re^{m \times m}$ and $S_r \in \Re^{p \times p}$ are design parameters which govern the reduced order motion. The square matrix S_2 must be non-singular. The parameter, S_1 can be calculated by applying linear feedback methods to the $(\tilde{A}_{11}, \tilde{A}_{12})$ subsystem. When ideal sliding motion occurs, $s(z_1, z_2) = 0$ and equation 6-13 can be rewritten as:

$$z_2(t) = -S_2^{-1}S_1 z_1(t) + S_2^{-1}S_r r(t)$$

Equation 6-14

The design parameter, S_r can be of any value as it does not affect the stability of the closed-loop system. In this case, it is calculated as follows [Edwards 1998],

$$S_r = -S_2(B_r^T A_{11s}^{-1}\widetilde{A}_{12})^{-1} B_r^T A_{11s}^{-1} B_r$$

Equation 6-15

where $A_{11s} = (\widetilde{A}_{11} - \widetilde{A}_{12}M)$ and M is a state feedback matrix. It is possible to choose S_2 so that it is invertible. Let

$$M = S_2^{-1}S_1$$

Equation 6-16

The system dynamic during the ideal sliding motion on ℓ_{IASMC} can be obtained by substituting equation 6-14 into the first equation in equation 6-11:

$$\dot{z}_1(t) = (\widetilde{A}_{11} - \widetilde{A}_{12}M)z_1(t) + (\widetilde{A}_{12}S_2^{-1}S_r + B_r)r(t)$$

Equation 6-17

The matrix M is seen to have the role of a state feedback controller for the z_1 subsystem and this matrix can be determined via a state feedback method if the pair $(\widetilde{A}_{11}, \widetilde{A}_{12})$ is completely controllable. The pair $(\widetilde{A}_{11}, \widetilde{A}_{12})$ is controllable if and only if (A, B) is controllable and (A, B, C) has no invariant zeros at the origin. The reachability condition for the controller is similar to the form used in the previous example in Chapter 3. Hence, the continuous (i.e. linear part) and discontinuous (i.e. non-linear part) part are computed as follows. Assume the following reachability condition.

$$\dot{s} = \Phi s - \rho \operatorname{sign}(s)$$

Equation 6-18

where Φ is a stable design matrix. Differentiating equation 6-13 and substituting from equation 6-11 yields

$$\begin{aligned}
\dot{s} &= S_1\dot{z}_1 + S_2\dot{z}_2 - S_r\dot{r} \\
&= S_1(\widetilde{A}_{11}z_1 + \widetilde{A}_{12}z_2 + B_r r) + S_2(\widetilde{A}_{21}z_1(t) + \widetilde{A}_{22}z_2(t) + B_2u(t)) - S_r\dot{r} \\
&= S_1(\widetilde{A}_{11}z_1 + \widetilde{A}_{12}z_2) + S_2(\widetilde{A}_{21}z_1 + \widetilde{A}_{22}z_2) + S_1B_r r + S_2B_2u - S_r\dot{r} \\
&= S\widetilde{A}x_{tol} + S_1B_r r + S_2B_2u - S_r\dot{r}
\end{aligned}$$

Equation 6-19

where $S = [S_1 \quad S_2]$. The proposed control law is given as:

$$u = u_L(x_{tol}, r) + u_N$$

Equation 6-20

where u_L and u_N represent the linear and non-linear components of the control law. With the reachability condition in equation 6-18, the linear part of the control can be derived from equation 6-18 and 6-19 as follows:

$$\begin{aligned} u_L(x_{tol}, r) &= \Lambda^{-1}(-S\tilde{A}x_{tol} - S_1 B_r r + S_r \dot{r} + \Phi s) \\ &= L_{x_{tol}} x_{tol} + L_r r + L_{\dot{r}} \dot{r} \end{aligned}$$

Equation 6-21

where
$$\begin{aligned} L_{x_{tol}} &= -\Lambda^{-1} S\tilde{A} + \Lambda^{-1}\Phi S \\ L_r &= -\Lambda^{-1} S_1 B_r - \Lambda^{-1}\Phi S_r \\ L_{\dot{r}} &= \Lambda^{-1} S_r \end{aligned}$$

Equation 6-22

where $\Lambda = S_2 B_2$. The linear control law will drive the system into sliding motion asymptotically. However, in order to make the controller robust against matched uncertainties and to achieve sliding in finite time, a nonlinear control component is required. The nonlinear smooth part of the control is given as:

$$u_N = \begin{cases} -\rho_N \Lambda^{-1} \dfrac{Ps}{\|Ps\| + \delta_c} & \text{if } s \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Equation 6-23

where P is a symmetric positive definite matrix satisfying the equation:

$$P^{-1}\Phi^T + \Phi P^{-1} = -Q$$

Equation 6-24

for some positive definite matrix Q. δ_c is a small positive constant known as the smoothing factor. It is used to eliminate the chattering in the otherwise discontinuous control action. This positive constant can be tuned during implementation. Here, ρ_N is defined as the positive scalar function:

$$\rho_N(e_y) = \rho_{c1} + \rho_{c2}\|e_y\| + \rho_{c3}\|e_y\|^2$$

Equation 6-25

where ρ_{c1}, ρ_{c2} and ρ_{c3} are positive design scalars and e_y is the engine speed error. The second and third terms in equation 6-25 will only be introduced in the application to improve the speed recovery. Note that $\delta_c = 0$, yields finite time convergence to the sliding manifold. For positive δ_c, convergence to a boundary layer of the sliding manifold is assured in finite time.

6.4.3 Observer Formulation

The above control law requires all the system states to be available. Since the required states are not available from the system under consideration, a sliding mode observer is employed to estimate these system states. The system obtained from dynamic identification is a square. The observer based on this square system from [Edwards 98] is used in this section. Consider the system in equation 6-1. The following assumptions are required for the observer design:

(a) C and B are full rank
(b) The invariant zeros of (A,B,C) are stable

Consider the regular form system in equation 6-3, the transformed system is similar to equation 6-5 and equation 6-6. The proposed observer has the structure as follows

$$\dot{\hat{x}} = A\hat{x} + Bu - GCe + v$$

Equation 6-26

where e is the state error and $G \in \Re^{nxp}$, the observer linear gain, is defined as

$$G = \begin{bmatrix} A_{12}C_2^{-1} \\ A_{22}C_2^{-1} - C_2^{-1}A_{22}^{\phi} \end{bmatrix}$$

Equation 6-27

where A_{22}^{ϕ} is a stable design matrix.

The gain G is chosen such that the closed loop observer matrix, $A_{closed} = A - GC$ has stable eigenvalues and satisfies the Lyapunov equation:

$$P_{obs}A_{closed} + A^T{}_{closed}P_{obs}^T = -Q$$

Equation 6-28

where Q is some positive definite matrix and P_{obs} is a Lyapunov matrix which satisfies the structural constraint

$$C^T F^T = PB$$

Equation 6-29

where $F \in \Re^{mxm}$ is a non singular design matrix and is given as $[P_{obs}C_2B_2]^T$ [Edwards 96].

The sliding surface for the observer is:

$$\ell_{observer} = \left\{ e \in \Re^n : \ s(e) = 0 \right\}$$
$$s(e) = Fe$$

Equation 6-30

The non-linear part of the observer is defined as:

$$v = \begin{cases} -\rho_o \dfrac{Fe}{\|Fe\| + \delta_o} & \text{if } Fe \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Equation 6-31

where ρ_o is a positive design scalar and δ_o is a small positive constant, the smoothing factor. This non-linear component can reject parameter uncertainties and maintains a sliding motion. This Section has described the formulation for the observer-based sliding mode controller design for idle speed control of an engine. The following Section will describe several design issues toward the practical implementation of the controller.

6.5 Diesel Engine Idle Speed IASM Controller Design

The following sub-section will describe the design issues which emerged during the practical design and implementation.

6.5.1 Model Modification

The identified model is first balanced to reduce the model sensitivity. The designed controller was written as an S-function for dSPACE implementation. The S-function was converted into C code using a specific function in the Matlab^{TM}/RealTime toolbox. This C code was further cross-compiled into an executable file to be run on a Texas Instrument TMS 320F240 DSP microcontroller. The balanced plant model equation has the following system matrices:

$$A = \begin{bmatrix} -0.0055 & -0.0026 \\ 0.0026 & -0.7170 \end{bmatrix}$$

$$B = \begin{bmatrix} 4.9617 \\ -1.1836 \end{bmatrix}$$

$$C = [4.9617 \quad 1.1836]$$

Equation 6-32

The matrices in equation 6-3 have the following values

$$A_{reg} = \begin{bmatrix} -0.6787 & -0.1632 \\ -0.1580 & -0.0438 \end{bmatrix}, \; B_{reg} = \begin{bmatrix} 0 \\ -5.1009 \end{bmatrix}, \; C_{reg} = [2.3025 \quad -4.5517]$$

6.5.2 Hyperplane Design

In Chapter 3, two methods have been described for the hyperplane design. For the practical controller design, the robust pole placement method is employed here. The design is carried out using the available toolbox from [Edwards 1998]. The poles are set to [-1.0;-1.2] and the resulting switching function has the value of $S = \begin{bmatrix}0.0346 & 0.0337 & 0.1000\end{bmatrix}$.

6.5.3 Controller/Observer Variables

The other design variables are set as follows. The designed observer dynamic was set to be faster than the controller dynamic to provide good estimates of the system states for use by the controller. The observer parameter, A_{22}^{ϕ}, was set to -20. The design matrix Φ is set to -1. The scalar terms for the discontinuous terms are: $\rho_o = 0.1$, $\rho_{c1} = 0.1$, $\rho_{c2} = 0$, $\rho_{c3} = 0$, $\delta_o = 0.1$ and $\delta_c = 0.1$. The second and third terms in equation 6-25 have not introduced in the application.

6.5.4 Test Setup

The observer/controller pair design is setup in the Matlab Simulink environment. The estimated states from the observer are fed to the controller algorithm. Both of the sliding mode observer and controller algorithms are written as C-Mex c-code s-functions. The resulting s-functions allow a faster simulation (compared to Simulink block design) and the s-function has a systematic structure which provides better understanding in the design. The s-function also allows he user to vary the step size of the simulation conveniently. The speed of the observer can be improved by decreasing the integration step size for the observer. The practical test is setup in such a way that the observer speed is faster than the controller by a factor of 10.

The design variables were set as follows: The design matrix Φ is set to -1. The scalar terms for the discontinuous terms are: $\rho_o = 0.1$, $\rho_{c1} = 0.1$, $\rho_{c2} = 0$, $\rho_{c3} = 0$, $\delta_o = 0.1$ and $\delta_c = 0.1$.

6.6 IASM Controller Implementation

For the IASM controller implementation, there is a need to break into the engine speed control loop. The previously mentioned PI controller is used to start the engine and then switched to the IASM controller for testing without shutting down the engine. The hardware requirements for the implementation have been mentioned in Chapter 2.

6.6.1 dSPACE (Digital Signal Processing and Control Engineering)

dSPACE is a package which interfaces Maltab and Simulink (i.e. using Matlab Real-Time Workshop) to the real application. It consists of specific software and dedicated hardware to run the Simulink structures on a microprocessor. The cross compilation of Simulink structures into an executable file which is then run on a microprocessor has been described in Chapter 2.

6.6.2 Controller Gain Tuning

To provide an online controller tuning facility, three control gains namely, $g\_L_{x_{tol}}$, g_L_r and $g\_L_{\dot{r}}$, were introduced to the respective control components in equation 6-22, $L_{x_{tol}}$, L_r and $L_{\dot{r}}$. The introduction of these gains provides a facility to allow the user to tune the control components individually. The diesel engine system appeared to be stable in tests using this new tuning concept. The design parameters, ρ_{c1} and ρ_o effectively tune the non-linear term. The gains $g\_L_{x_{tol}}$ and ρ_{c1} were tuned to a performance that could cope with large load variations, as well as small step loads. The gains for both the speed tracking demand gain, L_r and derivative of speed tracking demand gain, $L_{\dot{r}}$ were set to $g\_L_r = 0.001$ and $g\_L_{\dot{r}} = 0.001$ respectively. These gains were tuned to obtain a better speed tracking for different dynamic profile (Γ). The gain tuning process is summarised in Table 6-1. The tuning approach is effective and straightforward for the particular diesel engine under consideration. The summary of the gain tuning process is as follows:

Step	Description
1	**Tuning ρ_{c1}** Firstly, set g_L_r and $g\_L_{\dot{r}}$ to zero and then slowly increase ρ_{c1} to a small value. Adjust ρ_{c1} so that the controller maintains the engine reference speed.
2	**Tuning $g\_L_{x_{tol}}$ and ρ_{c1}** Gradually increases $g\_L_{x_{tol}}$ and adjusts this gain under different loading conditions. Adjusting ρ_{c1} can improve the speed recovery under loading condition.
3	**Tuning g_L_r and $g\_L_{\dot{r}}$** Tuning both of these gains to improve the engine speed track with the demand signal.

Table 6-1: IASM controller gain tuning algorithm.

The engine performance with the proposed controller gain tuning approach is described in the following Section.

6.6.3 Performance Tests

The designed controller was put under load rejection and speed tracking test. At zero loading condition, the speed was maintained at about 1.28 % of the reference speed. The nominal engine speed is 1500rpm. The load rejection property was tested by applying different loading condition (i.e. small, medium and large steps load) to the engine. Chapter

2 has mentioned the controller test criteria. To mimic the real life situation, different loading conditions are applied. Firstly, a small and medium steps load, i.e. load from 0kW to 60kW with each load increment of 10 kW or 20 kW, were applied to the engine. A large load, i.e. a 60kW step load, was also applied to the engine. Subplot 1 in Figure 6-1, Figure 6-2 and Figure 6-3 shows the engine speed response in the presence of respectively applied loads. The engine speed dipped when steps load were applied and the engine speed overshot when the loads were taken off. For the 10kW steps load, the engine speed shows a short period of oscillation when the load was getting above 40kW. Similarly, the 20kW steps loading also caused small magnitude oscillation at higher loading condition. During the step off load (i.e. from 60kW to 40kW), the engine speed hit the speed safety level (i.e. about 1650rpm) and the engine fuel was cut off and shutdown the engine automatically. Thus, the load decrement was reduced to 10kW from 60kW to 10kW. The subplot 1 in Figure 6-2 shows the resulting effect. For either of the loading conditions, the controller is returning the engine speed to the reference value in a reasonably short time. The maximum dip is 18.5 % when the load changes from 40 kW to 60 kW with a step load of 20 kW. During the large step load test, the speed dipped at maximum of about 36% of the nominal speed and the speed recovered within thirteen seconds with a small magnitude of oscillation which quickly disappeared.

6.6.4 Robustness Tests

A good speed controller design not only maintains the engine speed but also must show the robustness to operate over a wide operating envelope. For the purpose of this test, the engine reference speed was changed from 1500 rpm to 1350 rpm and 1200 rpm. At these different reference speed settings, the three different electrical loads, as mentioned earlier, were applied to observe the controller performance. The resulting performance is shown in the subplot 2 and 3 of Figure 6-1, Figure 6-2 and Figure 6-3. The controller could be seen maintaining the reference speed with small speed variation at zero load condition. At 10kW steps loading for both the 1350rpm and 1200rpm reference speed settings, the engine speed recovered quickly. At 20kW step loading, the engine speed suffered a sustained oscillation at 60kW at 1350rpm reference speed however, at 1200rpm reference speed, the speed recovered quickly when load was applied. When a large load (i.e. 60 kW) was applied to both the different speed settings, the controller was seen to stabilises the engine speed. A small transient was noticed at reference speed of 1350rpm.

6.6.5 Speed Tracking Tests

Finally, a performance test of the incorporated speed tracking system in the controller design was considered. The dynamic profile of speed tracking depends on the term Γ. The demand vector R was set to change in the fashion of step-changes between 1450 rpm and 1550 rpm. The resulting performance is shown in Figure 6-4. Three different settings of Γ were set at -0.1, -0.5 and -1.0. The higher the Γ value is, the faster the speed tracking rate is. At Γ equal to -0.1 and -0.5, the controller maintained the engine speed closely to the demand signal R at a slower rate. At Γ=-1.0, the controller tracked very rapidly and closely to the demand value. The tracking performance was determined by both the g_L_r and g_L_r settings.

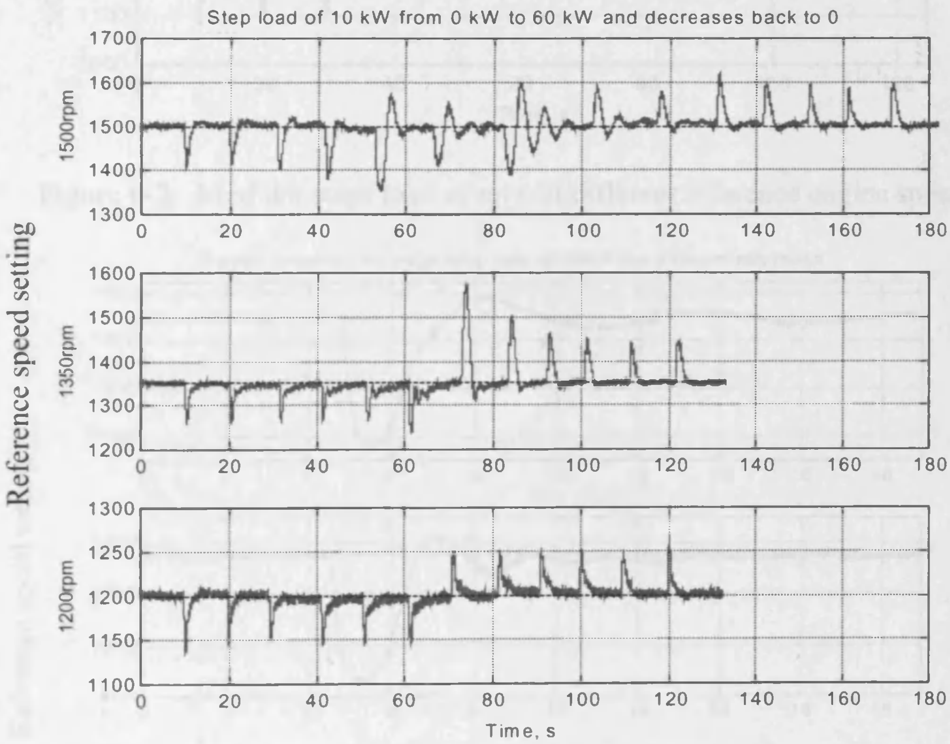


Figure 6-1: Small steps load at several different reference engine speeds.

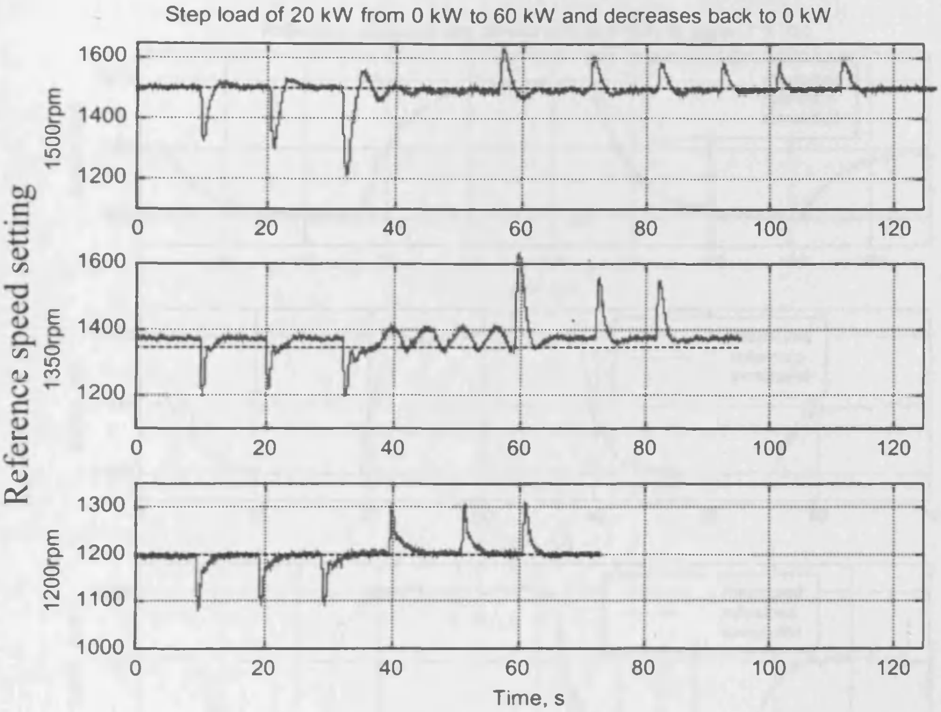


Figure 6-2: Medium steps load at several different reference engine speeds.

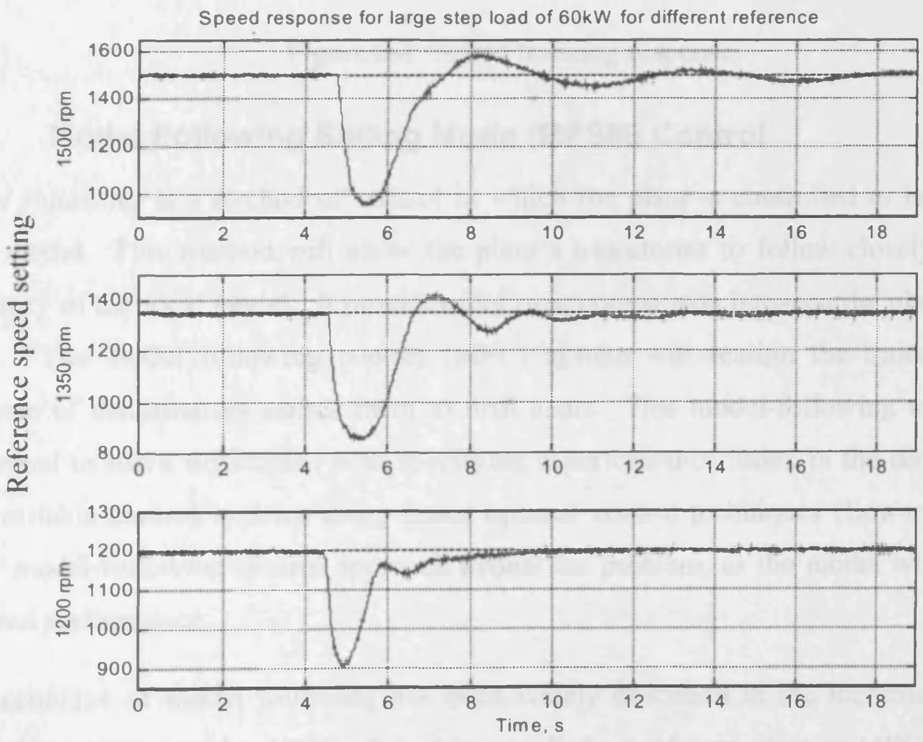


Figure 6-3: Large step load is applied at several different reference engine speeds.

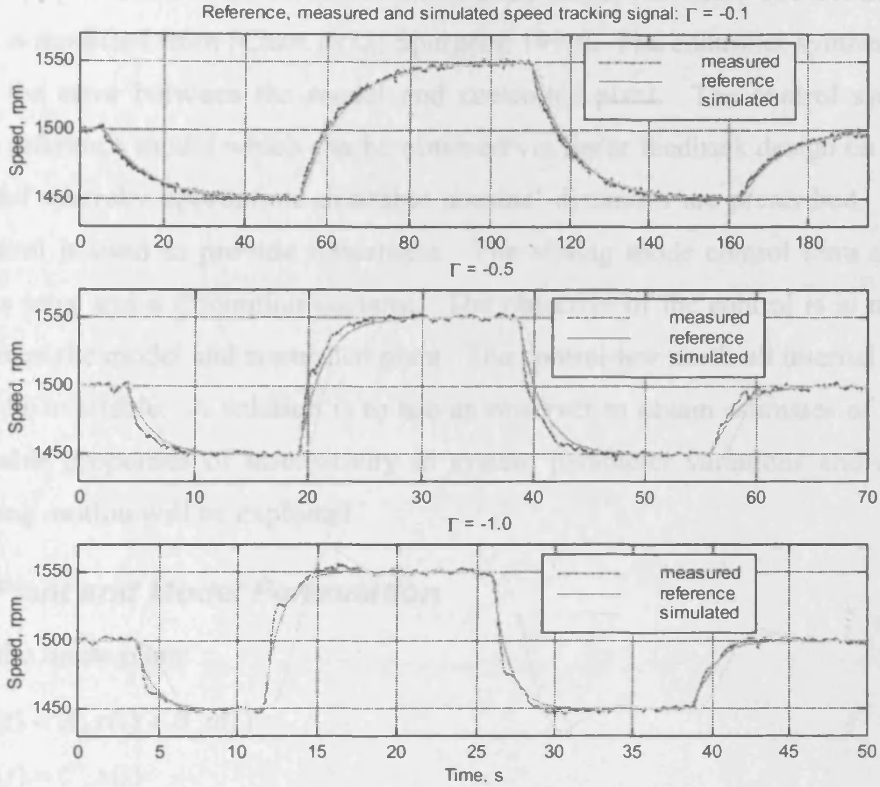


Figure 6-4: Speed tracking response.

6.7 Model Following Sliding Mode (MFSM) Control

Model following is a method of control in which the plant is controlled to behave like an ideal model. This method will allow the plant's trajectories to follow closely a particular trajectory of the ideal model. It provides real time comparison between the plant and model states. The model following control (MFC) system will realign the trajectories if the presence of disturbances causes them to drift apart. This model-following technique was developed to solve difficulties with specifying a performance index in the direct design of multivariable control systems using linear optimal control techniques [Edwards 1998]. A linear model-following control approach avoids the problem, as the model will specify the required performance.

The technique of model following has been widely described in the literature [Balestrino 1984, Chan 1973, Landau 1974]. Balestrino applied a nonlinear adaptive MFC to a problem of pendulum position control. Chan demonstrated the difference between perfect model following and real model following. Landau used a multivariable adaptive MFC technique for aircraft control.

The technique used in this research is based on the theory of VSS. The control technique used here is modified from [Chan 1973, Spurgeon 1990]. The controller synthesis is then to minimise the error between the model and controlled plant. The control system design requires a reference model which can be obtained via linear feedback design on the nominal plant model whereby appropriate desirable nominal dynamics are prescribed. The sliding mode control is used to provide robustness. The sliding mode control term consists of a continuous term and a discontinuous term. The objective of the control is to minimise the error between the model and controlled plant. The control law needs all internal states of the system to be available. A solution is to use an observer to obtain estimates of these states. The desirable properties of insensitivity to system parameter variations and disturbances when sliding motion will be exploited.

6.7.1 Plant and Model Formulation

Consider the linear plant:

$$\dot{x}(t) = A_p x(t) + B_p u(t)$$
$$y(t) = C_p x(t)$$

Equation 6-33

The corresponding ideal model is defined as:

$$\dot{w}(t) = A_m w(t) + B_m r(t)$$
$$y_m(t) = C_m w(t)$$

Equation 6-34

where $x \in \Re^n$ and $w \in \Re^n$ are the state vectors of the plant and the ideal model respectively, $u \in \Re^m$ is the control vector, $r \in \Re^r$ is a reference input and A_p, B_p, A_m and B_m are compatibly dimensioned matrices.

Assumptions:

1. The pair (A_p, B_p) is controllable.
2. The ideal model is stable; the eigenvalues of A_m have negative real parts.

A tracking error state, $e_{track}(t)$, is defined as the difference between the plant and model states:

$$e_{track}(t) = x(t) - w(t)$$

Equation 6-35

The objective of MFC is to guarantee that the error between the plant and model states tends asymptotically to zero, i.e. $\lim_{t \to \infty} e_{track}(t) = 0$. Differentiating the error equation 6-35 with respect to time yields:

$$\dot{e}_{track}(t) = \dot{x}(t) - \dot{w}(t)$$

Equation 6-36

From equation 6-33 and equation 6-34, the dynamic of the model-following error system of equation 6-36 becomes:

$$\dot{e}_{track}(t) = A_p x(t) - A_m w(t) + B_p u(t) - B_m r(t)$$

Equation 6-37

By adding and subtracting a term $A_m x$ in equation 6-37 yields

$$\dot{e}_{track}(t) = A_m e_{track}(t) + (A_p - A_m)x(t) + B_p u(t) - B_m r(t)$$

Equation 6-38

For perfect model following, the following condition holds [Erzberger 1968] for time, *t*.

$$\begin{aligned} \dot{x}(t) - \dot{w}(t) = 0 \quad &\Rightarrow \quad \dot{x}(t) = \dot{w}(t) \\ x(t) - w(t) = 0 \quad &\Rightarrow \quad x(t) = w(t) \end{aligned}$$

Equation 6-39

Assume some arbitrary term, feeding forward from the model states, is added to the control action and equation 6-33 becomes

$$\dot{x}(t) = A_p x(t) + B_p (u(t) + G_{fwd} w(t))$$

Equation 6-40

where the matrix G_{fwd} is some arbitrary gain. From the perfect model following condition in equation 6-39, the equation 6-34 and equation 6-40 become

$$A_p x(t) + B_p u(t) + B_p G_{fwd} x(t) = A_m w(t) + B_m r(t)$$

Equation 6-41

Rearranging the equation 6-41 to obtain an expression for the control gives

$$u(t) = B_p^+ (A_m w(t) + B_m r(t) - A_p x(t) - BG_{fwd} x(t))$$

Equation 6-42

where B_p^+ denotes the Moore-Penrose pseudo-inverse[1] of the matrix B_p. Substituting this control expression into equation 6-41 yields

[1] The Moore-Penrose pseudo inverse of a matrix can be solved by using Matlab command called 'pinv'. Further details are in [Zadeh 1963, MathWorks-4 1998]

$$A_p x(t) + B_p B_p^+ (A_m w(t) + B_m r(t) - A_p x(t) - BG_{fwd} x(t)) + BG_{fwd} x(t) - A_m w(t) - B_m r(t) = 0$$

Since $B_p B_p^+ B_p = B_p$ by definition, the above equation can be simplified as follows

$$(B_p B_p^+ - I)A_m w(t) - (B_p B_p^+ - I)A_p x(t) + (B_p B_p^+ - I)B_m r(t) = 0$$
$$(B_p B_p^+ - I)(A_m - A_p)x(t) + (B_p B_p^+ - I)B_m r(t) = 0$$

Equation 6-43

It can be seen that the direct use of model states w in the control loop has no effect on the condition for model following. Based on equation 6-39, the equation 6-43 is satisfied for all $x(t)$, $w(t)$ and r if

$$(B_p B_p^+ - I)(A_p - A_m) = 0$$

Equation 6-44

$$(B_p B_p^+ - I)B_m = 0$$

Equation 6-45

If equation 6-44 and equation 6-45 hold, consider a control law with the structure as follows:

$$u(t) = u_1(t) + u_2(t)$$

Equation 6-46

where

$$u_1(t) = B_p^+ (A_m - A_p)x(t) + B_p^+ B_m r(t)$$

Equation 6-47

$$u_2(t) = -Me_{track}(t)$$

Equation 6-48

Equation 6-44 and 6-45 are the conditions for perfect model following and equation 6-46 is the control law for implementing it. The generation of equation 6-46 requires the model states. By substituting the control law equation 6-46 into 6-38 gives

$$\dot{e}_{track}(t) = (A_m - B_p M)e_{track}(t)$$

Equation 6-49

Thus, the dynamics of the model following error system reduces to the closed-loop matrix $(A_m - B_p M)$. With proper selection of eigenvalues for M, the errors' settling rate can be controlled.

A well-known theorem from linear algebra states that for the system of simultaneous equations denoted by

$$HD = F$$

Equation 6-50

a solution for D exists if and only if rank $[H \;\; F]$ = rank $[H]$. Assume the following rank conditions hold

$$\text{rank}\,[B_p \quad A_p - A_m] = \text{rank}\,[B_p]$$
$$\text{rank}\,[B_p \quad B_m] = \text{rank}\,[B_p]$$

Equation 6-51

According to the theory, there thus exist matrices G_L and L of suitable dimensions such that

$$B_p L = A_m - A_p \qquad \Rightarrow \; A_m = B_p L + A_p$$
$$B_p G_L = B_m \qquad \Rightarrow \; B_m = B_p G_L$$

Equation 6-52

These effectively define the model dynamics. Hence, the equation 6-47 with u_1 can now be expressed as

$$u_1(t) = Lx(t) + G_L r(t)$$

Equation 6-53

Consider equation 6-34, at steady state whereby

$$w_{ss} = -A_m^{-1} B_m r$$

Equation 6-54

At steady state, the reference model output equation becomes

$$y_{mss} = C_m w_{ss}$$

Substituting from equations 6-54 and 6-52,

$$y_{mss} = -C_m A_m^{-1} B_m r$$
$$= -C_m A_m^{-1} B G_L r$$

If the model is to attain the reference, it follows that

$$G_L = inv(-C_m A_m^{-1} B)$$

Equation 6-55

The term L in equation 6-52 can be obtained as a linear feedback matrix by selecting appropriate eigenvalues for the reference model. Thus, the reference model matrix can be written as

$$A_m = A_p + B_p L$$

Equation 6-56

u_2 in equation 6-46 is incorporated in the control law to provide robustness. This term contains a continuous and a discontinuous control term. They are derived using the sliding mode concept in the next section.

6.7.2 Controller Formulation

In the control system design, the plant output is required to follow the output of the reference model in equation 6-34. The objective of the design is to choose a hyperplane matrix, S and related discontinuous control law such that the error states display a sliding motion. By considering the error states, equation 6-35, an error dependent switching function is defined as:

$$\ell_{MFSM} = \left\{ e_{track}(t) \in \Re^n : s(e_{track}(t)) = 0 \right\}$$
$$s(e(t)) = Se_{track}(t)$$

Equation 6-57

When a sliding motion takes place,

$$Se_{track}(t) = 0$$

Equation 6-58

By differentiating equation 6-57, substituting equation 6-38 into it and rearranging the equation gives:

$$\dot{s}(e_{track}(t)) = S\dot{e}_{track}(t) = S(A_m e_{track}(t) + (A_p - A_m)x(t) + B_p u(t) - B_m r(t))$$

Equation 6-59

Inserting u from equation 6-46 with u_1 from equation 6-53, the following is obtained

$$\dot{s}(e(t)) = S(A_m e_{track}(t) + B_p u_2(t))$$

Equation 6-60

The reachability condition for the controller component u_2 is similar to the form in Equation 3-14 and is given as:

$$\dot{s} = \Phi s - \rho \text{sign}(s)$$

Equation 6-61

where Φ is a stable design matrix and the eigenvalues of this matrix determine the speed at which the system goes into sliding motion. By considering both the equation 6-60 and 6-18, the control law u_2 in equation equation 6-48 can be re-defined into two parts: a linear and nonlinear part. Representation of u_2 is given as:

$$u_2 = u_l + u_n$$

Equation 6-62

where the linear control law, u_l will drive the system into sliding motion asymptotically. However, in order to make the controller robust against matched uncertainties and to achieve sliding in finite time, a non-linear control component u_n is required. By equating the equations 6-60 and 6-61, the linear control law, u_l becomes

$$u_l(t) = -(SB_p)^{-1}(SA_m - \Phi S)e_{track}(t)$$

Equation 6-63

The non-linear control component, u_n is approximated to reduce the chattering effect and it is given as

$$u_n(t) = -\rho_N(t, e_{track})(SB_p)^{-1}\frac{P_2 Se_{track}(t)}{\|P_2 Se_{track}(t)\| + \delta}$$

Equation 6-64

where δ is a small positive constant used to smooth the otherwise discontinuous control action. $P_2 \in \Re^{m \times m}$ is a symmetric positive definite matrix satisfying the Lyapunov equation

$$P_2\Phi + \Phi^T P_2 = -I$$

Equation 6-65

ρ_N is defined as:

$$\rho_N(e_y) = \rho_{c1} + \rho_{c2}\|e_y\| + \rho_{c3}\|e_y\|^P$$

Equation 6-66

where ρ_{c1}, ρ_{c2} and ρ_{c3} are positive design scalars and e_y is the system output error. It is assumed that SB_p is chosen to be non-singular. The complete model-following variable structure control scheme has the form of

$$u(t) = u_1(t) + u_l(t) + u_n(t)$$

Equation 6-67

The above control law requires all the system states to be available and thus, a similar observer used for IASM control implementation is employed here.

6.8 Diesel Engine Idle Speed MFSM Controller Design

This section will describe the design process used to implement the MFSM controller.

6.8.1 Model Modification

Similarly, the identified model is first balanced to reduce the model sensitivity. The designed controller was written as an S-function for dSPACE implementation. The observer parameter, A_{22}^{ϕ}, was set to –20. The balanced plant model equation has the following system matrices:

$$A_p = \begin{bmatrix} -0.0055 & -0.0026 \\ 0.0026 & -0.7170 \end{bmatrix}$$

$$B_p = \begin{bmatrix} 4.9617 \\ -1.1836 \end{bmatrix}$$
$$C = [\,4.9617 \quad 1.1836]$$

Equation 6-68

The reference model is chosen to have poles at $p = -0.6 \pm 0.001j$. The poles are chosen in such a way that the model improves the open-loop speed of response. The reference model is defined as

$$A_m = \begin{bmatrix} -0.15015 & -0.0802 \\ 0.1209 & -0.6985 \end{bmatrix}$$
$$B_m = \begin{bmatrix} 0.1011 \\ -0.0241 \end{bmatrix}$$

Equation 6-69

$$L = [-0.10 \quad -0.01]$$
$$G_{fwd} = [0.0204]$$

Equation 6-70

6.8.2 Test Setup

The desired poles of the sliding motion are set to [-0.35] and the resulting switching function has the value of $S = [-0.5054 \quad 2.1911]$. The design parameters were set as follows: $A_{22}^{\phi} = -20$, $\Phi = -0.35$, $\rho_o = 150$, $\rho_{c1} = 0.01$, $\rho_{c2} = 0.0005$, $\rho_{c3} = 0.0005$, $pow = 1.5$, $\delta_o = 0.1$ and $\delta_c = 0.1$. The initial test on the diesel engine showed engine speed dipped slightly for each added load. The results consistently show a steady state error in the speed signal. In the ideal sliding mode situation, the controller can guarantee asymptotic tracking with zero steady state error. When implemented using a continuous approximation such as in equation 6-64 using the smoothing factor, δ, a pseudo sliding is induced instead of ideal sliding. Thus, the controller can no longer guarantee zero steady state error. To reduce the steady state error, δ can be reduced but too small a value will yield undesirable chattering. One way to achieve zero steady state error is by introducing integral action in the controller. The resulting control law becomes

$$u(t) = u_1(t) + u_2(t) + K_{\text{int}} u_3$$

Equation 6-71

where $\dot{u}_3 = e(t)$. The anticipation of the integral action is injected into the control loop as shown bellow.

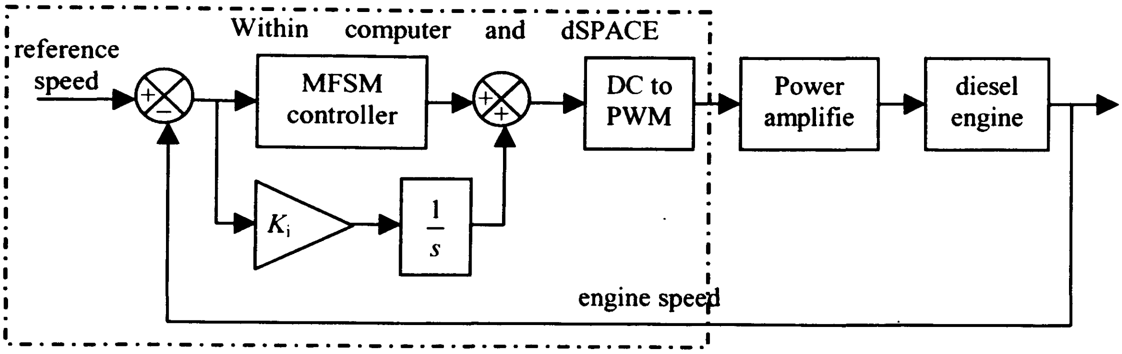


Figure 6-5: Incorporation of integral action in MFSM control strategy.

The signal information (i.e. scale and amplitude) of the control variable has been shown in Chapter 2, Figure 2.8. As mentioned earlier, the sliding motion is insensitive to the matched uncertainties. With the inclusion of integral action, this property of the sliding motion is assumed to hold. The new parameter settings are: $\rho_{c1} = 0.05, \rho_{c2} = 0.001$, $\rho_{c3} = 0$, $\rho_o = 150$, $\rho_o = 0.1$ and $K_{\text{int}} = -0.002$.

6.9 MFSM Controller Implementation

Similarly to the IASM controller implementation approach, a PI controller is employed to switch from one controller to another without shutting down the engine.

6.9.1 Controller Gain Tuning

The control law comprises of linear, u_l, non-linear, u_n and feedforward, u_1 terms. The tuning concept, which was employed in IASM controller, is considered here. Two scaling gains (g_u_l and g_u_1) were introduced for the linear and feedforward terms, whereas, the previous defined term, ρ_N was used to tune the discontinuous switching control term. During the practical implementation, the terms g_u_l, g_u_1 and ρ_N were adjusted to achieve a better engine performance against electrical load disturbances. Additional terms which relate to the speed error in equation 6-66 were included to improve the speed recovery in the presence of larger load disturbances. The gain settings were found to be: $g\_u_l = 1$ and $g\_u_1 = 1.13$. The summary of the gain tuning process is given as follows:

Step	Description
1	**Tuning ρ_{c1}** Set g_u_l and g_u_1 to zero. Set ρ_{c1} to a small value. Switch the control to the MFSM controller. Adjust ρ_{c1} so that the controller maintains the engine reference speed.
2	**Tuning g_u_l, g_u_1 and ρ_{c1}**

	Gradually increase g_u_l and g_u_I, and adjust the gains with different loading conditions. Tune the ρ_{c1}, ρ_{c2} and ρ_{c3} to improve the controller performance

Table 6-2: MFSM controller gain tuning algorithm.

The engine performance due to the mentioned controller gain tuning approach is described in the following Section.

6.9.2 Performance Tests

The two different controller settings were applied to the engine. The control law without the integral control action was firstly tested. The initial speed varied within ±1.639 % of the reference speed at zero load conditions. The first performance test was to apply a step load to the system, i.e.10 kW and 20 kW. Figure 6-6 and Figure 6-7 show the engine speed signal for the respective step loading. The engine speed dipped slightly for each added load. The speed dropped by about 30 rpm at a load of 60 kW. The results consistently show a steady state error in the speed signal.

With the additional integral action, the engine speed recovered immediately. Figure 6-8 and Figure 6-9 show the recovery of the engine speed with different step loads respectively. A short period of oscillation occurred when the load was stepped up from 40kW to 60kW. The speed dip can be improved by increasing the integral control gain but this is at the cost of increased amplitude and oscillation of the speed variation. By decreasing the overall value of ρ_N, the control objective can be achieved because some of the uncertain effects are now taken care of by the integral term.

6.9.3 Robustness Tests

To test controller robustness, the engine was set to operate at different reference speeds, i.e.1200 rpm and 1350 rpm. At these speed settings, the engine was subjected to step loading of 20kW and the resulting controller performance is shown in Figure 6-10. The designed controller demonstrated a small speed variation on the measured speed. The speed variation increases at the reference speed of 1200rpm. This is to be expected as it is further from the nominal operating point.

6.9.4 Speed Tracking Test

The speed tracking test involves adjusting the demand signal, $r(t)$ and the rate of tracking is determined by the reference model pole setting. The demand signal was set to vary within ±50 rpm. The resulting controller performance is shown in Figure 6-11. The first subplot shows the reference signal. The measured engine speed is plotted in the second subplot.

The resulting speed error is shown in the third subplot. The controller tracks closely to the demand signal although the speed signal seems to have large amplitude variation.

6.10 Discussion

The previous sections have presented two different controller design schemes with simulation, implementation and validation. Several issues have been raised during the implementation. Firstly, the identified model will be discussed.

6.10.1 Identified Model

The transition of the actual system to a model is not perfect, especially for a complex system like the diesel engine. In the transformation between high and low-resolution versions, some information must be lost (by definition). An information loss is acceptable if the model retains the main characteristics of the actual system. In general, this model only approximates the behaviour of the actual physical plant. The control system should, therefore, be insensitive to errors in the mathematical model of the actual system and the controller.

6.10.2 Controller Tuning Approach

The basic control law design for a sliding mode controller consists of linear and non-linear terms. The conventional way of tuning a sliding mode controller is to adjust the parameters involve in those two terms, as well as the eigenvalues for the hyperplane design. In the linear term, the general design parameters are Φ, S etc. The general design parameters in the non-linear term are ρ, P etc. The sliding mode controller design normally involves model matrix transformation, matrix inversion etc. The designed controller (in the Matlab Simulink environment) has to be converted to appropriate C-code to download onto the DSP processor. These processes are not transparent to the users when tuning the practical controller. At times, the user has to turn off the engine, go back to the design file and re-start the whole process.

Thus, to avoid the non-transparent controller tuning process, a new tuning approach was introduced in the test. The scaling gains were introduced to the control components and were adjusted to meet the plant performance. The objective of this approach is to ease the trouble of implementing and tuning a robust controller design. The approach has been shown to be effective on the particular diesel engine under consideration. However, this approach is appropriate for practical application on a range of systems.

6.11 Conclusion

This chapter has presented classical sliding mode controller design methods for engine idle speed control. The chapter started with a description of the control problem and the techniques used in the past. Two classical sliding mode controller design schemes have been proposed. Implementation procedure for each designed controller has been described. Issues such as integration step size, model modification and hyperplane design. For the controller implementation, a controller gain tuning algorithm has been formed. The performance tests, robustness tests and speed tracking tests have been carried out. The tuning approach and integral control action issues have been discussed. The controller gain tuning approach has been shown to be straightforward to apply.

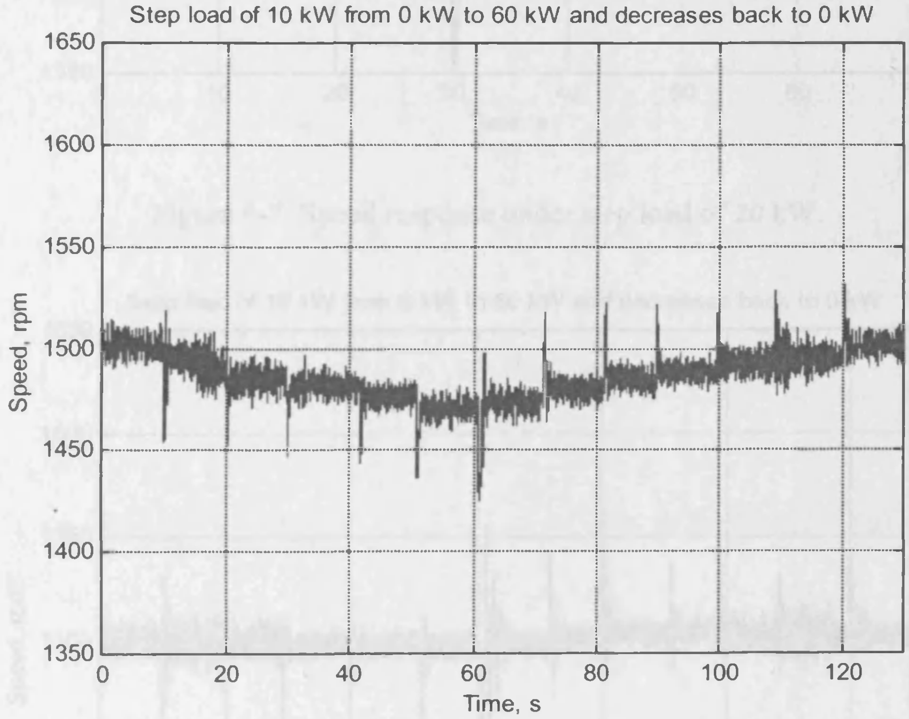


Figure 6-6: Speed response under step load of 10 kW.

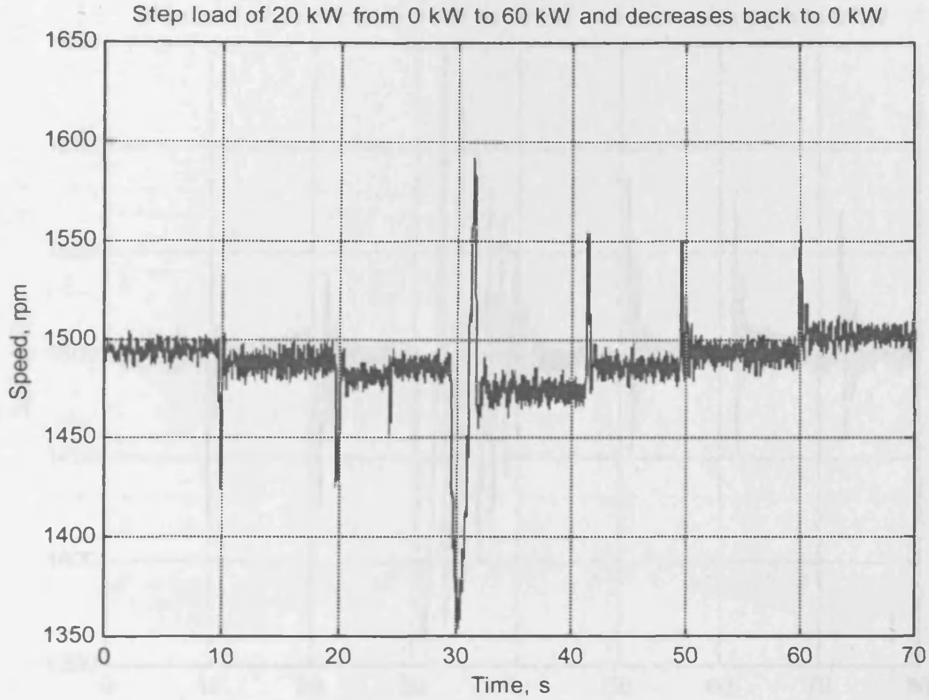


Figure 6-7: Speed response under step load of 20 kW.

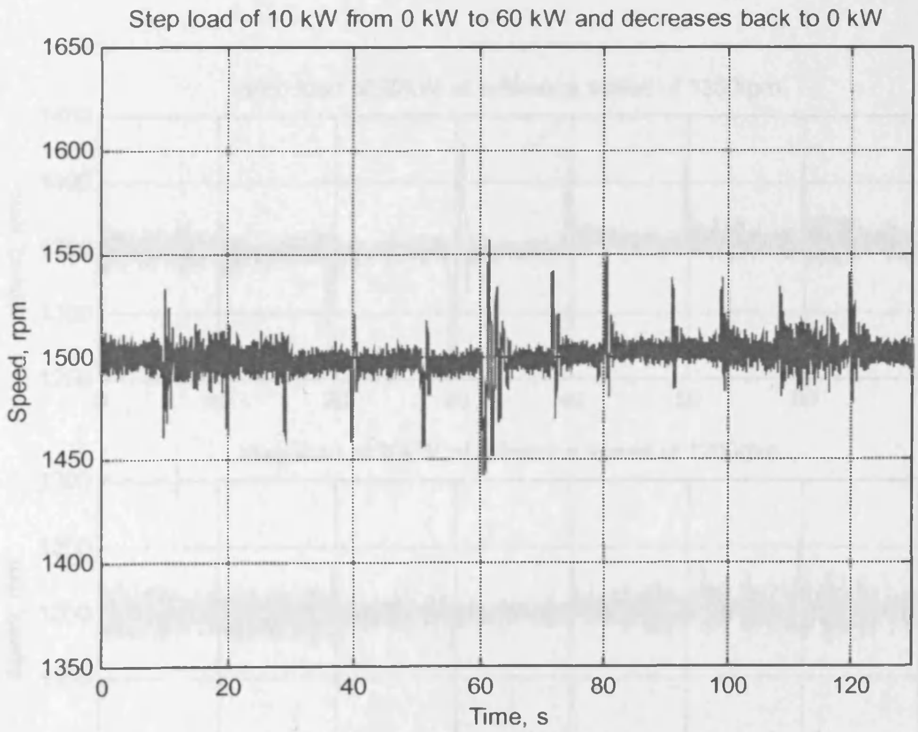


Figure 6-8: Speed response under step load of 10 kW with integral action control.

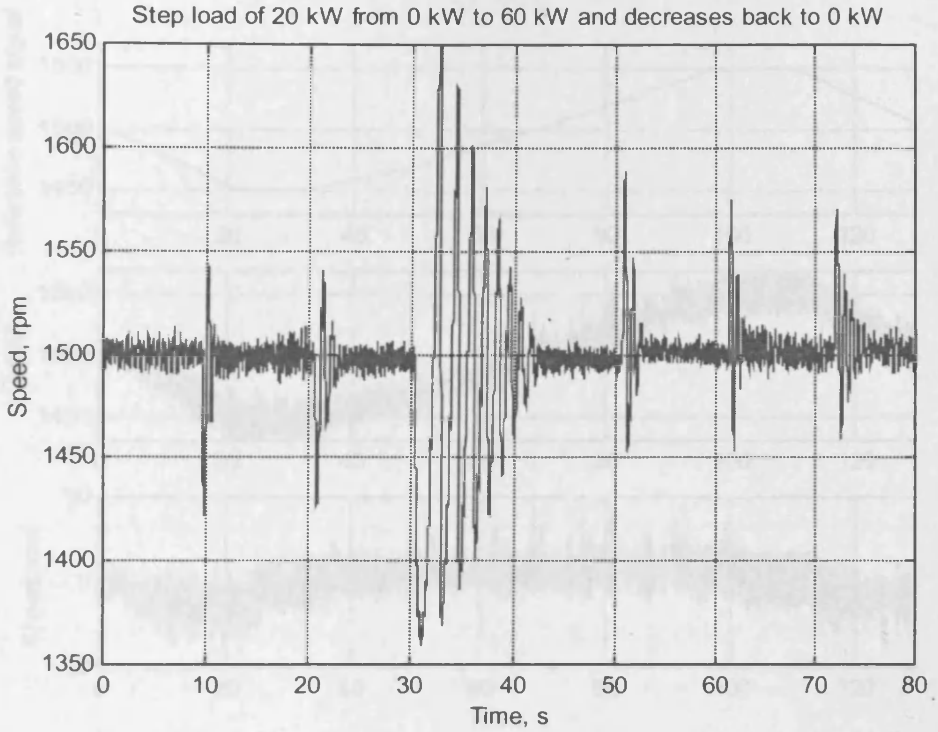


Figure 6-9: Speed response under step load of 20 kW with integral action control.

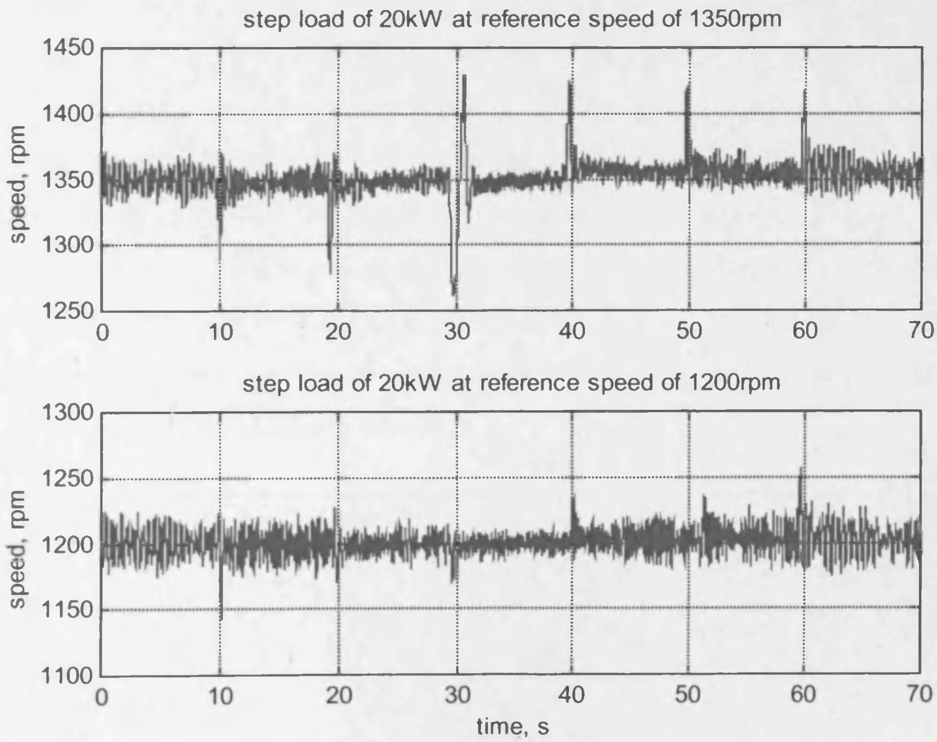


Figure 6-10: Control robustness test at various reference speed settings.

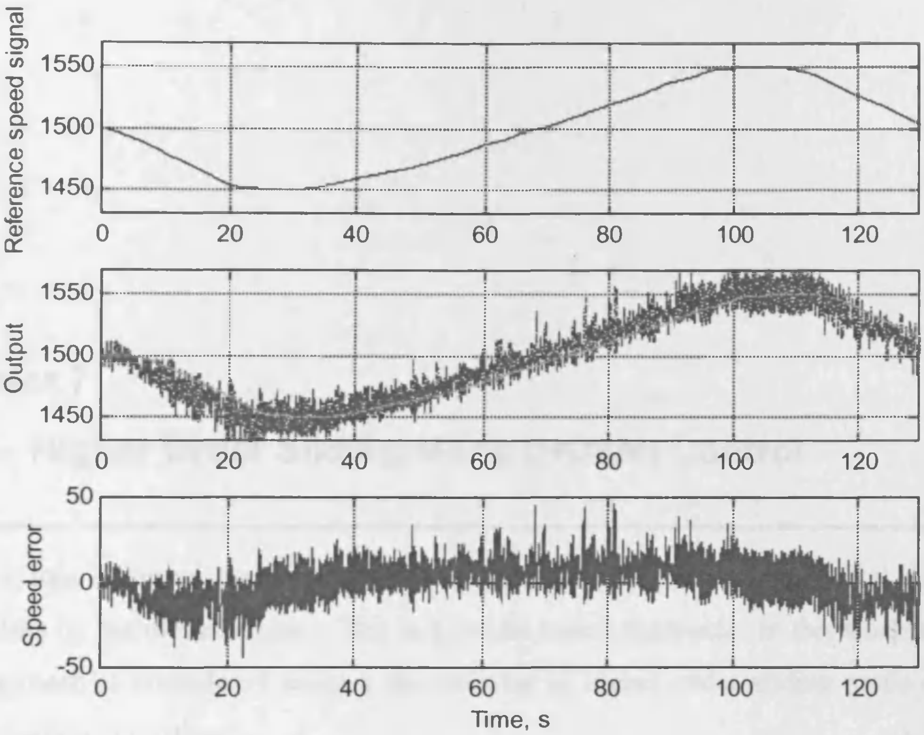


Figure 6-11: Speed tracking performance for MFSM controller.

Chapter 7

7 Higher Order Sliding Mode (HOSM) Control

The previous Chapter has presented the design and application of classical sliding mode controllers on the diesel engine. This is a model-based approach. In this Chapter, a model-free approach is considered using a special case of higher order sliding mode control, the super-twisting algorithm.

The organisation of this Chapter is as follows: A brief introduction to HOSM control is described in Section 7.1, and early work is reviewed. Section 7.2 shows the hardware requirements and setup for the implementation of the HOSM controller. The concept and definition of HOSM are given in Section 7.3 and the particular super-twisting algorithm is presented in Section 7.4. Section 7.5 discusses the engine controller problems for the HOSM controller design. Section 7.6 describes the need for modification to the original algorithm in order to improve the controller accuracy and robustness. A particular HOSM controller gain-tuning algorithm is established and presented in Section 7.7. Implementation of the HOSM is fully described in Section 7.8, as well as several practical test results. Several issues have been raised during the practical test, such as environmental factors and controller issues are discussed in Section 7.9.

7.1 Introduction

Chapter 3 has described the sliding mode concept which is based on variable structure systems. The previous simulation and implementation results have proved the high accuracy and robustness with respect to disturbances of sliding mode control. However, there is a drawback. The so called chattering effect. To avoid this chattering, many approaches exist in the literature. [Slotine 1984] changed the dynamics near to the sliding surface in order to avoid real discontinuity and at the same time to preserve the sliding mode properties. However, the accuracy and robustness of the sliding mode were partially lost. In the previous chapter a nonlinear approximation to the discontinuous term was used. However, again this compromises the inherent properties of sliding mode control.

Another approach proposed to eliminate the chattering problem uses higher order sliding modes (HOSM). The HOSM concept was invented in the 1980s with the motivation of tackling the chattering problem. A number of publications on HOSM were from Emerlyanov *et al* [Emel'yanov 1996], Levant [Levant 1993 1997] and Bartolini *et al* [Bartoline 1998b]. The HOSM generalises the basic sliding mode idea by acting on the higher order time derivatives of the sliding variable instead of influencing the first time derivative as happens in the standard ('classical') sliding mode control. Keeping the main advantages of the original approach, at the same time, HOSM totally mitigates the chattering effect.

Although there are a number of papers in the literature which consider theoretical issues in the area of HOSM and a number of simulation studies exist, very little is reported in the literature about applications and gain tuning procedures of such HOSM control to industrial processes. Thus, one of the research objectives is to practically apply the HOSM controller to solve the idle speed control problem. The design of the HOSM controller is based on a super-twisting 2-sliding algorithm.

7.2 Diesel Generator System, Hardware Setup and Governor

The hardware configuration for the test is similar to the one shown in Figure 5-8. The schematic shows that the engine speed is fed back to the control system (i.e. HOSM or PI controller) and is compared with the nominal speed of 1500 rpm.

7.3 Higher Order Sliding Modes (HOSM)

HOSM techniques have received much attention as it has potential to eliminate the chattering phenomenon, which is inherent in many classical sliding mode control configurations. The sliding order characterises the dynamic smoothness in the vicinity of the sliding mode.

The main problem in implementation of the HOSM is the increasing information demand. Generally, any r-sliding controller keeping $s^r = 0$ needs $s, \dot{s}, \ldots, s^{(r-1)}$ to be available. One of the known exclusions is a so-called "super twisting" 2-sliding controller [Levant 1993], which needs only measurement of s. Hence, this super twisting control scheme [Levant 1993 2001] is applied to the problem of speed control of a diesel generator. It is assumed that upper bounds on the non-linear dynamics are known. The sliding variable here is the speed error. The second order sliding mode control algorithm does not require measurement or estimation of any derivatives of the speed error. It thus operates on a measurement of speed error alone.

7.4 Statement of the Super Twisting Algorithm

Consider a system of the form

$$\dot{x}(t) = \phi(t,x) + \gamma(t,x)u(t)$$

Equation 7-1

where $x = s$ and $\phi(t,x)$, $\gamma(t,x)$ are smooth uncertain functions with $|\phi| \le \Phi_{HOSM} > 0$, $0 < \Gamma_m \le \gamma \le \Gamma_M$. The super twisting algorithm converges to the 2-sliding set $(s = \dot{s} = 0)$ in finite time. The trajectories of the super twisting algorithm are characterised by twisting around the origin on the phase portrait of the sliding variable. This algorithm does not need the sign of the time derivative of the sliding variable. Figure 7-1 shows the phase plot of the super-twisting algorithm. The trajectories perform an infinite number of rotations while converging in finite time to the origin. The vibration magnitudes along the axes as well as the rotation times decrease in as a geometric progression.

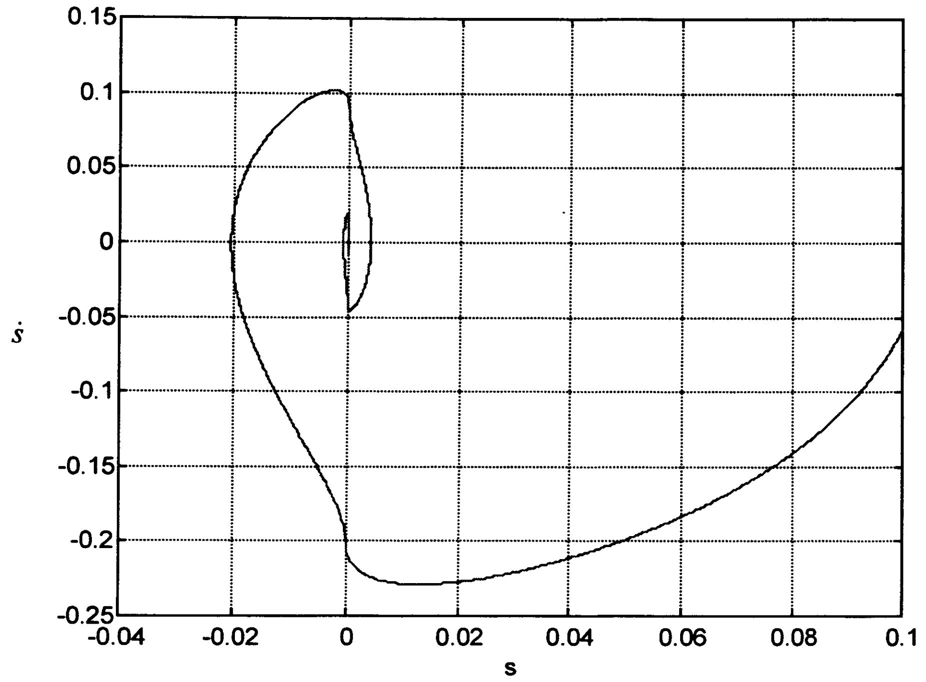


Figure 7-1: Phase plot of the super-twisting algorithm.

The super twisting algorithm defines the control law, $u(t)$ as a combination of two terms. The first is defined in terms of a discontinuous time derivative while the second is a continuous function of the sliding variable. The super twisting algorithm is defined as follows [Levant 1993]:

$$u(t) = u_1(t) + u_2(t)$$

Equation 7-2

$$\dot{u}_1(t) = \begin{cases} -u & |u| > 1 \\ -W_{HOSM}\,sign(s) & |u| \le 1 \end{cases}$$

Equation 7-3

$$u_2(\mathrm{t}) = \begin{cases} -\lambda_{HOSM}\,|s_0|^{\rho_{HOSM}}\,sign(s). & |s| > s_0 \\ -\lambda_{HOSM}\,|s|^{\rho_{HOSM}}\,sign(s). & |s| \le s_0 \end{cases}$$

Equation 7-4

Where W_{HOSM}, λ_{HOSM}, ρ_{HOSM} are variable control parameters and sufficient conditions for finite time convergence are:

$$W_{HOSM} > \frac{\Phi_{HOSM}}{\Gamma_m} > 0$$

$$\lambda_{HOSM}^2 \ge \frac{4\Phi\Gamma_M(W_{HOSM} + \Phi_{HOSM})}{\Gamma_m^3(W_{HOSM} - \Phi_{HOSM})}$$

$$0 < \rho_{HOSM} \le 0.5$$

Equation 7-5

The super-twisting algorithm does not need any information on the time derivative of the sliding variable. The choice of $\rho_{HOSM} = 0.5$ assures that sliding order 2 is achieved.

7.5 Engine Control Problem

In general, any r-sliding controller that keeps $s = 0$ needs $s, \dot{s}, \ddot{s} \ldots, s^{(r-1)}$ to be made available. In the case of engine speed control, this implies that acceleration should either be measured or else an observer constructed to estimate it. Because the "super-twisting" 2-sliding algorithm is used, no knowledge of the engine acceleration is required and hence speed control of the diesel engine is based on speed measurement alone, without an observer.

The main practical problem is the load disturbance on the diesel generator despite the existence of some unknown torque disturbances as well. The changes in load will cause the speed to vary, i.e. speed will drop with increasing load and vice versa. In order to maintain the generated electricity frequency, the problem is to maintain the generator speed at its nominal speed setting. A switching function, s, is defined as the difference between the measured engine speed, N_{mea}, and the desired nominal speed, N_{nom}.

$$s = N_{mea} - N_{nom}$$

Equation 7-6

It is assumed that the engine dynamics are such that the speed error dynamics, and thus the sliding mode dynamics are of the form

$$\dot{s} = \phi(t) + \gamma(t)u$$

Equation 7-7

where u is the applied input signal. The following bounds are assumed on the dynamics

$$0 < \Gamma_m \leq \gamma(t) \leq \Gamma_M$$

$$|\phi(t)| \leq \Phi_{HOSM} > 0$$

Similarly to [Levent 1993, Khalid 2001], the simplified algorithm from equation 7-2, 7-3 and 7-4 for systems linear in the control where $s_0 = \infty$ will be employed in the test. The super twisting algorithm can be simplified as follows:

$$u(t) = -\lambda_{HOSM} |s|^{\rho_{HOSM}} sign(s) + u_1$$

Equation 7-8

$$\dot{u}_1 = -W_{HOSM} sign(s)$$

Equation 7-9

where the conditions for finite time convergence correspond to equation 7-5. This will ensure convergence, in finite time, from any initial speed deviation. This control algorithm does not need any information on the time derivatives of the sliding variable nor any explicit knowledge of other system parameters. This method not only reduces the number of sensors used but also reduces the computational burden of the controller. Effectively the controller can be tuned via three parameters, ρ_{HOSM}, λ_{HOSM}, W_{HOSM} and this can be achieved using the dSPACE-Control Desk application. The following section describes the implementation findings and reports on the modifications required to improve the system performance.

7.6 Modification of the Super Twisting Algorithm

The initial implementation test result showed that the proposed HOSM controller design could not attain the nominal speed at high load conditions. The problem was solved by adding a linear feedback term, L_{add}, of speed error, which is effectively a linear term in s. The resulting performance is shown in Figure 7-2. This modification can be likened to using a linear term providing asymptotic decay on to the sliding manifold, which is often used in the implementation of first order sliding algorithms. Equation 7-8 becomes

$$u(t) = -\lambda_{HOSM}|s|^{\rho_{HOSM}} sign(s) - W_{HOSM} sign(s) + L_{add} s$$

Equation 7-10

The value of ρ_{HOSM} was set to be 0.5. By testing the ρ_{HOSM} at different values, for values less than 0.5, the disturbance rejection properties of the controller degraded. Figure 7-3 shows some sub-plots which show the engine speed response when the ρ_{HOSM} value is reduced. The HOSM control performance degrades with ρ values smaller than 0.5. The W_{HOSM}, λ_{HOSM} and L_{add} values were set to 0.1, 0.01 and -0.001 respectively for the current tests. Details for choosing the gain values are explained in the following section. The results have shown that with the setting of $\rho_{HOSM} = 0.5$, the system gave the best performance compared to other settings of ρ_{HOSM}.

7.7 HOSM Controller Gain-Tuning Algorithm for the Engine

The proposed HOSM controller gain tuning algorithm is simple and straightforward compared to the Ziegler-Nichols method of tuning a PID controller which involves measuring and calculation. There are two methods from Ziegler Nichols.

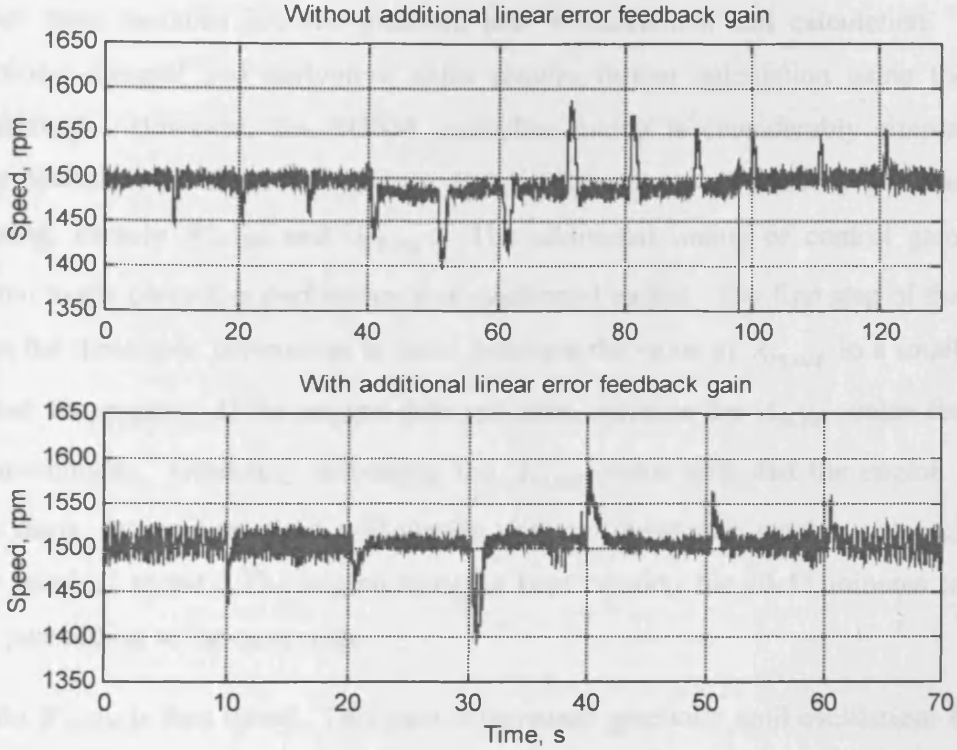


Figure 7-2: Controller performance with and without linear feedback element.

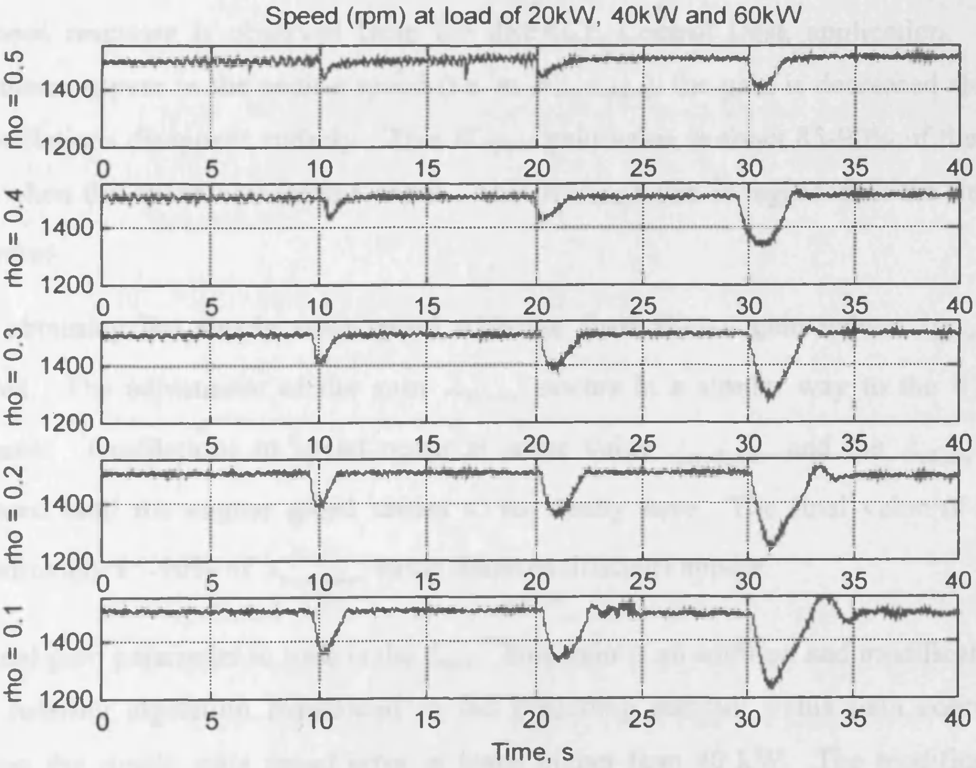


Figure 7-3: Engine response at different ρ_{HOSM} value.

Both of these methods involve practical test, measurement and calculation. The three proportional-integral and derivative gains require further calculation using the obtained measurement. However, the HOSM controller tuning is considerably simpler than the Ziegler-Nichols' PID controller tuning. Practically, only two controller gains are required for tuning, namely W_{HOSM} and λ_{HOSM}. The additional tuning of control gain, L_{HOSM} is subjected to the controller performance as mentioned earlier. The first step of the algorithm is to set the three gain parameters to zero. Increase the value of λ_{HOSM} to a small value and then start the engine. If the engine does not start, increase the λ_{HOSM} value further using small increments. Gradually increasing the λ_{HOSM} value will start the engine. Once the engine starts, the engine speed will slowly rise and eventually reach a value close to the engine nominal speed. The engine must be kept running for 10-15 minutes to warm up before proceeding to the next stage.

The gain W_{HOSM} is then tuned. This gain is increased gradually until oscillations first appear in the speed response. These can be noticed by listening to the response of the engine speed or looking at the engine speed indicator. In the process of finding the gain-tuning algorithm, the speed response is observed from the dSPACE Control Desk application. Once the oscillations appear in the engine speed (i.e. at $W_{oscillation}$), the gain is decreased slowly until the oscillations disappear entirely. This W_{HOSM} gain value is about 85-90% of the $W_{oscillation}$ value when the speed oscillations occur. This W_{HOSM} value is regarded as the final setting for W_{HOSM}.

After obtaining the steady state speed with the final W_{HOSM} gain value, λ_{HOSM} is then adjusted. The adjustment of the gain λ_{HOSM} occurs in a similar way to the W_{HOSM} gain parameter. Oscillations in speed occur at some value $\lambda_{oscillation}$ and the λ_{HOSM} value is decreased until the engine speed settles to its steady state. The final value of λ_{HOSM} is approximately 85-90% of $\lambda_{oscillation}$ value when oscillations appear.

The final gain parameter to tune is the L_{add}. This gain is an addition and modification to the super twisting algorithm mentioned in the preceding section. This gain contributes to reducing the steady state speed error at loads higher than 40 kW. The modification also helps in speed recovery during speed drop and speed overshoot. The setting of this gain is small and is about 0.001% of the maximum control signal range value for the particular engine. Another way to set this gain is by trial and error. Set the gain to a small value, apply high load (i.e. above 60% of the declared power value of a engine) to the engine and observe the engine speed. Repeat the steps until good performance (i.e. no steady state speed error) is obtained. The system performance will degrade if the L_{add} gain is set too high. The above algorithms on the HOSM gain tuning are summarised in Table 7-1.

Step	Description
1	Set all the gain parameters to zero. Increase λ_{HOSM} and start the engine. Warm the engine for 10-15 minutes.
2	Increase W_{HOSM} until oscillations appear in speed and then decrease W_{HOSM} until the speed reaches steady state. $W \approx$ 85% to 90% of $W_{oscillation}$
3	Increase λ_{HOSM} until oscillations appear in speed and then decrease λ_{HOSM} until the speed reaches steady state. $\lambda_{HOSM} \approx$ 85% to 90% of $\lambda_{oscillation}$
4	Set L_{add} gain to 0.001% of the highest control signal range value or by trial and error method.

Table 7-1: HOSM controller gain tuning algorithm for the engine.

The above algorithm can be applied with no calculation involved in the HOSM controller gain tuning process which is not the case for the Ziegler-Nichols method. There is no report in the literature on the HOSM controller gain tuning algorithm for a physical industrial process like the engine. This algorithm is established based on the particular engine. The established algorithm can serve as a guideline for a non-expert operator to tune the three controller gains conveniently.

7.8 HOSM Controller Implementation and Results

The hardware setup for the controller implementation is similar to the previous controller tests. The controller gain tuning has been described and the performance of the controller is now assessed.

7.8.1 Start-up of Engine

This controller is capable of starting the engine. There is an integral gain block in the design and thus this integral sum will start immediately. The initial engine speed error is large (i.e. 1500 rpm) and this causes a large integral sum of speed error. Thus, before the engine is switched on with HOSM, the integral action has to be reset. The starting performance is shown in Figure 7-4. The second subplot gives the required control action to perform the start-up of the engine. The initial control signal is rather high due to the large speed error.

At this moment, the throttle opening should be open at maximum to pump in more fuel to the engine. The transient response of the start-up phase is given in Table 7-2.

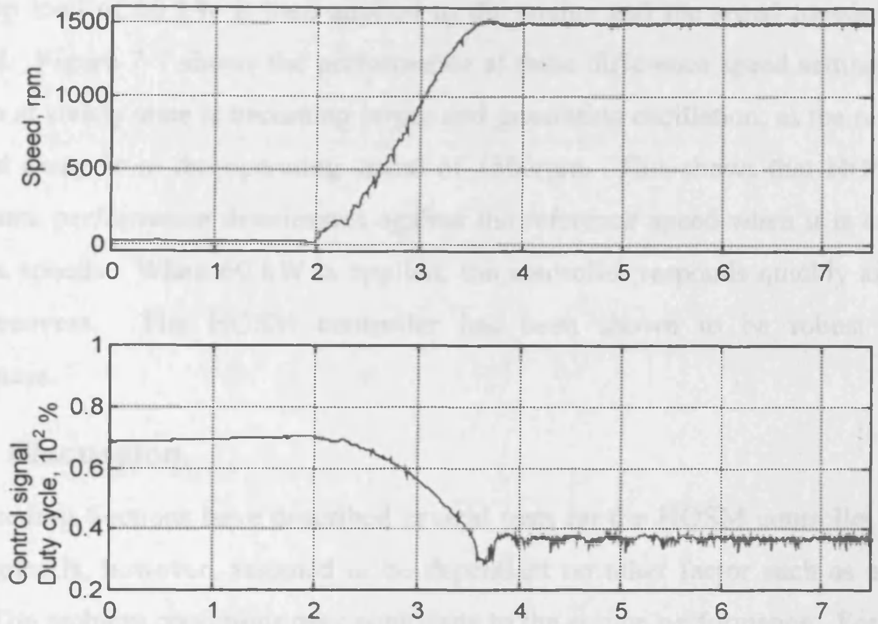


Figure 7-4: Start-up of the engine using the HOSM controller.

Speed transient response			
Overshoot (%)	Rise time (s)	Settling time (s)	Delay (s)
1.600	1.41	2.50	0.89

Table 7-2: Transient response during the start-up of the engine.

7.8.2 Performance Tests

When the controller is tuned to a level which gives the best overall performance (i.e. in term of disturbance rejection), the immediate observation is that the controller maintains the engine speed with small speed variation of 1.784% of the reference speed setting. For tracking and disturbance rejection tests, a small step load is applied to the engine and the performance is observed. The resulting performance is shown in Figure 7-5 when a step load of 20 kW is applied to the engine. The load is gradually increased up to 60 kW and then decreased back to 0 kW. The engine speed dips when the load steps in and vice versa. The maximum speed dip is about 13% when the load changes from 40 kW to 60 kW. The second subplot shows the required control action to maintain the engine speed. Figure 7-6 shows the controller performance under a large load of 60 kW and the control signal demand is given in the second subplot.

7.8.3 Robustness Tests

The robustness test is similar to the one performed in Chapter 5 on the classical sliding mode controllers. The engine reference speed is set to 1500 rpm, 1350 rpm and 1200 rpm. A large step load of 60 kW is then applied to the engine and the speed transient response is observed. Figure 7-7 shows the performance at three difference speed settings. The speed variation at steady state is becoming larger and generating oscillation, as the reference speed is drifted away from the operating speed of 1500rpm. This shows that HOSM controller steady state performance deteriorates against the reference speed when it is set at different reference speeds. When 60 kW is applied, the controller responds quickly and the engine speed recovers. The HOSM controller has been shown to be robust against load disturbances.

7.9 Discussion

The preceding Sections have described several tests on the HOSM controller. The engine performance is, however, assumed to be dependant on other factor such as environmental factor. The ambient conditions may contribute to the engine performance. For example, the starting characteristic may depend on several factors such as air temperature and starting air pressure.

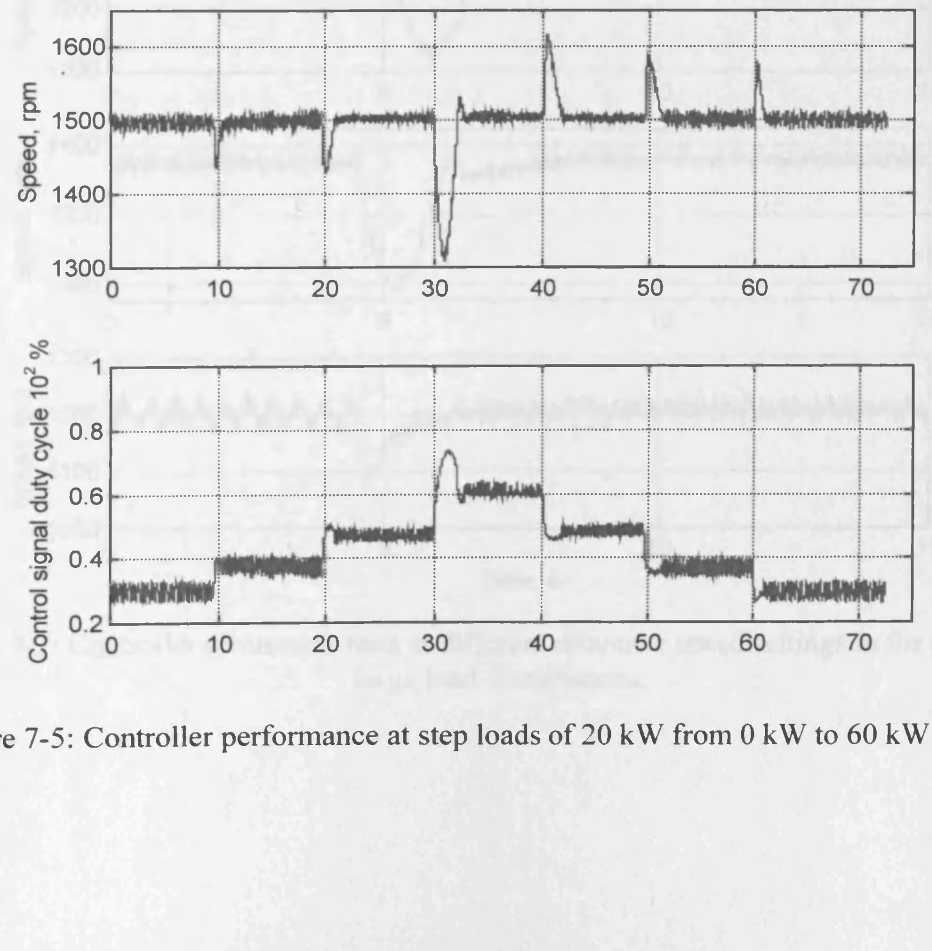


Figure 7-5: Controller performance at step loads of 20 kW from 0 kW to 60 kW to 0 kW.

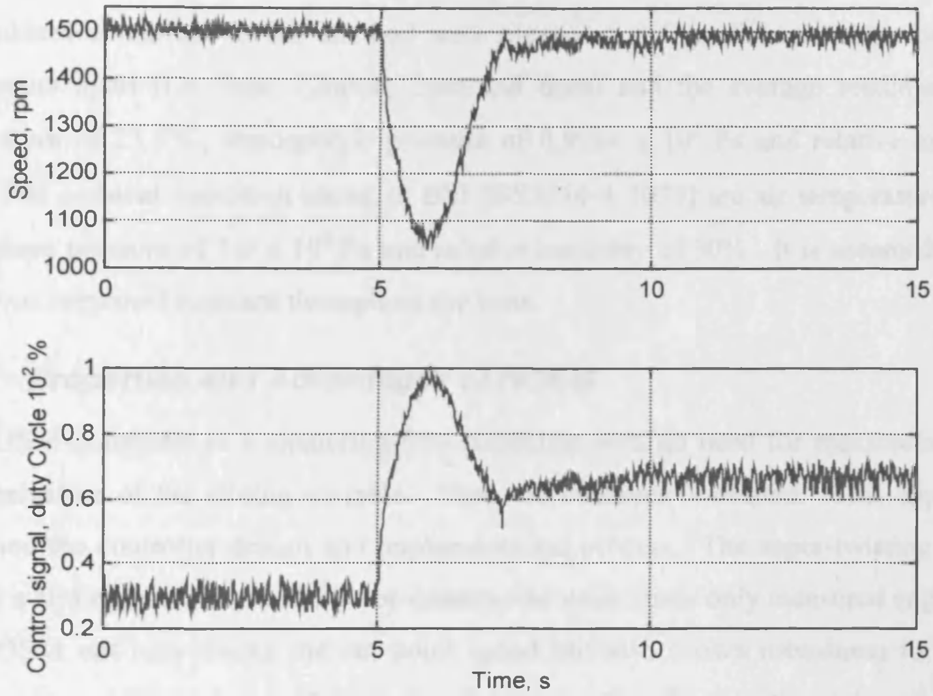


Figure 7-6: Speed and control signal response to a load of 60 kW at a speed of 1500 rpm.

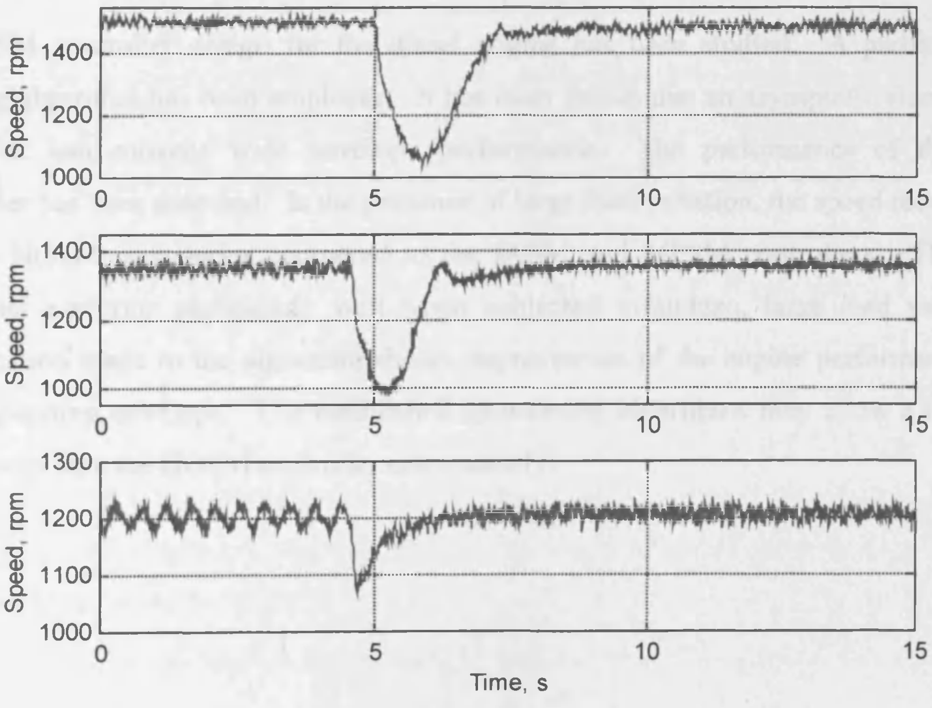


Figure 7-7: Controller robustness tests at different reference speed settings in the presence of large load disturbances.

7.9.1 Environmental Factors

The ambient conditions at the test bed were measured at four different times on a day at three hours apart (i.e. 9am, 12noon, 3pm and 6pm) and the average readings were air temperature of 23.5°C, atmospheric pressure of 0.9984×10^5 Pa and relative humidity of 53%. The ambient condition stated in BSI [BS5514-4 1979] are air temperature of 25°C, atmosphere pressure of 1.0×10^5 Pa and relative humidity of 30%. It is assumed that these conditions remained constant throughout the tests.

7.9.2 Properties and Advantages of HOSM

The HOSM controller is a chattering-free controller with no need for measurement of the time derivative of the sliding variable. Thus, no observer is needed. This approach has simplified the controller design and implementation process. The super-twisting algorithm enables a sliding mode controller to be constructed which uses only measured engine speed. The HOSM not only tracks the set point speed but also shows robustness to parameter variations (modeling error) and load disturbances. Most importantly, it is valid across a wide operating envelope and has minimal use of computer resources.

7.10 Conclusion

A HOSM controller design for the diesel engine has been studied. A particular super twisting algorithm has been employed. It has been shown that an asymptotic element in the controller can enhance wide envelope performance. The performance of the HOSM controller has been assessed. In the presence of large load variation, the speed recovery time for the HOSM is superior compared to the IASM and MFSM controllers. The HOSM controller performs particularly well when subjected to sudden, large load variation. A modification made to the algorithm shows improvement of the engine performance over a wide operating envelope. The established gain-tuning algorithms may allow a non-expert operator to tune the HOSM controller conveniently.

Chapter 8

8 Controller Performance Assessment and Comparison

The previous Chapters have described the design and implementation of several sliding mode controllers. This Chapter will focus on the controller performance assessment by comparing all the designed controllers particularly against the benchmark set by the relevant British Standards, and also against the performances obtained by commercial (COM) and classical PI controllers. Both the COM and PI controllers undergo the test criteria as described in Chapter 2.

This Chapter is organised as follows: A brief introduction to the performance assessment is given in Section 8.1. The first performance test is considered in Section 8.2; small and large loads are applied to the engine and the controller performance is studied. Controller robustness tests are presented in Section 8.3 and the controllers are tested across a wide operating envelope. Section 8.4 and Section 8.5 describe the fuel consumption and exhaust gas emissions produced by the different controllers. Engine steady state speed variation is observed in Section 8.6 and the speed variation is presented as a percentage of the reference engine speed setting. Section 8.7 describes the capability of the different controllers including the ability to handle the engine start-up phase. Particular speed transient and exhaust gas parameters are studied. Finally, particular issues of concern noted during the tests are described in Section 8.8.

8.1 Introduction

This chapter investigates the performance of the various designed controllers. The performance is compared with that obtained by a commercial genset (COM) controller and a classical proportional/integral (PI) controller. The influence of the sliding mode methodology and modification of the algorithm on overall generator performance, in particular in the presence of large load changes and in terms of fuel efficiency, exhaust emissions, starting speed transient response and steady speed variation, are assessed. The test criteria which have been described in Chapter 2, are taken as a guideline for these tests. The following Sections will deliver the details of the tests.

8.2 Performance Tests

The dynamic behaviour of interest is the speed change with load. The transient response is defined as the maximum deviation of speed after sudden load change from previous speed to steady state level and the speed change is expressed as a percentage of operating speed (i.e. 1500 rpm). These tests involve applying an electrical load to the system during steady state conditions and the speed transient responses are studied. Two objectives are set for the tests. The first objective is to simulate the situation of small varying demands by the consumer. This can be achieved by applying a small load (i.e. 20 kW) to the engine. The second objective of the tests is to assess the controllers in the presence of a large step change in load (i.e. 60 kW). This corresponds to a sudden high power demand by the consumer. The resulting performances are shown in Figure 8-1 and Figure 8-4 respectively.

Figure 8-1 shows the engine speed response when a step load of 20 kW (i.e. from 0 kW to 60 kW and then back to 60 kW) is added to the diesel engine. The speed recovers within three seconds for a step load of 20 kW for the HOSM controller. The IASM controller shows the ability to cope with step loads up to 60 kW. However, it exhibits a longer settling time and larger speed dips. During the step off load (i.e. from 60kW to 40kW), the engine speed hit the speed safety level (i.e. about 1650rpm) and the engine fuel was cut off and shutdown the engine automatically. Thus, the load decrement was reduced to 10kW from 60kW to 10kW. The subplot 2 in Figure 8-1 shows the resulting effect. The integral action MFSM copes reasonably well with the step load. The controller performance shows a smaller speed dip but suffers with oscillation when the load step is from 40 kW to 60 kW. The engine speed settles down after about 12 s. Both the PI and COM controllers perform better at small varying loads. Both of the controllers perform consistently with short recovery time.

During the large load test, the HOSM controller settles down in about 2.5 seconds. A large speed drop is observed due to the large change in load. The settling time for the HOSM controller is faster than the rest of the controllers; the COM took about 5 seconds settling time and the IASM needs a longer time than the COM due to low frequency oscillation in the speed signal. The PI and MFSM controllers struggle with such a large step load change and do not regain the set point; the speed settles down to about 1300 rpm and 1000 rpm respectively. The MFSM can handle a maximum large step load of 45 kW.

Overall, the HOSM shows a faster settling time to large step changes in load especially at the nominal genset speed, 1500 rpm. The COM controller copes well with load changes but takes a longer time to reach the nominal speed. The MFSM controller and the traditional PI controller cannot tolerate large step changes in load at any speed. A possible reason for the poor performance obtained from the PI controller is the limited integral term which was used to avoid control signal saturation. The IASM controller copes wells in both load situations but with a longer settling time. In the BSI standard, only 10 % nominal speed drop is allowed and the speed must recover in less than 8 seconds. The results show that the HOSM controller has an approximate 27% speed drop, 35% for the IASM controller and the COM controller has a 23% drop at a step load of 60 kW. This shows that none of the tested controllers are capable of meeting the 10% speed drop.

8.3 Robustness Tests

The objective of this test is to show the robustness of the controller performances at different operating conditions. Similarly to the previous test in Chapter 6 and 7, the reference speed of the diesel engine is set to 1350 rpm and 1200 rpm. These changes are done online during the test. Similar test procedures to those used in the preceding section are employed here: the engine is subject to a large speed load of 60 kW and a series of 20 kW step load; the speed response is studied. Figure 8-2, Figure 8-3, Figure 8-5 and Figure 8-6 show the speed response for the mentioned tests respectively. The instant observation of the test is that all controllers, except the PI controller, maintain the engine reference speed setting. The HOSM shows a faster speed recovery than the other controllers at 1350 rpm but the COM controller recovers faster at 1200 rpm. The IASM controller shows similar performance and copes reasonably well with electrical load but with a longer settling time compared to the HOSM controller. The modified MFSM controller tracks the speed well at the reference setting. However, the controller cannot cope with a large step load of 60 kW at 1350 rpm. The particular PI controller doesn't maintain the new speed setting. This test shows the

HOSM controller is robust and capable of controlling diesel engine speed over a wide operating envelope. The analysis shows the HOSM performs better at a reference speed of 1350 rpm. The HOSM, MFSM, IASM and COM controller outperform the PI controller at all reference speed settings. During large loading conditions, the HOSM, IASM and COM controllers perform better than the MFSM and PI controllers.

8.4 Fuel Consumption

The fuel consumption of the controllers is now considered. Two methods of fuel consumption calculations are presented here, *volume per second* ($10^{-9}m^3/s$) and *volume per unit of power* ($10^{-9}m^3/kW$). The test is designed to run with the engine operating at different loading conditions. For each of the tests, three sets of data were recorded and the average values are considered. Upon each load change, the measurements are only made when the speed reaches steady state. Table 8-1 shows the test results.

		Electrical load						
Controller type		**0 kW**	**10 kW**	**20 kW**	**30 kW**	**40 kW**	**50 kW**	**60 kW**
HOSM		0.5979	1.2474	1.9056	2.5674	3.2334	3.9632	4.7794
	*	0.5979	0.1247	0.0953	0.0856	0.0808	0.0793	0.0797
IASM		0.5913	1.1880	1.8668	2.5502	3.2361	3.9307	4.7337
	*	0.5913	0.1188	0.0933	0.0850	0.0809	0.0786	0.0789
MFSM		0.5584	1.2130	1.8549	2.5502	3.2029	3.8819	4.6691
	*	0.5584	0.1213	0.0927	0.0850	0.0801	0.0776	0.0778
COM		0.5979	1.2224	1.9561	2.5760	3.2112	3.9036	4.7337
	*	0.5979	0.1222	0.0978	0.0859	0.0803	0.0781	0.0789
PI		0.6012	1.4149	2.0920	2.7012	3.3824	4.1199	4.9814
	*	0.6012	0.1415	0.1046	0.0900	0.0846	0.0824	0.0830

Table 8-1: Fuel consumption at different loading conditions for different controllers.

The rows with a single star represent units of $10^{-9}m^3/kW$. From the results, the HOSM IASM, MFSM and COM controllers show good performance compared to the PI controller. The MFSM controller has an advantage at higher load conditions whereas the IASM controller gives better performance in the low load situation. The fuel consumption in *volume per second* is plotted against the applied load to give a better visual comparison as in Figure 8-7. The PI controller is consistently showing higher fuel consumption. The rest of the controllers show similar characteristics of fuel consumption against load.

8.5 Exhaust Gas Emissions

The objective of this test is to measure the quality of the exhaust emissions from the genset. The exhaust is measured in opacity % unit which is measuring the relative light intensity through the exhaust smoke. The principle of the measurement has been described in Chapter 2. The opacity readings are recorded for load changes from 0 kW to 60 kW at each step change of 10 kW. At every load condition, the steady state is achieved before data recording. Three data sets (each of 60 seconds) are recorded for each load condition and the average values of the readings are plotted in Figure 8-8. The results show that all controllers except the PI controller have similar smoke opacity readings. The PI controller emits rather thick smoke during 20 kW and 30 kW load. At higher loading condition, both the IASM and MFSM show a lower exhaust opacity emission and the other three controllers follow in a similar fashion at a higher opacity %.

8.6 Steady State Behaviour – Speed Band

This test is to determine the width of the envelope of variation of the engine speed under steady state conditions. It is expressed as a percentage of the reference speed (i.e. 1500 rpm). Several steady state loading conditions were applied to the engine. The power level categories which have been described in Chapter 2 Section 2.8.1 are shown in Table 8-2. A total of three data sets for each load were recorded and the average speed variation values measured. The data recording was only started when the speed reached its steady state. The speed variation (i.e. in percentage of the nominal speed value) results are tabulated in Table 8-2.

The results show the COM controller has a better performance in terms of speed variation at higher load condition. However, the COM controller shows consistently large speed variation at loads of 20 kW and 30 kW. Among the other controllers, the PI controller performs the worst. At the low loading condition, the IASM controller shows a better performance over the rest of the controllers. The BSI standard for speed variation is 0.8% and 1.0% for $\geq 25\%$ and $< 25\%$ power respectively. All the controllers show that they are struggling to meet BSI standards.

	Power Category						
	$\leq 25\%$ of 65kW		$> 25\%$ of 65kW				
Controller type	**0 kW**	**10 kW**	**20 kW**	**30 kW**	**40 kW**	**50 kW**	**60 kW**
HOSM	1.784	2.197	1.888	1.611	1.622	1.872	1.975
IASM	1.280	1.242	1.356	1.367	1.628	1.655	2.469
MFSM	1.639	1.758	1.845	1.297	1.568	1.530	1.790
COM	1.156	1.590	2.311	1.964	0.933	1.123	1.194
PI	1.557	2.951	2.724	3.109	2.800	2.778	3.678

Table 8-2: Tabulation of steady state speed variation (%) for different controllers.

8.7 Start-up of Diesel Generator

Among the designed sliding mode controllers, the HOSM controller has the capability of starting the engine. Thus, this section concentrates on the three controller performances (i.e. HOSM, COM and PI controllers) for the engine-starting phase. Two engine parameters were investigated during the test, namely, engine speed and exhaust emission. During the test, three sets of data are recorded for each controller and average values are calculated. The performance results are shown in Table 8-3. The respective parameters are plotted in Figure 8-9

	Speed transient response				Exhaust emission	
Controller type	**Overshoot (%)**	**Rise time (s)**	**Settling time (s)**	**Delay (s)**	**100% opacity (s)**	**Settling time (s)**
HOSM	1.600	0.141	0.250	0.089	0.067	1.617
COM	0.000	0.252	0.318	0.148	0.028	0.292
PI	1.530	0.190	0.371	0.116	0.080	0.300

Table 8-3: Genset response during starting on different controllers.

The transient properties have been defined in Chapter 2. The results show that the HOSM controller has 1.6% overshoot while the COM controller shows heavy damping, which has 0% overshoot. The HOSM provides the best transient responses in rise, delay and settling time. For the exhaust emission test during starting, the COM controller has a good response compared to the HOSM and PI controllers. Overall, the HOSM performs best in speed transient response and the speed overshoot is within the 2% range mentioned in the BSI standard.

8.8 Discussion

The overall performance of the designed sliding mode controllers has been assessed with comparison to a commercial controller and a PI controller. There are several issues concerning the outcome of the test results. Environmental factors may affect the overall results. The ambient conditions and other factors may contribute to the genset performance. For example, the starting characteristic may depend on several factors such as air temperature, temperature of the reciprocating internal combustion (RIC) engine, starting air pressure , condition of the starter battery, viscosity of the oil, total inertia of the generating set and quality of the fuel and the state of the starting equipment. The quality of exhaust emission may depend on the air intake constituents and the type of fuel used.

The ambient conditions at the test bed were measured prior to the test and the measurement is tabulated in Table 8-4, as well as the ambient condition stated in BSI [BS5514-4 1979]. The HOSM, PI, COM are tested on the same day. In fact the previous results on exhaust emission show that these three controllers give a rather similar opacity % reading. This results indicate that the ambient conditions may be an issue in comparing different controllers if the test is performed at different ambient conditions.

8.9 Conclusion

The sliding mode controllers performances for the particular diesel engine have been assessed and studied. For the HOSM controller, which uses a particular super twisting algorithm, it has been shown that an asymptotic element in the controller can enhance wide envelope performance. The IASM controller has been shown to track the speed well both at low and high load conditions. The performance of the MFSM can only be improved with the use of integral action to solve the steady state error problem. This controller is capable of tracking the reference speed at low and medium load conditions. i.e. up to 45 kW. The HOSM, COM, MFSM and IASM controllers appear to outperform a classic PI controller across the whole operating envelope. The particular HOSM controller performance is seen to be broadly similar in some tests to that achievable by a current commercial engine control system. The HOSM controller performs particularly well when subjected to sudden, large load variation and in terms of transient response during engine start. For the steady state speed variation test, the IASM performs better at low and medium load conditions whereas at medium to higher load condition, the HOSM controller stands out among the rest of the controllers under consideration.

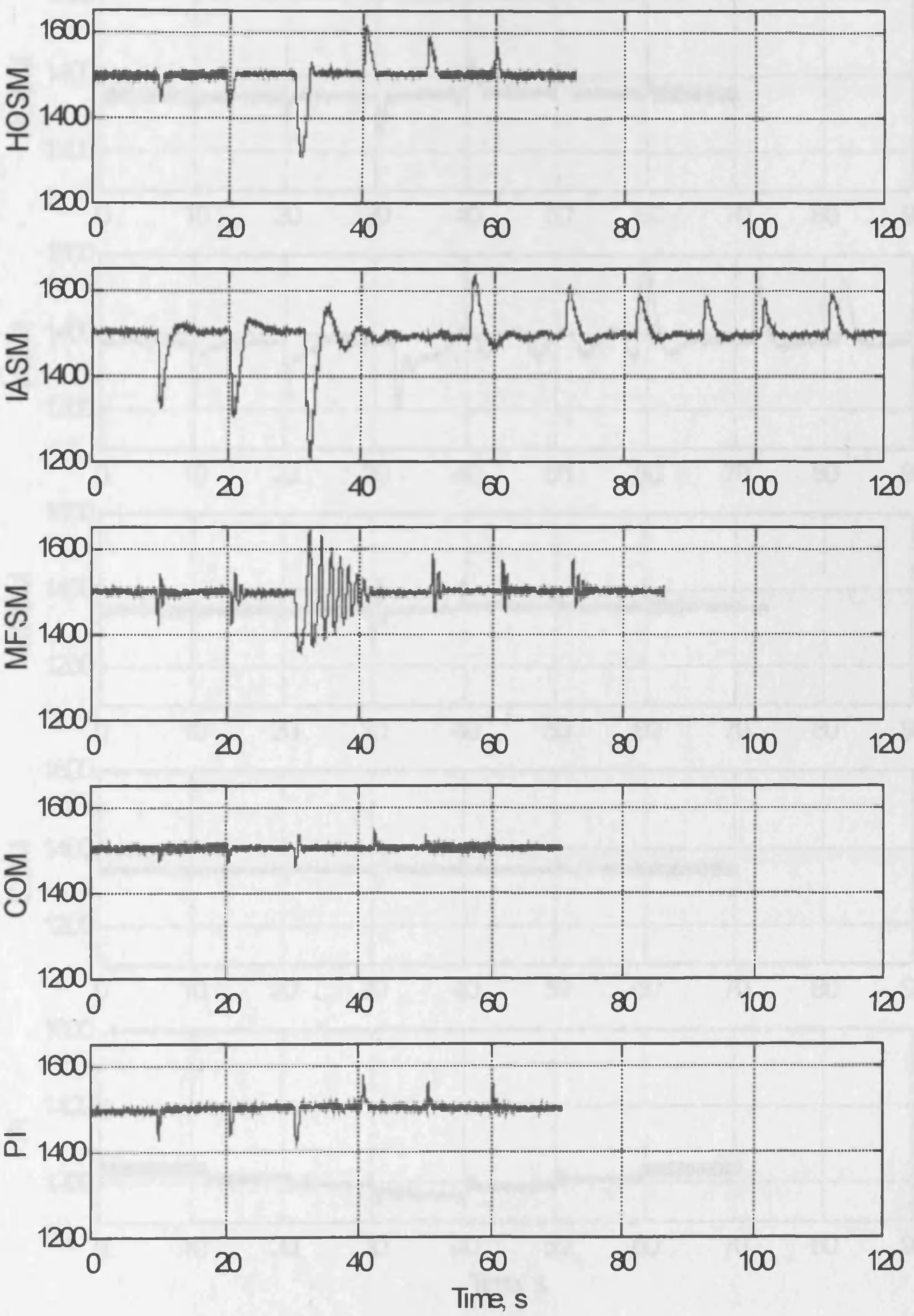


Figure 8-1: Controller performance to a step load change of 20 kW up to 60 kW and step down back to 0 kW at 1500 rpm.

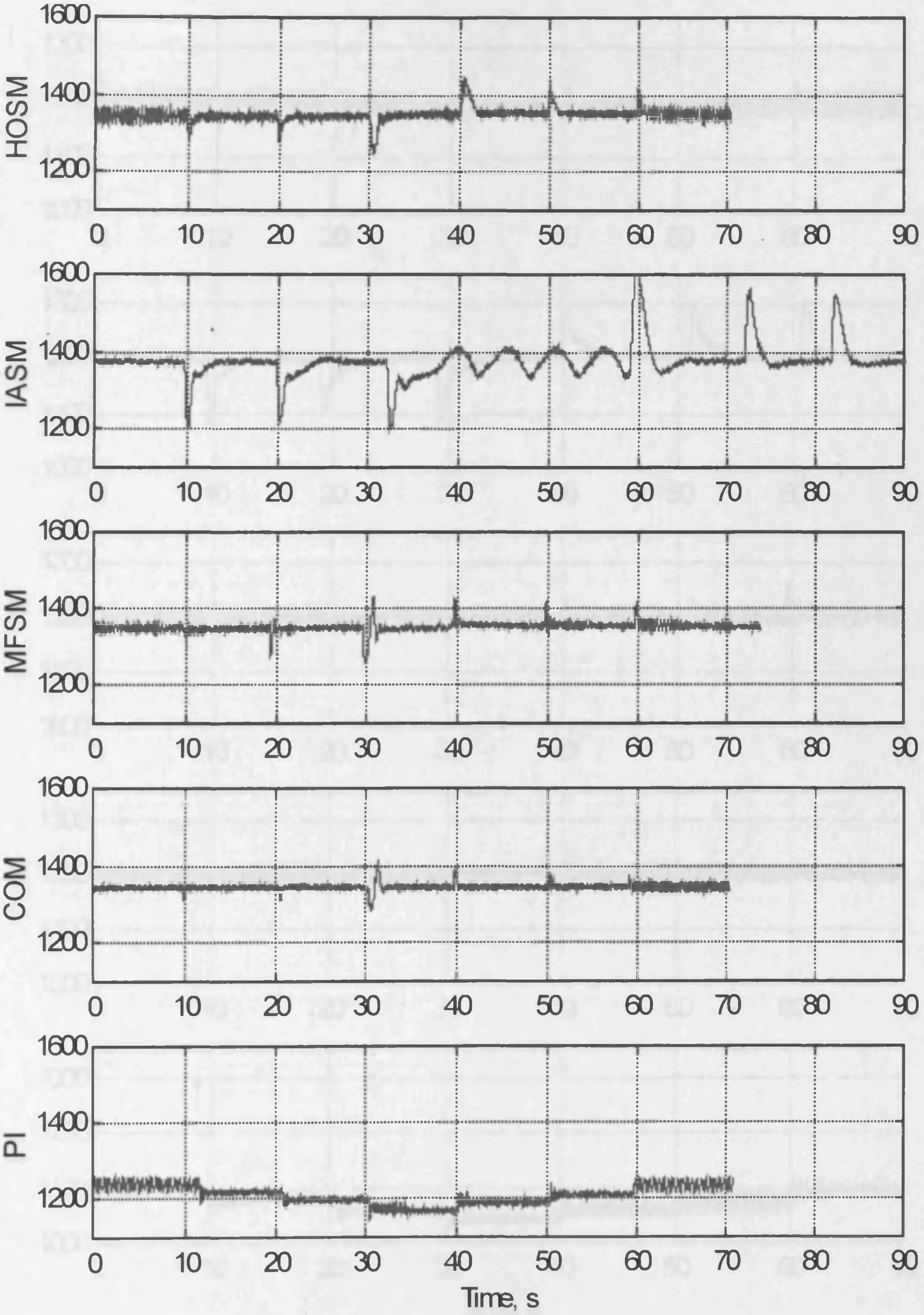


Figure 8-2: Controller performance to a step load change of 20 kW up to 60 kW and step down back to 0 kW at 1350 rpm.

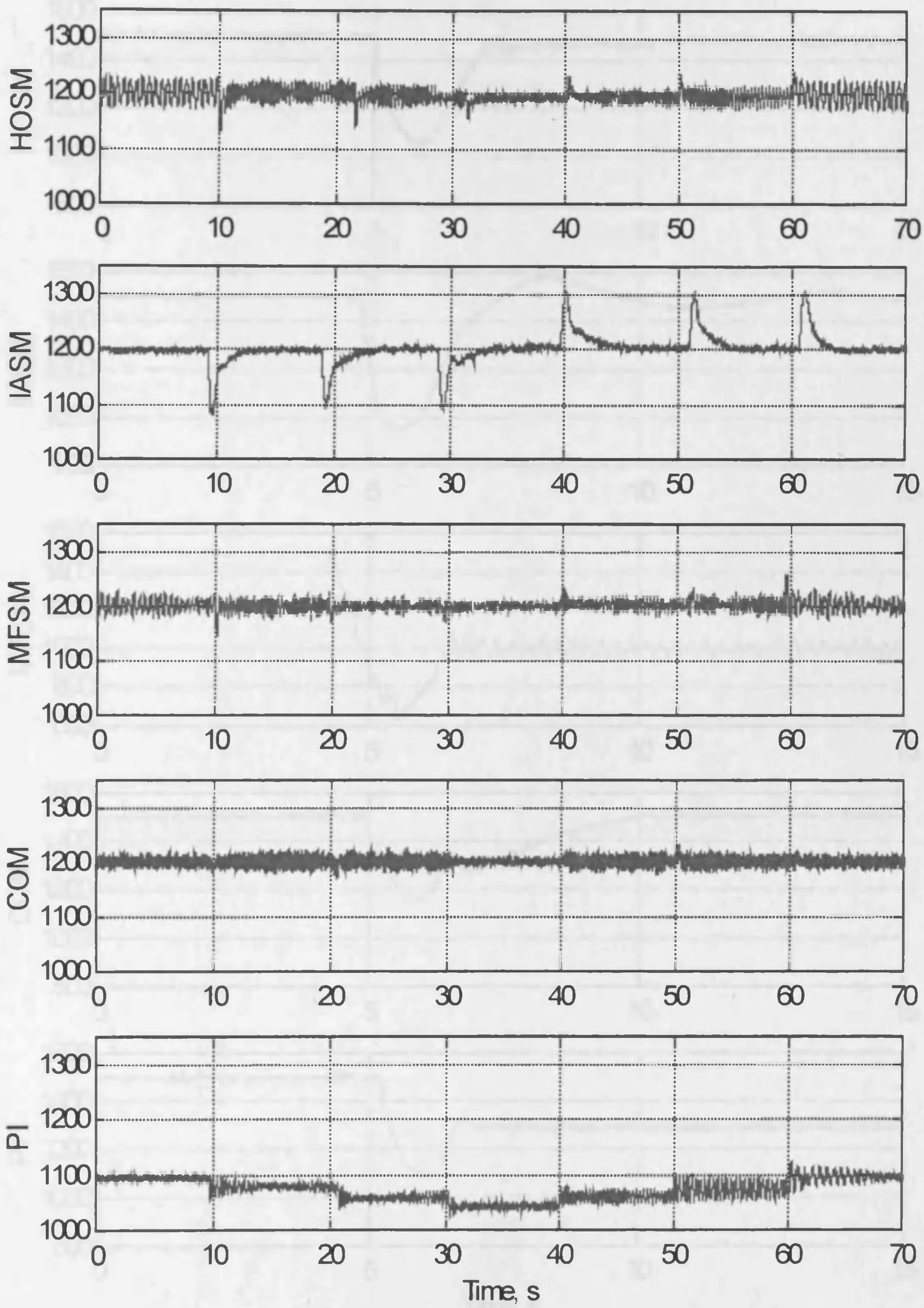


Figure 8-3: Controller performance to a step load change of 20 kW up to 60 kW and step down back to 0 kW at 1200 rpm.

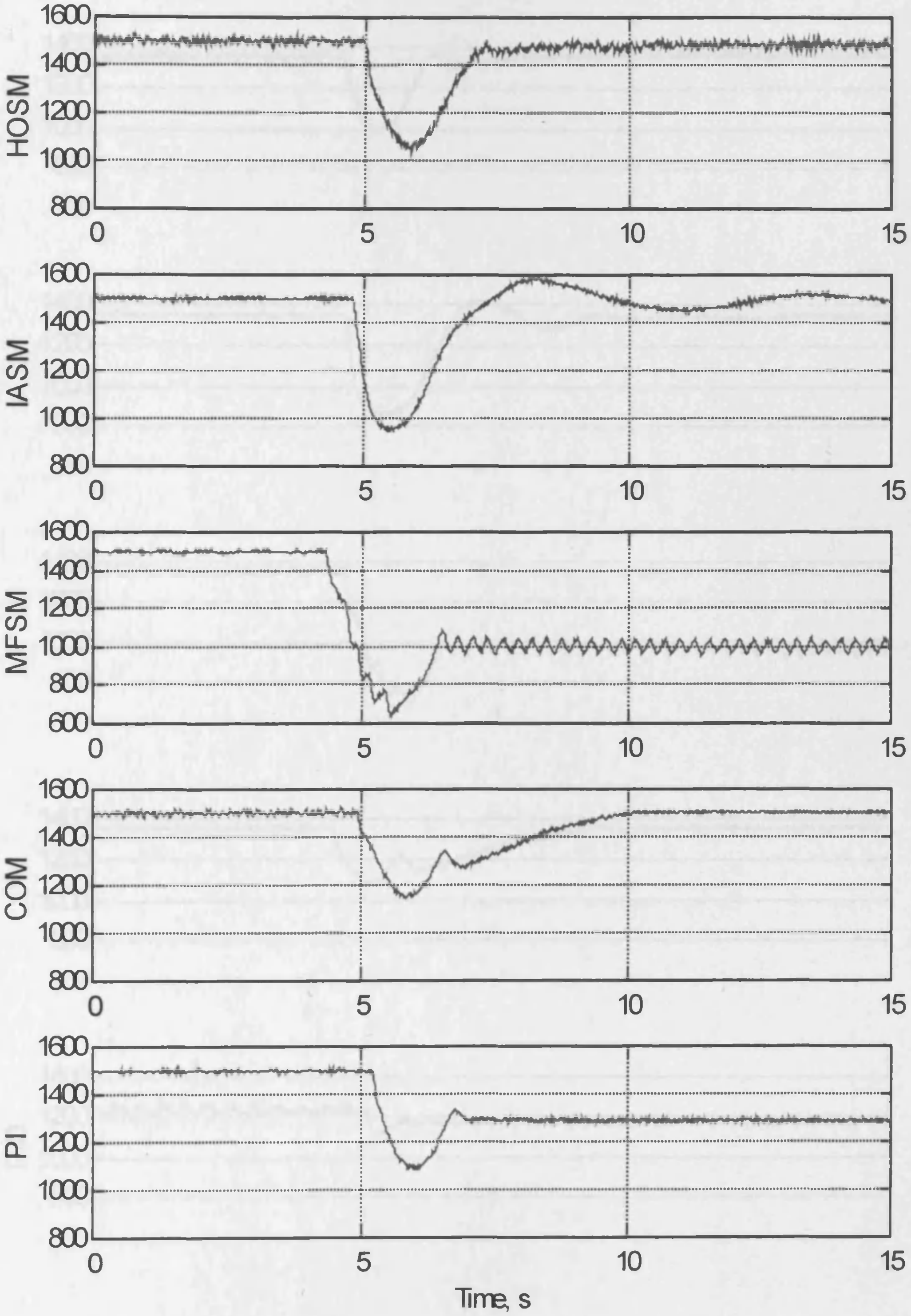


Figure 8-4: Comparison of the different controllers to a step change in load of 60 kW at 1500 rpm.

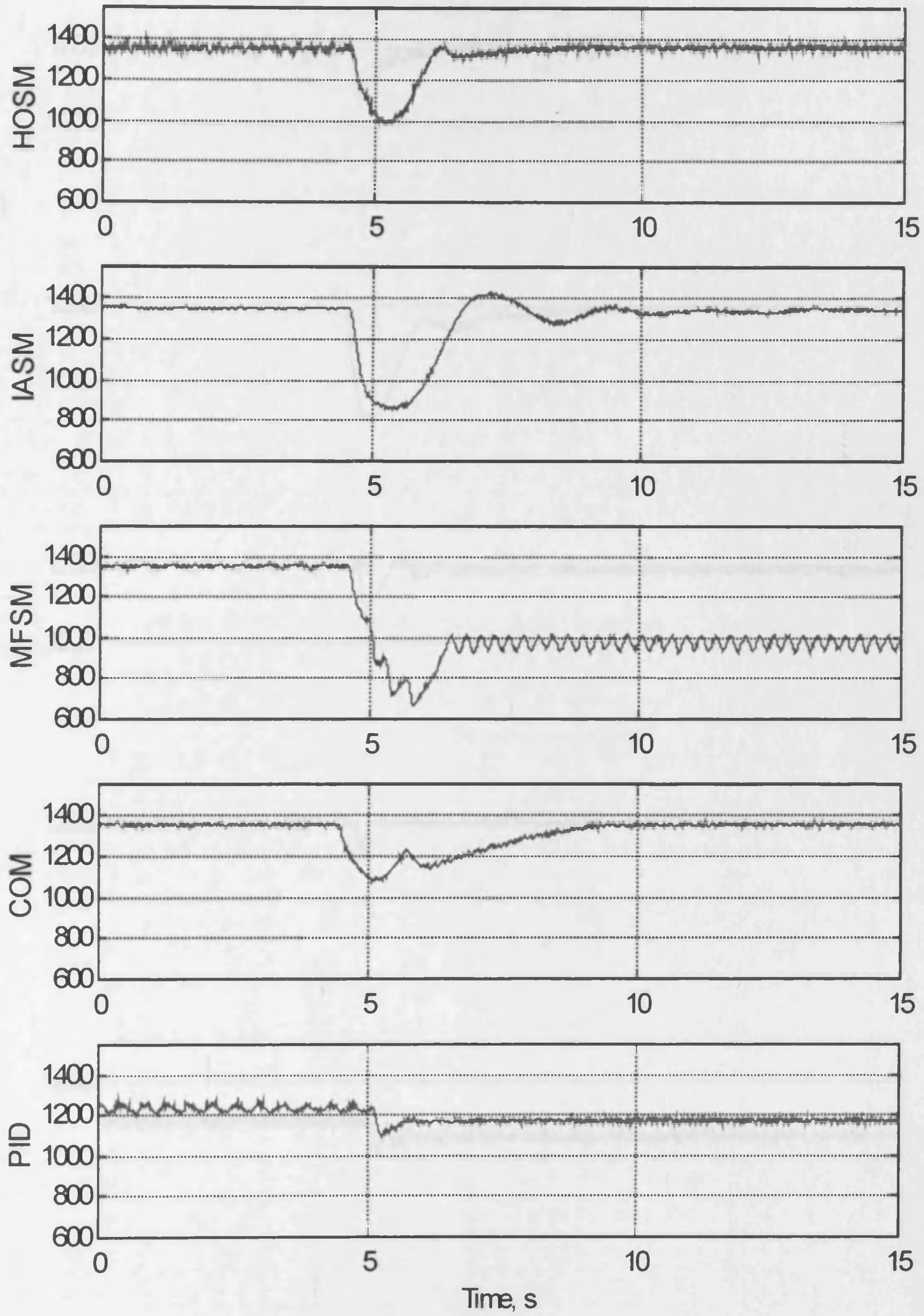


Figure 8-5: Comparison of the different controllers to a step change in load of 60 kW at 1350 rpm.

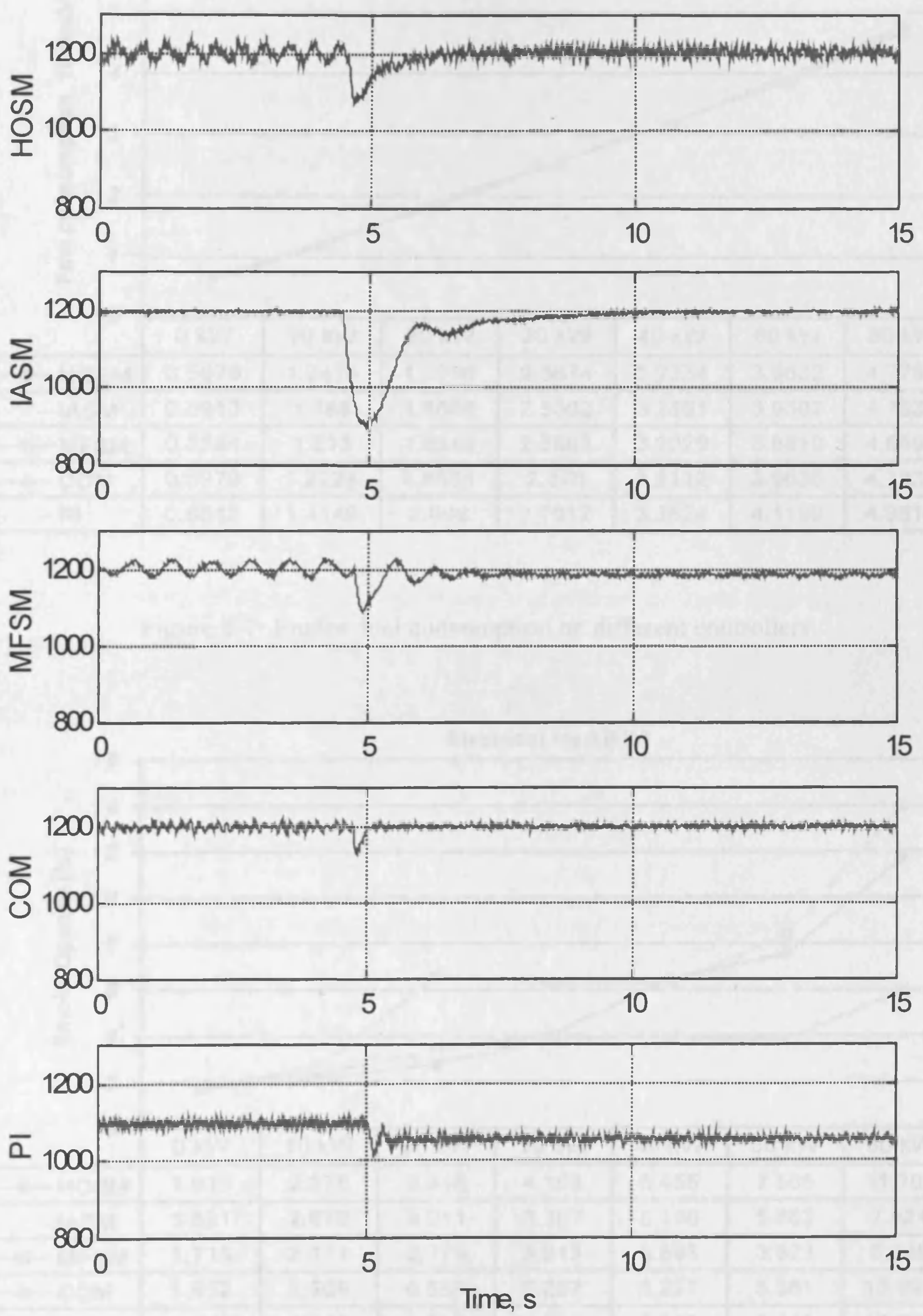
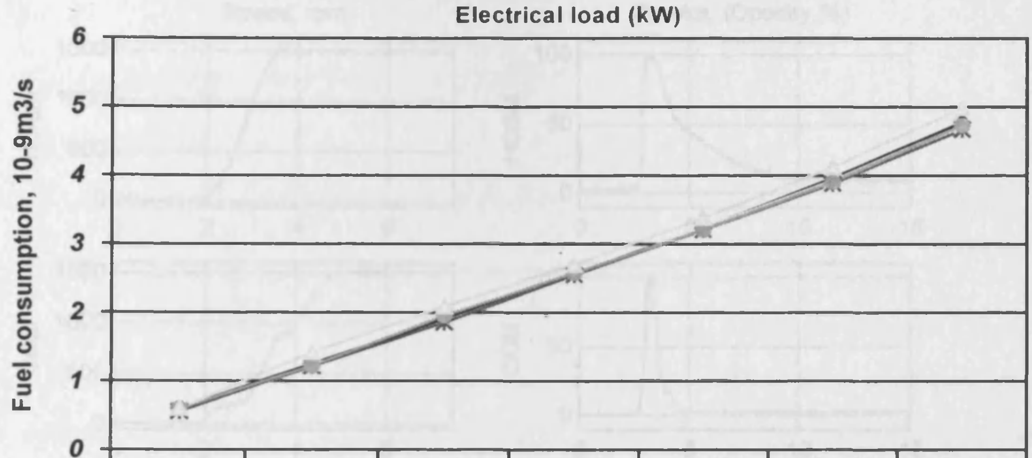
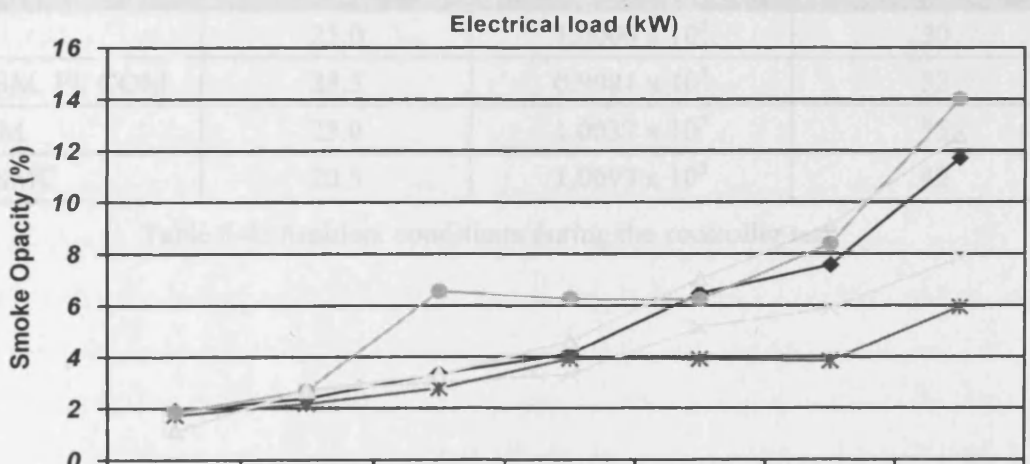


Figure 8-6: Comparison of the different controllers to a step change in load of 60 kW at 1200 rpm.



	0 kW	10 kW	20 kW	30 kW	40 kW	50 kW	60 kW
HOSM	0.5979	1.2474	1.9056	2.5674	3.2334	3.9632	4.7794
IASM	0.5913	1.188	1.8668	2.5502	3.2361	3.9307	4.7337
MFSM	0.5584	1.213	1.8549	2.5502	3.2029	3.8819	4.6691
COM	0.5979	1.2224	1.9561	2.576	3.2112	3.9036	4.7337
PI	0.6012	1.4149	2.092	2.7012	3.3824	4.1199	4.9814

Figure 8-7: Engine fuel consumption on different controllers.



	0 kW	10 kW	20 kW	30 kW	40 kW	50 kW	60 kW
HOSM	1.913	2.375	3.348	4.108	6.455	7.565	11.701
IASM	1.821	2.678	3.011	3.307	5.136	5.883	7.824
MFSM	1.718	2.171	2.779	3.913	3.898	3.823	5.936
COM	1.852	2.669	6.558	6.267	6.227	8.381	13.991
PI	1.118	2.706	3.316	4.591	7.012	9.119	12.569

Figure 8-8: Opacity (%) reading for different controllers at different loading conditions.

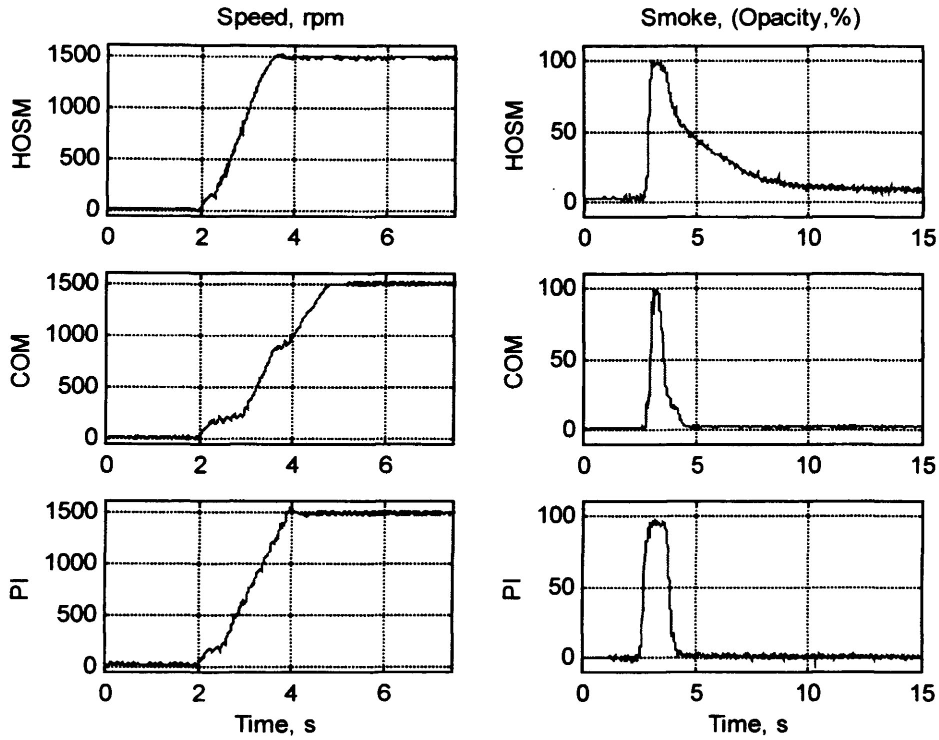


Figure 8-9: Engine start up: Speed and exhaust emission response.

Controllers	Air temperature °C	Atmospheric pressure Pa	Relative humidity %
BSI	25.0	1.0000×10^5	30
HOSM, PI, COM	23.5	0.9984×10^5	53
IASM	23.0	1.0037×10^5	55
MFSMC	20.5	1.0093×10^5	42

Table 8-4: Ambient conditions during the controller test.

Chapter 9

9 Conclusion and Future Work

9.1 Conclusion

The research undertaken within this thesis has led to various conclusions.

9.1.1 Sliding Modes (Variable Structure Systems)

Firstly, the variable structure system will be mentioned. The VSS presented in Chapter 3 concluded that sliding modes are a possible solution for both fault diagnosis and control systems because of their advantageous properties. The basic concepts of VSS have been presented accompanied by model simulation and results. The simulation results of sliding modes have been shown to be robust against uncertainties. The order of the system is reduced when the sliding mode occurs. The inherent chattering in sliding mode control has been discussed and possible solutions to this problem have been presented with simulation results. The results have shown that the solutions are capable of removing the high frequency component in the control signal.

9.1.2 Robust Fault Diagnostic System

The proposed fault diagnostic scheme using a sliding mode observer has been designed and implemented. The scheme was tested using real engine data under normal and fault conditions. The final diagnostic model has been implemented and a low coolant flow condition was simulated. The diagnostic technique has shown a good correlation to the particular fault situation. The overall diagnostic results have shown that the sliding mode technique is robust in re-constructing system parameters, as well as the system states. One particular issue of concern in the design is that a reasonably detailed understanding of the particular model under consideration is essential. Overall, the developed diagnostic system has been shown to be realistically applicable as an on-board fault diagnostic system and it is capable of continuously monitoring the engine for its day-to-day operating conditions. Most importantly, the diagnostic scheme is cost effective as only low cost sensors were involved in the implementation.

9.1.3 Dynamic Identification

A closed-loop event-based system identification procedure has been performed and a valid engine model has been developed for use in sliding mode controller design. The event-based technique requires additional hardware design and software programming to generate event-based data. Since it is a 'black-box' type of identification, the process is iterative. Several lengths and amplitudes of PRBS signal have been injected to the engine. Technical knowledge of the Matlab Identification Toolbox is required. A state-space model has been generated and has been validated on several tests; such as data correlation, comparing the MSFV value and comparing different model structures. The identified model is a second order state space model.

9.1.4 Classical Sliding Mode Control

Two different control algorithms (IASM and MFSM) have been proposed under the classical sliding mode control. Both of the algorithms require system states to be available and hence, a sliding mode observer has been introduced. During the implementation, both of the controller algorithms have been written as *s*-functions which provide faster simulation compared to the conventional Matlab Simulink block design. The observer speed has to be fast and the simulation step time for the observer has been set to be faster by a factor of 10 compared to the controller. The practical implementation also studied the limitations of Matlab Simulink during real time application as some of the Simulink blocks that deal with absolute time cannot be employed in the design. During practical tests, the IASM and MFSM controllers have shown great tracking capability and good load disturbance rejections properties. The modification (with integral action) to the original MFSM algorithm has eliminated the steady state error. However, for large loads, the performance of the MFSM remains an issue. A new method of sliding mode controller tuning has been proposed and the tuning algorithm has been established. This method has been shown to be straightforward to tune the controller and with a basic understanding of the controller concept, a non-expert operator can conveniently tune the sliding mode controller.

9.1.5 HOSM Control

HOSM control design is a model-free sliding mode approach. A particular super-twisting algorithm has been described as well as the control algorithm. The modification of the control algorithm (giving the asymptotic decay element) has shown to improve the controller performance. The controller has shown robustness across a wide operating envelope and it is capable of handling large load disturbance. The established tuning procedure is simple and straightforward. This controller has also shown the capability of starting the engine. Without the need of a system model, the design is simpler, uses less computer resource and is easy to implement. Hence, the objectives 4, 5 and 6 set out in Chapter 1 were met.

9.1.6 Control Community and Industry

The overall results have shown the potential of robust sliding mode techniques in various automotive applications. The on-board fault diagnostic system is economically applicable and the robustness of the diagnostic system has been shown. The last objective specified in Chapter 1 was to bridge the gap between the control community and industry by raising the acceptance of robust sliding mode control strategies by providing industry with simple designs and workable implementation procedures. This has been accomplished. The controller tuning-algorithms established in this work are straightforward and can allow a non-expert to maintain and tune the controller conveniently. Good performance can be achieved.

9.2 Future Work

Despite the fact that this research work has come to an end, other research possibilities have been identified.

9.2.1 Fault Diagnosis

Chapter 4 has presented a diagnostic technique and results. The immediate results show that estimates of particular sub-component parameters can be obtained without knowledge of the danger level to the system. Fault classification may be proposed to categorise each of the sub-component parameters into a sub-level. Each of these sub-levels may represent some system conditions such as normal condition, moderate fault, danger, high danger and effect system shutdown. One immediate and possible solution to this is to employ a fuzzy-based classification system. The nature of the fuzzy logic technique means that it can classify the observed parameters into categories and that overlapping of each category is allowed. This feature provides information on how the parameter can progress from a normal condition to a dangerous situation and is particularly useful for the engine operator. The fuzzy-based design technique is transparent to the user as it provides reasoning on every signal parameter, as well as fault conditions.

The diagnosis techniques which have been proposed only target the sub-components of a particular diesel engine cooling system, such as the thermostat valve, fuel pump and radiator. There is potential to monitor a non-sub-component parameter such as the coolant composition. This fault can be simulated by changing the amount of anti-freeze in the coolant mixture. The normal composition for the engine coolant is 50 % water and 50 % ethyline glycol (anti-freeze solution). The change in mixture composition will change the heat transfer coefficient of the mixture and therefore, change the heat dissipation rate.

By reducing the anti-freeze, it increases the heat transfer coefficient and hence the engine runs in a cooler condition. This undesired condition might affect engine emissions and performance, as the engine is not running at its optimum condition. To simulate this fault, two different compositions with 10% and 25% of anti-freeze may be proposed for acquiring fault data. The following equation is proposed.

$$\dot{T}_B = \frac{1}{(mc)_B}\left[(hA)_{cyl\_to\_block}(T_{gas} - T_B) - (\dot{m}c)_2(T_2 - T_1) - (hA)(T_s - T_{amb}) - \varepsilon\sigma A(T_s^4 - T_{amb}^4)\right]$$

The details of the equation can be found in [Twiddle 2001]. The engine block heat coefficient is of interest here. A similar diagnostic technique, the non-linear sliding mode observer scheme, may be employed to reconstruct the parameter. It has been mentioned in the Conclusion section that there is a limitation in the setup where assumptions have to be made on the parameters which can not be realistically measured.

The robust diagnostic scheme may be applied into different areas and different type of fault environments. It has been concluded that a model of the system under consideration is essential for the realisation of the sliding mode observer diagnostic scheme. One of the possible research areas may be the engine combustion system which has a strong influence on the emissions. No particular research has been proposed in this field but it could well prove a fruitful field.

9.2.2 Control System

The proposed classical sliding mode control scheme was presented in Chapter 6. A particular hyperplane design using the robust pole assignment technique has been employed in the desgin. It could be proposed to venture into other techniques such as a linear quadratic approach (which was mentioned in Chapter 3), H_∞ approach etc. This could provide a wider range of design techniques to the end user.

The particular IASM and MFSM controllers have shown their robustness against load disturbance and have been shown to be valid across a wide operating envelope. However, the speed dips remain an issue. Large load disturbances may also be an issue for the MFSM. A different design approach may be used as mentioned in the previous paragraph to help investigate this issue.

The basic ingredients for the engine combustion are air and fuel. With a proper mixture of both elements, the exhaust emissions may be reduced. Thus, the control of exhaust emissions can feasibly be done by means of controlling the air-fuel ratio. This effectively changes the combustion stochiometric condition. Some research has been published on this topic and has shown the use of the air/fuel control method but very few mention the application of sliding mode concepts on a diesel engine. Puleston *et al* [Puleston 2002] mentioned the use of sliding mode technique on air-fuel ratio and speed control. The results were simulated from an engine model. This available engine model could be used as the start of a project. The proposed further work thus gives a chance to apply the technique to the available diesel engine. The additional instruments mentioned in Chapter 1 and Appendix A have provided the information on air drawn into the engine, exhaust gas opacity and fuel flow rate. The hardware setup and the additional parameter information obtained during the tests of this thesis will facilitate the proposed work. The established fuel flow control (in Chapter 6 and 7) could contribute to this future work. This research would be environmentally friendly in reducing exhaust emissions.

9.2.3 Microprocessor/Controller Based Technique

Computing technology has advanced and many applications which were previously impossible, can now be realised with relative ease. Micro-processors have been around for more than three decades since the first processor was developed in 1971 by Fedrico Faggin [Wilson 2001]. With this available technology, it will be possible to develop a micro-controller based fault diagnostic system and control system. In most situations, the engine runs without a computer beside it. In many cases, the engine only provides temperature information to the operator. With the proposed sliding mode observer scheme built into a micro-controller, it will help the operator not only to monitor the additional engine parameters but also to better understand the engine operation. The application of microcontroller fault diagnostic techniques are advantageous and cost effective. The research may be started with a basic 16 bit PIC microcontroller kit for the micro-processor based development. This technique is also applied to the sliding mode controller design. Besides the development hardware tool kits, knowledge of assembly language and the C++ programming language are essential to the design.

9.2.4 Web-based Fault Diagnostic System and Control system

The last area which may prove fruitful is web-based fault diagnostic and control system design using sliding mode techniques. It would be possible to use the World Wide Web to perform some applications, thus providing remote control from any corner of the world. The LabView package from National Instruments provides the technology and is capable of carrying out the task.

References

Airikka P., 2004 "The PID controller: Algorithm and implementation," IEE: Computing and control engineering magazine, December/January 2003/4, pp.6-11.

Armstrong-Helouvry, B., Dupont, P. and Canudas De Wit, C., 1994, "A survey of models, analysis tools and compensation methods for the control of machines with friction," Automatica, vol.30, no.7, pp.1083-1138.

Autoshop, 1998, "Cooling System Overview, Automotive 101," http://www.autoshop-online.com/auto101/cooling.html.

Balestrino, 1984, "Nonlinear adaptive model-following control," *Automatica*, vol.20, no.5, pp.559-568.

Banisoleiman, K., 1990, "The interaction of diesel engine turbo-charging and tuned inlet manifold systems under steady speed and transient operation," Turbocharging and turbochargers, ch.33, pp.171-187.

Bartolini, G, 1995, p.45

Bartolini, G., Coccoli, M. and Ferrara, A., 1998a, "Vibration damping and second order sliding modes in the control of a single finger of the AMADEUS gripper," *International Journal of Systems Science*, vol.29, no.5, 497-512.

Bartolini, G., Ferrara, A. and Usai, E., 1998b, "Chattering avoidance by second-order sliding mode control," *IEEE Transaction on Automatic Control*, vol.43, no.2, pp.241-246.

Bartolini, G. and Punta, E., 2000, "Chattering elimination with second-order sliding modes robust to Coulomb friction," *Journal of Dynamic Systems, Measurement, and Control*, vol.122, pp.679-686.

Bhatti, A.I., 1998, "Thesis: Advanced sliding mode controllers for industrial applications," University of Leicester, Leicester, LE1 7RH, England.

Bhatti, A.I., Spurgeon, S.K., Dorey, R. and Edwards, C., 1999a, "Sliding mode configurations for automotive engine control," *International Journal of Adaptive Control and Signal Processing*, vol.13, pp.49-60.

Bhatti A.I., Twiddle J.A., Spurgeon S.K. and Jones N.B., 1999b, "Engine coolant system fault diagnostics with sliding mode observers and fuzzy analyser," *IASTED International Conference on Modelling, Identification and Control*, Innsbruck, Austria

Briggs, P.A.N. and Godfrey, K.R., 1966, "Pseudorandom signals for the dynamic analysis of multivariable systems," *Proceeding of IEE*, vol.113, pp.1259-1267.

Brown, D. and Carbon, G., 1977, "Survey of excitation techniques applicable to the testing of automotive structures," Society of Automotive Engineers, SAE paper: 770029.

BS5514-1, 1987, "Specification for reciprocating internal combustion engine: Performance. Part 1. Specification for standard reference conditions and declarations of power, fuel consumption and lubricating oil consumption," by British Standards Institution.

BS5514-4, 1979, "Specification for reciprocating internal combustion engine: Performance. Part 4. Speed governing," by British Standards Institution.

Butts, K.R., Sivashankar, N. and Sun, J., 1999, "Application of l1 optimal control to the engine idle speed control problem," *IEEE Trans. on Control System Tech.*, vol.7, no.2, p.258-270.

Cartona, E., Onder, C.H., & Guzzella, L., 2001, "Model-based temperature control for improved fuel economy of SI engines," *3rd IFAC Workshop*, vol.1, pp. 227-232.

Challen, B. and Baranescu, R., 1999 "Diesel engine reference book," Oxford: Butterworth-Heinemann.

Chan, Y.T., 1973, "Perfect model following with a real model," Jt Autom Control Conf of the Am Autom Control Counc, 14th, Annu, Prepr, Columbus, Jun 20-22. No.10-5, pp 287-293

Chen J. and Patton R.J., 1999, "Robust model-based fault diagnosis for dynamic systems," Kluwer Academic Publishers

Chiang E.C., Ursini V.J. and Hohnson J.H., 1982, "Development and evaluation of a diesel powered truck cooling system computer simulation program," Society of Automotive Engineer, SAE Paper: 821048

Chiang, L.H., Russell, E.L. and Braatz, R.D., 2001, "Fault detection and diagnosis in industrial systems," Springer.

Chin, Y.K. & Coats, F.E., 1986, "Engine Dynamics: Time-based versus crank-angle based," Society of Automotive Engineer, SAE Paper: 860412.

Clarke, D.W., 1984, "PID algorithms and their computer implementation," *Trans. Inst. M. C.*, 6(6), pp. 305-316.

Crestchic, 1998, "Load modules: Installation and operational manual," by Crestchic Ltd., Power plant providing and commissioning, Paget St., Burton-on-Trent, Staffordshire, England, DE14 3TQ.

Dalton, T. and Patton, R., 1998, "Model-based fault diagnosis of two-pump system," *Trans. Inst. M. C., 20*(3), pp. 115-124.

DeCarlo, R.A., Zak, S.H. and Matthews, G.P., 1988, "Variable Structure Control of Nonlinear Multivariable Systems: A Tutorial," Proceeding of the IEEE, vol.76, no.3.

De Santis, R.M., 1994, "A novel PID configuration for speed and position control," *Journal of Dynamic Systems, Measurement, and Control*, vol.116, pp. 543-549.

DSPACE, 1999, "ControlDesk Experiment Guide," dSPACE GmbH.

Dutton, K., Thompson, S. and Barraclouh, B., 1997, "The art of control engineering," Addison-Wesley.

Edwards C. and Spurgeon S.K., 1996, "Robust output tracking using a sliding-mode controller/observer scheme," *International Journal of Control*, vol.64, no.5, pp.967-983.

Edwards C. and Spurgeon S.K., 1998, "Sliding mode control: theory and applications," UK: Taylor & Francis.

Edwards C., Spurgeon S.K. and Patton R.J., 1999, "Sliding mode observer for fault detection and isolation," *Automatica*, vol. 36, pp. 541-553

Emel'yanov, S.V., S.K. Korovin and Levant, A., 1996, " High order sliding model in control systems," *Computational Mathematics and Modelling*, vol.7, no.3, pp. 294-318.

Enviro, 2001, "Flow Transducer, User Instruction Manual," by Enviro Systems Ltd., Grange-over-Sands, Cumbria, LA11 6BE, England.

Erzberger, H., 1968, "On the used of algebraic methods in the analysis and design of model following control systems," TN D-4663, NASA.

Farnell-1, 2001, "Hall effect sensor, UGN3020, Product data sheet," by Farnell Electronics.

Farnell-2, 2001, "Hall effect sensor, UGN3503N, Product data sheet," by Farnell Electronics.

Fleming, W.H., 1998, "Future directions in control theory – a mathematical perspective," Report of the panel on future directions in control theory, SIAM reports on issues in the mathematical sciences, society for industrial and applied mathematics, Philadelphia.

Flower, J.O. and Windett, G.P., 1976-a, "Dynamic measurements of a large diesel engine using p.r.b.s. techniques. Part 1: Development of theory for closed-loop sampled system," International Journal of Control, vol.24, no.3, 379-392.

Flower, J.O. and Windett, G.P., 1976-b, "Dynamic measurements of a large diesel engine using p.r.b.s. techniques. Part 2. Instrumentation, experimental techniques and results," International Journal of Control, vol.24, no.3, 393-404.

Frank, P. M., 1990, "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.

Funahashi, Y. and Nakamora, K., 1974, "Parameter estimation of discrete-time systems using short-periodic pseudo-random sequences," *International journal of control*, vol.19, pp.1101-1110.

GAC 1999a, "ADC100 series, electric actuator manual: PIB 2040," Governors America Corp., 720, Silver St., Agawam, MA 01001-2907, USA.

GAC 1999b, "ESD5500 series, speed control unit manual: PIB 1002E," Governors America Corp., 720, Silver St., Agawam, MA 01001-2907, USA.

Gertler, J. J., 1988, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control System Magazine*, December issue, pp. 3-11.

Gertler, J. J., 1998, "Fault detection and diagnosis in Engineering Systems," New York: Marcel Dekker.

Gertler, J., Costin, M., Fang X., Kowalczuk, Z., Kunwer, M. & Monajemy, R., 1995, "Model-based diagnosis for automotive engines - algorithm development and testing on a production vehicle," *IEEE Trans. on control system technology, 3*(1), pp. 61-69.

Goh, K.B., Spurgeon, S.K. and Jones, N.B., 2002, "Fault diagnostics using sliding mode techniques," *Control Engineering Practice*, Vol. 10 (2), pp. 207-217.

Governors America Corporation, 2001, "ADC 100 Series Electrical Actuator user manual,"

Guzzella, L. and Amstutz, A., 1998, "Control of Diesel Engines," *IEEE Control Systems*, pp. 53-71.

Hendricks, E., Jensen, M., Chevalier, A. and Vesterholm, T., 1994, "Problem in event based engine control," Proceeding of the American Control Conference, Baltimore, Maryland, United States, vol.2, pp.1585-1587.

Hermans F.J.J. and Zarrop M.B., 1996, "Sliding mode observer for robust sensor monitoring," *In Proceeding of the 13th IFAC World Congress*, pp. 211-216

Heywood , J.B., 1988, "Internal combustion engine fundamentals," London: McGraw-Hill.

Isermann, R, 1984, " Process fault detection based on modeling and estimation methods – a survey" *Automatica*, vol.20, no.4, pp.387-404.

Isermann, R and Freyermouth, B., 1991, "Process fault diagnosis based on process model knowledge – Part 1: Principles for fault diagnosis with parameter estimation," *Journal of Dynamic Systems, Measurement and Control*, vol.113, pp.620-633.

Isermann R. and Balle P., 1997, "Trends in the application of model-based fault detection and diagnosis in technical processes," *Control Engineering Practice*, vol. 5(5), pp. 709-719

Jimenez_Garcia, S., Magana, M.E., Benitez-Read, J.S. and Martinez-Carballido, J., 2000, "Modelling, simulation, and gain scheduling control of large radiotelescopes," Simulation Practice and Theory, vol.8, issue 3-4, pp.141-160.

Jones, N.B., Spurgeon, S.K., Pont, M.J., Twiddle, J.A., Lim, C.L., Parikh, C.R. and Goh, K.B., 2000a, "Diagnostic schemes for biomedical and engineering systems" MEDSIP 2000 Conference. No. 476, pp.1-8.

Jones N.B., Spurgeon S.K., Pont M.J., Twiddle J.A., Lim C.L., Parikh C.H. & Goh K.B., 2000b, "Aspects of diagnostic schemes for biomedical and engineering systems," Proceeding of IEE. Science, Measurement and Technology, no.147, pp.357-362

Jones, V.K., Ault, B.A., Franklin, G.F. and Powell, J.D., 1995, "Identification and air-fuel ratio control of a spark ignition engine," *IEEE Trans. on control system technology, 3*(1), pp. 14-21.

Khan, M.K., Spurgeon, S.K. and Puleston, P.F., 2001, "Robust speed control of an automotive engine using second order sliding modes," *Proceedings of the European Control Conference*, pp.974-978.

Krishnaswami, V., Luh, G.C. and Rizzoni, G., 1995, "Nonlinear parity equation based residual generation for diagnosis of automotive engines fautls," Control Engineering Practice, vol.3, no.10, pp. 1385-1392.

Landau, I.D. and Courtiol, B., 1974, " Design of multivariable adaptive model following control systems," *Automatica*, vol.10, pp.483-494.

Levant, A., 1993, "Sliding order and sliding accuracy in sliding model control," *International Journal of Control*, vol.58, no.6, pp.1247-1263.

Levant, A., 1997, "Higher order sliding: collection of design tools," *Proceeding of the European Control Conference, Brussels.*

Levant, A., 2001, "Universal SISO sliding mode controllers with finite-time convergence," *IEEE Transactions on Automatic Control*, vol.46, no.9, pp.1447-1451.

Lilly , L.C.R., 1984, "Diesel engine reference book," London: Butterworths.

Lim, C.L., 2002, "Thesis: Models of human knee movements for internal signal estimation," University of Leicester, Engineering department, Leicester, LE1 7RH, England.

Ljung, L., 1999, "System identification toolbox for use with Matlab: User's guide," The Math Works Inc.

Ljung, L, 1987, "System identification: theory for the user," London: Prentice-Hall.

Ljung, L. and Glad, T., 1994, "Modeling of dynamic systems," Englewood Cliffs, New Jersey: Prentice Hall.

Ljung, L. and Mchelvey, T., 1996, "Subspace identification from closed-loop data," *Signal processing*, vol.52, no.2, pp. 209-215.

Luenberger, D.G., 1971, "An introduction to observer," IEEE Transcations, vol.A.C.16, 596-602.

MathWorks-1, 1999, "Control system toolbox for use with Matlab: User's guide," by The MathWorks Inc., 24, Prime Park Way, Natick, MA 01760-1500.

MathWorks-2, 1999, "Real-time workshop for use with Simulink: User's guide," by The MathWorks Inc.

MathWorks-3, 1999, "Signal processing toolbox for use with Simulink: User's guide," by The MathWorks Inc.

MathWorks-4, 1998, " Matlab Function Reference (Volume 1: Language)," by The MathWorks Inc.

Miao, W. and Francesco, C., 2002, "Design and implementation of a gain scheduling controller for a water level control system," ISA Transactions, vol.41, Issue 2, pp.323-331.

Miguel, L.J., Mediavilla, M. and Peran, J.R., 1996, "Fault diagnosis system based on sensitivity analysis and fuzzy logic," *26th International Symposium on Multiple-valued Logic*, Santiago De Compostela, Spain. pp.50-55.

Moon, J.F., 1974, "Rudolf Diesel and the diesel engine," London: Priory Press.

Nasa, 1986, "Space shuttle: Challenger," http://www.ksc.nasa.gov/shuttle/missions/51-l/mission-51-l.html

Natke, H.G., 1988, "Application of system identification in engineering," New York: Springer-Verlag Wien.

Ogata, K., 1997, "Modern Control Engineering," 3rd Edition, Upper Saddle River, N.J.; London: Prentice Hall.

Parikh, C.R., 2001, "Thesis: The design of classifiers for condition monitoring and fault diagnosis applications," University of Leicester, Engineering department, Leicester, LE1 7RH, England.

Pasadyn, A.J., Qin, S.J. and Valle-Cervantes, S., 1999, "Closed-loop and open-loop identitfication of an industrial wastewater reactor," Proceedings of the American Control Conference, vol.6, pp.3965-3969.

Patton, R.J., Frank, P.M. and Clark, R.N., 1989, "Fault diagnosis in dynamic systems, theory and application," Control Engineering Series. Prentice Hall, London.

Perkins, 1992, "Perkins' Electropaks 1004-4TG test manual," Perkins International Ltd., Perterborough PE1 5NA.

Perkins, 1994, "Electropaks 1004TG Technical Data," Perkins International Ltd., Peterborough PE1 5NA

Perkins, 1995, "Workshop manual phaser and 1000 series. Publication No. TPD 1312E issue 1," Technical publications, Perkins International Ltd., Peterborough PE1 5NA 1-04-1995.

Perkins, 1998, "User's hankbook: 1000 Series, Models AA to AH and YA to YD: TPD 126EIS," Perkins International Ltd., Peterborough PE1 5NA

Popken, D. A., 1999, "Application of system identification techniques to simulation model abstraction, " Conference on enabling technology for simulation science III, Orlando, Florida, United States, vol. 3696, ch.33, pp.76-87.

Proakis, J.G. and Manolakis, D.G., 1996, "Digital signal processing: Principles, algorithms and application," 3rd edition, Upper Saddle River, N.J.: Prentice Hall.

Puleston, P.F., Monsees, G. and Spurgeon, S.K., 2002, "Air-fuel ratio and speed control for low emission vehicles based on sliding mode techniques," *Journal of Systems Control Engineer*, vol. 216(2), pp. 117 – 124.

Schoukens, J., Pintelon, R., Ouderaa, E. and Renneboog, J., 1988, "Survey of excitation signals for FFT based signal analyzers," IEEE Transactions on Instrumentation and Measurement, vol. 37, no. 3, pp. 342-352.

Schwarzenbach, J. and Gill, K.F., 1992, " System modelling and control," 3rd edition, London: Edward Arnold.

Slotine, J.J.E., 1984, "Sliding controller design for nonlinear system," *International Journal of Control*, vol.40, no.2.

Soderstrom, T. and Stoica, P., 1989, "System identification," London: Prentice Hall.

Soderstrom, T., 1981, "Identification of stochastic linear system in presence of input noise," *Automatica*, vol.17, no.5, 713-725.

Soderstrom, T., Gustavsson, I. and Ljung, L., 1975, "Identifiability condition for linear systems operating in closed loop," *International Journal of Control*, vol.21, no.2, pp.243-255.

Sokolov, V.F., 1996, "Closed-loop identification for the best asymptotic performance of adaptive robust control," *Automatica*, vol.32, no.8, pp.1163-1176.

Spurgeon, S.K., Goh, K.B. and Jones, N.B., 2002, "An application of higher order sliding modes to the control of a diesel generator set (Genset)," *7th International Workshop on Variable Structure Systems*, VSS2002, Sarajevo, July 17-19, 2002.

Spurgeon, S.K., Yew, M.K., Zinober, A.S.I. and Patton, R.J. (1990): 'Model-following control of time-varying and nonlinear avionics systems' In 'Deterministic control of uncertain systems', Chapter 5, Edited by A.S.I. ZINOBER. (Peter Peregrinus Ltd).

Takahashi, T., Ueno, T. and Yamamota, A., 1985, "A simple engine model for idle speed control, " *SAE Transactions*.

Tao, J., 1995, "Constant speed control of a motor driven mechanism system," Mech. Mach. Theory, vol.30, no.5, pp.737-748.

Telonic, 2001, "Celesco Model 107 In-line Opacity Meter: User Instruction Manual," by Telonic Berkeley Inc., P.O. Box 277, Laguna Beach, CA 92652, United States.

Thornhil, M. and Thompson, S, 1999, "Adaptive fuzzy logic control of engine idle speed," Proceeding of Instn. Mech. Engrs., vol.213, part I, pp.145-155.

Twiddle, J.A., 2001, "Thesis: Fuzzy systems in real-time condition monitoring and fault diagnosis, with a diesel engine case study," University of Leicester, Engineering department, Leicester, LE1 7RH, England.

Tzafestas, S.G., 1977, "Multivariable control system identification using pseudo-random test inputs," *Automatic control theory and application*, vol.5, no.3, pp.58-65.

Unique, 2001, "Engine Cooling System are and Maintenance," http://www.4unique.com/cooling/cooling_tutorial.htm

Utkin, V.I., 1978, "Sliding modes and their application in variable structure systems," MIR publishers, Moscow.

Utkin, V.I., 1977, "Variable Structure Systems with Sliding Mode," IEEE Transaction on Automatic Control, vol. AC22, no.2, pp.212-222

Utkin, V.I., 1992, "Sliding modes in control and optimization," London: Springer-Verlag.

van den Hof, P.M.J. and Schrama, R.J.P., 1995, "Identification and control – closed-loop issues," Automatica, vol.31, no.12, pp.1751-1770.

Vantine, K.J., Christen, U., Glover, K. and Collings, N., 2001, "Analysis of an event-based diesel engine model for control purposes", *IFAC, pp.361-366.*

Verhaegen, M., 1993, " Application of a subspace model identification technique to identify LTI systems operating in closed-loop," *Automatica*, vol.29, no.4, pp.1027-1040.

Verbruggen, H.B., 1975, "Refreshing articles: Pseudo random binary sequences," *Journal A*, *16*(4), 205-207.

Walcott, B.L. and Zak, S.H., 1986, "Observation of dynamical systems in the presence of bounded nonlinearities/uncertainties," Proceeding of 25th Conference on Decision and Control, pp.961-966.

Wilson, R., 2001, "Inventor recalls birth of the MPU, Eletronic Engineering Times," 12 Dec 2001. http://www.eetimes.com/story/OEG20011212S0047.

Yang H and Saif M, 1995, "Nonlinear Adaptive Observer Design for Fault Detection," *In Proceedings of the American Control Conference*, pp. 1136-1139.

Yang H and Saif M, 1995, "Fault Detection in a Class of Nonlinear System via Adaptive Sliding Observer." *In Proceedings of the IEEE transaction on Systems Man and Cybernetics*, vol. 3, pp. 2199-2204

Young, K.D., 1977, "Asymptotic stability of model reference systems with variable structure control," *IEEE Transactions on Automatic Control*, April 1977, technical note and correspondence, pp.279-281.

Young, K.D. 1978, "Design of variable structure model-following control system," IEEE Transaction on Automatic Control, vol.AC23, no.6, pp.1079-1085.

Zadeh, L.A. and Desoer, C.A., 1963, "Linear system theory – the state space approach," McGraw-Hill.

Ziegler, J.G. & Nichols, N.B., 1942, "Optimum setttings for automatic controller," *Trans. ASME 64*, pp. 759-768.

Zinober, A.S.I., 1990, "Deterministic control of uncertain systems," London: Peter Peregrinus Ltd.

Zinober, A.S.I., 1994, "Variable Structure and Lyapunov Control," London: Springer-Verlag.

Appendices

This Appendix provides extra relevant information to this thesis. Information includes data acquisition issues, hardware calibration, hardware design and modification etc. The organisation of the Appendices is as follows: Appendix A gives greater details of the diesel engine under consideration. The instruments involved in the tests are described in detail. Several issues regarding the data acquisition for the diesel engine tests are discussed in Appendix B. Appendix C describes the calibration processes for the electronic fuel meter, opacity meter and air flow meter. The calibration results are presented and the relationship between the respective parameters is written in polynomial format. Appendix D describes the specification and implementation of the GAC control system. Particular hardware requirements to run the controller are described, as well as the initial GAC controller test results. Lastly, Appendix E presents the technique used for event-based identification and gives details of the hardware design.

A Diesel Engine and Test Instrumentation

A.1 Introduction

In Chapter 2, an overview of the research test diesel engine and test instruments were given. The hardware setup for both the CMFD system and the control system design for the engine was shown. The appendix here provides further details of the engine and instrumentation.

A.2 Diesel Engine Test-bed

The test engine is a Perkins 1000 Series, four litre, four-cylinder, turbo charged diesel engine. It drives an alternator at a speed of 1500 rpm. The alternator generates maximum power of 65 kW and dissipates the generated power via an electrical resistor load bank. The load value starts from as small as 1 kW up to 65 kW. The change of load facility is used to assess the controller designs as mentioned in Chapters 5, 6 and 7.

A.3 Sensors

Sensors are mentioned in many parts of this thesis. Most importantly, they measure and provide the acquired signal. In this research, a number of sensors are used. Table A-1 names particular sensors and their functionality.

Sensor type	Functionality in the research tests
Hall effect	The sensor uses a changing magnetic flux density principle. Applies to the fuel flow transducer to measure fuel flow rate.
Magnetic	Detects teeth on the flywheel of the engine. A series of pulses corresponds to the number of teeth on the flywheel to give engine speed information.
Reflective	Detects holes (aligned to TDC) on the aluminum disc (Chapter 5) to provide event-based data.
k-type thermocouples	These are small and flat and used to measure the temperature of an engine surface.
Platinum resistance thermometers	These are thin in shape and used to measure air temperature near the radiator.
Pressure transducers	These are used to measure internal component pressure.

Table A-1: Sensors or transducers and their applications.

Many of the sensors are used in the CMFD system and during the control system design test. The sensor must be robust in term of sensitivity and reliability.

A.4 Instrumentation on the Diesel Engine

Chapter 2 has described the use of dSPACE DS1103 DSP data acquisition unit and also various sensors in the test. Additional instruments that involve in the test are fuel flow meter, air flow meter, frequency-to-voltage converter, exhaust gas meter and resistive load bank. The details of sensor channel numbers on the physical test-bed connector are given in Table A-2.

Chan #	Description	Calibration
1	Engine speed	2000 rpm/V. Input 2 F-V at 126 teeth per rev
2	Coolant temperature, pump inlet	Slope 17.540 degC/V, intercept 1.163 degC
3	Coolant temperature, engine outlet under thermostat	Slope 17.562 degC/V, intercept 0.882 degC
4	Coolant temperature, radiator inlet	Slope 17.553 degC/V, intercept 1.828 degC
5	Air temperature, inlet elbow	Slope 17.534 degC/V, intercept –1.166 degC
6	Air temperature, induction manifold	Slope 17.540 degC/V, intercept –0.408 degC
7	Spare	
8	Engine oil pressure, main gallery	0-10V=0-6 bar
9	Air pressure, induction manifold	0-10V=0-2.5 bar
10	Coolant pressure, pump inlet	0-10V=-1-1.5 bar
11	Coolant pressure, pump outlet	0-10V=0-2.5 bar
12	Coolant pressure, radiator inlet	0-10V=0-2.5 bar
13	Coolant pressure, under thermostat	0-10V=0-2.5 bar
14	Exhaust gas pressure, before turbo	0-10V=0-6 bar
15	Air pressure, inlet pipe	0-10V=-1-1.5 bar
16	Exhaust gas pressure, after turbo	0-10V=0-2.5 bar
17	Spare	
18	Spare	
19	Spare	
20	Spare	
21	Surface temperature cylinder head LHS	I 0-200DegC=0-1v
22	Surface temperature cylinder block LHS	J 0-200DegC=0-1v
23	Surface temperature cylinder head RHS	K 0-200DegC=0-1v
24	Surface temperature cylinder block RHS	L 0-200DegC=0-1v
25	Oil temperature, oil rail	M 0-200DegC=0-1v
26	Exhaust temperature front	N/A
27	Exhaust temperature rear	N/A
28	Exhaust temperature before turbo	N 0-1000DegC=0-5v
29	Exhaust temperature after turbo	O 0-1000DegC=0-5v
30	Air temperature, intake to enclosure	A 0-200DegC=0-1v
31	Air temperature before radiator, top LHS	B 0-200DegC=0-1v
32	Air temperature before radiator, bottom LHS	C 0-200DegC=0-1v
33	Air temperature before radiator, top RHS	D 0-200DegC=0-1v
34	Air temperature before radiator, bottom RHS	E 0-200DegC=0-1v
35	Air temperature after radiator, top LHS	F 0-200DegC=0-1v
36	Air temperature after radiator, bottom LHS	G 0-200DegC=0-1v
37	Air temperature after radiator, top RHS	H 0-200DegC=0-1v
38	Air temperature after radiator, bottom RHS	M 0-200DegC=0-1v

Table A-2: Sensor channels and their specification.

A.5 Cost of the Instruments

The total cost of the test instruments is approximately £40,000. The total cost is broken up in the following Table A-3.

Instrument	Cost (£)
FGW P70 diesel generator set (70kVA)	7050.00
Crestchie resistive load bank with cable and controller	2350.00
Instrumentation for FGW PTD	1645.00
Experimental fuel pump (LUCAS)	1762.00
Customisation	600.00
DSPACE data acquisition system	10302.00
Matlab Toolbox (Total)	4060.00
Sensors	1450.00
Computer set	NK
Laboratory installation	NK
Equipment: Power supply unit, frequency to voltage converter, oscilloscope, electrical tools	NK
Modification work & devices	NK
Woodward STI-125 control system set	NK

Table A-3: Cost of the test instruments. (NK=not known)

A.6 Conclusion

This appendix provided details of the Perkins diesel engine described in many chapters in this thesis. It also shows how the test instruments are inter-connected within the test system. The next appendix presents details of data acquisition for the experiments in this thesis using dSPACE and signal processing issues.

B Data acquisition from the Perkins diesel engine

B.1 Introduction

Appendix A provided details of the diesel engine. This appendix deals with issues relating to data acquisition for the diesel engine cooling system and control system. It covers various issues such as data collection, sampling rate, initial signal processing, procedure for acquiring data using dSPACE and advantages of digital over analog signal processing.

B.2 Prior to Data Collection

The engine was run without a load for a considerable time so as to reach steady state operation such as to warm up the engine. At this point, changing the resistive load on the alternator coupled to the diesel engine altered the load on the engine. The increased load meant that the temperature of coolant coming out of the engine block rises and the data is acquired for approximately 20 minutes, allowing the temperature to reach a steady value.

B.3 Sampling Rate – Initial Signal Processing

During the data collection process, the data sampling rate was set to 1000 Hz and was down sampled to 10 Hz after filtering through a 6th order Butterworth filter. The filter is used to eliminate high frequency components that appear in the data signal. Possible sources may come from electrical noise.

B.4 Issues Related to Digital Signal Processing (DSP)

Digital signal processing is very important when dealing with real data. The signals that are obtained from the engine are analogue in nature. Subsequent recording of the data into a computer convert it into digital signals. Conversion from analogue to digital signals raises factors such as sampling rate, anti-aliasing effects and quantisation. Many real data contain noise and the handling of this noise must also be considered.

B.4.1 Sampling Frequency

This is the conversion of a continuous-time signal (analog) into a discrete-time signal (digital) obtained by taking "samples" of the continuous-time signal at discrete-time instants. The selection of the sampling frequency is important to avoid possible loss of signal information or aliasing effects. If the maximum frequency of the analogue signal is F_{max}, the chosen sampling frequency, F_s, is selected so that $F_s > 2F_{max}$. The sampling frequency is also known as the Nyquist-Shannon frequency.

Figure B-1 illustrates the phenomenon of different sampling rates. The choice of sampling frequency must be made with care as poor selection adversely affects the accuracy of the results. The top graph uses a suitable frequency. The sampled data points well represent the curve of the signal. The bottom-left graph suffers loss of important information in the signal as the sampling interval is being too large. Conversely in the bottom-right graph, no information would be lost but computation time is increased unnecessarily as the sampling frequency is higher than needed.

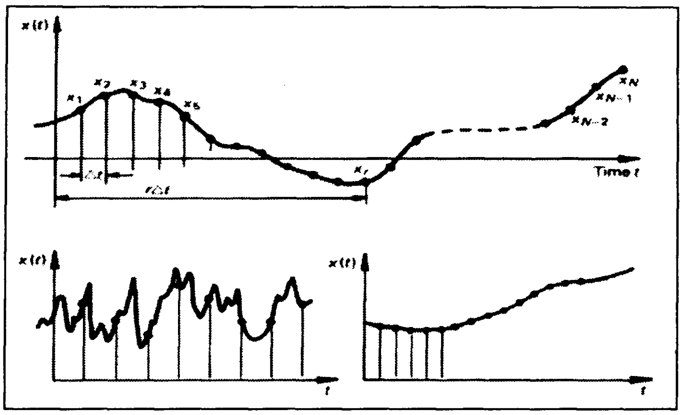


Figure B-1: Sampling of a continuous signal (top) suitable sampling rate (bottom left) low sampling rate and (bottom right) high sampling rate.

B.4.2 Quantisation

This is the conversion of a discrete-time continuous-valued signal into a discrete-time, discrete-valued (digital) signal. The value of each signal is represented by a value selected from a finite set of possible values. In the ADC process, the accuracy depends on the accuracy of an n-bit of ADC board. The 16-bit dSPACE board has 65536 (2^{16}) steps or different sets of possible values can be selected. For a slow cooling systems dynamics, where the temperature range is limited (maximum 400° C) it is sufficiently accurate to capture the true signal value with a very low quantisation error (±1/65536) of the amplitude range.

B.4.3 Digital Filter

Many signals inherently contain noise and so do the research test data. The diesel engine is running at 1500 rpm. The noise source may come from electrical noise. In the test, the data is sampled at 1000 Hz and down sampled to 10Hz after passing through a low pass Butterworth digital filter. The choice of Butterworth filter is due to its less mathematical complexity and no ripples in the filtered response [Twiddle 2001]. This filter is of 5^{th} order and allows sharp transition from the pass band to stop band at a cut off frequency of 2 Hz.

B.5 Data Acquisition Procedure using dSPACE

The procedure for acquiring data using dSPACE is summarised in the following steps. Please refer to dSPACE ControlDesk manuals [dSPACE 1999] for more detailed information.

1. Develop Simulink model (.mdl file) for capturing and processing data.
2. Using 'ADC' and 'Slave DSP I/O' block in dSPACE, integrate dSPACE with the Simulink model.
3. Build real time code to generate '.ppc' file using Real-Time Workshop.
4. Download the code onto dSPACE processor board as detailed out in the manuals.
5. Use ControlDesk to create a new experiment.
6. Import '.ppc' file that contains description of the Simulink model along with 'variables file .trc'. Outputs at different stages of the Simulink model are now visible. The variables preceded by 's:' are final outputs from the Simulink model.
7. Link these variables with GUI for 'on-line' display by creating a new Layout. Drag variables for display on Plotter on appropriate axis. To control time of acquiring data, add Capture Settings block to layout and set parameters.
8. To adjust the control gains online (in real-time), add Slider to the GUI setting.
9. From Instrumentation, Animation menu and click on Start button to start acquiring data.
10. Save data in appropriate format.

B.6 Advantages of Digital over Analog Signal Processing

There are a few advantages of using digital signal processing [Proakis 1996]. First, a digital programmable system allows flexibility in reconfiguring the digital signal processing operation by simply changing the program rather than changing the hardware configuration in an analog signal process. Accuracy in digital signal processing is a big advantage. Digital processing provides precision in the definition of filter characteristics. Tolerances in an analog circuit make it difficult to control accuracy. Digital signals are also easily stored and transportable to any remote laboratory.

B.7 Conclusion

This appendix provided details about methods used to process real data and considered a few factors such as sampling rate and various issues related to digital signal processing. A brief procedure on obtaining data using dSPACE was presented.

C Equipment Calibration

C.1 Introduction

Equipment calibration involves determining the input-output relationship of the equipment relative to a standard device that has the same unit value. Chapter 2 mentioned a number of equipments requiring under calibration. This appendix provides further details on the equipment calibration processes.

C.2 Fuel Flow meter

The Electronic Fuel Meter (provided by Perkins Engine Limited) is manufactured by Enviro Systems Limited. It is typically used for monitoring fuel consumption on combustion engines. There are two Hall-Effect sensors mounted inside an aluminum block. On the side of the aluminum block are two oval shape gears [Enviro 2001], one of them with a rod shaped magnet mounted on it.

When the gears move, the sensors on the other side will pick up pulses due to changes in magnetic flux density. These pulses are fed through an integrated circuit within the sensor itself and produce a series of pulses at the output pin. The number of pulses over a certain time period determines the rate of fuel flow into the engine. The relationship between the pulses and fuel consumption rate can be written as:

$$\text{Fuel flow rate} = f(\text{pulse frequency})$$

Equation C-1

Note: At the beginning of the test, the two original sensors [Farnell-1 2001] were found to be not functioning as no pulses were detected. They were replaced by [Farnell-2, 2001] which have the same function as the original sensor, but a different electrical specification. The latter is more sensitive and is suitable for this test, as the pulse signal is rather weak. The following steps are proposed to carry out the calibration test.

C.2.1 Parameters to Measure

Time taken ($T_{ss\#\#}$) for the diesel engine to consume a fixed volume of fuel at various load settings. When recording the time, data acquisition starts simultaneously to record down the output signals from the electronic fuel meter over the period. Ambient temperature is measured near the header fuel tank using a lab thermometer, see Figure C-1 (≈ fuel temperature).

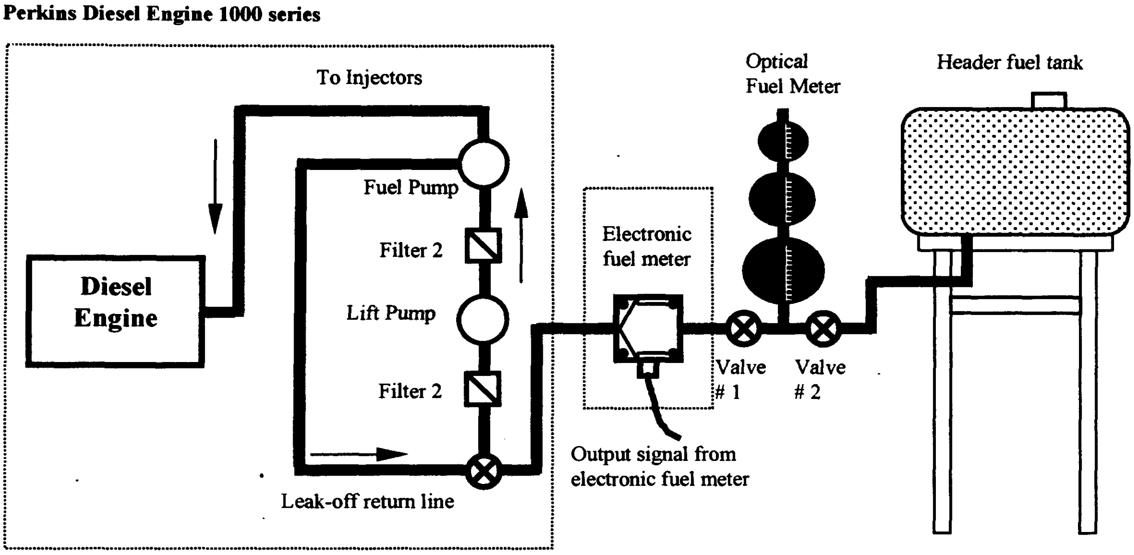


Figure C-1: A simplified schematic shows installation and setup of the optical fuel meter and electronic fuel meter together with header fuel tank to the diesel engine. (Not to scale)

C.2.2 Data acquisition

There are two output signals from the electronic meter (one from each sensor) and one is 180° out of phase with the other. These data were recorded using the similar procedure in Appendix B.5. The initial data were recorded at a sampling rate of 1kHz and then down-sampled to 100Hz before being saved to hard disk. Thus, the data was effectively captured at 100Hz. At the same time, the size of data to be saved to the local disk drive was reduced.

Figure C-1 shows the installation of test equipment, electronic and optical fuel meter, header fuel tank to the engine. The optical fuel meter was installed in series with the electronic flow meter. The fuel rate could be measured simultaneously when the data acquisition program was triggered to record the output from the electronic meter.

Two sets of data were recorded. First, the pulse reading when a step load (step change of 20 kW) was applied and secondly, pulse reading at steady state. For the optical fuel meter, time was recorded during steady state for a fixed volume of fuel consumed.

C.2.3 Results

A simple Simulink model is designed to count the number of pulses in 10 seconds. Ten seconds is chosen as the fuel rate is rather low and a longer period will give a reasonably accurate count and average out the overall result. Figure C-2 shows the rate of pulses (convert to pulse/s) plotted against applied load from the information in Table C-1. The recorded pulses are shown in Figure C-3 and Figure C-4.

Load applied (at steady state)	Number of pulses in one second
0	1.6
20	3.6
40	6.1
60	10.2

Table C-1: Number of pulses corresponding to applied load

The relationship between the pulse frequency and load is found to be,

$$\text{Pulse frequency}\,(pulse/s) = 0.0004 * load^2 + 0.0905 * load + 1.6$$

Equation C-2

It is noticed that the coefficient of the non-linearity term is small (≈0).

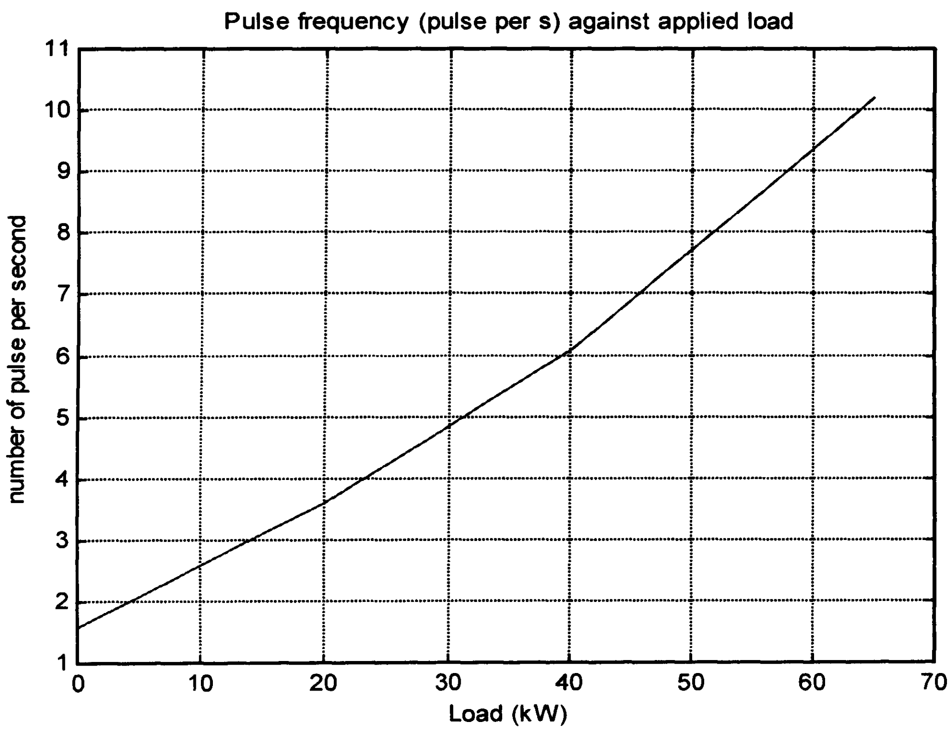


Figure C-2: Plot of number of pulses recorded in 10 seconds against applied load

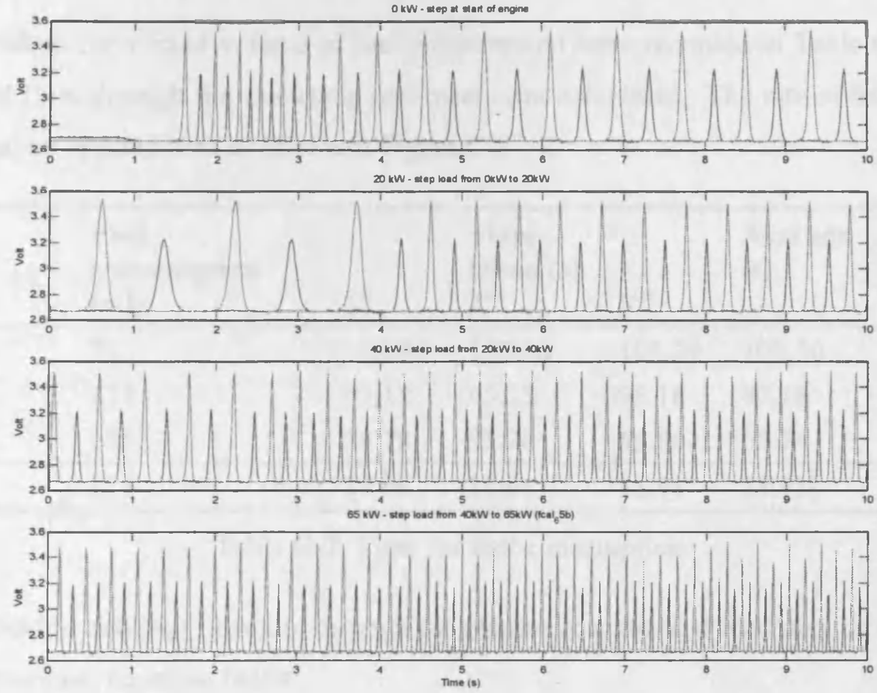


Figure C-3: Changes of frequency of pulse when step load is applied

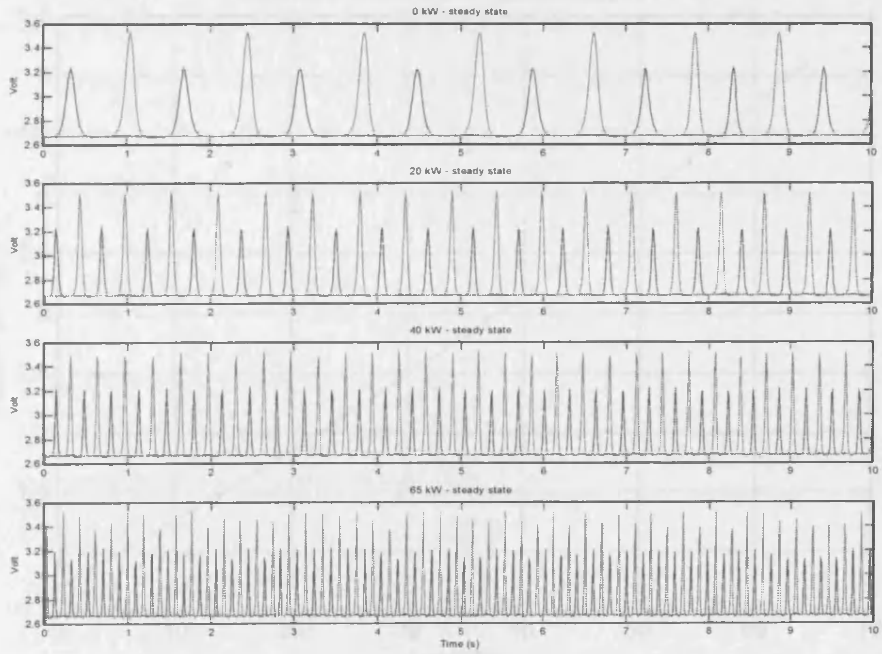


Figure C-4: Pulses at steady state

C.2.4 Electronic fuel meter

The time taken for a fixed volume of fuel consumption were recorded in Table C-2 and the rate of fuel flow through the electronic fuel meter was calculated. The rate of fuel flow was plotted against applied load as shown in Figure C-5.

Load applied (kW)	Fuel consumption (ml)	Time taken (s) 1st	2nd	3rd	Average (s)	Flow Rate (ml/s)
0	75	105.33	105.28	105.29	105.30	0.71225
20	175	95.12	95.25	95.18	95.18	1.83862
40	175	55.78	55.78	55.66	55.74	3.13958
65	175	33.96	33.91	33.91	33.93	5.15768

Table C-2: Time for fuel consumption

The relationship between flow rate through the electronic meter and load may be represented by the polynomial equation below.

$$\text{Flow rate}\,(mm^3/s) = 0.0001 * load^2 + 0.0536 * load + 0.7122$$

Equation C-3

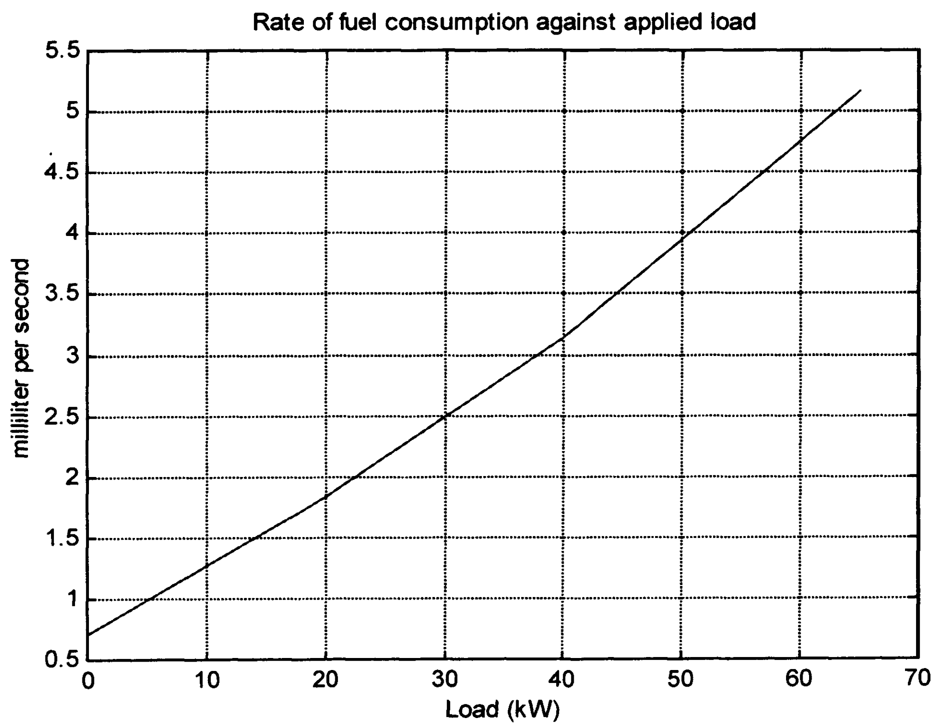


Figure C-5: Plot of number of pulses recorded over 10 seconds against applied load

With the above results, the relationship between fuel flow rate and number of pulses was represented by the following polynomial equation.

$$\text{Fuel flow rate}\,(mm^3/s) = 0.0006 * p^3 - 0.0164 * p^2 + 0.6335 * p - 0.2651$$

Equation C-4

where, p = rate of pulse (pulse/s)

C.2.5 Fuel Specification

The fuel is from Texaco Gas Oil
At ambient temperature at 24°C
Density = 0.866 kg/m^3

C.2.6 Assumption

Density of fuel does not change through out the test.

Fuel temperature (≈ ambient temperature) is constant.

C.3 Exhaust Gas Meter Calibration

Manufacturer : Telonic Berkeley Inc. (Model: Celesco Model 107)

The Opacity meter measures the relative light absorption of the smoke discharged from the diesel engine exhaust. Measurement is done by passing LED light pulses, I_o, through the engine exhaust stream and using a photoelectric detector to detect the loss in light transmission due to smoke. This reduces the intensity of the light reaches the photodiode to I. The relative light energy loss is translated into both opacity and smoke density (k) value. Definitions of opacity and smoke density [Telonic 2001] are shown below.

Light intensity reduction can be express as, $\dfrac{I}{I_o} = e^{-naQL}$

where n = number density of smoke particles

a = average particle projected area

Q = average particle extinction coefficient

L = light beam path length within the smoke

k = naQ

I_o = intensity of LED source

I = intensity of light reaching photodiode

$$\frac{I}{I_o} = e^{-kL} = \frac{T}{100} = 1 - \frac{N}{100}, \quad N = \left(1 - \frac{I}{I_o}\right)100 = Opacity(\%), \quad k = \frac{\ln\left(1 - \frac{N}{100}\right)}{L} = SmokeDensity(m^{-1})$$

Figure C-6 shows the actual equipment being employed in the calibration process and Figure C-7 show the emitter and detector inside the smoke meter device.

C.3.1 Parameter to Measure

Smoke meter reading (opacity, %)

Engine speed reading (rpm)

Ambient temperature (°C)

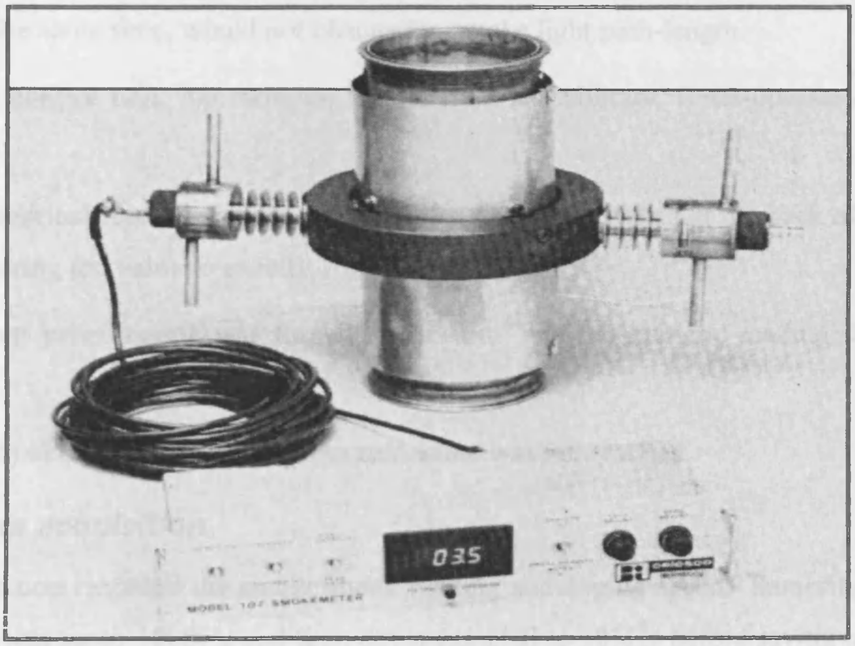


Figure C-6: Smoke meter

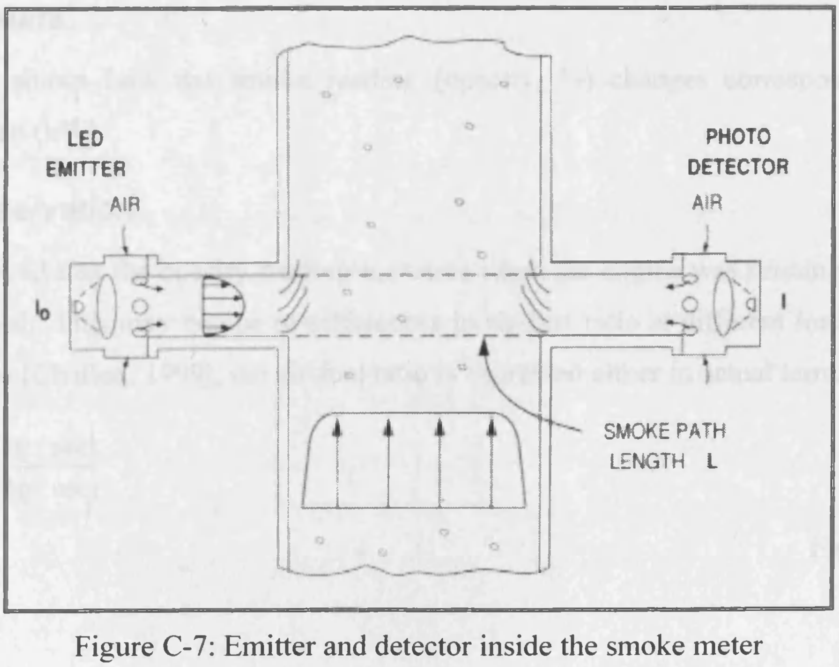


Figure C-7: Emitter and detector inside the smoke meter

C.3.2 Procedures

The Celesco Model 107 consists of two basic units:

1. The sensor unit - which was installed in-line in the engine exhaust gas flow line, contains a light source and a photo-detector module.
2. A rack-mounted, all solid-state control unit contains the operating controls, digital display, power supplies and signal processing electronics.
3. A purge air system was installed to protect the detector sensor from smoke deposition and at the same time, would not change the smoke light path-length.
4. On the control unit, the switches were turned to Calibrate, Reset-operate and opacity position.
5. The numerical reading showed around 100. Tuned the switch at the back of the control unit to bring the value to exactly 100.
6. The front panel switch was turned to 'operate' and the numeric reading showed zero value.
7. Steps (4) to (6) were repeated if the zero value was not reached.

C.3.3 Data acquisition

Data acquisition recorded the smoke meter reading and engine speed. Sampling frequency of the data was set to 1000Hz and then down-sampled to 100Hz before saving to hard disk. Data acquisition steps followed Appendix B.5.

C.3.4 Results

Figure C-8 shows how the smoke reading (opacity, %) changes corresponding to the electrical load (kW).

C.3.5 Observation

It was observed that the opacity reading increases when the engine was running with higher electrical load. This may be due to differences in air-fuel ratio at different load conditions. According to [Challen, 1999], the air-fuel ratio is expressed either in actual terms,

$$A/F = \frac{\dot{m}_a(kg/\sec)}{\dot{m}_f(kg/\sec)}$$

Equation C-5

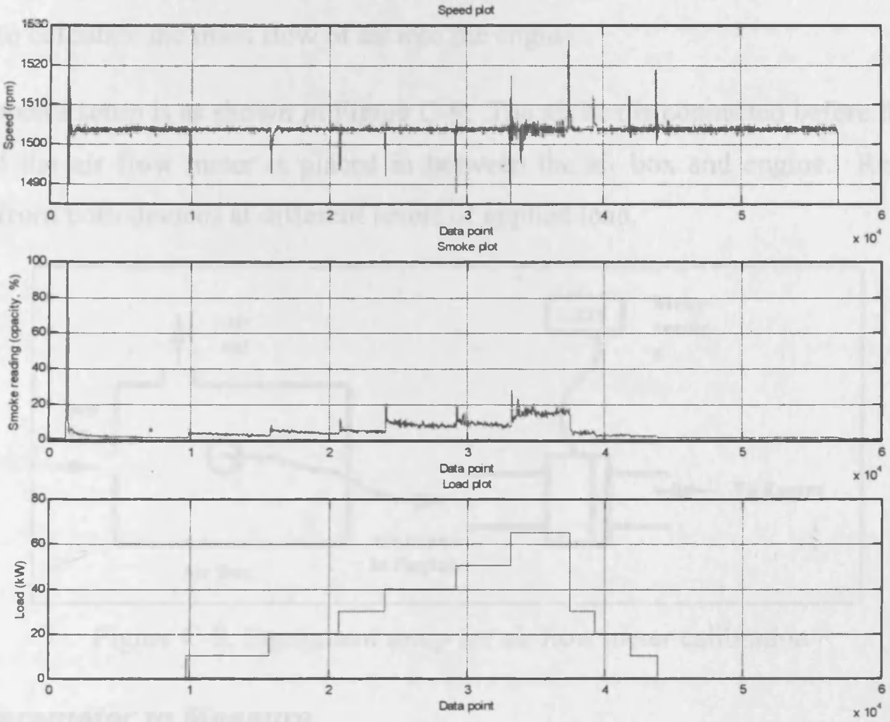


Figure C-8: Graph of speed reading, smoke reading, load profile

where $\dot{m}_a$ is rate of trapped airflow to the engine or relative to the stoichiometric air fuel ratio for the particular fuel, i.e. excess air factor, ε,

$$\varepsilon = \frac{(A/F)_{actual}}{(A/F)_{stoichiometric}}$$

Equation C-6

In practice, for most hydrocarbon fuels, $(A/F)_{stoichiometric} \approx 14.9$. The liming relative air fuel ratio for smoke-free combustion is in the range of $1.2 < \varepsilon < 1.6$.

It was also noticed that during the start of the engine, the opacity value at one time reached 100%. The fuel pump throttle was at its maximum during the start of engine and pumped maximum possible fuel to the system. The different air-fuel ratio at this point may not produce smoke-free exhaust and hence a high opacity value was observed.

C.4 Electronic Air Flow Meter Calibration

The air flow meter is used to measure mass flow rate of air in kg/s draws into engine. With this reading, air/fuel ratio can be measured. The meter gives readings in term of a voltage. In order to calibrate the air flow meter, an air box is needed to measure the actual airflow rate into the engine. The air box is based on the Bernoulli effect and also uses a continuity equation to calculate the mass flow of air into the engine.

The equipment setup is as shown in Figure C-9. The air box is connected before the air flow meter and the air flow meter is placed in between the air box and engine. Readings are recorded from both devices at different levels of applied load.

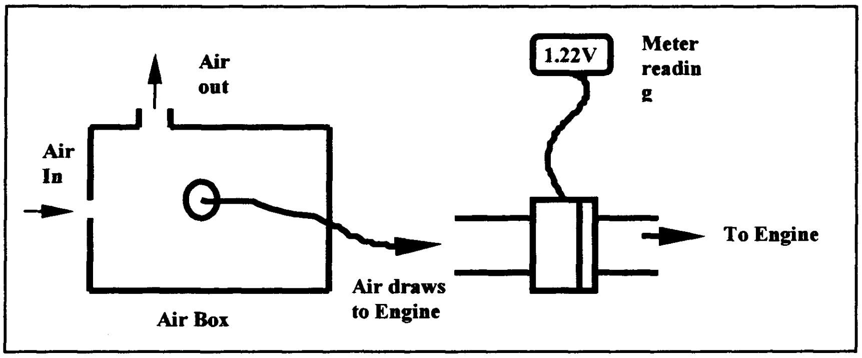


Figure C-9: Equipment setup for air flow meter calibration

C.4.1 Parameter to Measure

Air flow meter reading (voltage, V)
Engine speed reading (rpm)
manometer reading from drum (mm)
Ambient temperature (℃)
load (kW)
Air mass flow rate (kg/s) = m

DAQ recorded the above measurands at sampling frequency of 1000Hz and then down-sample to 100Hz before saving into hard drive. The procedures in appendix B.5 were followed.

C.4.2 Results

Atmospheric pressure at start, P = 759.0mmHg
Ambient temperature at start, T_{amb} = 289K
Δh at start = 0mm
Table C-3 shows the recorded reading from the test. The air flow rate is calculated from the

formulae as follows:- $m = 1.4221 * 10^{-6} * d^2 \sqrt{\frac{P\Delta h}{T_{amb}}}$

Operation	**Load (kW)**	**AFM (V)**	**Δh (mm)**	**Air flow rate, m (kg/s)**
Startup	0	-	-	
Steady	0	4.27	54	0.0553
Steady	20	4.42	67	0.0616
Steady	40	4.62	86	0.0698
Steady	50	4.72	98	0.0745
Steady	60	4.83	112	0.0797

Table C-3: Tabulation of test reading.

For air box

A graph of *m* against load is plotted. Air mass flow rate is linear to applied load. The relationship is shown in the following equation and the plot is shown in Figure C-10.

$$\text{Air flow rate}, m \ (kg/s) = 0.0003 * load + 0.0553$$

Equation C-7

For air flow meter

Figure C-11 shows the relationship between the meter reading and applied load. A polynomial equation of degree 2 is to describes the relationship. The non-linear coefficient (0.0001) is small and it is close to a linear relation.

$$\text{Air flow rate}, m \left(kg/s\right) = 0.0001 * load^2 + 0.0059 * load + 4.2698$$

Equation C-8

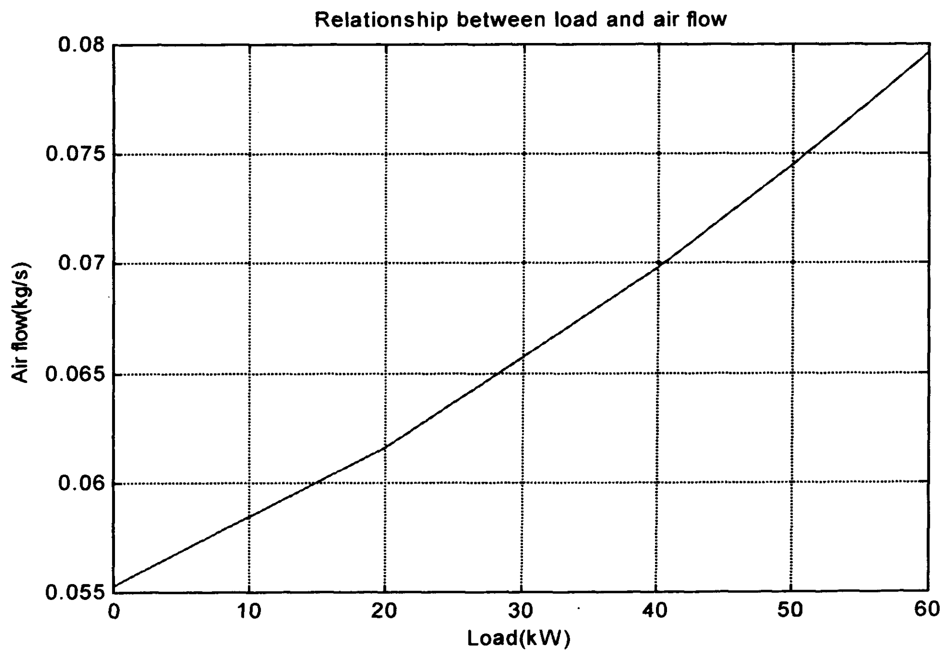


Figure C-10: Air mass flow rate (using air box) against applied load.

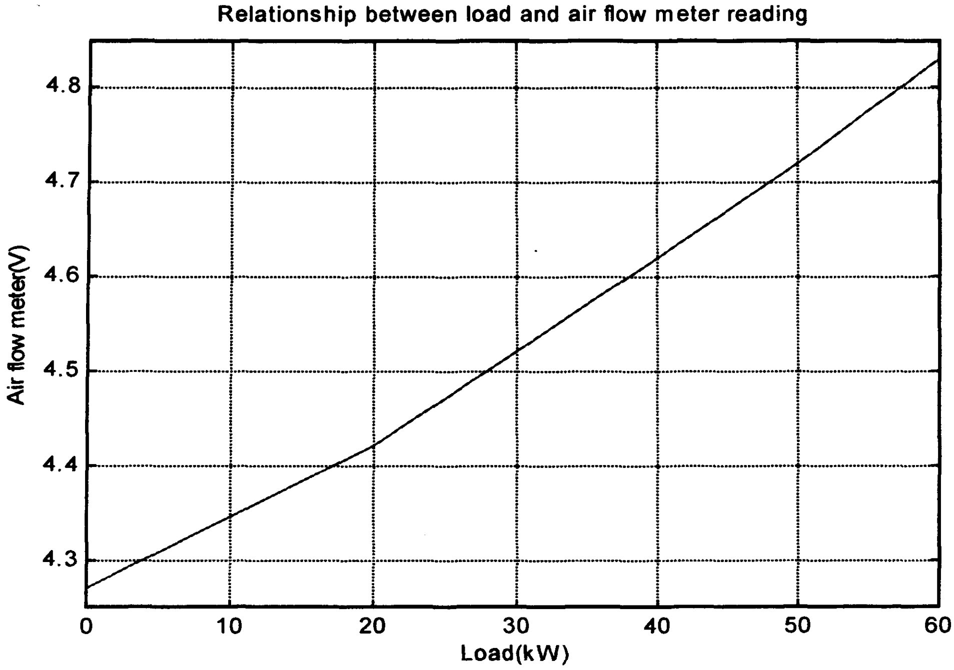


Figure C-11: Air meter reading against applied load.

Air mass flow rate

Having obtained the relationship between air mass flow, *m* against load and meter reading (volt) against load, the air mass flow in terms of voltage, V was obtained. The equation was a polynomial equation of degree 3. Figure C-12 shows this relationship.

$$\text{Air flow rate}, m\,(kg/s) = 0.019 * V^3 - 0.2523 * V^2 + 1.1571 * V - 1.7657$$

Equation C-9

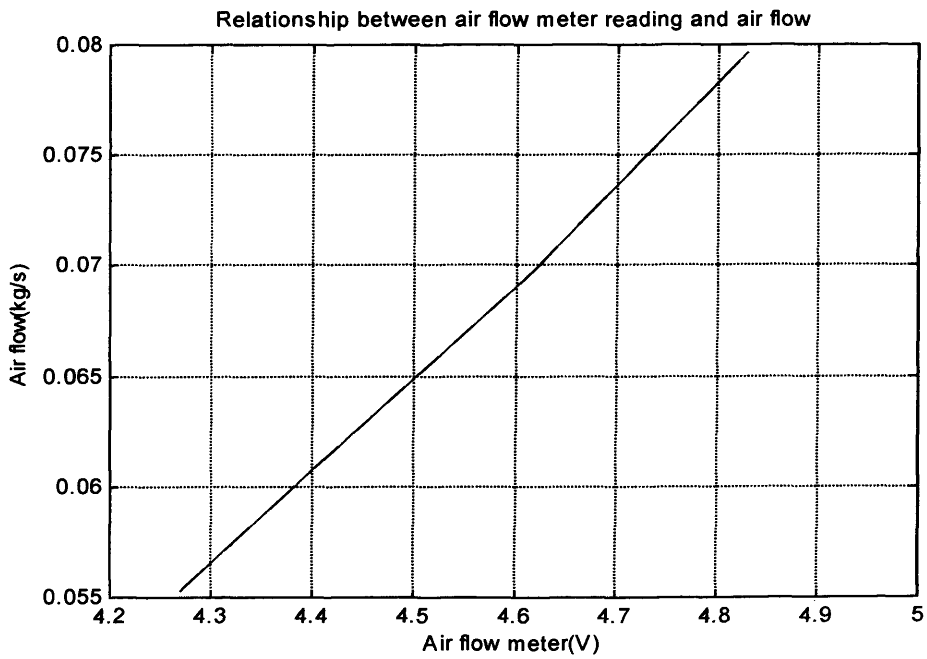


Figure C-12: Air flow, *m* against air flow meter reading, Voltage.

A 3-D graph in Figure C-13, shows the relationship between load (kW), meter reading (Voltage) and air mass flow (kg/s).

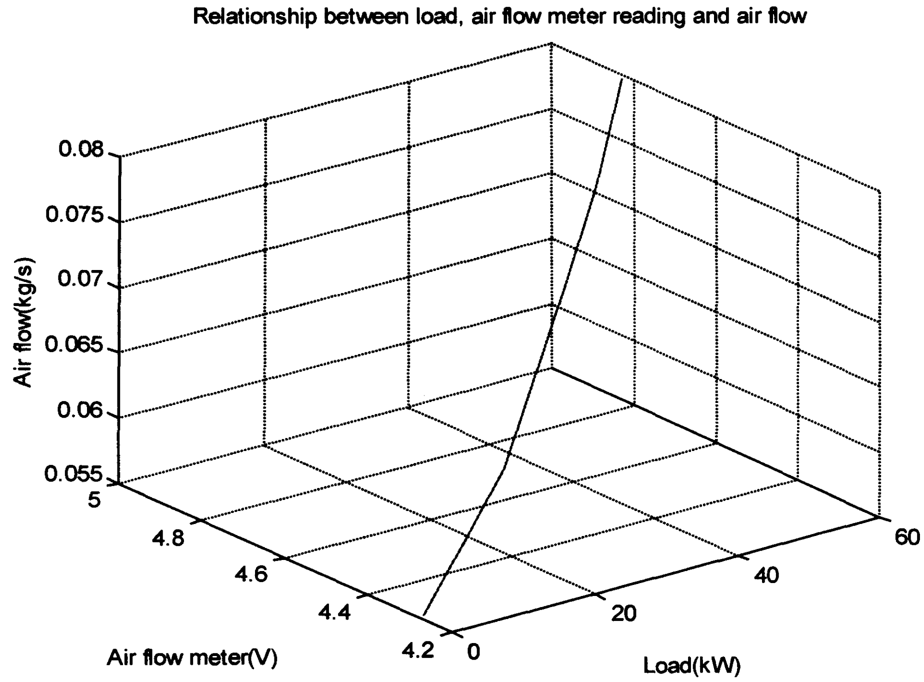


Figure C-13: Plot of air flow, air flow meter value and applied load.

C.4.3 Reference

Air box orifice equation information

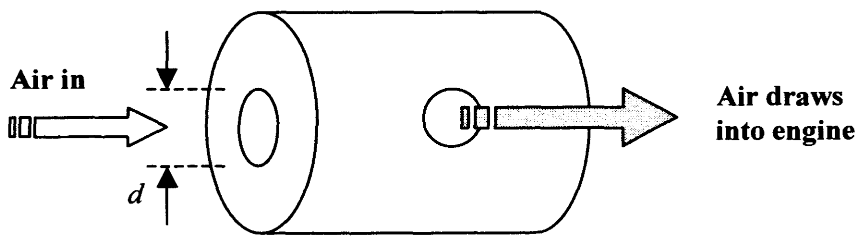


Figure C-14: Air box – direction of air flow

Figure C-14 shows the direction of air flow. On the other side of the air box, there is a 'manometer' type of device inclined at 20° and this device is used to measure the pressure difference by measuring the vertical change in water level (Δh). The formula to calculate the mass flow is as follows.

$$m = 1.4221 * 10^{-6} * d^2 \sqrt{\frac{P\Delta h}{T_{amb}}}$$

where, m = air mass flow rate (kg/s)

d = orifice diameter (mm)

P = atmosphere pressure (mmHg)

T_{amb} = ambient temperature (K)

Δh = vertical change in water level (mm)

C.5 Conclusion

This appendix provided details of the calibration process, the equipment setup and calibration results. The calibrated equipment provides monitoring parameters for the controller design in Chapter 6 and 7. These parameter values are useful to the proposed work in Chapter 9.

D Specification, setup and implementation of GAC control system

D.1 Introduction

The GAC control system is used in the research test. This appendix outlines the GAC control system and gives preliminary tests on this system. It also gives a brief benchmark of this control system and hardware setup for the identification process.

D.2 Hardware setup

The control system consists of a control unit (ESD5500) and an actuator (ADC100). The ESD5500 is a closed loop speed controller with speed pulse feedback to the controller. Output of the control unit to the actuator is a PWM signal. The ADC100 is an electromagnetic servo device. The movement of the actuator shaft depends on the change in current from the speed control unit, which effectively changes the magnetic force to the actuator. The movement of the shaft drives the throttle of the fuel pump and controls the volume of fuel going into the engine.

Figure D-1 shows the control system setup on the diesel engine. The ADC100 actuator is mounted directly to the Stanadyne fuel injection pump. The speed information from the MPU goes to the ESD5500 control unit.

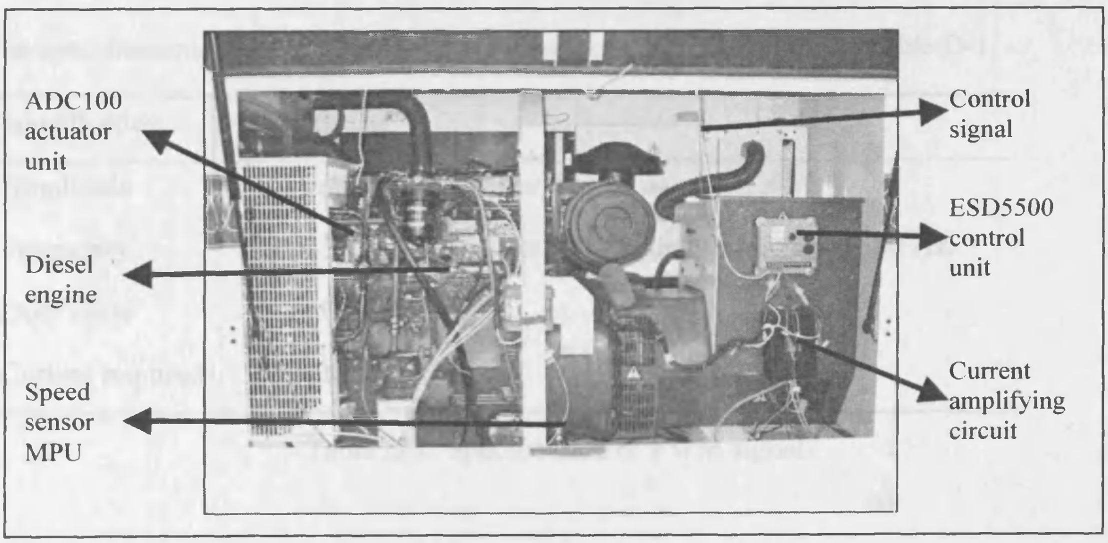


Figure D-1: Hardware setup for GAC control system.

For the identification process, it is necessary to break into the signal circuit. i.e. a designed controller signal (from dSPACE) drives the actuator directly. The control signal from the ESD5500 control unit to the actuator unit is a PWM signal. This signal is studied in order to obtain its specification for system identification and controller design implementation.

D.3 PWM Signal Specification

Pulse-width modulated (PWM) signal is a square-wave signal, which specifies the width of each pulse as a fraction of a period. The pulses start at the beginning of each period and effectively it is the duty cycle of each pulse. The duty cycle value ranges from zero to one.

The ADC100 actuator requires a PWM signal. A test was carried out to identify this signal's properties. Both the ADC100 and ESD5500 were under test. The measured parameters are frequency, amplitude and duty cycle of the PWM signal. Step loads were applied during the test. The load was gradually increased, each time with 10 kW and gradually decreased with the same step load.

Figure D-2 and Figure D-3 show the PWM signal properties. It was found that the duty cycle of the PWM signal has a value of about 0.35 at the no load condition. This value gradually increased when the load increased. At maximum load, the duty cycle is about 0.7. The frequency of the PWM signal ranges from about 320 Hz to 430 Hz and the frequency increases with increases in load. The operating frequency is 500Hz, which is mentioned in [GAC-2 1999]. The shaft rotation is proportional to the amount of actuator current (control signal) [GAC-1 1999]. Figure D-3 shows the relationship between frequency and duty cycle. It was noticed that the signal frequency is close to linear with the signal duty cycle from 0.3 to 0.6. The amplitude of the PWM signal is approximately 12 volts.

The specifications of the PWM signal are summarised in the following Table D-1.

Specification	Value
Amplitude	~12 volts (engine battery)
Frequency	From 320 Hz to 430 Hz. Operating frequency = 500 Hz.
Duty cycle	From 0.3 to 0.6
Current required	1.5 A to 2.7 A (max)

Table D-1: Specification of PWM signal.

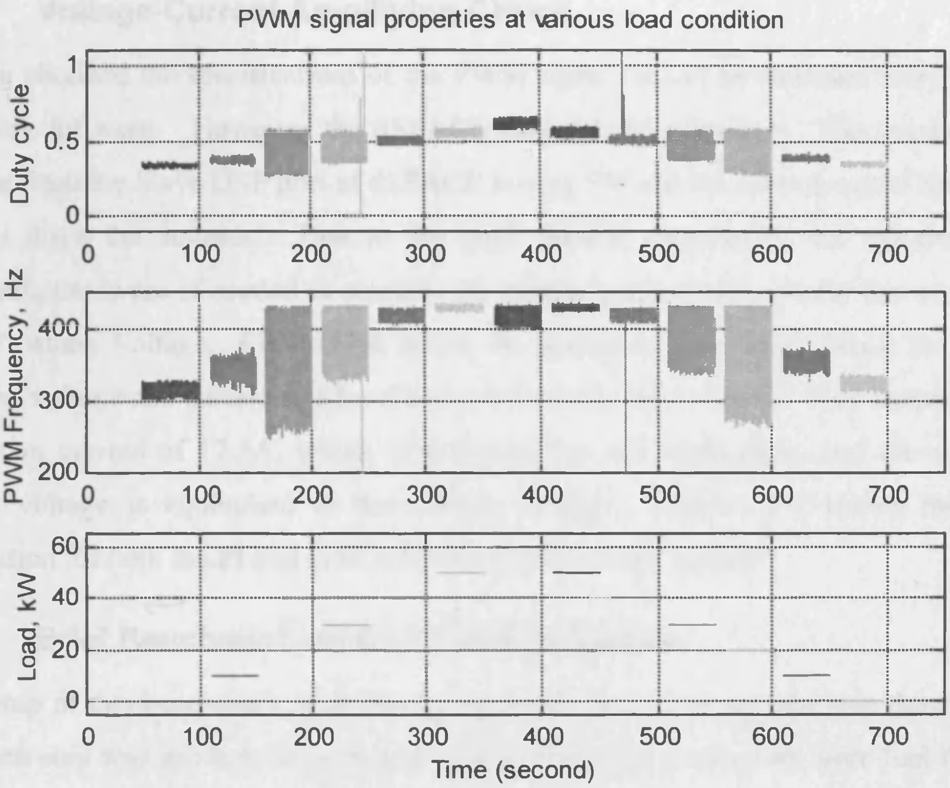


Figure D-2: Properties of PWM signal: duty cycle and frequency at different load.

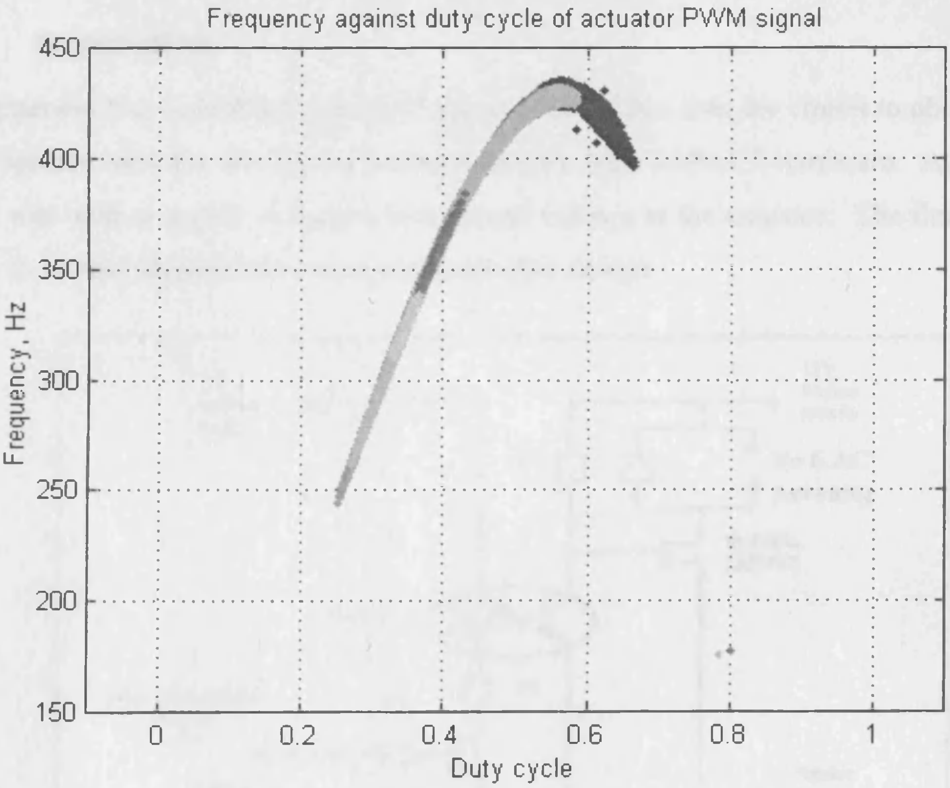


Figure D-3: Relationship between frequency and duty cycle of PWM signal.

D.4 Voltage-Current Amplifying Circuit

Having obtained the specifications of the PWM signal, it can be modeled using the Matlab Simulink software. However, the dSPACE output has limitations. The maximum output voltage from the Slave DSP port of dSPACE is only 5V and the current out of the port is too low to drive the actuator. Due to the high current demand by the actuator, separate electronic hardware is needed to amplify the supply current and operate the actuator at the engine battery voltage. Figure D-4 shows the proposed electronic circuit for amplifying both the voltage and current to a level suitable to drive the actuator. This circuit can drive a maximum current of 17.5A, which is sufficient for the application, and the amplitude of output voltage is equivalent to the battery voltage. Figure D-5 shows the hardware connection for both the PI and GAC control system, on the engine.

D.5 Brief Benchmark on GAC Control System

The setup of this benchmark is similar to Appendix D.3. Step up and step down load at 10 kW each step was applied during the test. The measured parameters were fuel flow pulses, smoke meter reading, air draw into engine (volts), load and engine speed. These parameters are shown in Figure D-6, Figure D-7 and Figure D-8 respectively.

D.6 Conclusion

This appendix has studied the new GAC system and broken into the circuit to obtain a PWM signal specification for driving the actuator directly from dSPACE hardware. An additional circuit was built to supply sufficient current and voltage to the actuator. The final appendix relates to system identification work and controller design.

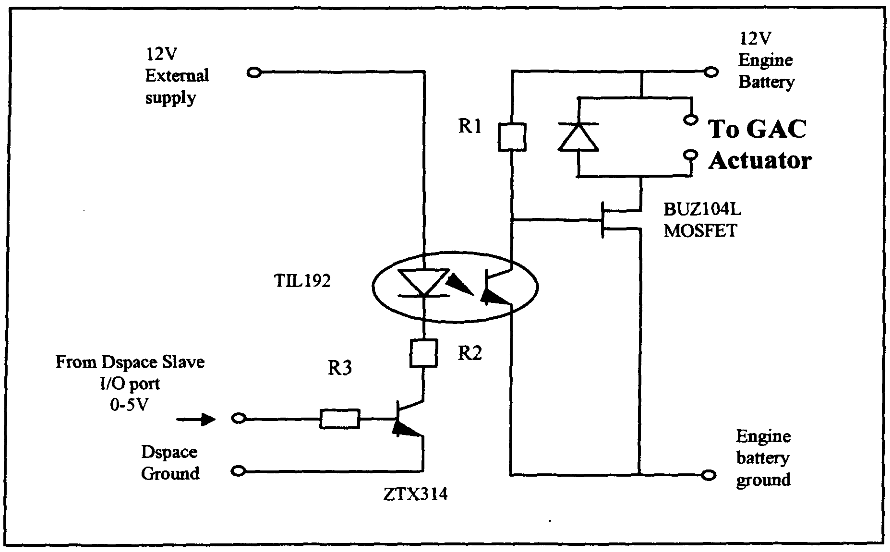


Figure D-4: Voltage-current amplifying circuit.

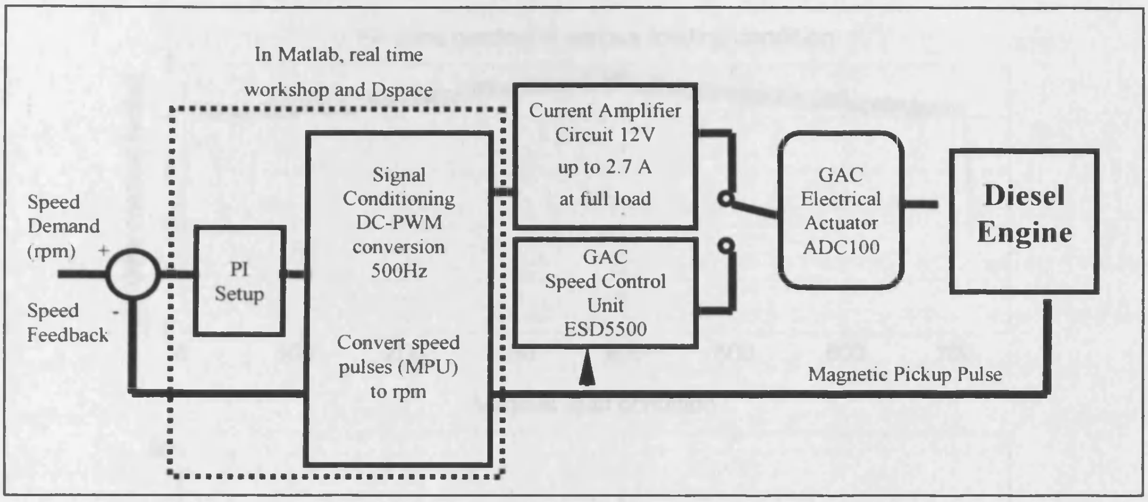


Figure D-5: Hardware setup for both the PI and GAC control system.

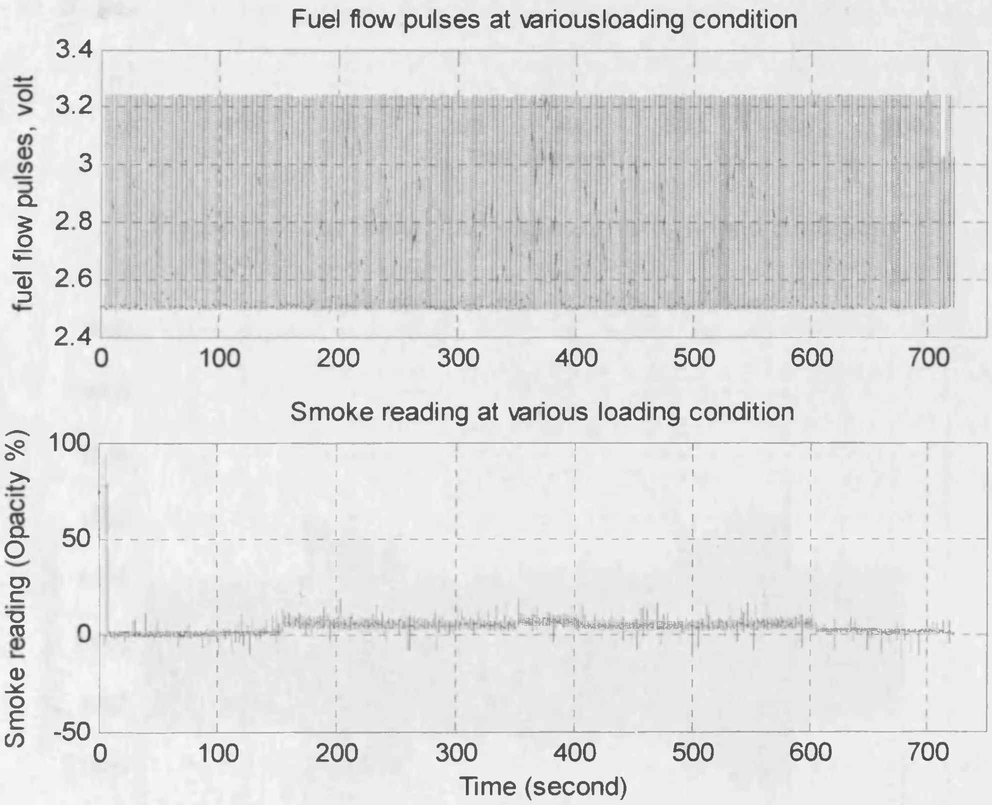


Figure D-6: Fuel flow pulses and smoke reading at different load conditions.

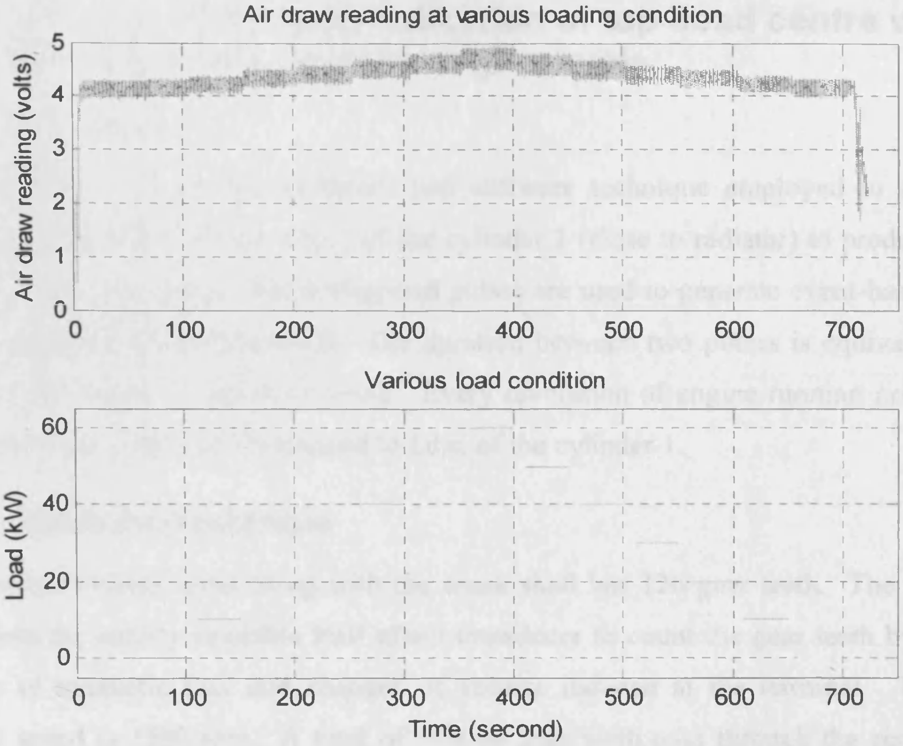


Figure D-7: Air drawn into engine at various load conditions

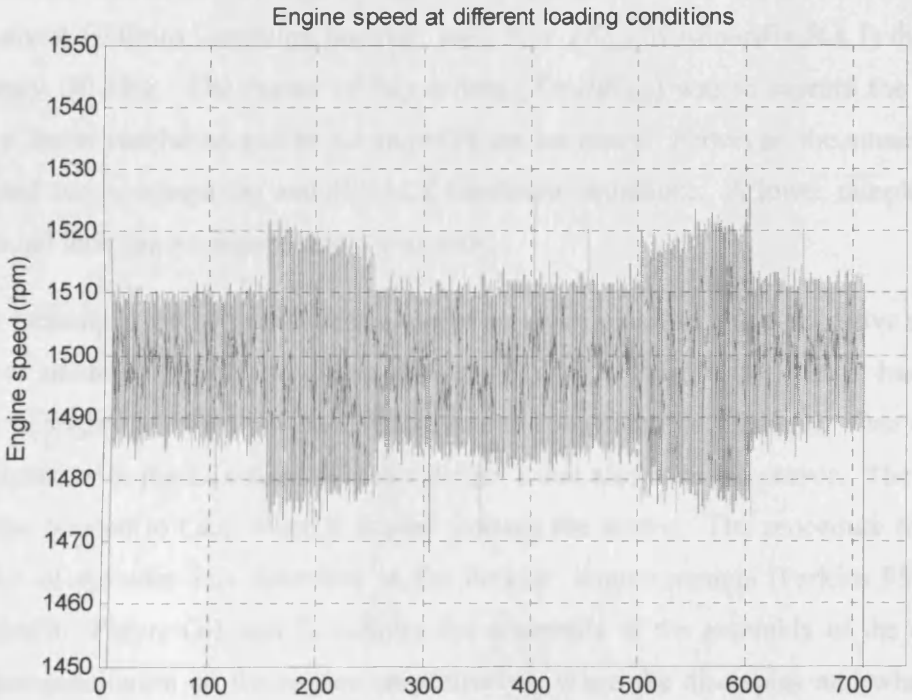


Figure D-8: Engine speed response with loading condition.

E A system for online indication of top dead centre with respect to the 4-stroke engine cycle

E.1 Introduction

This appendix describes the hardware and software technique employed to produce an indication of top-dead-centre (t.d.c.) of the cylinder 1 (close to radiator) to produce a series of event triggering pulses. These triggered pulses are used to generate event-based data for system identification in Chapter 5. The duration between two pulses is equivalent to 180 degree of the engine combustion cycle. Every revolution of engine running produces two pulses and these pulses are referenced to t.d.c. of the cylinder 1.

E.2 Hardware Technique

The fly-wheel which turns along with the crank shaft has 126 gear teeth. The initial plan was to use the readily available Hall effect transducer to count the gear teeth by means of changes of magnetic flux and changes of voltage induced at the terminal. The engine nominal speed is 1500 rpm. A total of 189000 gear teeth pass through the sensor in one minute or 3150 gear teeth in one second. So, every 1575 teeth passing the sensor is equivalent to 180 degree of engine combustion cycle. The sampling frequency was initially set to about 10 times (sampling theorem says, $F_s > 2F_{max}$ in Appendix B.4.1) the gear teeth frequency, 30 kHz. The reason of this setting (F_s=$10F_{max}$) was to capture the peak of the pulse at better resolution and hence improve the accuracy. However, the situation was not permitted due to computing and dSPACE hardware limitations. A lower sampling rate was considered later but no improvement was seen.

A new technique was proposed with a use of an aluminum disc and a reflective sensor. The disc was mounted to one end of the crankshaft (near to the radiator) and it has two holes, each 180 degree apart on the disc. Each hole is about 2mm in diameter. One of the holes was aligned with the t.d.c mark of the cylinder 1 and also with the sensor. The subsequent hole also pointed to t.d.c. when it passed through the sensor. The procedure for obtaining the t.d.c. of cylinder 1 is described in the Perkins' engine manual [Perkins 95], operation 17A-01A/B. Figure G.1 and G.2 shows the schematic of the assembly of the devices and hardware installation on the engine respectively. When the disc spins and when the t.d.c. mark is at the same level as the level pointer, the cylinder 1 (cylinder that nearest to the radiator) is at its t.d.c. A reflective object sensor (OPB704) was used to capture the hole-pulse signal. With this technique, a lower sampling frequency was considered as the signal rate is lower compared to the previous technique. The engine runs at 1500 rpm. With the two holes on the disc, they increase the signal to 3000 pulses per minute or 50 pulses per second. The location of the sensor is very critical. Placing the sensor was done by trial-and-error and it was a time consuming process throughout.

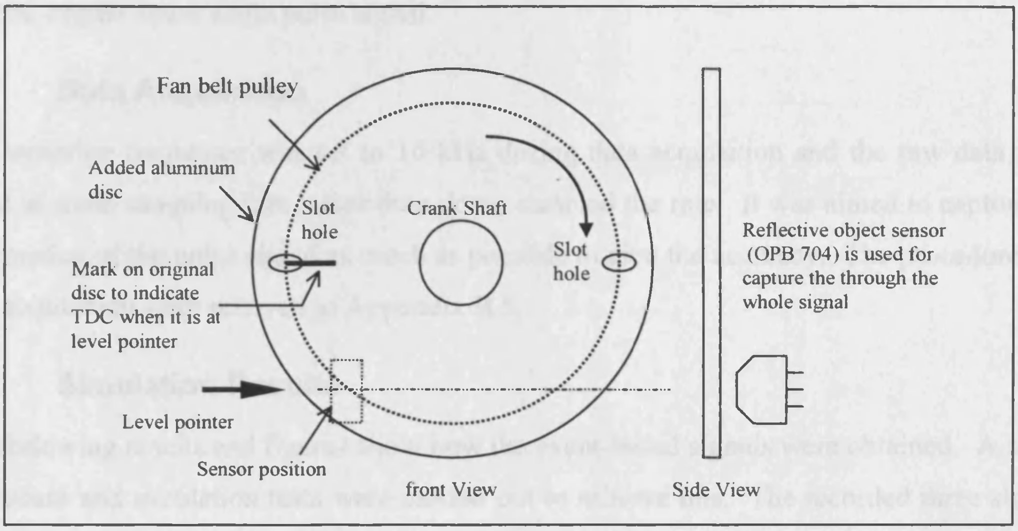


Figure E-1: Setup for 180 degree pulse trigger system

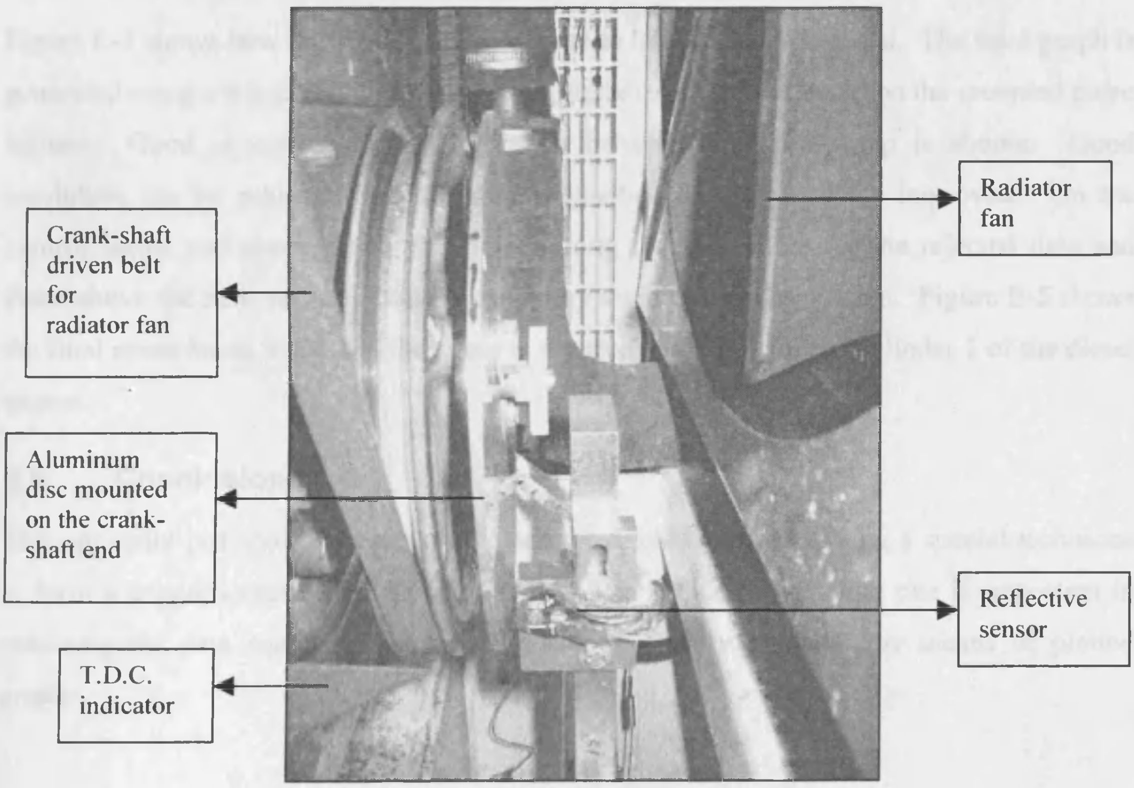


Figure E-2: Hardware installation on the engine.

E.3 Parameter to Measure

1) Control signal – its unit is equivalent to duty cycle value
2) Engine speed signal (rpm)
3) 180 degree crank angle pulse signal

E.4 Data Acquisition

The sampling frequency was set to 10 kHz during data acquisition and the raw data were saved at same sampling rate rather than down sampled the rate. It was aimed to capture the information of the pulse signal as much as possible to give the accuracy. The procedures for data acquisition were referred to Appendix B.5.

E.5 Simulation Results

The following results and figures show how the event-based signals were obtained. A series of process and simulation tests were carried out to achieve this. The recorded three signals were plotted in Figure E-3. The data were recorded without an intermediate filter. It was shown that in one second, a total of 50 pulses recorded on the graph. This figure shows time-based data set.

Figure E-4 shows how the pulses series data use to trigger the time-based. The third graph is generated using a trigger system written in a Matlab m-file. It is based on the recorded pulse signals. Good consistency of the duration between each pulse gap is shown. Good resolution can be achieved and the data acquisition accuracy can be improved. On the control signal and speed graph, the points along the zero value are the rejected data and those above the zero value are data obtained by the pulse trigger system. Figure E-5 shows the final event-based value and the event is referred to the t.d.c of the cylinder 1 of the diesel engine.

E.6 Conclusion

This appendix provides the technique to acquire event-based data using a special technique to form a trigger system for data acquisition. The choice of sampling rate is important in obtaining the data and the data process demonstration was shown by means of plotted graphs.

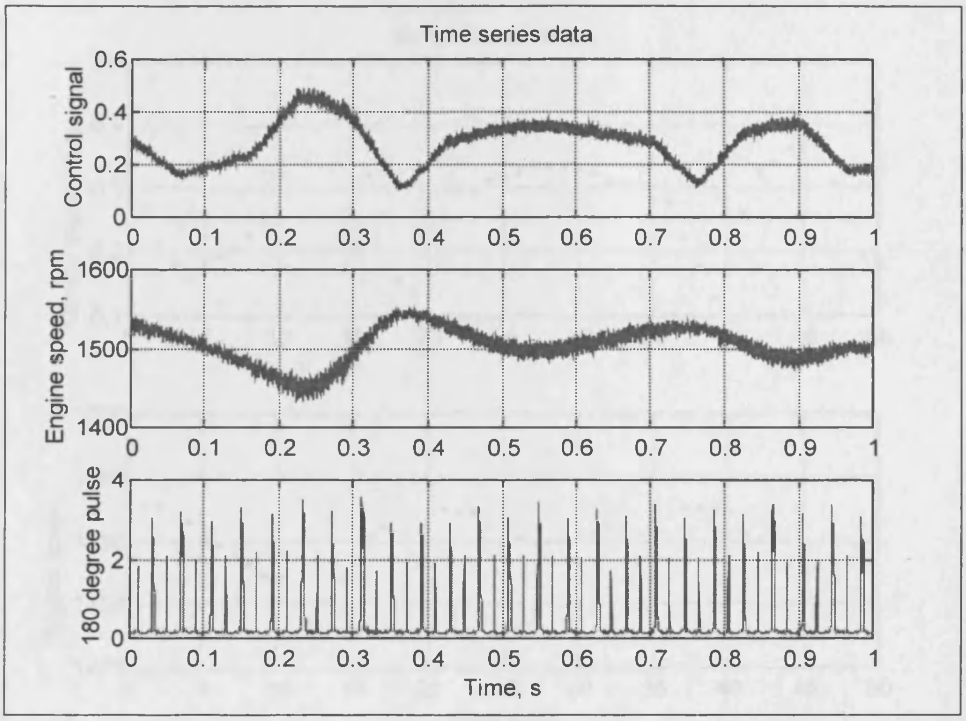


Figure E-3: Raw data recorded during the test.

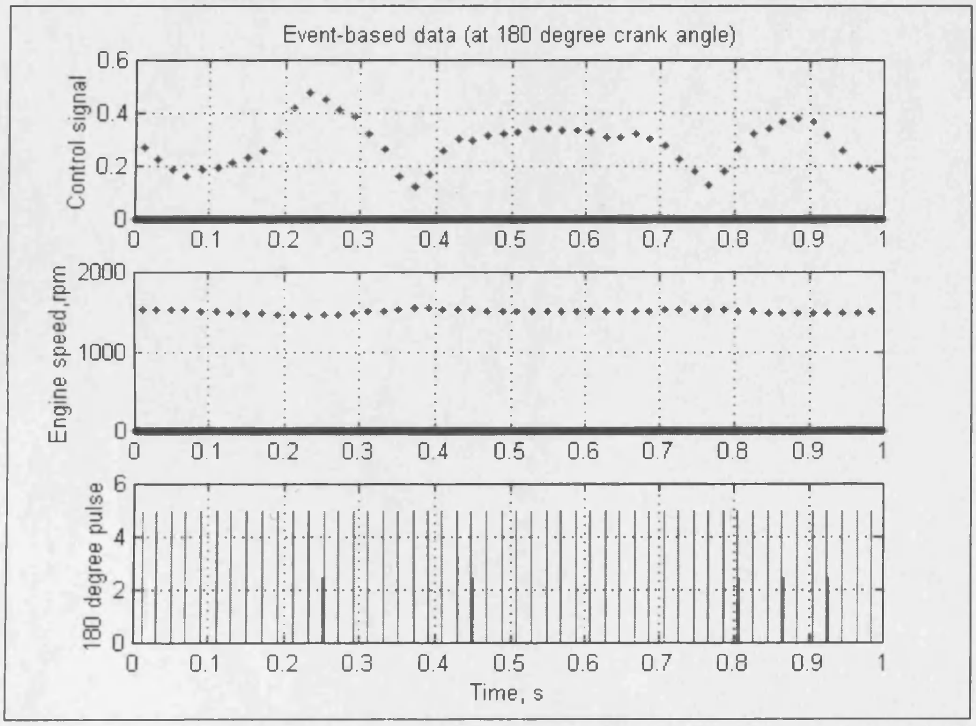


Figure E-4: Wanted and unwanted data based on the pulses.

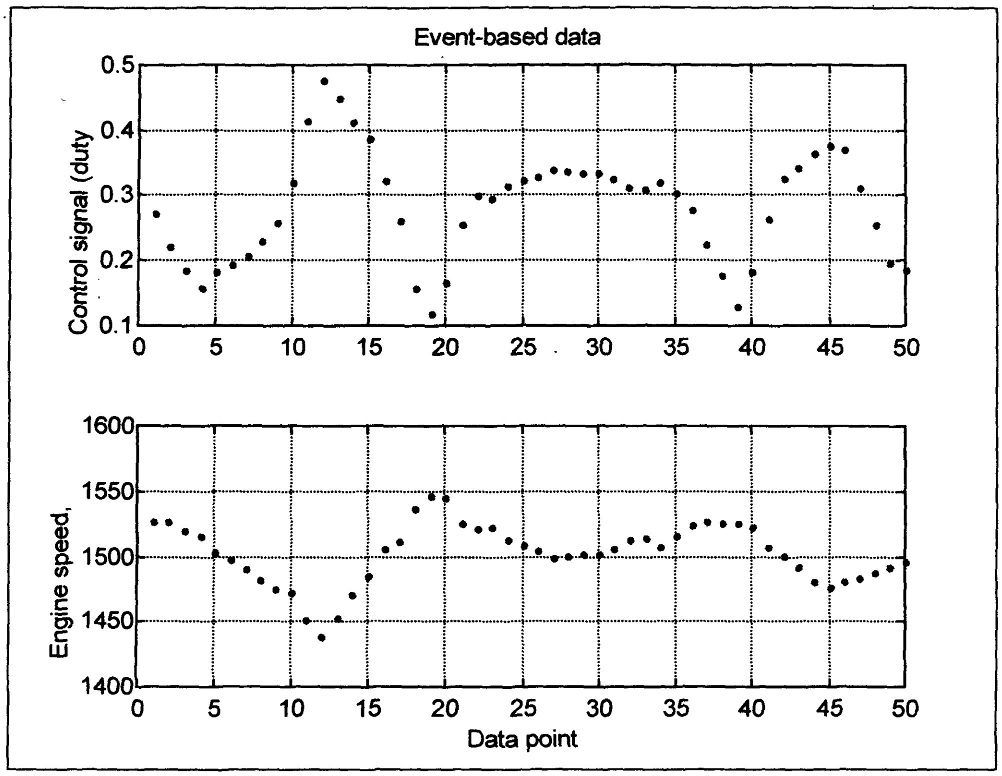


Figure E-5: Event-based (at cylinder 1 T.D.C.) data.