Institutionen för systemteknik
Department of Electrical Engineering

Master Thesis

Model Based Diagnosis of the Intake Manifold Pressure on a Diesel Engine

Master thesis performed in Vehicular Systems
at the Institute of Technology in Linköping
by

Christoffer Bergström and Gunnar Höckerdal

LiTH-ISY-EX–09/4290–SE

Linköping 2009

Linköpings universitet
TEKNISKA HÖGSKOLAN

Model Based Diagnosis of the Intake Manifold Pressure on a Diesel Engine

Master thesis performed in Vehicular Systems
at the Institute of Technology in Linköping
by

Christoffer Bergström and Gunnar Höckerdal

LiTH-ISY-EX–09/4290–SE

Supervisor: **Johan Wahlström**
                  ISY, Linköpings universitet
               Carl Svärd
                  Scania CV AB

Examiner: **Erik Frisk**
                  ISY, Linköpings universitet

Linköping, 31 August, 2009

Titel Title	Modellbaserad laddtrycksdiagnos för en dieselmotor Model Based Diagnosis of the Intake Manifold Pressure on a Diesel Engine

Författare Author	Christoffer Bergström and Gunnar Höckerdal

Sammanfattning
Abstract

Stronger environmental awareness as well as actual and future legislations increase the demands on diagnosis and supervision of any vehicle with a combustion engine. Particularly this concerns heavy duty trucks, where it is common with long driving distances and large engines. Model based diagnosis is an often used method in these applications, since it does not require any hardware redundancy.

Undesired changes in the intake manifold pressure can cause increased emissions. In this thesis a diagnosis system for supervision of the intake manifold pressure is constructed and evaluated. The diagnosis system is based on a Mean Value Engine Model (MVEM) of the intake manifold pressure in a diesel engine with Exhaust Gas Recirculation (EGR) and Variable Geometry Turbine (VGT). The observer-based residual generator is a comparison between the measured intake manifold pressure and the observer based estimation of this pressure. The generated residual is then post treated in the CUSUM algorithm based diagnosis test.

When constructing the diagnosis system, robustness is an important aspect. To achieve a robust system design, four different observer approaches are evaluated. The four approaches are extended Kalman filter, high-gain, sliding mode and an adaption of the open model. The conclusion of this evaluation is that a sliding mode approach is the best alternative to get a robust diagnosis system in this application. The CUSUM algorithm in the diagnosis test improves the properties of the diagnosis system further.

Nyckelord
Keywords   Model Based Diagnosis, MVEM, Observer, CUSUM

Abstract

Stronger environmental awareness as well as actual and future legislations increase the demands on diagnosis and supervision of any vehicle with a combustion engine. Particularly this concerns heavy duty trucks, where it is common with long driving distances and large engines. Model based diagnosis is an often used method in these applications, since it does not require any hardware redundancy.

Undesired changes in the intake manifold pressure can cause increased emissions. In this thesis a diagnosis system for supervision of the intake manifold pressure is constructed and evaluated. The diagnosis system is based on a Mean Value Engine Model (MVEM) of the intake manifold pressure in a diesel engine with Exhaust Gas Recirculation (EGR) and Variable Geometry Turbine (VGT). The observer-based residual generator is a comparison between the measured intake manifold pressure and the observer based estimation of this pressure. The generated residual is then post treated in the CUSUM algorithm based diagnosis test.

When constructing the diagnosis system, robustness is an important aspect. To achieve a robust system design, four different observer approaches are evaluated. The four approaches are extended Kalman filter, high-gain, sliding mode and an adaption of the open model. The conclusion of this evaluation is that a sliding mode approach is the best alternative to get a robust diagnosis system in this application. The CUSUM algorithm in the diagnosis test improves the properties of the diagnosis system further.

Acknowledgments

First of all, we would like to thank our excellent supervisor at Scania CV AB, Carl Svärd, for all long inspiring discussions and for his great support during the work. We would also like to thank our supervisor Johan Wahlström and our examiner Erik Frisk at Linköping University for always taking their time to support us during our work.

The staff at NESD, NESE and a lot of people in building 101 deserve a thank for their involvement and help. We would also like to thank our fellow master thesis workers, Johan Björling, Josef Dagson, Oskar Franke and Samuel Nissilä Källström, for all motivating and humorous discussions regarding our projects and everything else.

Finally, our families deserve a big thank for always supporting and guiding us in both good and bad times!

<div align="right">

Christoffer Bergström and Gunnar Höckerdal
Södertälje, June 2009

</div>

Contents

Chapter 1

Introduction

1.1   Background

Stronger environmental awareness and future legislations increase the demands
on lowered emissions from any vehicle with a combustion engine. To meet these
requirements on heavy duty trucks, truck manufacturers equip their vehicles with
emission reduction systems like Exhaust Gas Recirculation (EGR). These measures
are though not sufficient. It is also required to diagnose and supervise the engine
systems that affect the formation of emissions. Such a system is called an On
Board Diagnostic system (OBD).

Examples of systems that need to be diagnosed in the engine are the tur-
bocharger, consisting of a compressor and a turbine, and the rest of the gas flow
system. Faults in these kinds of systems lead to higher emissions since the air
mass flow into the cylinders will be affected.

In this masters thesis the objective is to design a diagnosis system for detection
of under-boost and over-boost, of significant magnitude, in the intake manifold.
Under-boost and over-boost are symptoms of any fault causing the measured boost
pressure to deviate from the estimated boost pressure. A fault is an actual physical
defect, and a symptom is the visible result of a fault. Reasons for under-boost can
be leakage or a malfunctioning turbocharger and a reason for over-boost may be
a restriction in the cylinder air intake. A more exhaustive analysis is found in the
fault tree analysis in Appendix D.

The goal of this master thesis work is to construct and analyse a diagnosis
system for supervision of the intake manifold pressure on a diesel engine with EGR
and Variable Geometry Turbine (VGT), and to decide if it is possible to implement
it in an OBD. The aim of the diagnosis system is to detect under-boost and over-
boost in the intake manifold pressure, and to isolate them from each other. An
efficient design method for the diagnosis system is to use a residual generator based
on an observer. This observer is based on a Mean Value Engine Model (MVEM)
of the intake manifold pressure. In the work with the diagnosis system design, it
is assumed that all components after the exhaust manifold work as they would
in the fault-free case. This means that all possible faults in components after the

exhaust manifold are neglected.

1.2 Design Process

The diagnosis system constructed in this work consists of a diagnostic observer, a residual generator and a diagnosis test. The diagnostic observer is based on an MVEM. The residual generator is a comparison between the measured intake manifold pressure and the observer based estimation of this pressure. Finally the diagnosis test uses the generated residual to make diagnosis statements.

Figure 1.1 gives an overview of the diagnosis system design process. First a model of the intake manifold pressure is derived and validated. To achieve a robust system, the next step in the design is to evaluate different observer approaches to find the most appropriate design for the diagnostic observer. To complete the diagnosis system, a diagnosis test is constructed. For the test to be able to make diagnosis statements, a comparison between test quantities and thresholds is needed. The last part of the design work is to perform an evaluation of the diagnosis system, to make sure that it works properly in both stationary and transient conditions.
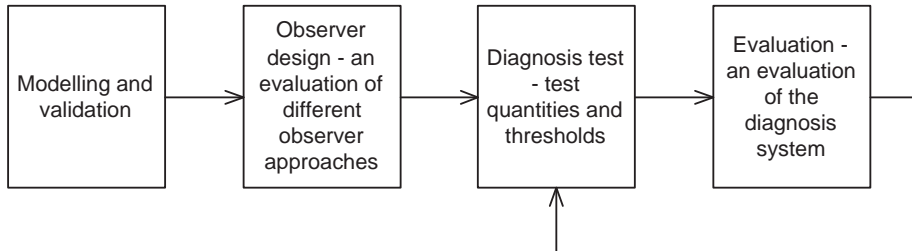
Figure 1.1. The different steps in the process of the diagnosis system design.

1.3 Problem Description

The problem to be investigated in this thesis work is as follows:

Construct and analyse a model-based diagnosis system for supervision and detection of under-boost and over-boost in the intake manifold pressure in a five cylinder diesel engine with EGR and VGT, and investigate if it can be implemented in the Engine Control Unit (ECU).

1.4   Contributions

The main contributions of this thesis work are

- a simplification of the dynamic model of the intake manifold pressure described in [16],

- a transformation of the dynamic model from differential and algebraic equations to a state-space model,

- an evaluation of four different observer designs based on the model,

- a diagnosis system based on the diagnostic observer for supervision of the intake manifold pressure,

- an evaluation of the performance of the diagnosis system using real stationary and transient measurements.

1.5   Outline

Chapter 1 , **Introduction**, gives the problem description and the background to this thesis.

Chapter 2 , **Detection Theory**, gives a short introduction to fault detection and model-based diagnosis.

Chapter 3 , **Modelling**, describes the MVEM for the gas flow in a diesel engine with EGR and VGT.

Chapter 4 , **Model Validation and Sensitivity Analysis**. The model is validated and a sensitivity analysis is carried out.

Chapter 5 , **Observer Design**. Three different observer designs based on the model derived in **Chapter 3** are investigated.

Chapter 6 , **Construction of the Diagnosis Test**. Two different observer methods are used for implementation of a diagnosis system. Further, the two diagnosis systems are evaluated.

Chapter 7 , **Conclusions and Future Work**. Results and conclusions of the work are presented and possible future work is stated.

Chapter 2

Detection Theory

This master thesis considers model based diagnosis of a technical system. In this chapter, an introduction to model based diagnosis including fault detection and a short description of diagnosis system design are given. Note that in general it is desirable to isolate different faults from each other, but in this case it is only wished to detect any fault that results in under-boost or over-boost and then isolate these two symptoms from each other.

2.1   Definitions

In this section some definitions and concepts, that are commonly used in the diagnosis area are presented. The IFAC (International Federation of Automatic Control) Technical Committee SAFEPROCESS has suggested preliminary definitions of some terms [12], but the explanations below are as they should be interpreted in this thesis.

Fault
An unpermitted deviation of at least one property or variable of the system that results in an unacceptable behaviour.

Symptom
The actual visible effect of a fault.

Disturbance
An unknown and uncontrolled input to the system.

Observation
Consists of the known input signals and the measured signals.

Residual
A comparison between two signals that describe the same quantity.

Test Quantity
A quantity that shall be small in the fault-free case, and large otherwise.

Threshold
A limitation of how large the test quantity is allowed to be before an alarm is generated.

Fault Detection
Determination of if there are any fault present in the system.

Fault Isolation
Determination of the location of the fault, i.e. which component or components that have failed.

Diagnosis
Conclusion of what symptom or what symptoms that can explain the deviant system behaviour.

Alarm
An announcement of that the diagnosis system has detected a fault.

False Alarm
An alarm that is generated even though there is no fault present.

2.2   Diagnosis System

The diagnoses shall be produced by the diagnosis system, which acts on the observations from the system to be diagnosed, in this case the diesel engine. Based on these observations, the diagnosis system makes a diagnosis statement containing information about if there is a fault present and also which symptom this fault causes [12]. In Figure 2.1 the structure of the diagnosis application is shown.
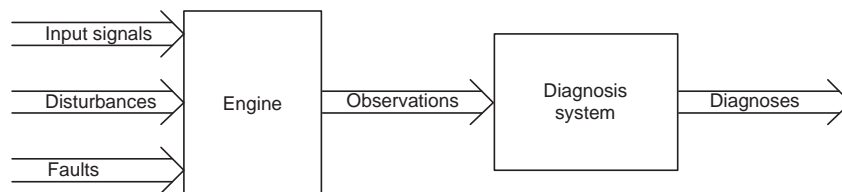
Figure 2.1. Structure of the diagnosis application. The diagnosis system consists of a model-based observer working as a residual generator, and a diagnosis test.

The diagnosis system can be considered as a function from the observations to the diagnosis statement, i.e. the diagnoses in Figure 2.1. In some simple cases it is easy to illustrate a diagnosis test by typing the observations and the respective diagnoses in a table. This is done in Example 2.1.

┌── **Example 2.1: A simple diagnosis system** ──────────────┐

Consider the water pump system below. It consists of a water tank, an electric water pump and a valve. The water pump (P) pumps water from the tank and with the valve (V) set to open, the water flows through the pipe and leaves it at point O. If the valve instead is set to closed, there is no water flow at O.
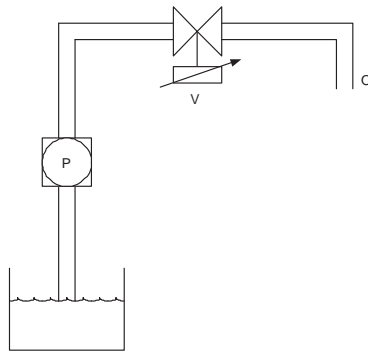
Figure 2.2. An overview of the water pump system.

Assume that only three faults can occur: the valve stuck in open position ("V SO"), the valve stuck in closed position ("V SC") and the pump is broken ("P broken"). The input signal to the system is the desired valve position (open or closed) and the observations are the desired valve position and if water flows at point O or not.

Table 2.1. A simple diagnosis system.

Desired valve position	Water observation	Diagnosis statement
open	water flow	"no fault", "V SO"
open	no water flow	"V SC", "P broken" "V SC and P broken" "V SO and P broken"
closed	water flow	"V SO"
closed	no water flow	"no fault", "V SC", "P broken" "V SO and P broken" "V SC and P broken"

A diagnosis system can then look like Table 2.1. For an example, assume that the observations are that the desired valve position is open and that water flows at point O. Then the diagnosis statement is either "no fault" or "valve stuck open".

Each of the statements in Table 2.1 is called a diagnosis and consists of either single faults or multiple faults. In this thesis only the symptoms under-boost and over-boost are considered and therefore these two symptoms are the only possible diagnosis. In many simple cases the diagnosis system can sufficiently well be represented by a table like the one in Example 2.1. However, the diesel engine is a more complex system which, in this thesis, is diagnosed using a model-based diagnosis system described in the following chapters, with a focus on the actual test in Chapter 6.

2.3   Model-Based Diagnosis

The diagnosis system considered in this thesis is based on a physical model of the diesel engine. This model is described in Chapter 3. Compared to using traditional diagnosis with hardware redundancy, the model based diagnosis has a lot of advantages [12].

In model based diagnosis, a sensor output can be compared to the modelled output, instead of being compared to a redundant physical sensor. This implies that no extra hardware is needed. Another advantage is that diagnosis of this kind may have a higher diagnosis performance and detect smaller faults in less time. With a model it is also possible to compensate for disturbances, which yields higher accuracy.

The main disadvantage of model based diagnosis is that reliable models are needed [12]. This results in more complex design procedures when the models are constructed. Another possible drawback is the computational effort needed. This is though not a general disadvantage, since it depends on the complexity of the model.

When constructing the diagnosis system, robustness is an important aspect. An efficient method to achieve this is to base the diagnosis system on an diagnostic observer. In this work, the diagnostic observer is based on the physical model of the engine.

2.4   Construction of Diagnosis Tests

An important part of a diagnosis system is the actual diagnosis test, which is used to achieve robustness. The diagnosis test evaluates the test quantity and alarms if it exceeds a certain threshold. This test quantity is normally based on a residual, produced by a residual generator. In this thesis, the residual generator builds on the model-based state observer and produces a residual according to $r = y - \hat{y}$, where y is the measured quantity and $\hat{y}$ is the estimated quantity.

To be able to make reliable decisions based on the residual, it is sometimes necessary to apply some post treatment to it. This is done to lower the noise of the signal and to obtain a trade-off between detection performance and detection time. For this purpose a simple low-pass filter or a threshold can be applied to the residual [12]. However, in this thesis the often used CUSUM algorithm is applied, see Section 6.2.

Now the entire diagnosis system consists of a diagnostic observer, a residual generator and a diagnosis test. An overview of a diagnosis system is shown in Figure 2.3.
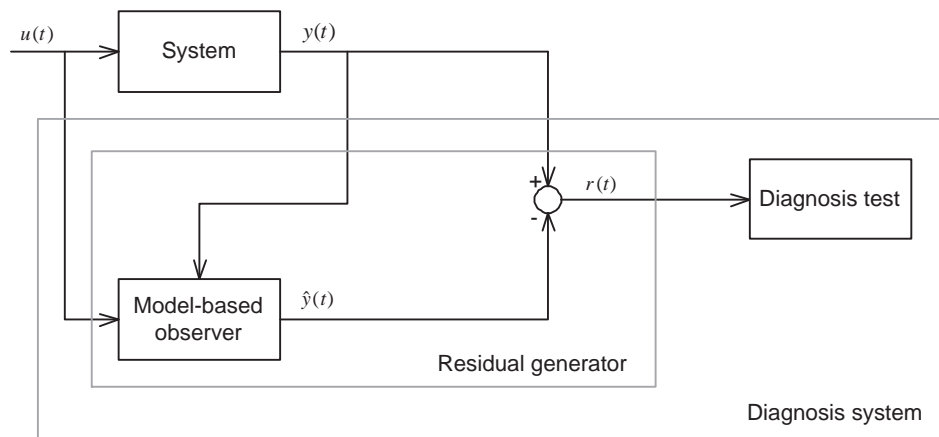

Figure 2.3. An overview of the model based diagnosis system. The system output, $y(t)$, is measured with a sensor and compared with the estimated model output, $\hat{y}(t)$. The result of this comparison, the residual $r(t)$, is then used in a diagnosis test to be able to detect if any fault has occurred.

Chapter 3

Modelling

Since the purpose of this work is to supervise the intake manifold pressure, it is desirable to generate a residual reacting on deviations in this pressure. One of the existing sensor signals in the ECU is the intake manifold pressure. In order to generate a pressure residual the intake manifold pressure is also modelled using other sensor signals. The first step in the construction of the diagnosis system is therefore to investigate what sensors the engine is equipped with, in order to decide what parts of the engine that need to be modelled. An objective is to keep the model complexity at a low level, so only if a needed quantity is not measured with a physical sensor, this quantity has to be modelled based on other sensor signals.

In this chapter a MVEM for the gas exchange in a diesel engine with EGR and VGT is developed. This model is used as a basis for the further design of the diagnosis system. The model is derived from [1], [14] and [16]. The intake manifold pressure will be supervised and therefore, just in order to keep the model complexity at a low level, this is the only state considered. The model is fed with a set of known signals, either measured signals or actuator signals.

Further, the model needs to be parameterised using data from a real engine. This is a new engine with new software, which makes it hard to obtain enough proper data. This problem is solved by using parameters from models of similar engines and then tuning them ad hoc.

3.1 Model Structure

The diagnosis system is based on a physical Mean Value Engine Model (MVEM). The structure of the model is shown in Figure 3.1. The considered engine is a five cylinder, 9.3 liter diesel engine with EGR and VGT. Note that the structure in Figure 3.1 is a substantial simplification of a real diesel engine, for example neither the turbo intercooler nor the EGR cooler is present.

This engine has a throttle in the air intake in order to minimise the NO_x emissions at low engine speeds. This throttle is normally open and activated only

in certain operating points, and therefore the diagnosis system can simply be shut off in these points and the modelling of the throttle thereby also be neglected.

These simplifications of the model are done mainly because it is of major interest to get a model with as low complexity and computational demand as possible, which makes the model-based diagnosis system easier to implement in an OBD system. Further, the coolers do not affect the intake manifold pressure in particular, but the temperature and density. Here the intake manifold temperature is a known measured signal and given as an input signal to the model, and therefore it is no need to model the temperature changes over the two coolers.
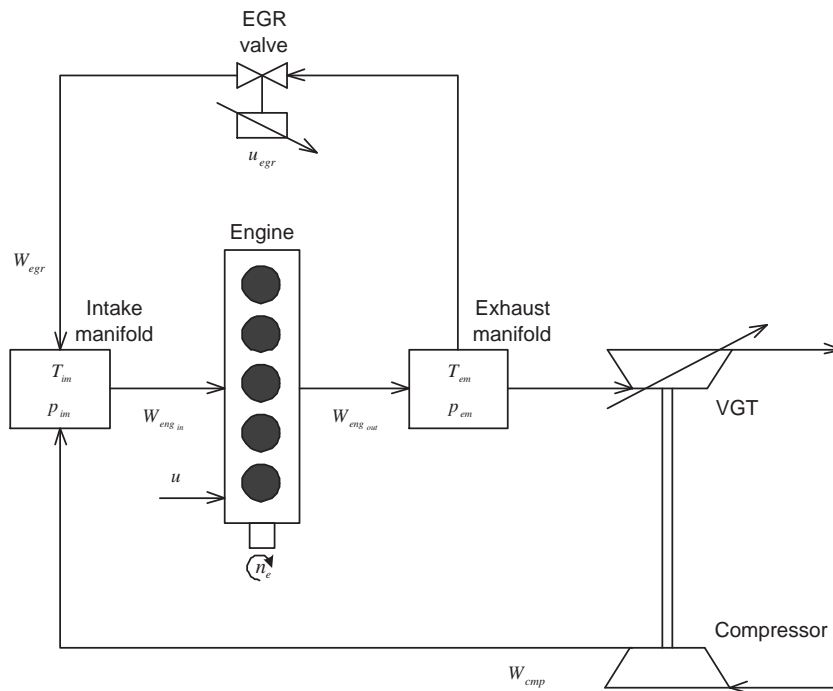
Figure 3.1. A simplified model structure of the five cylinder, 9.3 liter diesel engine with EGR and VGT.

3.2 Known Signals

The intake manifold pressure, p_{im}, is one existing sensor signal. To achieve the redundancy needed for the observer-based residual generation, $r = p_{im} - \hat{p}_{im}$, the intake manifold pressure is modelled as a state. As a basis for the model design, an analysis is performed to investigate which existing sensor and actuator signals in the ECU that are needed as known signals to the simplest possible model, see Figure 3.1. The following signals are needed:

Sensor Signals

- T_{amb}
  the ambient temperature is given in °C,

- W_{cmp}
  the air mass flow through the turbo compressor is given in kg/min,

- T_{im}
  the intake manifold temperature is given in °C,

- n_e
  the engine rotational speed is given in rpm,

- p_{amb}
  the ambient pressure is given in Pa,

- p_{em}
  the exhaust manifold pressure is given in Pa,

Actuator Signals

- u_δ
  the injected amount of fuel is given in $mg/cycle$,

- u_{egr}
  the EGR valve position is given in percent (completely open when $u_{egr} = 100\%$ and completely closed $u_{egr} = 0\%$).

Some of the signals have to be transformed into SI units before they are fed into the model. Though, the engine rotational speed, n_e, is intended to be in rpm in the model equations.

3.3   Intake Manifold

The most frequently used model for the intake manifold pressure, p_{im}, is an isothermal model, which assumes that the temperature in the manifold is constant, $T_{im} = T_{in} = T_{out}$. To determine the pressure, the ideal gas law, $pV = mRT$, is differentiated,

$$\frac{dp_{im}}{dt} = \frac{RT_{im}}{V_{im}}\frac{dm}{dt}. \tag{3.1}$$

The intake manifold can be viewed as a thermodynamic control volume, with constant volume V_{im}, that stores mass and energy [1]. The mass variations in the volume are determined by the inlet and outlet air mass flows, $\dot{m}_{in}$ and $\dot{m}_{out}$. These flows describe the dynamic behaviour according to

$$\frac{dm}{dt} = \dot{m}_{in} - \dot{m}_{out}. \tag{3.2}$$

The mass flow into the cylinders, $W_{eng_{in}}$, and the mass flow through the EGR system, W_{egr}, will be described in Sections 3.4 and 3.6 respectively. By combining (3.2) and (3.1) and consider $\dot{m}_{in} = W_{cmp} + W_{egr}$ and $\dot{m}_{out} = W_{eng_{in}}$, the intake manifold pressure can be modelled as

$$\dot{p}_{im} = \frac{R_a T_{im}}{V_{im}} \left(W_{cmp} + W_{egr} - W_{eng_{in}} \right), \tag{3.3}$$

where R_a is the ideal gas constant for air. This model is a simplification since the temperature is assumed to be constant and therefore the energy conservation is neglected [1].

The air mass flow through the compressor, W_{cmp}, is a measured quantity and works as an input signal to the model, see Section 3.2.

3.4   Cylinders

The total air mass flow from the intake manifold into the cylinders, $W_{eng_{in}}$, depends on several parameters, but the intake manifold pressure, p_{im}, the engine rotational speed, n_e, and the intake manifold temperature, T_{im}, are the most important ones [1]. These quantities are used when the flow is modelled using the volumetric efficiency [6], η_{vol}, according to

$$W_{eng_{in}} = \eta_{vol} \frac{p_{im} n_e V_d}{120 R_a T_{im}}. \tag{3.4}$$

This model covers the whole engine, and therefore the constant V_d is the total engine displaced volume, not the volume per cylinder.

The volumetric efficiency is a measurement of the effectiveness of the engine's ability to induct new air [1], [6]. In reality it depends on many engine parameters, but in most cases it is accurate enough to approximate η_{vol} as dependent on the intake manifold pressure, p_{im}, and the engine rotational speed, n_e. A frequently used approach in engine mean value modelling is to represent η_{vol} with the black box model

$$\eta_{vol} = c_{vol1} + c_{vol2}\sqrt{p_{im}} + c_{vol3}\sqrt{n_e}, \tag{3.5}$$

according to [1]. The constants c_{vol1}, c_{vol2} and c_{vol3} can be determined by solving a least square problem, combining (3.4) and (3.5), using stationary measurements.

Before the combustion the fuel is injected and mixed with the air. The injected amount of fuel, u_δ, is an actuator signal which decides the injected mass of fuel in mg per cycle and cylinder. The total fuel mass flow then becomes

$$W_f = \frac{10^{-6}}{120} u_\delta n_e n_{cyl}, \tag{3.6}$$

where n_{cyl} is the number of cylinders.

3.5 Exhaust Manifold

After the combustion the exhaust gas mixture is pushed out from the cylinders into the exhaust manifold. If valve overlap is neglected, the total mass flow out from the cylinders, $W_{eng_{out}}$, must be equal to the total mass flow into the cylinders in order to reach equilibrium. Since the air and fuel mass flows are the only contributions on the inlet side, $W_{eng_{out}}$ can be expressed as

$$W_{eng_{out}} = W_{eng_{in}} + W_f. \tag{3.7}$$

For the modelling of mass flow through the EGR system in Chapter 3.6 the temperature in the exhaust manifold has to be modelled. The exhaust gas temperature can be modelled in several ways [1], [14]. One common approach, also used here, is to model the cylinder out temperature, T_{cyl}, using the ideal Otto cycle, see [14]. A cycle calculation, described in Appendix A, gives

$$T_{cyl} = \eta_{oc} T_1 \left(\frac{p_{em}}{p_{im}}\right)^{1-\frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma-1}}\right)^{\frac{1}{\gamma}}, \tag{3.8}$$

where η_{oc} is a compensation factor for a non ideal cycle. The specific energy contents of the charge per unit mass is described by

$$q_{in} = \frac{W_f q_{HV}}{W_{eng_{out}}} \left(1 - x_r\right), \tag{3.9}$$

where q_{HV} is the heating value of diesel. The residual gas fraction, x_r, is the proportion of burned gas that never leaves the cylinder before new air is inducted and can be modelled as

$$x_r = \frac{1}{r_c} \left(\frac{p_{em}}{p_{im}}\right)^{\frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma-1}}\right)^{-\frac{1}{\gamma}}. \tag{3.10}$$

Due to a relatively high compression ratio, this fraction is rather small (a few percent) and sometimes, in naturally aspirated diesel engines, x_r is approximately constant since the intake is unthrottled [6]. This is not the case in this thesis, and therefore the residual gas fraction is modelled according to (3.10).

Further, the temperature of the gas mixture before compression, T_1, is calculated according to

$$T_1 = x_r T_{cyl} + (1 - x_r) T_{im},$$ (3.11)

and the compensation factor, η_{oc}, can be expressed by solving (3.8)

$$\eta_{oc} = \frac{T_{cyl}}{T_1} \left(\frac{p_{em}}{p_{im}}\right)^{\frac{1}{\gamma} - 1} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma - 1}}\right)^{-\frac{1}{\gamma}}.$$ (3.12)

The problem with the temperature modelling is that it contains one unknown variable more than there are equations, since (3.12) is equivalent to (3.8). Analysis show that the compensation factor is rather constant during a whole driving cycle and therefore, in the following, η_{oc} is approximated to be constant. Thereby the number of unknown variables is reduced and the system of equations can be solved explicitly.

However, the temperature in the exhaust manifold is not equal to the cylinder out temperature. Due to heat transfer and cooling, the temperature will drop significantly. A simple model for this process is the first temperature drop model in [1],

$$T_{em} = T_{amb} + (T_{cyl} - T_{amb}) e^{\frac{h_{tot} A}{W_{eng_{out}} c_p}},$$ (3.13)

where c_p is the specific heat capacity during constant pressure. Here all heat transfer contributions have been lumped together to one total heat transfer coefficient, h_{tot}. The constant A is the inner wall area of the exhaust manifold.

3.6   EGR System

Exhaust Gas Recirculation (EGR) is used in diesel engines to reduce the creation of NO_x emissions. Burned gas is returned from the exhaust system into the intake manifold and mixed with the air. The fraction of oxygen in the inducted air mixture is then decreased, which in turn decreases the combustion temperature and by that also the creation of NO_x.

In the intake manifold pressure model (3.3) the EGR mass flow is needed. This is modelled as a compressible flow restriction with variable area [16],

$$W_{egr} = \frac{A_{egr}(u_{egr})p_{em}\Psi_{egr}}{\sqrt{R_e T_{em}}}, \tag{3.14}$$

where A_{egr} is the effective flow area of the EGR. This flow model assumes that inverse flow can not occur when $p_{em} < p_{im}$. The function Ψ_{egr} is modelled as a parabolic function,

$$\Psi_{egr} = 1 - \left(\frac{1 - \Pi_{egr}}{1 - \Pi_{egropt}} - 1 \right)^2. \tag{3.15}$$

The pressure ratio, Π_{egr}, is limited by the sonic condition,

$$\Pi_{egropt} = \left(\frac{2}{\gamma_e + 1} \right)^{\frac{\gamma_e}{\gamma_e - 1}}, \tag{3.16}$$

and by $1 < \frac{p_{im}}{p_{em}}$, i.e. no reverse flow can occur. This yields that the pressure ratio can be expressed as

$$\Pi_{egr} = \begin{cases} \Pi_{egropt} & \text{if } \frac{p_{im}}{p_{em}} < \Pi_{egropt}, \\ \\ \frac{p_{im}}{p_{em}} & \text{if } \Pi_{egropt} \leq \frac{p_{im}}{p_{em}} \leq 1, \\ \\ 1 & \text{if } 1 < \frac{p_{im}}{p_{em}}. \end{cases} \tag{3.17}$$

The effective EGR flow area, A_{egr}, is determined by

$$A_{egr}(u_{egr}) = A_{egrmax} f_{egr}(u_{egr}) \tag{3.18}$$

where A_{egrmax} is the maximum EGR area and f_{egr} is a polynomial function of u_{egr},

$$f_{egr}(u_{egr}) = \begin{cases} c_{egr1} u_{egr}^2 + c_{egr2} u_{egr} + c_{egr3} & \text{if } u_{egr} \leq -\frac{c_{egr2}}{2c_{egr1}}, \\ \\ c_{egr3} - \frac{c_{egr}^2}{4c_{egr1}} & \text{if } u_{egr} > -\frac{c_{egr2}}{2c_{egr1}}, \end{cases} \tag{3.19}$$

according to [16]. The constants c_{egr1}, c_{egr2} and c_{egr3} can be determined by solving a least square problem using stationary measurements.

3.7 Parameter Estimation

The different parameters in this model are

$$A, \; A_{egrmax}, \; c_{egr1}, \; c_{egr2}, \; c_{egr3}, \; c_{vol1}, \; c_{vol2}, \; c_{vol3}, \; h_{tot}.$$

It is desirable to calculate these parameters using least square optimisation using data from stationary measurements, i.e. data collected when the engine is driven in stationary conditions in an engine test cell. It is important to get as many stationary points as possible, to get a good optimisation, as well as it is important to get the correct signals measured. If any measured signal is missing in the set of data, the parameter estimation of the model becomes harder. Sometimes, when not all wanted quantities are measured, it is needed to calculate them instead, on the basis of physical relations to other measured quantities.

Since the engine is new, it is not completely calibrated and it is not decided how the final engine will be equipped. Therefore it is not possible to obtain sufficiently proper measurements for a parameter estimation. However, iterative simulations show that it is possible to use parameters from an older model of a similar engine and then, if it is necessary, tune them manually to fit the new model.

Chapter 4

Model Validation and Sensitivity Analysis

The performance of the model-based diagnosis system mainly depends on the accuracy of the model. To ensure that the model in Chapter 3 describes the reality sufficiently well, it therefore needs to be validated. The model is divided into submodels for a more systematic validation. The validated submodels are the models describing $W_{eng_{in}}$ and W_{egr}. Since the model of W_{egr} includes T_{em}, which is a complex model in it self, T_{em} is validated separately. Finally the entire model, p_{im}, is validated.

Last in this chapter a sensitivity analysis is presented. It contains an analysis of the model's sensitivity to both parameter errors and errors in the input signals.

The data used for these purposes is measurements from the five cylinder diesel engine with EGR and VGT driven by a World Harmonised Stationary Cycle (WHSC) in an engine test cell. The measurements consist of steps in injected amount of fuel, u_δ, and EGR valve position, u_{egr}, see Figure 4.1. Also the engine rotational speed, n_e, is allowed to change. The model is fed with the actual signals from the measurements and then, the predicted intake manifold pressure is compared with the measured one.

4.1 Validation Prerequisites

Not all models in Sections 3.3-3.6 can be validated because of limitations in where sensors can be placed in a physical engine and which quantities that really can be measured with sensors. An additional limitation is that the engine intake air mass flow, $W_{eng_{in}}$, and the EGR mass flow, W_{egr}, are not measured in this case, which makes it necessary to calculate these quantities from other measurements in order to validate corresponding models. This can be done with the expected EGR fraction, x_{egr}, which is calculated in the ECU according to
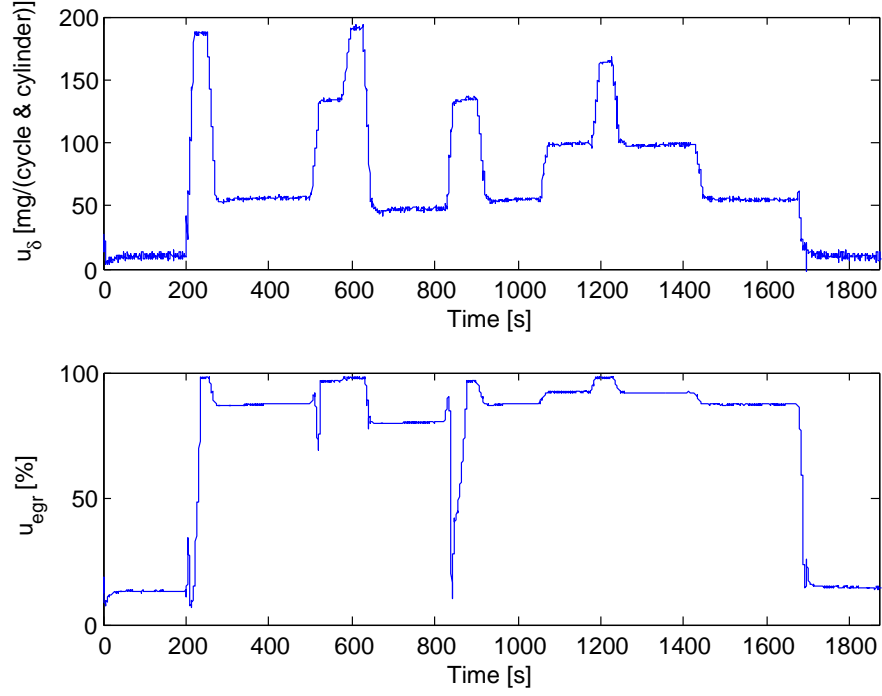
Figure 4.1. Steps in u_δ and u_{egr}.

$$x_{egr} = \frac{W_{egr}}{W_{eng_{in}}}, \tag{4.1}$$

where

$$W_{eng_{in}} = W_{cmp} + W_{egr}. \tag{4.2}$$

By combining (4.1) and (4.2), the mass flows can then be expressed as

$$W_{eng_{in}} = \frac{1}{1 - x_{egr}} W_{cmp}, \tag{4.3a}$$

$$W_{egr} = \frac{x_{egr}}{1 - x_{egr}} W_{cmp}, \tag{4.3b}$$

It is important to notice the fact that these assumptions only are valid in stationary conditions, i.e. when (4.2) is valid. The results from the validations will be presented in absolute and relative error in the intake manifold pressure.

$$\text{Absolute error} = |measurement - model|$$

$$\text{Relative error} = \frac{|measurement - model|}{measurement}$$

4.2   Cylinders

The validation in this section concerns the equations (3.4) and (3.5). The result of the validation of the engine intake mass flow can be seen in Figure 4.2. Because of the calculations and assumptions made in (4.3) and (4.2), this validation can only be performed during stationary conditions. It seems like the mass flow model describes the reality very well and differences can only be seen in a few operating points. One such condition is when the engine idles, at the start and end of the cycle. The idling modes are therefore disregarded in the final evaluation result, see Section 4.6.

Since the volumetric efficiency, η_{vol}, only is present in the model for $W_{eng_{in}}$ according to (3.4), also the efficiency model in (3.5) can be considered to work properly.
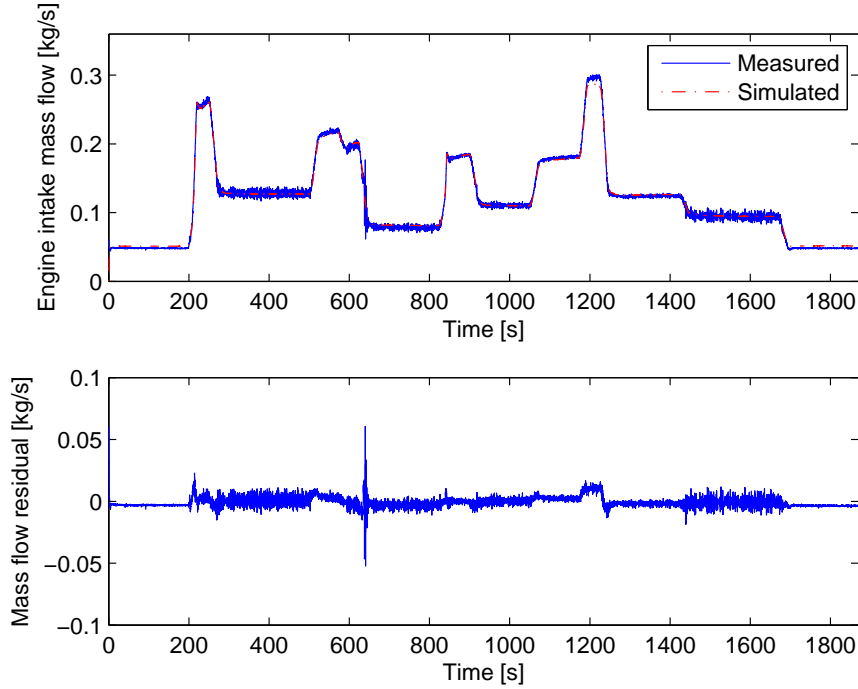
Figure 4.2. Validation of $W_{eng_{in}}$. The model can only be evaluated in the stationary points because of the assumptions made in the calculations.

4.3   Exhaust System

The exhaust system temperature, T_{em}, given from the model equations (3.6)-(3.13), needs to be filtered before the validation in order to compensate for the slow dynamics in the temperature sensor. The result is shown in Figure 4.3. As can be seen, the temperature model describes the measurement values very good in some operating points and less good in others. This is not a problem in the end, the magnitudes of the relative errors are acceptable and in the total model the exhaust gas temperature just affects the pressure through the EGR mass flow. A fault in T_{em} contributes with the error $\mathcal{O}(\frac{1}{\sqrt{n}})$ in W_{egr} according to (3.14).

Further, this is the result from an ideal Otto cycle and therefore the expectations on the temperature model are not very high. But, as can be seen in Section 4.6, the errors in the final temperature output are sufficiently small.
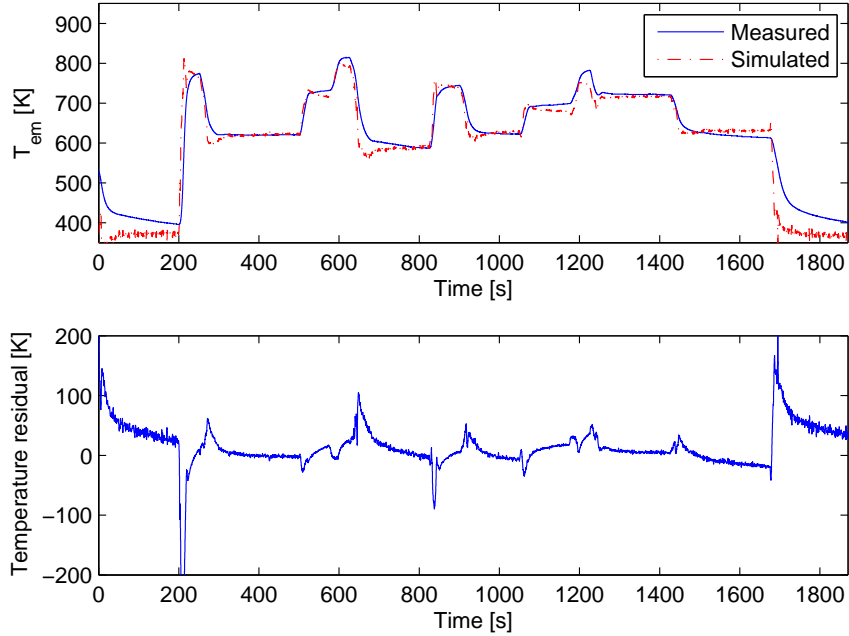
Figure 4.3. Validation of T_{em}. The model describes the temperature well, except in idling mode and some transients.

4.4 EGR-System

The validation of the EGR system is almost equivalent to the validation of the intake mass flow. Due to the same reason as in that case, it can only be evaluated in the stationary conditions. The result can be seen in Figure 4.4. The dynamic behaviour in the EGR valve is neglected in the model, but the reproduction of the EGR mass flow is still accurate. It shall also be noticed that this is not the most accurate way to validate a model. Since W_{egr} does not exist in the test cell data set, it is calculated according to (4.3b).

Figure 4.4. Validation of W_{egr}. The model can only be evaluated in the stationary points because of the assumptions made in the calculations.

4.5   Intake Manifold

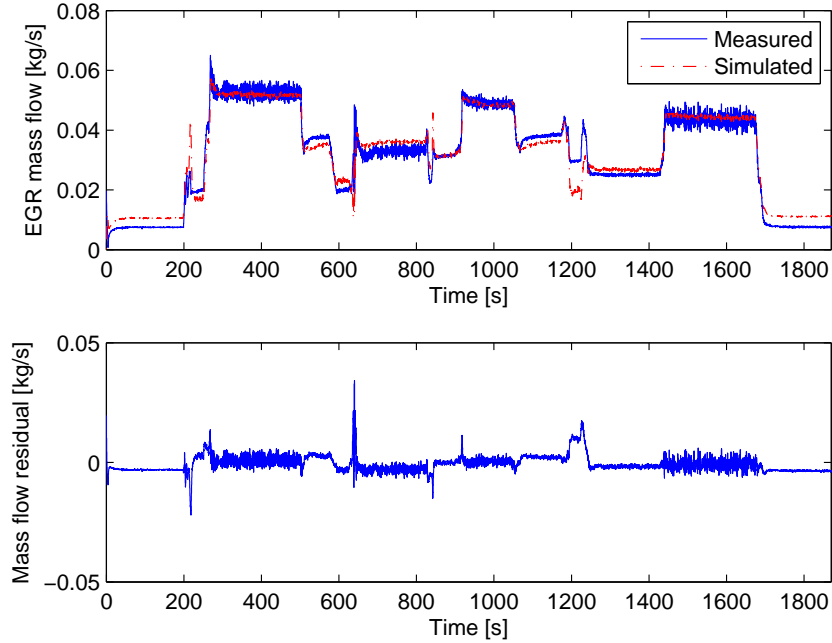Finally it is time to see how well the modelled intake manifold pressure agrees with the measured intake manifold pressure. In Figure 4.5 it is illustrated that the model depicts the transients well, but there is an offset which appears to be a constant bias at approximately three percent. This is not a problem when it comes to diagnosis, since if it is known that the model has a constant fault, this can be taken care of in a diagnosis test. Since the prediction error, $p_{im} - \hat{p}_{im}$, is used as residual, it is not an issue if the model does not totally agree with the reality. This is instead a question of how small faults that are of interest to detect. This can be adjusted by the thresholds of the diagnosis test.

On the other hand, when the bias is constant in each operating point, it would be possible to make an adjustment of this model error. This means that the bias can be eliminated with some kind of adaption or observer. An investigation of these methods is carried out in Chapters 5 and 6.
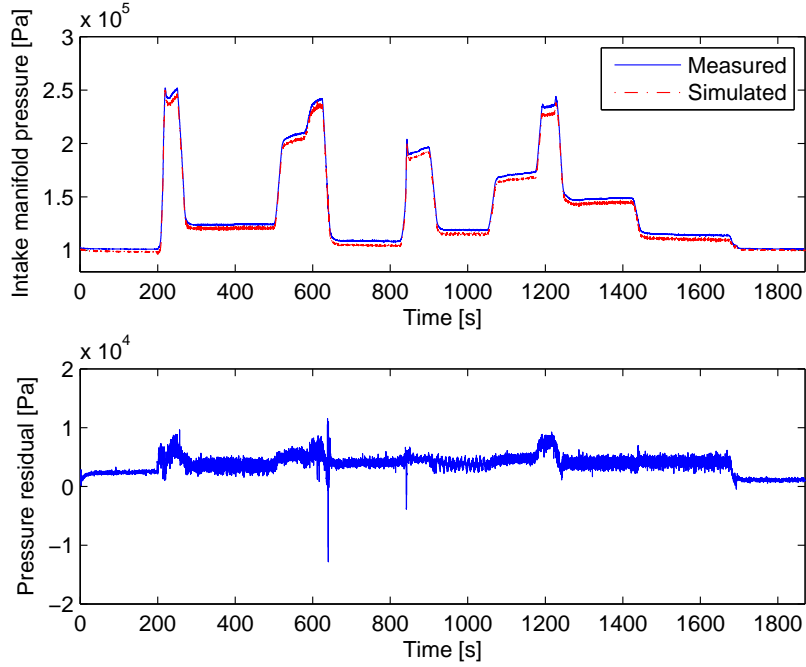
Figure 4.5. Validation of p_{im}. The model can only be evaluated in the stationary points because of the assumptions made in the calculations.

4.6 Results

The results of the model validation are shown in Table 4.1. The presented values are mean values during the non idling phases of the driving cycle. In the figures earlier in this chapter, it corresponds to the time interval 200 seconds to 1,700 seconds. The main reason for not including the idling modes in the evaluation is that the model does not agree very well with the reality in this operating point, and therefore it is not desirable to run the diagnosis test in this point. The reason why the model is less accurate in idling modes is not obvious, but no significant diagnosis performance will be lost when neglecting idling modes.

The errors are small and only the EGR system tends to give a relative error that is a factor two larger than the others.

Table 4.1. Results from the validations. The values presented are mean values during the WHSC, except the idling modes.

Model	Mean absolute error	Mean relative error [%]
Engine	0.0039 kg/s	3.3
Exhaust system	16 K	2.4
EGR system	0.0025 kg/s	7.2
Intake manifold	0.047 bar	3.4

4.7   Sensitivity Analysis

To achieve a robust diagnosis system, it is desirable to obtain a model with low sensitivity to faults in model parameters. An analysis to investigate this is presented in this chapter. To be able to better form an opinion on what faults a diagnosis system based on this model can detect, an analysis of the model's sensitivity to faults in the input signals, presented in Section 3.2, is also made and presented in this chapter.

For this analysis the model output is compensated for the bias fault discussed in Section 4.5, i.e. the modelled intake manifold pressure, $\hat{p}_{im}$, is adjusted three percent. This is done in order to give a more correct comparison of the effects of faults in the input signals.

4.7.1   Sensitivity Analysis in Respect to Model Parameter Uncertainties

When the model is used in a series of trucks, it is important that it works properly, irrespective of the accuracy of the hardware setup. In this section, the sensitivity of the model is studied in respect to the model parameters. The results are shown in Table 4.2. The considered variations are 20% in each direction, i.e. the parameters have faults representing 20% decreased values and 20% increased values respectively. Then the absolute and relative errors in the intake manifold pressure, p_{im}, are calculated.

Here all the EGR parameters are evaluated at the same time by only adjusting A_{egrmax}. This is also done in the volumetric efficiency model where the three parameters c_{vol1}, c_{vol2} and c_{vol3} are equally adjusted at the same time. The parameters that affect the model output the most are the η_{vol} parameters and the engine displaced volume, V_d. These parameters equally affect the model, which depends on that these parameters just exist in the same manner in (3.4). The rest of the parameters affect the model only slightly. This is a good property of the model since it increases the robustness against individual variations.

Further, a parameter that is not considered here is the volume of the intake manifold, V_{im}. The reason for this ignorement is that the parameter does not affect the pressure equilibrium in (3.3). A change in V_{im} only contributes to a changed pressure derivative, but the equilibrium will still remain around zero and not give any changes to the pressure in the stationary conditions.

Table 4.2. Results from the sensitivity analysis in respect to the model parameters. The values presented are mean values during the entire cycle, except the idling modes. The +20% column represents a parameter fault that increases the parameter value with 20% and the -20% column represents values decreased with 20%.

Parameter	+20%		-20%	
	Abs. err. [bar]	Rel. err. [%]	Abs. err. [bar]	Rel. err. [%]
$c_{vol1,2,3}$	0.055	3.8	0.067	3.8
A_{egrmax}	0.0081	0.56	0.026	2.0
η_{oc}	0.016	1.2	0.0072	0.51
V_d	0.055	3.8	0.067	3.8
r_c	0.0094	0.71	0.011	0.81
$h_{tot}A$	0.0099	0.74	0.010	0.76

4.7.2 Sensitivity Analysis in Respect to Input Signal Disturbances

A study is carried out by adjusting the different input signals, described in Section 3.2. The input signals are adjusted one by one and the variations are ±20%. The results are illustrated in Table 4.3.

Table 4.3. Results from the sensitivity analysis in respect to the model's inputs. The values presented are mean values during the entire cycle, except the idling modes. The + 20% column represents a sensor fault that increases the measurements with 20% and the - 20% column represents measurements decreased with 20%.

Input	+20%		-20%	
	Abs. err. [bar]	Rel. err. [%]	Abs. err. [bar]	Rel. err. [%]
n_e	0.057	3.9	0.062	3.6
p_{em}	0.25	17.3	0.25	17.7
W_{cmp}	0.043	2.3	0.047	3.1
T_{im}	0.046	2.6	0.060	4.1
T_{amb}	0.010	0.8	0.010	0.7
u_δ	0.012	0.9	0.0082	0.6
u_{egr}	0.010	0.8	0.012	0.9

As can be seen, the model is more sensitive to faults in some of the sensors. The pressure sensor in the exhaust manifold, p_{em}, is the most obvious one when it comes to the model's sensitivity to sensor faults. Faults in the rotational speed, n_e, compressor mass flow, W_{cmp}, and intake manifold temperature, T_{im}, sensors cause deviations of the same magnitude in the intake manifold pressure. The rotational speed sensor, on the other hand, is a reliable sensor and such large faults as ±20% are not to expect. The mass flow sensor in the compressor is sensitive to where it is placed. Small differences in position can result in large variations in the measured mass flow through the compressor. However the model appears not to be that sensitive to this mass flow measurement. Also faults in the sensor for the

intake manifold temperature, T_{im}, give detectible deviations in the intake manifold pressure. The last three input signals in Table 4.3 give very small deviations in the modelled pressure.

4.7.3   Conclusions

As discussed in Section 3.7 the engine considered is new and it is therefore hard to obtain proper data for a least square optimisation of the model parameters. Instead model parameters from similar engine models are used and tuned ad hoc. This method turns out to work well as shown in the validation result in Section 4.6. Table 4.1 shows that the relative errors are small.

The sensitivity analysis of the model parameter uncertainties shows that the parameter errors has to be large to cause the estimated pressure to deviate from the measured pressure sufficiently for a fault detection, i.e. false alarm. It is reasonable to alarm for a deviation in the estimated pressure around 20 % to minimise the probability for false alarm.

Finally the input signal disturbances sensitivity analysis shows that the diagnosis system will be more sensitive to faults in n_e, p_{em}, W_{cmp} and T_{im} than to faults in T_{amb}, u_δ and u_{egr}.

Chapter 5

Observer Design

The next step in the design of the diagnosis system is to construct a diagnostic observer for residual generation [12]. This observer is based on the engine model described in Chapter 3. To get an observer that is robust against model errors, noise, and individual variations, a common approach is to utilise a state-feedback. Such an observer also gives better convergence and stability properties. In this case there only exists one state, p_{im}, which also is measured. The difference between the measured and estimated signal, $p_{im} - \hat{p}_{im}$, is used as a feedback signal.

In this chapter, the construction of such an observer, based on the physical mean value model from Chapter 3, is described. First the model is modified to simplify the further work. Then different methods to compute the feedback gain is investigated. Three observer approaches are considered, Extended Kalman Filter (EKF) [8], because it is a commonly used observer, high-gain [2], since it is easy to understand and to implement, and sliding mode [15], because it is a commonly used observer design for diagnosis applications.

5.1   Conversion of the Model to a State-Space System

The modelling in Chapter 3 results in a non-linear, semi-explicit Differential-Algebraic Equation (DAE), i.e. a system consisting of both differential and algebraic equations. However, the considered observer design methods require a state-space system, i.e. in this case one explicit differential equation, and one measurement equation, see [8], [2], [15]. Hence, the DAE is first transformed into a state-space system. The state-space system is then discretised to fit the chosen EKF design, see Section 5.2.1.

5.1.1 Transforming the DAE into an State-Space System

After the modelling work performed in Chapter 3, the system of equations, see Appendix C, can be written in the semi-explicit form

$$\dot{x}_1 = f(x_1, x_2, z), \tag{5.1a}$$

$$0 = g(x_1, x_2, z), \tag{5.1b}$$

where x_1 is the only state considered (p_{im}), x_2 is all unknown variables ($W_{eng_{in}}$, W_f, $W_{eng_{out}}$, T_{cyl}, q_{in}, x_r, T_1, T_{em}, W_{egr}) and z is the known signals described in Section 3.2. The function g consists of the algebraic equations (3.4)-(3.19), with the assumption that η_{oc} in (3.12) is constant. Also the measurement of p_{im},

$$y = x_1 = p_{im}, \tag{5.2}$$

is included in the function g. This means that f consists of equation (3.3), while g is representing one equation for each unknown variable and one equation for the measurement (5.2), which makes it an overdetermined system. Also (5.2) gives that this one-state system is observable according to [3].

The system (5.1) is a semi-explicit DAE and to rewrite it to a state-space system, it is desirable to solve (5.1b) explicitly for the unknown variables, x_2. The difficulties with finding that solution is to solve the algebraic loop in the modelling of the exhaust temperature, i.e. (3.8)-(3.11). One possibility is to use a fixed point iteration. One drawback of that approach is that the computational effort then becomes large. Yet one drawback is that the model then can not be rewritten into a state space system as desired. Because of this the fixed point iteration is dismissed.

Instead MATLAB's Symbolic Toolbox is used to analytically solve (5.1b) and an explicit solution is obtained. In this way the unknown variables can be expressed in terms of the state x_1 and the known variables z. As mentioned above, the problem with getting explicit expressions for x_2 occurs in the expression for T_{em}, i.e. (3.8)-(3.13). The Symbolic Toolbox solves this system of equations and gives an explicit expression for T_{em}.

As a result, the system of equations can then be written as

$$\dot{x}_1 = f(x_1, x_2, z), \tag{5.3a}$$

$$x_2 = g_1(x_1, z), \tag{5.3b}$$

$$0 = g_2(x_1, z) = y - x_1, \tag{5.3c}$$

where g_1 contains the equations describing the unknown variables and g_2 is the measurement equation (5.2). By substituting (5.3b) into (5.3a), the system (5.3) can be rewritten to

$$\dot{x}_1 = \tilde{f}(x_1, z), \tag{5.4a}$$

$$0 = g_2(x_1) = y - x_1. \tag{5.4b}$$

For the case in this thesis work the new expression for T_{em}, together with (3.4)-(3.7) and (3.14)-(3.19), are inserted into (5.3a), while the measurement equation (5.3c), is left unchanged since it already only contains x_1.

The model is now expressed as a state-space system according to (5.4).

5.1.2 Discretising the State-Space Model

For the further work on observer design, diagnosis test design and implementation, the model is discretised. This is done because one of the observers to be evaluated is a discrete EKF, which needs a discretised model. To be consequent, the discretisation is used for the approximation of the pressure time derivative even in the other two observer designs. For the discretisation the Euler method is used [11]. The forward Euler method together with the expression for the pressure time derivative from the state-space system (5.4a), yields

$$x_1(t + T_s) = x_1(t) + T_s \tilde{f}(x_1(t), z(t)), \tag{5.5}$$

where T_s is the sampling time. The forward Euler method is used because it is an explicit method easy to implement, it demands low computational effort and its performance is assessed to be good enough for the purpose of this thesis, since the difference in the behaviour between the forward Euler method and a more complicated implicit method, is negligible when using the sampling time $T_s = 0.01$ seconds. Further evaluations of other discretisation methods are therefore not done.

Today there are two kinds of repeatedly running loops in the Scania ECU, 100 Hz and 20 Hz. The 100 Hz loop is chosen for the continued work on the thesis to assure the best possible simulation results. Though in Section 6.5.4 a comparison between the 100 Hz and 20 Hz loops is done to assure optimisation of the OBD computation exploitation.

5.1.3 Behaviour of the Discretised State-Space Model

To evaluate the impact on the model, caused by the modifications in Sections 5.1.1 and 5.1.2, a simulation with the modified model is done. This simulation result is

seen as the nominal behaviour of the state-space model, and used for evaluation of
the different observer design methods in Section 5.2. The evaluations of the open
model and the observers are made on the same WHSC as in Chapter 4, but with
a fault on the p_{im} sensor of 10 kPa (ca 5-10 % fault) after 600 seconds and a fault
of 20 kPa (ca 10-20 % fault) after 1,200 seconds. The reason to add these faults
is to compare and evaluate how the different observers compensate for faults of
different magnitude.

In Figure 5.1, the measured pressure is shown together with the simulated
pressure from the open model with the Euler approximation. Also the residual be-
tween the measured and modelled pressure is shown in Figure 5.1. The behaviour
of the simulated pressure is not affected by the transformation from a DAE to a
state-space system, but the Euler discretisation causes, together with the sampling
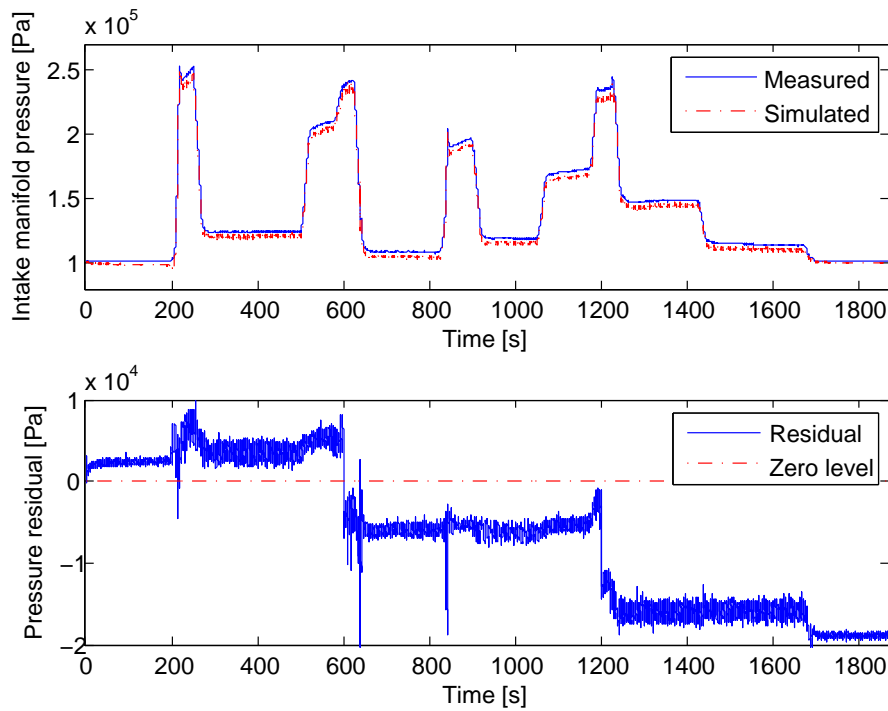time $T_s = 0.01s$, the variance to increase slightly.

Figure 5.1. The upper figure shows $\hat{p}_{im}$ from the open model with the Euler derivative
approximation together with the measured p_{im}. The lower figure shows the residual
between them. A fault on the p_{im} sensor of 10 kPa is applied after 600 seconds and a
fault of 20 kPa is applied after 1,200 seconds.

5.2 Different Design Methods for the Observer

To get a robust observer for the residual generation, the open model is complemented with a feedback. Another important property of this observer, is to improve the diagnosis performance. For example, a good detection ability and robustness against individual variations are desirable properties. There are many possible choices for the design of observer feedbacks. On the model in this thesis work, the use of three different feedback designs are evaluated. Extended Kalman Filter (EKF) [8] is one of them, because it is a very commonly used observer. Next comes the high-gain method [2], which is based on an easily understood theory and is simple to implement. The third method is the sliding mode [15], a commonly used design for diagnostic observers [9], [17].

5.2.1 Extended Kalman Filter (EKF)

The EKF is, as the name extended Kalman filter indicates, an extension of the regular Kalman filter, to be applicable also on nonlinear systems [8]. The methodology used here [5], is to linearise the model in every sampling point and then implement a Kalman filter for every linearisation.

A discretised model is considered and written as

$$x_{t|t-1} = f(x_{t-1|t-1}, w_t), \tag{5.6a}$$
$$y_t = h(x_{t|t}) + e_t, \tag{5.6b}$$

where the measurement error, e_t, and the model error, w_t, are additive white noises.

The EKF is in this work applied according to the algorithm described in [5]:

1. Initiate the filter with the initial information:

   $\hat{x}_{0|-1} = x_0$ and $P_{0|-1} = \Pi_0$,

   where x_0 is the initial state estimate and Π_0 is the covariance of x_0. Let $t = 0$.

2. Measurement update phase:

   $\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - h(\hat{x}_{t|t-1}))$,
   $P_{t|t} = (I - K_t H_t)P_{t|t-1}$
   $, K_t = P_{t|t-1}H_t^T(H_t P_{t|t-1} H_t^T + R_t)^{-1}$,

   where

   $H_t = (\nabla_x h(x)|_{x=x_{t|t-1}})$.

3. Time update phase:

$$\hat{x}_{t+1|t} = f(\hat{x}_{t|t}, 0),$$
$$P_{t+1|t} = F_t P_{t|t} F_t^T + G_t Q_t G_t^T,$$

where

$$F_t := (\nabla_x f(x, 0)|_{x=x_{t|t}}) \text{ and } G_t := (\nabla_w f(\hat{x}_{t|t}, w)|_{w=0}).$$

4. Let $t := t + 1$ and repeat from 2.

To implement the EKF, some assumptions are made and some parameters are decided. The only feedback signal is described in (5.2), which gives $H_t = 1$. Since the model only has one state, F_t is also scalar. An analytical expression for F_t can easily be found by deriving the Right Hand Side (RHS) in (5.5) with respect to the state, $x_1(t)$, using MATLAB's Symbolic Toolbox.

As mentioned above, Kalman filters are based on the assumption that the measurement and model errors are additive white noises. From this follows that Q_t and R_t are the variances of the model noise and the measurement noise respectively. It is hard to verify that these conditions, demanded for the EKF to work properly, are fulfilled and to find the correct G_t, Q_t and R_t. Therefore the observer's performance is evaluated for different approaches when it comes to finding and estimating G_t, Q_t and R_t.

One intuitive approach is to approximate R_t by finding the variance of a stationary measurement sequence of p_{im} using MATLAB. One way to decide G_t, in analogy to the calculation of F_t, is to derive the RHS in (5.5) in respect to each input signal respectively, i.e. the model noise is assumed to appear as the measurement noise in the model input signals. Then Q_t is set to a diagonal matrix with the approximated variance of each input signal respectively.

Another approach is to see G_t, Q_t and R_t as design parameters. A simpler implementation of the EKF is to set $G_t = 1$ and then have only one design parameter, $\frac{R_t}{Q_t}$.

In Figure 5.2 a simulation of the EKF based observer is plotted. It is obvious that the variance of $\hat{p}_{im}$ is lower in the EKF compared with in the open model in Figure 5.1, but the signal still has an offset.

A drawback of an EKF based observer is its high demand on calculation power. This demand can possibly be lowered by implementing some kind of offline solution, for example by mapping the gain in different operating points, or by just setting a constant gain. Depending on how G_t, R_t and Q_t are chosen, this offline solution could be reasonable. But then a big part of the reason for using an EKF is also lost. Figure 5.3 shows one example of how the EKF feedback gain varies.

Further, the only possible feedback signal for the observer is the residual, $p_{im} - \hat{p}_{im}$. A drawback of using this residual for the feedback is that the test quantity is based on the same signal. It is therefore desirable to minimise the observers impact on the residual and at the same time achieve the observer's robustness improving properties. This trade-off is not satisfied using the EKF observer. Figure 5.2 illustrates that, even though it is possible to see steps in the residual at $t = 600$ seconds and $t = 1,200$ seconds, it is still close to zero, i.e. the proportion between

the residual in the faulty case and in the fault free case are of the same magnitude using the EKF as when using the open model.
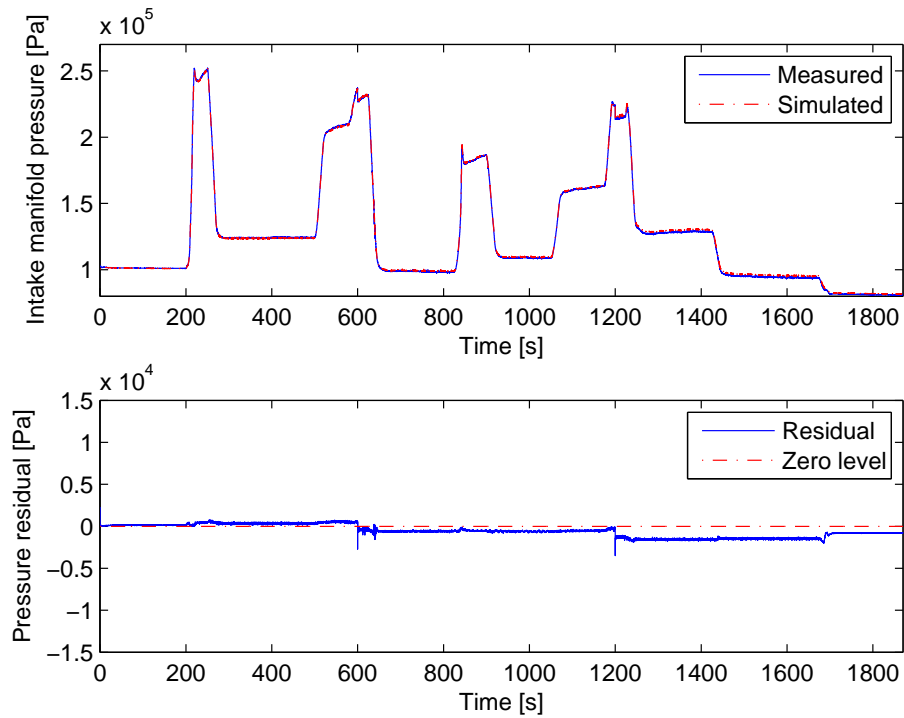

Figure 5.2. This figure shows the behaviour of the EKF observer, with R_t and Q_t chosen as the variance of each input signal respectively and G_t is found via derivation using MATLAB's Symbolic Toolbox. A fault on the p_{im} sensor of 10 kPa is applied after 600 seconds and a fault of 20 kPa is applied after 1,200 seconds.
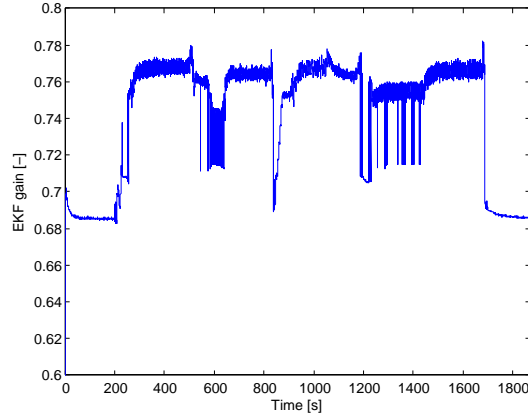
Figure 5.3. This figure shows how the EKF observer's gain varies over time. In this implementation of the EKF, R_t and Q_t are chosen as the variance of each input signal respectively and G_t is found via derivation using MATLAB's Symbolic Toolbox.

5.2.2 High-Gain Observer

The high-gain technique [2] is an often useful and robust design method for observers for nonlinear systems. Here follows a description of how the high-gain technique is applied to the model in this work [10].

The non-discretised model can be written as

$$\dot{x}_1 = \tilde{f}(x_1, z), \tag{5.7a}$$

$$y = x_1, \tag{5.7b}$$

according to (5.4). The observer is achieved by implementing a feedback from the output signal, y,

$$\dot{\hat{x}}_1 = \tilde{f}(\hat{x}_1, z) + h(y - \hat{x}_1), \tag{5.8}$$

where $\tilde{f}$ is the nominal model of the true function $\tilde{f}_{true}$. The observer gives the estimation error

$$\tilde{x}_1 = x_1 - \hat{x}_1, \tag{5.9}$$

which satisfies

$$\dot{\tilde{x}}_1 = -h\tilde{x}_1 + \delta(x_1, \hat{x}_1), \tag{5.10}$$

where $\delta(x_1, \tilde{x}_1) = \tilde{f}_{true}(x_1, z) - \tilde{f}(\hat{x}_1, z)$. To achieve asymptotic error convergence,

$\lim_{t \to \infty} \tilde{x}_1 = 0$, in absence of the disturbance δ, h must be a positive integer. When δ is present, an additional goal in the design of the observer gain is to reject the effect of the δ term. The transfer function from δ to $\tilde{x}_1$,

$$H_0(s) = \frac{1}{h + s},\tag{5.11}$$

should ideally then be identically zero. This is not possible, but by choosing $h >> 0$, H_0 gets arbitrarily close to zero. This means that the gain (the observers only design parameter), h, can be chosen so that the residual gets arbitrarily small.

For the actual implementation of the high-gain observer it is discretised using the Euler method described in Section 5.1.2.

In Figure 5.4, the behaviour of the intake manifold pressure using the high-gain observer is illustrated. It appears that the result of this observer is similar to the EKF implementation - the variance can be reduced but the offset stays. To not reduce the test quantity to much, the gain must be set not very high but to a magnitude between 1 and 10.

Advantages of this method, compared to the EKF observer, is that it is easy to implement and has a lower demand on computation power. There are not as many computations needed in the high-gain observer as there are in the EKF algorithm in Section 5.2.1.

As in the EKF case, the high-gain observer has a strong influence on the residual, which decreases the diagnosis system's detection ability. Also in this case is the proportion between residual in the faulty case and in the fault free case of the same magnitude as in the open model. This means that also the high-gain method does not improve the diagnosis system's properties.
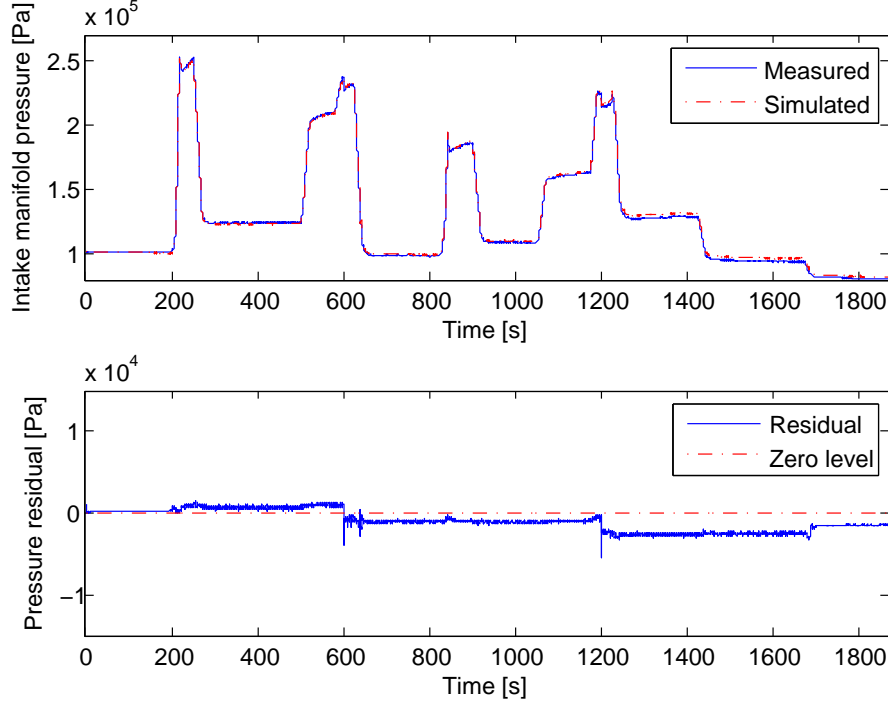
Figure 5.4. In this figure the behaviour of the high-gain observer is shown. The feedback gain, h, is here set to 5. A fault on the p_{im} sensor of 10 kPa is applied after 600 seconds and a fault of 20 kPa is applied after 1,200 seconds.

5.2.3   Sliding Mode Observer

Sliding mode is an observer design method [15] useful for both linear and nonlinear models. Properties of this method are that it gives a finite time convergence for all observable states, its implementation is fairly intuitive and simple, and it is robust against parameter uncertainties [7].

Based on the continuous system given in (5.7), and the nominal model function $\tilde{f}$, the sliding mode observer is designed according to [7],

$$\dot{\hat{x}}_1 = \tilde{f}(\hat{x}_1, z) + \lambda \mathrm{sgn}(x_1 - \hat{x}_1), \tag{5.12a}$$

$$\hat{y} = \hat{x}_1, \tag{5.12b}$$

where λ is the feedback gain, and sgn is the sign function. Further, this gives that the dynamics of the residual, $r = x_1 - \hat{x}_1$, is

$$\dot{r} = \tilde{f}_{true}(x_1, z) - \tilde{f}(\hat{x}_1, z) - \lambda \text{sgn}(x_1 - \hat{x}_1), \quad (5.13a)$$

$$\hat{y} = \hat{x}_1. \quad (5.13b)$$

With Lyapunov stability theory it can be proven that the residual will converge to zero in finite time, and through this also that the observer is stable.

The sign function in the sliding mode method implies that the correcting term on the derivative has a limited magnitude, no matter the size of the residual. Because of this, this observer has the advantage of only making small adjustments on the simulated quantity, $\hat{x}_1$, and thereby only has limited influence on it. But the sign function gives a discontinuity which cause the residual oscillate when it is close to zero.

To improve the sliding mode observer and get rid of the discontinuity, the feedback signal is smoothened by replaces the sign function with a hyperbolic tangent function. This means that the estimated pressure, $\hat{p}_{im}$, becomes less oscillating when hyperbolic tangent function is used instead of the sign function, which reduces the variance of the residual.
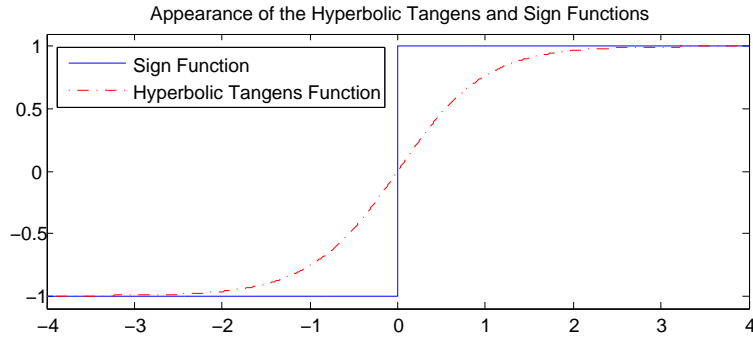

Figure 5.5. The behaviour of the sign function and the hyperbolic tangent function around zero.

Now the new sliding mode observer can be expressed as

$$\dot{\hat{x}}_1 = \tilde{f}(\hat{x}_1, z) + \lambda \tanh(\frac{x_1 - \hat{x}_1}{c}), \quad (5.14a)$$

$$\hat{y} = \hat{x}_1, \quad (5.14b)$$

where c (~ 1000) is a constant used to improve the smoothening of the residual.

The effect of using the observer feedback with the hyperbolic tangent function instead of the sign function is shown in Figure 5.6. It can be seen that the difference

in magnitude is not very large, but there is an improvement in the residual variance when using the hyperbolic tangent feedback. The hyperbolic tangent feedback pulls the residual closer towards zero, but the residual never changes sign compared to the open model residual in Figure 5.1.
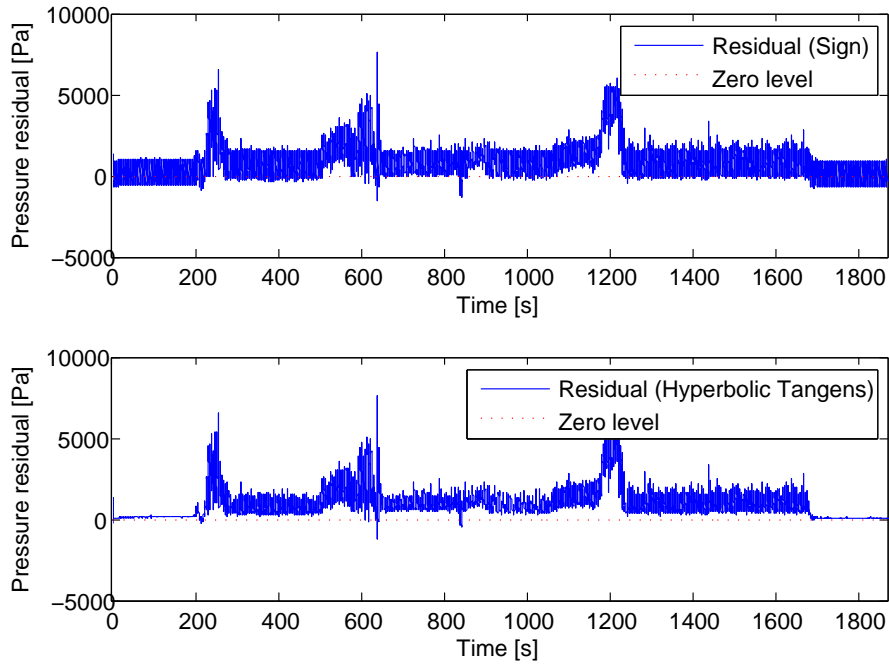

Figure 5.6. This figure shows the difference, for the residual, between using the sign feedback and the hyperbolic tangent feedback. This is a fault free case, i.e. the residual is close to zero and its sign is changing frequently

Figure 5.7 shows the appearance of $\hat{p}_{im}$ using the sliding mode observer according to (5.14). The faults are now seen much clearer in the residual than in the EKF and high-gain cases. The reason for this is that the sliding mode observer feedback is mostly based on the sign of the residual, not on the magnitude of the residual. This is a benefit that often is used in diagnosis contexts, the feedback size does not depend on the residual magnitude. This causes the proportion between the residual in faulty case and in fault free case to increase using the sliding mode method compared to using the open model, the EKF observer or the high-gain observer.
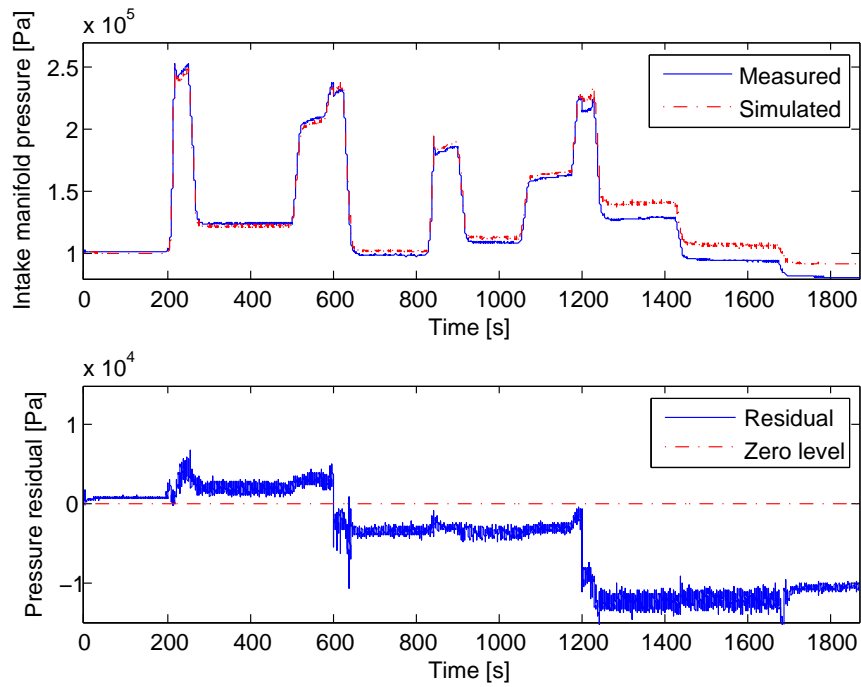
Figure 5.7. In this figure the behaviour of the sliding mode observer, using a hyperbolic tangent function for the feedback, is shown. The feedback gain, λ, is here set to 80,000. A fault on the p_{im} sensor of 10 kPa is applied after 600 seconds and a fault of 20 kPa is applied after 1,200 seconds.

Chapter 6

Construction and Evaluation of the Diagnosis System

In this chapter two different designs for the diagnosis system are described and evaluated. One uses the sliding mode observer, described in Section 5.2.3, as residual generator and the other one uses the open model for residual generation. Also the open model can be seen as an observer, only without the feedback, and it is from now on denoted the open model observer.

In order to make reliable decisions, the residuals from the two approaches need to be post filtered. The diagnosis systems therefore, beside the respective observer and residual generator, also consists of a diagnosis test. This test consists of two CUSUM algorithms, see Section 6.2, one for under-boost testing and one for over-boost testing.

6.1   Choice of Observer

To decide which observer to use in the design of the final diagnosis test, the different alternatives and their characteristics must be compared and evaluated.

By an empirical evaluation of the stability of the open model observer, it is shown that it is stable in itself for different initial values, $\hat{p}_{im}(t_0) > 0$, and that its settling time is short. This means that no feedback is needed to ensure stability in this case. Also the accuracy of the open model is assessed to be good. Therefore, the open model observer is an attractive candidate for the final diagnosis test design.

An investigation of the three different feedback observers shows that the EKF and high-gain observers have almost equivalent properties. The only things they contribute with is to lower the variance and the mean of the residual. These features could just as well be implemented as a part of the diagnosis test, with a low-pass filter or adaptive thresholds. The sliding mode observer, on the other hand, has other more appropriate diagnosis properties. The hyperbolic tangent function used there, implies that the correcting term on the derivative has a limited

magnitude, no matter the magnitude of the residual. Because of this, this observer only makes smaller adjustments on the estimated quantity, $\hat{p}_{im}$, and thereby the effect on the residual is limited. The sliding mode methodology therefore increases the robustness against model parameter uncertainties and individual variations between engines.

Further, the only possible feedback signal for the observers is the residual on which the actual diagnosis test is based. This fact leads to that the feedback observers compensate for an increased residual magnitude, which affects the diagnosis properties. The sliding mode observer is the method that affects the residual the least because of its saturated feedback. The properties of the sliding mode observer makes it the most suitable feedback observer for diagnosis, of the considered ones.

The conclusion of this evaluation is that the most interesting observers to base the diagnosis test on are the sliding mode observer and the open model observer.

6.2   The CUSUM Algorithm

After generating a residual it is often necessary to apply some kind of post filtering to the signal in order to be able to make a reliable decision based on it. A possible approach here is to use a low-pass filter applied to the residual. A frequently used algorithm when it comes to diagnosis is the CUmulative SUM (CUSUM) algorithm [13]. Both the sliding mode diagnostic observer and the open model diagnostic observer are complemented with this CUSUM algorithm. Here follows a short description of the CUSUM algorithm, based on [4] and [12]. For a complete analytical derivation of this algorithm, see [12].

The test quantity, $T[k]$, is based on the cumulative sum, $g[k]$, and is updated like

$$g[k] = g[k-1] + s[k], \tag{6.1a}$$

$$T[k] = g[k] - \min_{0 \leq i < k} (g[i]), \tag{6.1b}$$

where the score function, $s[k]$, has the following properties

$$E\{s[k]\} < 0, \text{ fault free}, \tag{6.2a}$$

$$E\{s[k]\} > 0, \text{ faulty}. \tag{6.2b}$$

When a fault occurs, $s[k]$ changes sign in the mean. The behaviour of $g[k]$ is then a negative drift, in the mean, in the fault free case and a positive drift in the faulty case.

Since there are two fault symptoms that should be detected in this case, there are also two CUSUM tests, one for under-boost and one for over-boost. The score function for the over-boost test is

$$s_{ob}[k] = r[k] - \nu_{ob}, \tag{6.3}$$

where $r[k] = p_{im}[k] - \hat{p}_{im}[k]$ and ν_{ob} is a positive drift parameter to decide how large the residual needs to be for the test quantity to increase. A rule of thumb is that ν_{ob} shall be chosen in the same order of magnitude as the residual in the fault free case [12].

Now the final over-boost test can be written as

$$T'_{ob}[k] = \max[0, T'_{ob}[k-1] + r[k] - \nu_{ob}], \tag{6.4}$$

where

$$T'_{ob}[0] = 0. \tag{6.5}$$

The under-boost test is analogically derived. Since the residual is negative in this case, the residual in the score function has to be subtracted instead of added. The score function for the under-boost test is then

$$s_{ub}[k] = -r[k] - \nu_{ub}, \tag{6.6}$$

and the final under-boost test can be written as

$$T'_{ub}[k] = \max[0, T'_{ub}[k-1] - r[k] - \nu_{ub}], \tag{6.7}$$

where

$$T'_{ub}[0] = 0. \tag{6.8}$$

The test's equations (6.4) and (6.7) are alarming when the test quantities for under- and over-boost falls bellow, or exceeds respectively, a threshold. A method to increase the robustness and decrease the probability of false alarm, is to require several threshold exceedances before the test alarms.

6.3   The Sliding Mode Based Diagnosis System

The observer used for the sliding mode based diagnosis system is constructed as described in Section 5.2.3, using the state-space structure and the discretisation method described in Section 5.1. To complete the diagnosis system, two CUSUM

algorithms are added - one for detection of under-boost and one for detection of over-boost.

The complete sliding mode based system, including the diagnostic observer, the residual generator, and the CUSUM tests, can be formulated according to

$$\hat{p}_{im}[k] = \hat{p}_{im}[k-1] + T_s(\tilde{f}(\hat{p}_{im}[k-1], z[k-1]) + \lambda \tanh(r[k-1])), \quad (6.9a)$$
$$r[k] = p_{im}[k] - \hat{p}_{im}[k], \quad (6.9b)$$
$$T'_{ub}[k] = \max[0, T'_{ub}[k-1] - r[k] - \nu_{ub}], \quad (6.9c)$$
$$T'_{ob}[k] = \max[0, T'_{ob}[k-1] + r[k] - \nu_{ob}], \quad (6.9d)$$
$$T'_{ub}[0] = 0, \quad (6.9e)$$
$$T'_{ob}[0] = 0, \quad (6.9f)$$

where T_s is the sampling time, $\tilde{f}$ is the model of the boost pressure time derivative, and z is a vector containing all seven of the model's input signals. The alarming part of the CUSUM test is designed as follows

if $T'_{ub}[k] > J_{ub}$ **then**
    $ALARM_{ub} = 1$
else
    $ALARM_{ub} = 0$
end if
if $T'_{ob}[k] > J_{ob}$ **then**
    $ALARM_{ob} = 1$
else
    $ALARM_{ob} = 0$
end if

Here $ALARM = 1$ means that an under-boost or an over-boost is detected, and $ALARM = 0$ that no abnormality in the boost pressure has been detected by the diagnosis test.

6.4   The Open Model Based Diagnosis System

In the implementation where the open model is the basis of the diagnostic observer, some supplements are first added on the observer to improve its properties when it comes to robustness against model parameter uncertainties.

In simulations of the open model an offset can be seen, Figure 4.5. This offset can be described as an almost constant share of the estimated pressure, and therefore be diminished trough

$$\hat{p}_{im}^1 = \hat{p}_{im}^0 \theta, \quad (6.10)$$

where $\hat{p}_{im}^0$ is the result of the open model simulation, θ is an adjusting term

($\sim 0.9 - 1.1$) and $\hat{p}_{im}^1$ is the adjusted estimated pressure. This adaption can also be used as basis for the sliding mode observer. The reason for not evaluating this approach in this thesis is that it is desirable to obtain a diagnosis system as simple as possible.

To further improve the performance of the observer, and reduce the offset even more, different θs can be applied for different operating points. A commonly used definition of operating point by Scania is engine speed, n_e, and engine load, L. But to minimise the number of input signals, the operating points are here described by engine speed, n_e, and injected amount of fuel, u_δ. The load, L, and injected amount of fuel, u_δ, do also correlate well with each other.

A least square calibration and mapping of θ against the operating points is individual for every truck, and could be done in connection with all other calibrations and settings by the construction of the truck.

The adjusting term, θ, can be used to handle ageing trough a low frequent recalibration and updating online. By keeping the first calibration, θ^0, as a norm and compare it with all updated terms, θ^l, a new test quantity could be achieved. This means that with this model based diagnostic observer, four tests could run instead of two. CUSUM is used, as in the sliding mode observer, to detect under-boost and over-boost.

The final structure of the open model based system, including the diagnostic observer, the residual generator, and the CUSUM tests, looks like

$$\hat{p}_{im}^0[k] = \hat{p}_{im}^0[k-1] + T_s(\tilde{f}(\hat{p}_{im}^0[k-1], z[k-1]), \tag{6.11a}$$

$$\hat{p}_{im}^1[k] = \hat{p}_{im}^0[k]\theta^l(n_e[k], u_\delta[k]), \tag{6.11b}$$

$$r_1[k] = p_{im}[k] - \hat{p}_{im}^1[k], \tag{6.11c}$$

$$r_2[k] = \theta^l - \theta^0, \tag{6.11d}$$

$$T_{ub}'[k] = \max[0, T_{ub}'[k-1] - r_1[k] - \nu_{ub}], \tag{6.11e}$$

$$T_{ob}'[k] = \max[0, T_{ob}'[k-1] + r_1[k] - \nu_{ob}], \tag{6.11f}$$

$$T_{ub}'[0] = 0, \tag{6.11g}$$

$$T_{ob}'[0] = 0, \tag{6.11h}$$

where T_s is the sampling time, $\tilde{f}$ is the model of the boost pressure time derivative, and z is a vector containing all seven of the model input signals. The alarming parts of the CUSUM test and the ageing test are designed as follows

if $T_{ub}'[k] > J_{ub}$ **then**
    $ALARM_{ub} = 1$
else
    $ALARM_{ub} = 0$
end if
if $T_{ob}'[k] > J_{ob}$ **then**
    $ALARM_{ob} = 1$

else
    $ALARM_{ob} = 0$
end if
if $r_2[k] < J_{ageing,ub}$ **then**
    $ALARM_{ageing,ub} = 1$
else
    $ALARM_{ageing,ub} = 0$
end if
if $r_2[k] > J_{ageing,ob}$ **then**
    $ALARM_{ageing,ob} = 1$
else
    $ALARM_{ageing,ob} = 0$
end if

Here $ALARM = 1$ means that an under-boost or an over-boost is detected, and $ALARM = 0$ that no abnormality in the boost pressure has been detected by the diagnosis test.

6.5 Evaluation and Comparison of the Two Diagnosis Systems

In this section the two alternative diagnosis systems, presented in Sections 6.3 and 6.4, are evaluated. This is done with

- a power function analysis,

- an analysis of the behavior with respect to different faults in the input signals,

- a validation using data from a transient cycle logged on a commercial vehicle,

- an analysis of how different simulation step sizes affect the behaviour of the two diagnosis systems.

In this section, the faults considered are sensor faults on the intake manifold pressure sensor and on the input signals to the open model. This is a relevant approach since the purpose with the work is to investigate the detection capabilities of the diagnosis systems in respect to under-boost and over-boost. A limited amount of measurement data on the considered engine makes it hard to do these evaluations on faulty data, i.e data corresponding to leakages or other non sensor faults.

6.5.1 The Power Function Analysis

One of the methods used to evaluate the diagnostic observers and their respective CUSUM algorithms is the power function, β. A power function shows the probability that a fault of a certain size is detected.

For this power function analysis an additive sensor fault, ϕ, on the intake manifold pressure is considered, i.e. $r = p_{im} + \phi - \hat{p}_{im}$, where $\phi = 0$ corresponds to

the fault free case, $\phi < 0$ corresponds to that an under-boost should be detected, and $\phi > 0$ corresponds to that an over-boost should be detected. The power function is then written as

$$\beta(\phi) = P(T'_{ob}[k] > J_{ob}|\phi) + P(T'[k]_{ub} > J_{ub}|\phi). \tag{6.12}$$

For a more detailed description and an analytical derivation of the power function see [12].

To get a good picture of the power function, it is desirable to have a large amount of data. The amount of data in this work is limited. The statistics for the power functions are based on the WHSC data presented in Section 4. A WHSC consists of twelve different operating points. Each of these twelve operating points is seen as one single test case. Statistics are gathered by running the diagnosis systems on the WHSC with different fault sizes, ϕ, and see, for each operating point, if the test quantities, $T'[k]_{ub}$ or $T'[k]_{ob}$, are increasing or not. For each fault size, the fraction of operating points where one of the test quantities is increasing is the value of the power function. The power functions for the two respective diagnostic observers are shown in Figures 6.1 and 6.2.
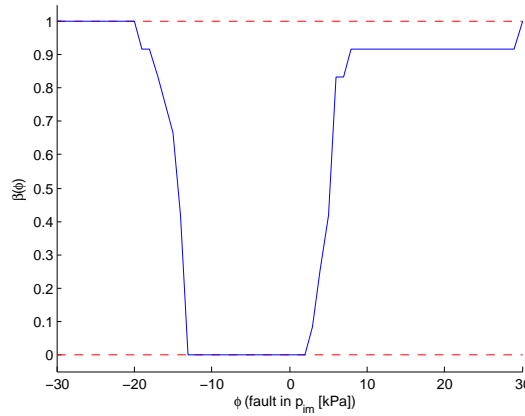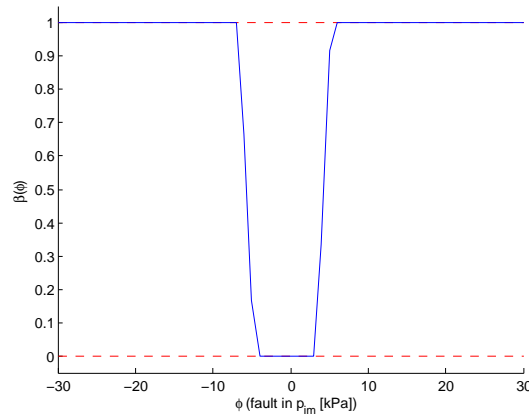

Figure 6.1. The power function for the sliding mode diagnostic observer. In the fault free case, the intake manifold pressure varies between 100 kPa (1 bar) and 350 kPa. For this analysis the drift parameters $\nu_{ub} = \nu_{ob} = 5$ kPa. The late saturation, to the right in the figure, is caused by the inaccuracy of the model in idling mode.

The power function for the sliding mode based diagnosis system, Figure 6.1, is centering around $\phi \approx -6$ kPa instead of $\phi = 0$ kPa. This is caused by the offset of the open model. Since the modelled intake manifold pressure is constantly smaller than the measured intake manifold pressure, it is easier to detect faults that give rise to over-boost, i.e. positive faults on the intake manifold pressure sensor. The sliding mode based diagnosis system performs differently well in different operating

points, because of the varying accuracy of the open model in different operating points.

The power function for the open model based diagnosis system, Figure 6.2, is better centering around $\phi = 0$ kPa than the sliding mode based one. This because of the adjusting parameter $\theta(n_e, u_\delta)$, which is eliminating the offset from the open model when it is well calibrated. Since this diagnostic observer do not have any feedback, it do not compensate for smaller faults like the sliding mode observer do, and the interval where the power function is zero then becomes smaller than in the case with the sliding mode diagnostic observer. Here the operating point dependency does not appear, since the open model based observer is calibrated for each operating point.

Figure 6.2. The power function for the open model diagnostic observer. In the fault free case, the intake manifold pressure varies between 100 kPa (1 bar) and 350 kPa. For this analysis the drift parameters $\nu_{ub} = \nu_{ob} = 5$ kPa.

6.5.2   Faults of Different Characteristics in the Input Signals

In this section the effects on the pressure estimation result caused by faults, of different characteristics, in the input signals are investigated. For all the input signals, faults appear as steps and as ramps of two different lengths. Both positive and negative faults are implemented. These tests are based on the WHSC data, see Section 4. The results for each individual type of fault show to be similar to each other and therefore, just a few of the resulting simulations from this evaluation are presented in Figures 6.3-6.5. The results presented here correspond to faults in T_{im}, p_{em} and n_e. See the figure captions for more information about the type of the fault in each case.

Besides that the pressure estimation has different sensitivity for changes in different input signals, as seen in Table 4.3, it can here also be seen that faults appear differently in different operating points. An evident example of this is
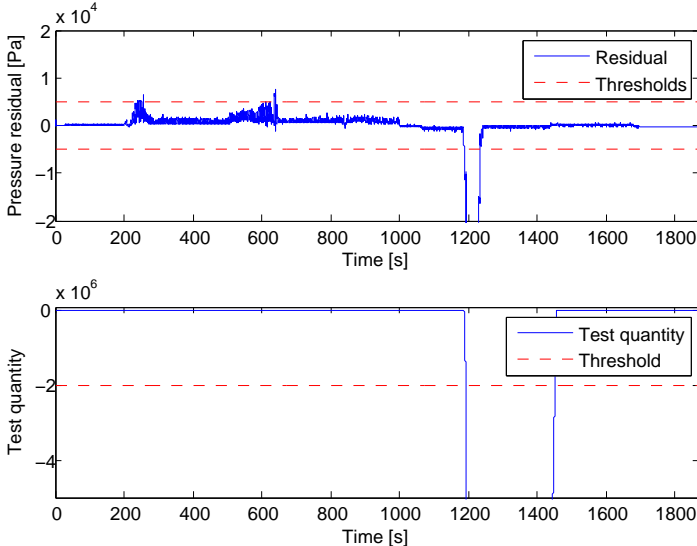
(a) Shows the behavior of the residual of the sliding mode observer for a positive step in T_{im} of 100 K at time 1,000 s.
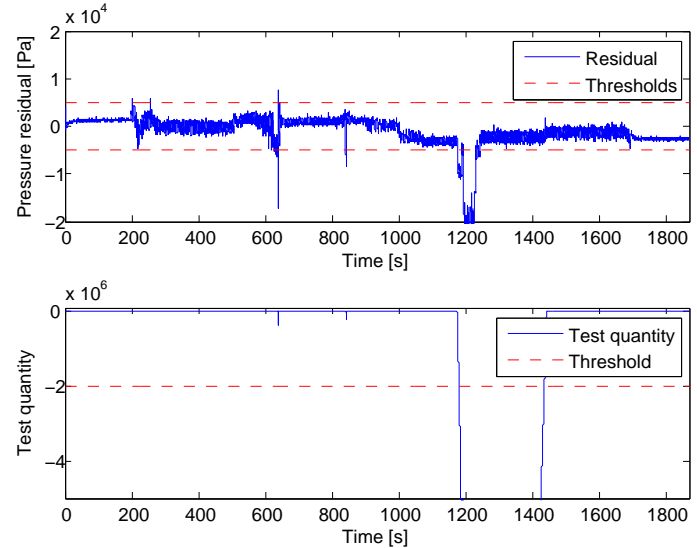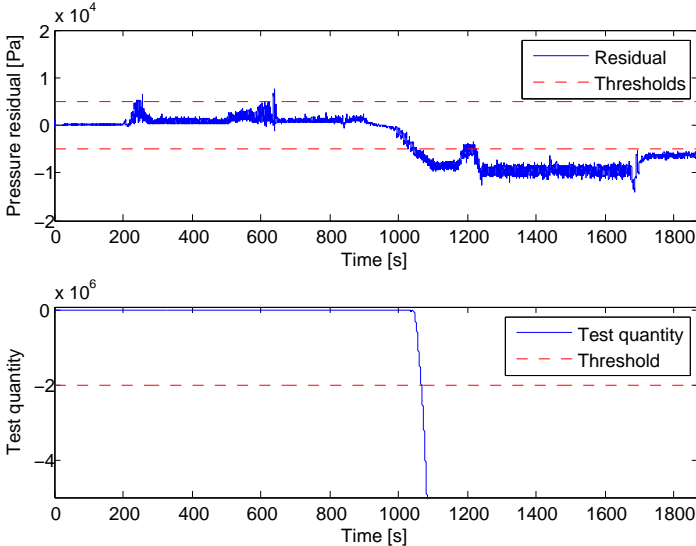

(b) Shows the behavior of the residual of the open model observer for a positive step in T_{im} of 100 K at time 1,000 s.

Figure 6.3. These two figures show how a fault in T_{im} affect the magnitude of the residual depending on operating point. Otherwise the effects of the fault are intuitive - the positive step in T_{im} can be seen in the residual and causes an under-boost alarm.

a fault in T_{im}, see Figure 6.3. The diagnosis test detects the fault in the T_{im} input signal only in one operating point. This is not a great problem since getting no false alarms is assessed more important than missed detections. Here it is a trade-off between how many times or during how long time the test quantity needs to exceed the threshold limit to allow an alarm. On the other hand, with more detailed informations on the different faults, this kind of behaviour can be used in isolation purposes. This is not studied in this thesis.

When it comes to how the shape of the faults affects the residual, $p_{im} - \hat{p}_{im}$, the results are very intuitive - an increasing fault causes an increasing residual and a step shaped fault causes a step in the residual. It is however important to notice that there exists an operating point dependency, as mentioned in the power function analysis in Section 6.5.1. This means that if the test detects a fault in one operating point, it will not necessarily lead to a detection in another operating point.
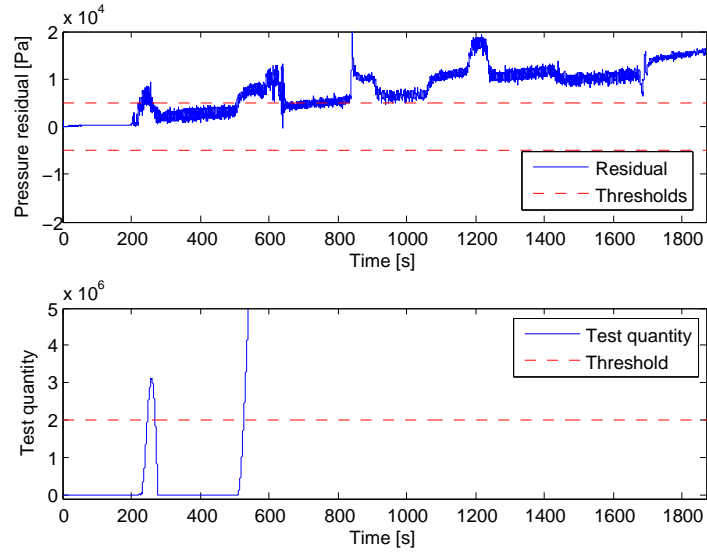
(a) Shows the behavior of the residual of the sliding mode observer for a positive integrating fault in p_{em} of 100 Pa/s between 900 s and 1,100 s.
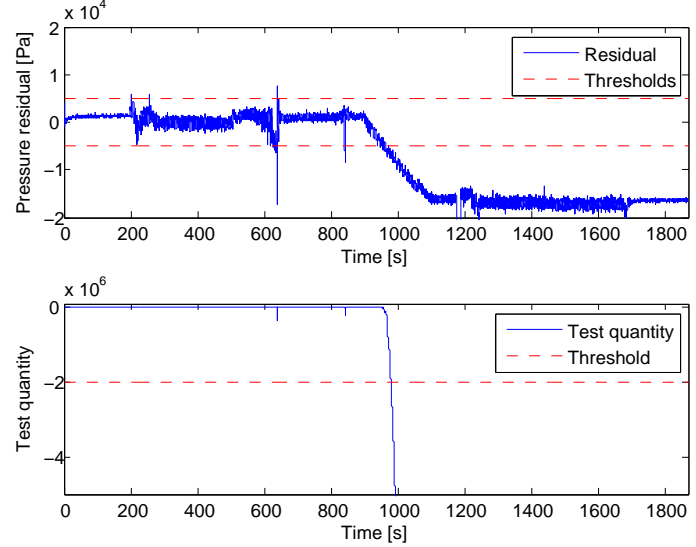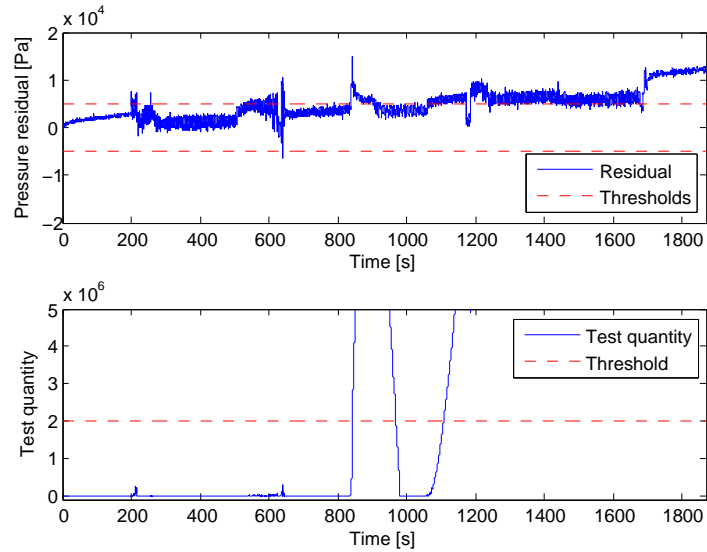

(b) Shows the behavior of the residual of the open model observer for a positive integrating fault in p_{em} of 100 Pa/s between 900 s and 1,100 s.

Figure 6.4. It is seen how the sliding mode observer at first obstruct the increasing residual to a certain level causing it to alarm later than the open model observer.

(a) Shows the behavior of the residual of the sliding mode observer for a positive integrating fault in n_e of 0.4 rpm/s over the whole cycle.

(b) Shows the behavior of the residual of the open model observer for a positive integrating fault in n_e of 0.4 rpm/s over the whole cycle.

Figure 6.5. It is seen how the sliding mode observer at first counteract the increasing residual to a certain level causing it to alarm later than the open model observer. Also here the operating point dependency can be seen as the test quantity returns to zero.

6.5.3 Validating the Diagnosis Systems on Transient Cycle Data

Until now, all the evaluations have been made on the stationary cycle, WHSC. The real driving performance of a truck on the road is not equal to this cycle, due to several obvious factors, like varying air temperature, wind speed and road inclination. Therefore it is of interest to see how these diagnosis systems act on situations that are more similar to ordinary usage.

As have been mentioned before, the considered engine in this thesis is in its developing phase and it is hard to get proper measurement data from it. The engine does not even exist in trucks yet and therefore the data used here is taken from a truck with a similar engine configuration. The data used for the evaluation in this section is from a Scania test truck with a five cylinder and 9.3 liter Euro 5 diesel engine with single step EGR and VGT.

The evaluation of the two different diagnosis systems, the sliding mode and the open model, will be made parallel to each other. Just as in the previous sections, the only faults considered are positive and negative drifts in the intake manifold pressure sensor. The model parameters are chosen to be the same as in the evaluations using the WHSC, described in Section 3.7. In the case with the open model diagnosis systems, the test quantity exceeds its threshold even though there is no fault present, as can be seen in Figure 6.6. This implies that the drift parameter, ν, or the threshold, J, are too low. As indicated in Section 6.2, these parameters are for the test designer to choose. The earlier decided values now seem to be too small and if they are adjusted the result can look like Figure 6.7
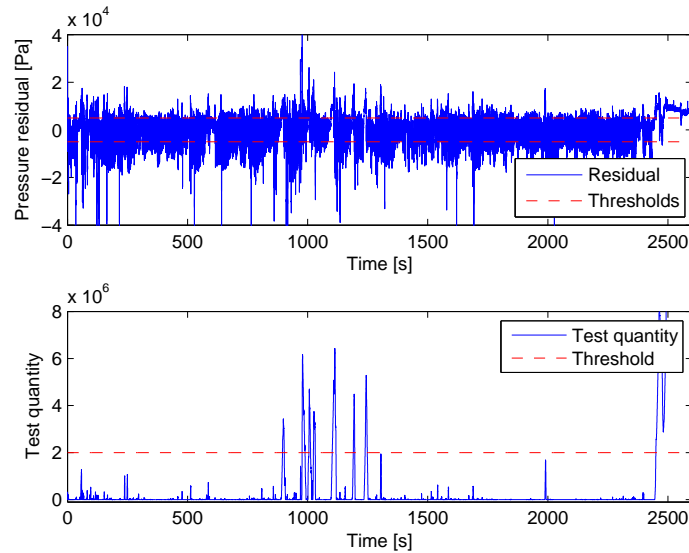

Figure 6.6. The behaviour of the open model diagnostic observer and its CUSUM test in the fault free case. Here the over-boost symptom is considered. It seems like the drift parameter or the test threshold, are to low.

Now the test do not produce any false alarm, even if it is close at the time 1,000 seconds. Such peaks can on the other hand be handled by constraints on how many times or how long time the test quantity needs to exceed the threshold before it is allowed to make an alarm.
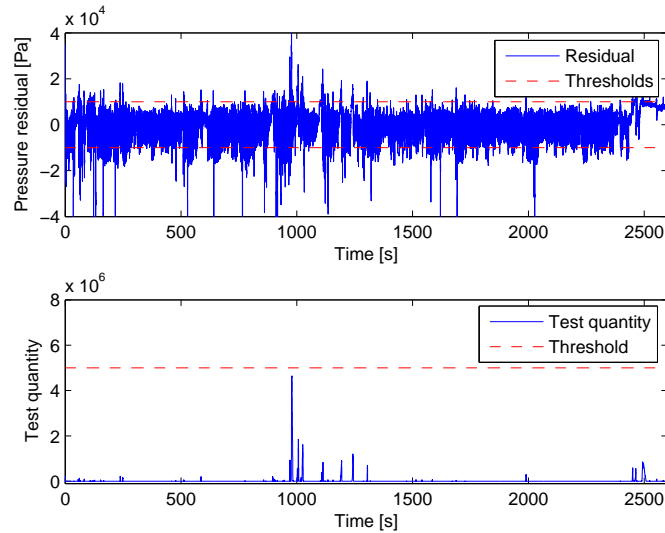

Figure 6.7. Now the drift parameter is set to 10 kPa and the threshold is adjusted to $5 \cdot 10^6$.

There are three types of faults examined as in the previous evaluations, one integrating fault that is building up during the entire cycle and one that just is building up during a shorter time and also a step somewhere during the cycle. In Figure 6.8, the result from the first type is illustrated when there is a positive drift in the pressure sensor. It is clearly shown that the fault is detected after about 1,000 seconds, i.e. when the fault has reached a size of 10 kPa. There are no larger differences in how fast the tests are going to react to the fault. The sliding mode observer test reacts faster, but then its test quantity is decreased below the threshold again. The open model based test exceeds its threshold for definitive before the sliding mode based test. But this is a very short time in the context and do not play any decisive role in the evaluation.

In the case with a negative drift instead, the open model based test is significantly faster than the other one, as can be seen in Figure 6.9. This coincides with the reasoning around the differences in the power function's centering properties, conducted in Section 6.5.1.

The second integrating fault, that is building up during a shorter period, affects the intake manifold pressure similar as the fault that is building up during the entire cycle, assumed that the fault reaches a value that is large enough to be detected. Also here it is a difference in reaction time of the tests when there is a negative drift present in the pressure sensor. Of course it is a smaller difference this time since the drift is made with a steeper slope.
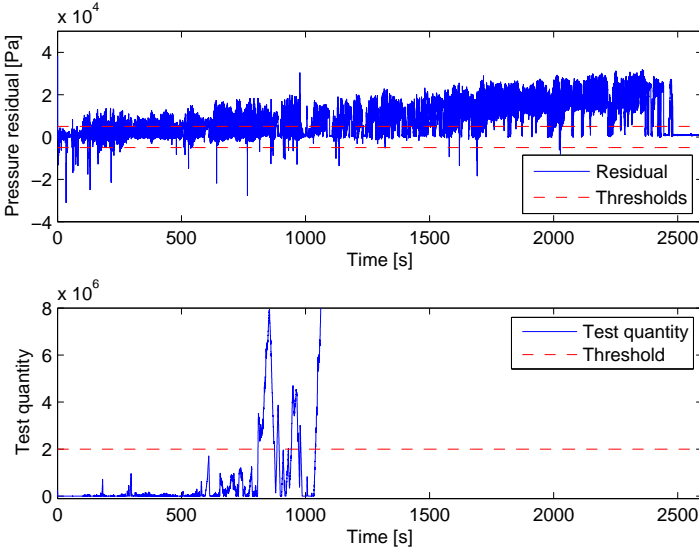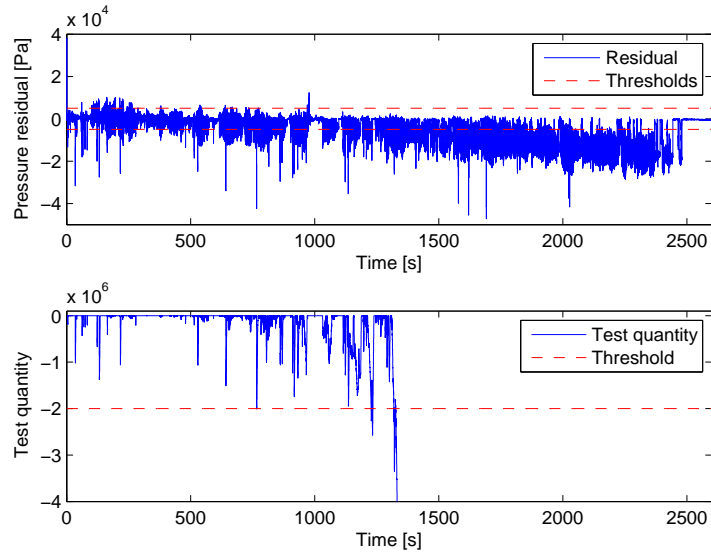
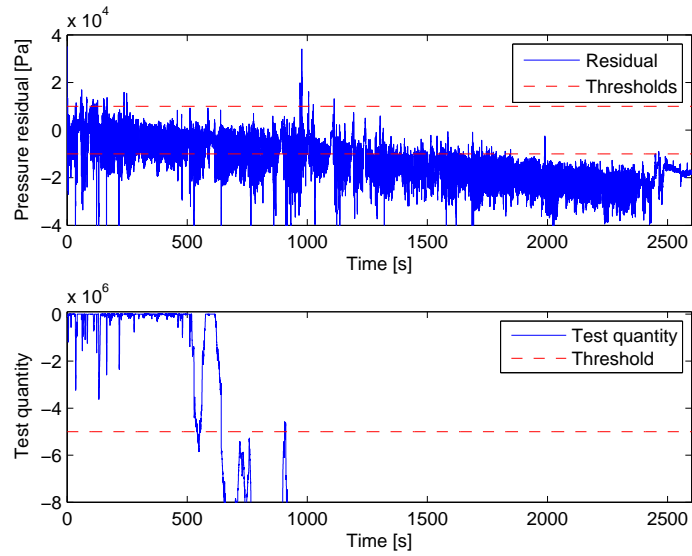(a) Sliding mode diagnostic observer and its CUSUM test.
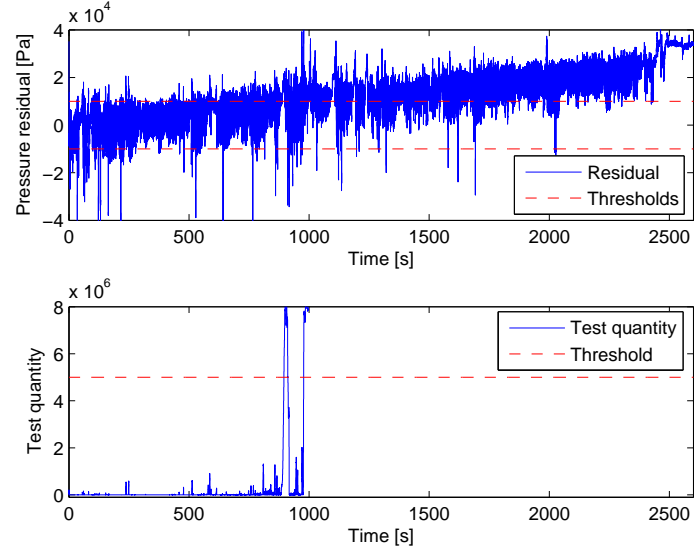
(b) Open model diagnostic observer and its CUSUM test.

Figure 6.8. Simulation result when the sensor fault is integrated with 10 Pa/s. It is clearly seen that the tests alarm after about 1,000 seconds, i.e. when the fault is 10 kPa (0.1 bar). In the fault free case the intake manifold pressure varies between 100 kPa (1 bar) and 350 kPa.

(a) Sliding mode diagnostic observer and its CUSUM test.

(b) Open model diagnostic observer and its CUSUM test.

Figure 6.9. Simulation result when the sensor fault is integrated with -10 Pa/s. Here it is obvious that there exists a difference in the detection time between the two tests. In the fault free case the intake manifold pressure varies between 100 kPa (1 bar) and 350 kPa.

The last type of fault is a step in the intake manifold pressure sensor, positive and negative steps respectively. The step is made after 1,000 seconds in the simulation and the result for a positive step of size 10 kPa for the open model based test is shown in Figure 6.10.
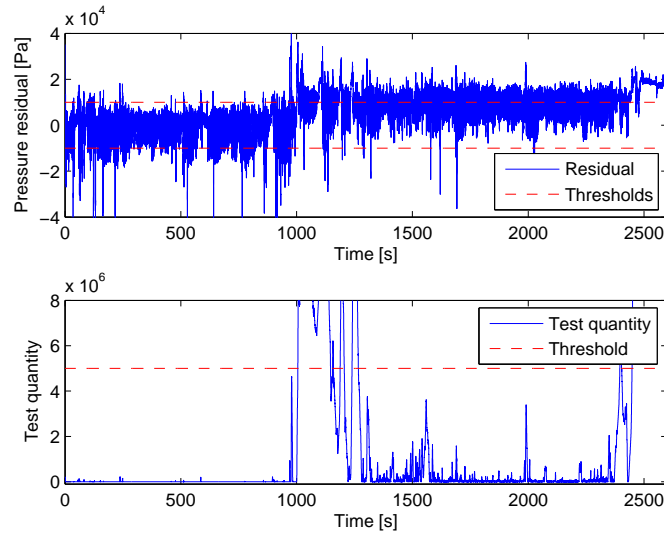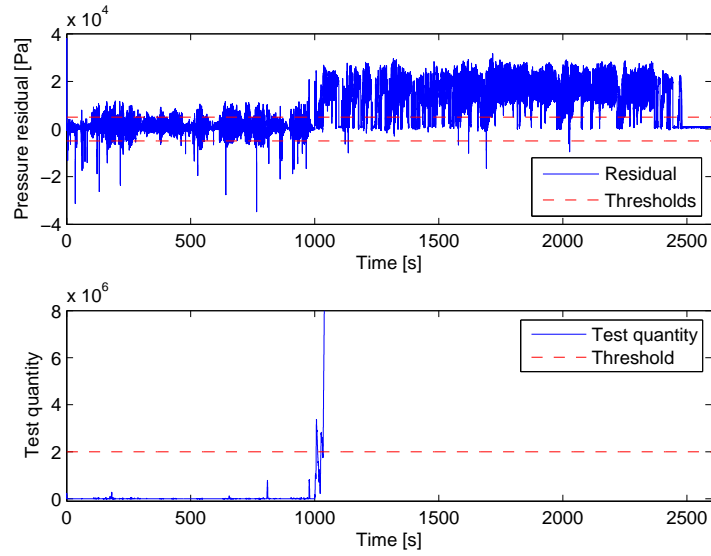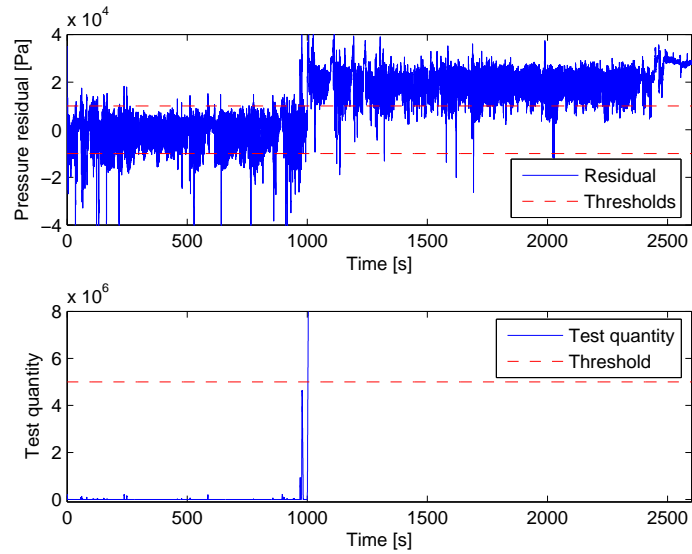

Figure 6.10. Simulation result for the open model based test with a pressure sensor fault corresponding a step of size 10 kPa at the time 1,000 s.

It seems like the drift parameter is too narrow and the test quantity first increases when the fault occurs, but then it returns to beneath the threshold. It would probably be impossible to make any conclusion on this test because the residual is close to its limits. This fault is too small for the sliding mode based diagnosis test to react. If the fault is increased to 20 kPa, the results of both tests are illustrated in Figure 6.11. Now the drift parameter and the test threshold are obvious exceeded, in both cases.

When there instead is a negative fault of size 10 kPa in the pressure sensor, it is harder for the sliding mode based test to detect the fault. This according to the same reasons as before. The test quantity is repeatedly increasing and decreasing since the residual is to small. Again when the magnitude of the fault becomes larger, it is also more obvious that the test will alarm.

(a) Sliding mode diagnostic observer and its CUSUM test.


(b) Open model diagnostic observer and its CUSUM test.

Figure 6.11. Simulation result when the sensor fault is a step of size 20 kPa. In the fault free case the intake manifold pressure varies between 100 kPa (1 bar) and 350 kPa.

6.5.4   A Simulation Frequency Analysis of the Diagnosis Systems

The performances of the diagnosis systems depend on the step size of their observers. What this dependency looks like is investigated in this section. The Engine Control Unit (ECU) has two possible frequencies for recurring simulations - 100 Hz and 20 Hz. It is desirable to implement the diagnostic observers with as large step size as possible, to keep the computational effort as low as possible. This investigation is therefore done to see if it is possible to increase the step size from 0.01 to 0.05 seconds for the two diagnostic observers.

The observers are simulated with the step sizes 0.01 and 0.05 seconds, using the WHSC presented in Chapter 4 with no faults implemented. The results are presented in Figures 6.12 and 6.13 and in Table 6.1.
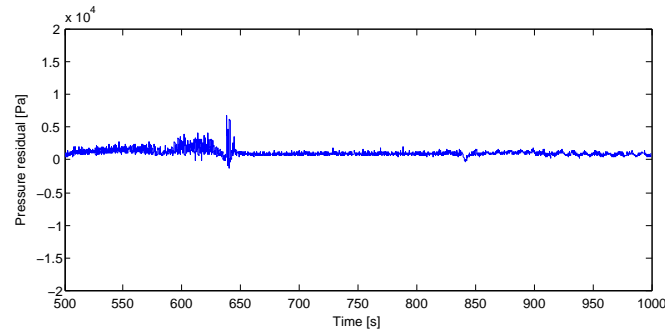
(a) The residual from the sliding mode observer with the step size 0.01 seconds.
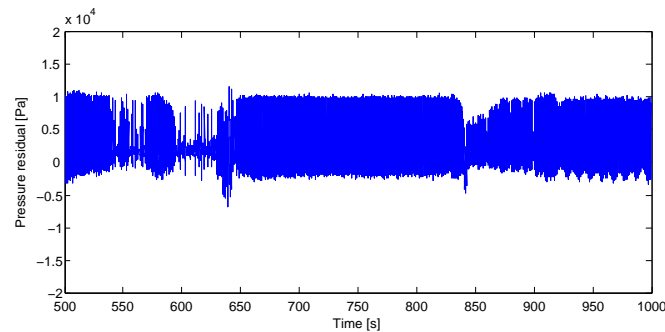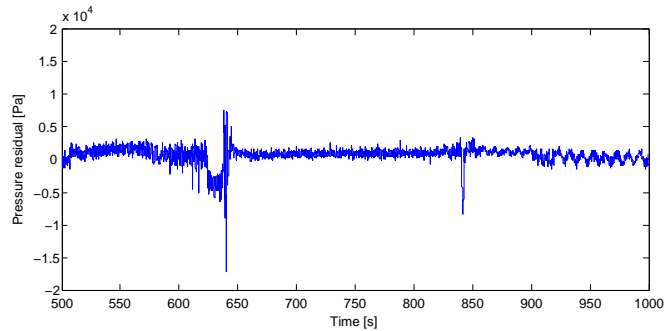

(b) The residual from the sliding mode observer with the step size 0.05 seconds.
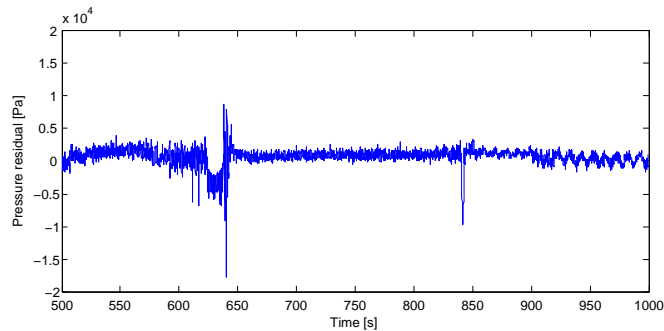
Figure 6.12. When increasing the step size from 0.01 to 0.05 seconds, it is obvious that the behaviour of the sliding mode observer gets worse.

The oscillations in the longer step size simulation of the sliding mode diagnosis test is probably caused by that the hyperbolic tangent feedback has a shorter

rising time than the open model. A step size of 0.05 seconds is not adequate for this observer. The adjusting term from the feedback gets to large in the 0.05 seconds step size, and causes the residual to change sign instead of just lowering its magnitude, and the purpose of using the hyperbolic tangent instead of the sign function is thereby lost.

(a) The residual from the adapting open model observer with the step size 0.01 seconds.

(b) The residual from the adapting open model observer with the step size 0.05 seconds.

Figure 6.13. The impact from increasing the step size for the open model observer from 0.01 to 0.05 seconds is negligible.

By comparing Figures 6.12 and 6.13 it is seen that the sliding mode observer using the step size 0.01 seconds seems to give the residual with the smallest variance and, even though this 0.01 seconds step size sliding mode observer has a somewhat larger error than the two open model observers, see Table 6.1, its residual is more stable and illustrative.

Table 6.1. Absolute and relative errors of the four alternative observer implementations.

Diagnostic observer	Absolute error [Pa]	Relative error [%]
Sliding mode, 100 Hz	1057	0.73
Sliding mode, 20 Hz	4776	3.60
Open model, 100 Hz	935	0.65
Open model 20Hz	961	0.67

Chapter 7

Conclusions and Future Work

To summarise the thesis work and see if the problem described in Section 1.3 is solved, this chapter presents the conclusions that can be drawn from the results of the work. Also new questions that have appeared during the work are described and summarised in this chapter.

7.1 Conclusions

In this section conclusions are drawn on which of the two evaluated diagnostic observers - the sliding mode observer or the open model observer - that is the most appropriate one to implement in an On Board Diagnostic system (OBD). The conclusions build on the results from Chapter 6.

From the two power functions in Section 6.5.1 it is hard to draw any conclusion on which observer to use. The fact that the open model has a slimmer area where the power function is zero and that it is better centered around $\phi = 0$ does not necessarily speak for it to be used. These properties are both easily adjusted with changes in the drift parameter ν and the threshold J. For example ν could be set differently for the under-boost and the over-boost CUSUM tests. The conclusion from the power function analysis is that both diagnostic observers could be used for fault detection.

Also in the evaluation of the effects of faults in the input signals, see Section 6.5.2, the resulting conclusion is that both alternatives would work. The behaviour of the residuals are similar between the two observers. The open model observer detects smaller faults but also here parameters in the observers (thresholds and the feedback constant) can be adjusted.

In the transient cycle evaluation, Section 6.5.3, it is shown that the open model based diagnosis test needs to be modified quite much to be able to make reliable conclusions. The drift parameter and the test quantity threshold needs to be increased to not produce any false alarm. This is caused by the difficulties to

calibrate the adaption parameters, θ, which are affected by the engine's different operating points. To calibrate these parameters as well as possible, as many operating points as possible needs to be executed, which can be hard and time consuming. In the figures in Section 6.5.3 this effect can be seen as a strong variance in the residual, even if there is no fault present. The conclusion after this evaluation is therefore that the sliding mode based diagnosis test is to prefer.

The last analysis done, the frequency analysis in Section 6.5.4, clearly shows that it is only the open model observer that is appropriate for 20 Hz simulation frequency and that the sliding mode observer demands the 100 Hz simulation frequency, which speaks for the open model to be used.

To detect small deviation in the pressure is not the most important property for the diagnostic observers in this thesis work. Both observers are able to detect what can be assessed to be small enough deviations in the intake manifold pressure. Reasonable fault magnitudes that should be detected are from 20% and up ($\sim$ 20-60 kPa). What is more important is to avoid false alarms. This together with the simplicity of the implementation of the sliding mode observer and the fact that it does not need any calibration or individual parameter settings, which also implies robuster performance, are great advantages of the sliding mode observer.

So the final conclusion is that, even though the open model has lower computational power demand, because of the possibility for it to be run with a lower simulation frequency, the sliding mode diagnostic observer would be the most appropriate one for implementation in a truck OBD. After comparing the sliding mode diagnosis system with other programs running in the ECU, at 100 Hz, it is evaluated that the computational power demanded by the diagnosis system can be managed by the ECU.

Is the problem stated in Section 1.3 solved?

Yes, it is possible to construct a model based diagnosis test for supervision of the intake manifold pressure on the Scania DC9 107 engine, that can detect under-boost and over-boost and that it is simple enough for implementation in the truck ECU. The best method found in this thesis work for this purpose is the sliding mode approach.

7.2 Future Work

In this section follows questions that have appeared during the thesis work and not been answered. Since there has not been enough time to investigate and answer these questions in the thesis work, they are seen as possible future works.

One interesting thing to do would be to implement the final diagnostic observer in a truck ECU and evaluate how the test works in an OBD.

Another interesting evaluation to do would be to run the test on known faulty data from trucks with under-boost or over-boost, i.e. engines with gas leakage or clogging.

The results from the evaluation of faults in the different input signals in Sec-

tion 6.5.2 shows that different faults causes different behaviour in the residual during different operating points. This could possibly be used for isolation of faults. A further investigation of this would be interesting.

The use of the ECU computation power could possibly be better optimised by lowering the simulation frequency for the sliding mode observer. Perhaps it is possible to accomplish this by adjusting the feedback constant λ or the constant c in (5.14). An evaluation of this would also be interesting.

An evaluation of how the adaption of the open model based observer could be improved would be interesting. It could also be investigated how this method could be used for handling of ageing of different components in the engine.

Bibliography

[1] Lars Eriksson and Lars Nielsen. *Modeling and Control of Engines and Drivelines.* LiU-Tryck, 2008.

[2] Farzad Esfandiari and Hassan K Khalil. *Output Feedback Stabilization of Fully Linearizable Systems, volume 56 of International Journal of Control, pages 1007-1037.* Taylor & Francis, London, 1992.

[3] Torkel Glad and Lennart Ljung. *Reglerteori - Flervariabla och onlinjära metoder.* Studentlitteratur, 2nd edition, 2003.

[4] Fredrik Gustafsson. *Adaptive Filtering and Change Detection.* John Wiley & Sons, Ltd, 2000.

[5] Gustaf Hendeby. *Performance and Implementation Aspects of Nonlinear Filtering.* Linköping studies in science and technology. dissertations. no. 1161, February 2008.

[6] John B. Heywood. *Internal Combustion Engine Fundamentals.* McGraw-Hill Book Co, 1988.

[7] M. Dejmai J-P. Barbot and T. Boukhobza. *Sliding Mode Control in Engineering, pages 103-130 Sliding Mode Observers.*

[8] Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory, volume 64 of Mathematics in Science and Engineering.* Academic Press, Inc, 1970.

[9] Sarah K. Spurgeon Keng Boon Goh and N. Barrie Jones. Fault diagnostics using sliding mode techniques, in control engineering practice, v 10, n 2, pages 207-217, feb. 2002. Technical Report ISSN: 0967-0661 CODEN: COEPEL, Control and Instrumentation Research Group, Department of Engineering, Leicester University, UK, 2002.

[10] Hassan K. Khalil. *High-gain observers in Nonlinear Feedback Control, chapter in New Directions in Nonlinear Observer Design (Lecture Notes in Control and Information Sciences).*

[11] Lennart Ljung and Torkel Glad. *Modellbygge och simulering.* Studentlitteratur, 2nd edition, 2004.

[12] Mattias Nyberg and Erik Frisk. *Model Based Diagnosis of Technical Processes.* LiU-Tryck, 2008.

[13] E. S. Page. *Continuous Inspection Schemes, volume 41 of Biometrika, pages 100-115.* Biometrika Trust, 1954.

[14] Pål Skogtjärn. Modelling of the exhaust gas temperature for diesel engines. Master's thesis, Linköpings Universitet, SE-581 83 Linköping, 2002.

[15] V. I. Utkin. *Variable structure systems with sliding modes, volume 22 of IEEE Transactions on Automatic Control, pages 212-222.* 1977.

[16] Johan Wahlström and Lars Eriksson. Modeling of a diesel engine with VGT and EGR including oxygen mass fraction. Technical Report LiTH-R-2747, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, 2006.

[17] Qing Wu and Mehrdad Saif. Robust fault detection and diagnosis in a class of nonlinear systems using a neural sliding mode observer, in international journal of systems science, volume 38, issue 11 january 2007 , pages 881 - 899. Technical Report DOI: 10.1080/00207720701628889, School of Engineering Science, Simon Fraser University, Vancouver, V5A 1S6, Canada, 2007.

Appendix A

Otto Cycle Calculations

A four stroke diesel engine, like the DC9 107 engine described in this thesis, operates in four different strokes during one single engine cycle [1]. These strokes are intake, compression, expansion and exhaust.

Intake The intake valve is open and a new air mixture is inducted in the cylinder as the piston moves from Top Dead Center (TDC) downwards to Bottom Dead Center (BDC).

Compression The intake valve is closed and the newly inducted air mixture is compressed to a higher pressure and temperature as the piston moves from BDC upwards to TDC. The fuel is injected directly into the cylinder under high pressure and the mixture is ignited due to the high compression heat.

Expansion The combustion is leading to a volume expansion which makes the piston to travel from TDC to BDC. This results in the work produced by the engine. Close before the BDC the exhaust valve is opened and the burned gases are blown out into the exhaust system by the pressure differences.

Exhaust The exhaust gases are pushed out from the cylinder into the exhaust manifold by the piston moving towards TDC. When the piston reaches the top, the intake valve is opened while the exhaust valve is closed, and the engine cycle restarts from the intake stroke.

In the modelling of the cylinder out temperature, T_{cyl}, an ideal Otto cycle is used, just to simplify the temperature model. Equations (3.8)-(3.12) describe the actual temperature in the cylinder outlet and they are derived from a thermodynamic study of an ideal Otto cycle. A p-V diagram of such a cycle can be seen in Figure A.1. The study is performed by breaking down the cycle into different parts, according to the different strokes declared above.
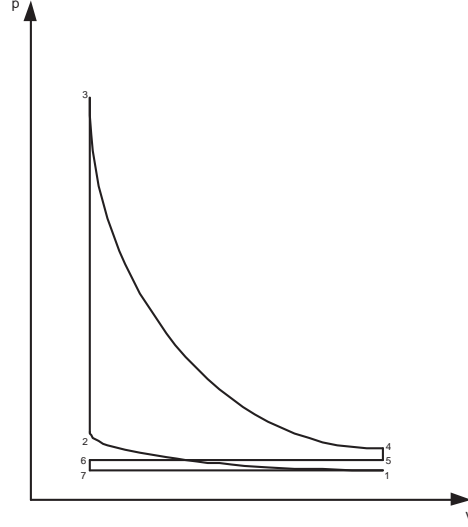
Figure A.1. A p-V diagram for an ideal Otto cycle.

Compression (1-2) This is an adiabatic and reversible process, i.e. an isentropic process, where the law

$$pv^\gamma = constant \tag{A.1}$$

holds [1]. This yields

$$p_1 v_1^\gamma = p_2 v_2^\gamma \Leftrightarrow p_2 = p_1 \frac{v_1}{v_2}^\gamma = p_1 r_c^\gamma, \tag{A.2}$$

where r_c is the engine compression ratio. Further, the ideal gas law, $pV = RT$, inserted in (A.2) gives

$$\frac{RT_2}{v_2} = \frac{RT_1}{v_1} r_c^\gamma \Leftrightarrow T_2 = \frac{v_2}{v_1} r_c^\gamma T_1 = T_1 r_c^{\gamma-1}. \tag{A.3}$$

Combustion (2-3) The combustion is an isochoric process (constant volume) where heat is transferred from the fuel to the fluid [1]. The specific energy contents of the charge per unit mass is

$$q_{in} = \frac{W_f q_{HV}}{W_{eng_{in}} + W_f} \left(1 - x_r\right), \tag{A.4}$$

which is the same expression as in (3.9). This can then be rewritten to

$$q_{in} = c_v \left(T_3 - T_2 \right) \Rightarrow T_3 = T_2 \left(1 + \frac{q_{in}}{c_v T_2} \right). \tag{A.5}$$

Since the volume is constant during this particular process, the ideal gas law gives that $\frac{p}{T} = \frac{R}{v} = constant$. This yields the pressure-temperature ratio

$$\frac{p_3}{T_3} = \frac{p_2}{T_2} \Leftrightarrow p_3 = p_2 \frac{T_3}{T_2}. \tag{A.6}$$

Expansion (3-4) Also the expansion is an isentropic process, just like the compression. With similar calculations as in (A.1)-(A.3), this gives

$$p_4 = p_3 \frac{1}{r_c^\gamma}, \tag{A.7}$$

$$T_4 = T_3 \frac{1}{r_c^{\gamma-1}}. \tag{A.8}$$

Blowdown (4-5) The blowdown refers to when the exhaust valve is opening and the cylinder pressure decreases to the exhaust pressure. Further, the remaining gas in the combustion chamber is assumed to experience an isentropic expansion process [1]. This yields

$$p_5 = p_{em}, \tag{A.9}$$

$$T_5 = T_{cyl} = T_4 \left(\frac{p_{em}}{p_4} \right)^{1 - \frac{1}{\gamma}}. \tag{A.10}$$

Exhaust (5-6) The burned gases are pushed out into the exhaust manifold without any heat transfer, i.e. $T_6 = T_{cyl}$, and under constant pressure, $p_6 = p_{em}$. When the exhaust valve is closed, not all the gas has left the combustion chamber. The remaining gas at this point is called the *residual gas*, with mass m_r. The fraction of residual gas is denoted x_r. This is the fraction of residual gases among the total mass of gases, m_t, in the cylinder,

$$x_r = \frac{m_r}{m_t} = \frac{1}{r_c} \left(\frac{p_{em}}{p_4} \right)^{\frac{1}{\gamma}}. \tag{A.11}$$

Intake valve opening (6-7) When the intake valve opens, some of the residual gas flows out from the cylinder into the intake manifold. This results in a decreased cylinder pressure, $p_7 = p_{im}$. Also the temperature starts to decrease.

Intake (7-1) The cylinder volume is filled with a new air mixture at constant pressure, $p_1 = p_{im}$. The temperature depends on both the intake manifold temperature and the residual gas temperature, due to the residual gas fraction. Therefore, the temperature during the intake process can be expressed as

$$T_1 = x_r T_{cyl} + (1 - x_r) T_{im}. \tag{A.12}$$

Combinations of the equations above give

$$T_{cyl} = T_1 \left(\frac{p_{em}}{p_{im}} \right)^{1 - \frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma - 1}} \right)^{\frac{1}{\gamma}} \tag{A.13}$$

$$x_r = \frac{1}{r_c} \left(\frac{p_{em}}{p_{im}} \right)^{\frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma - 1}} \right)^{-\frac{1}{\gamma}} \tag{A.14}$$

These equations, and also (A.4), are the same expressions as (3.8)-(3.11), except that the compensation factor η_{oc} is not included in the general ideal Otto cycle.

Appendix B

Notation

Here the symbols, their subscripts, and some abbreviations used in this thesis are listed.

Table B.1. Symbols used in the thesis.

Variable	Description	Unit
A	Area	m^2
c_p	Spec. heat capacity, constant pressure	$J/(kg \cdot K)$
c_v	Spec. heat capacity, constant volume	$J/(kg \cdot K)$
m	Mass	kg
n_{cyl}	Number of cylinders	$-$
n_e	Engine rotational speed	rpm
p	Pressure	Pa
q_{HV}	Heating value of diesel	J/kg
q_{in}	Specific energy content	J/kg
R	Gas constant	$J/(kg \cdot K)$
r_c	Compression ratio	$-$
T	Temperature	K
u_{egr}	EGR valve position	$\%$
u_δ	Injected amount of fuel	$mg/cycle$
V	Volume	m^3
W	Mass flow	kg/s
x_r	Residual gas fraction	$-$
γ	Specific heat capacity ratio	$-$
η	Efficiency	$-$

Table B.2. Subscripts used in the thesis.

Subscript	Description
a	Air
amb	Ambient
cmp	Compressor
cyl	Cylinder
e	Exhaust
egr	Exhaust gas recirculation (EGR)
em	Exhaust manifold
eng_{in}	Engine cylinder in
eng_{out}	Engine cylinder out
f	Fuel
im	Intake manifold
ob	Over-boost
oc	Otto cycle
ub	Under-boost
vol	Volumetric

Table B.3. Abbreviations used in the thesis.

Subscript	Description
BDC	Bottom Dead Center
EGR	Exhaust Gas Recirculation
RHS	Right Hand Side
TDC	Top Dead Center
VGT	Variable Geometry Turbine

Appendix C

Compilation of the Model Equations

Here follows a compilation of the model equations in the open model used in this thesis.

$$\dot{p}_{im} = \frac{R_a T_{im}}{V_{im}} \left(W_{cmp} + W_{egr} - W_{eng_{in}} \right) \tag{C.1}$$

$$W_{eng_{in}} = \eta_{vol} \frac{p_{im} n_e V_d}{120 R_a T_{im}} \tag{C.2}$$

$$\eta_{vol} = c_{vol1} + c_{vol2}\sqrt{p_{im}} + c_{vol3}\sqrt{n_e} \tag{C.3}$$

$$W_f = \frac{10^{-6}}{120} u_\delta n_e n_{cyl} \tag{C.4}$$

$$W_{eng_{out}} = W_{eng_{in}} + W_f \tag{C.5}$$

$$T_{cyl} = \eta_{oc} T_1 \left(\frac{p_{em}}{p_{im}} \right)^{1-\frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma-1}} \right)^{\frac{1}{\gamma}} \tag{C.6}$$

$$q_{in} = \frac{W_f q_{HV}}{W_{eng_{out}}} \left(1 - x_r\right) \tag{C.7}$$

$$x_r = \frac{1}{r_c} \left(\frac{p_{em}}{p_{im}}\right)^{\frac{1}{\gamma}} \left(1 + \frac{q_{in}}{c_v T_1 r_c^{\gamma-1}}\right)^{-\frac{1}{\gamma}} \tag{C.8}$$

$$T_1 = x_r T_{cyl} + (1 - x_r) T_{im} \tag{C.9}$$

$$T_{em} = T_{amb} + (T_{cyl} - T_{amb}) e^{\frac{h_{tot} A}{W_{eng_{out}} c_p}} \tag{C.10}$$

$$W_{egr} = \frac{A_{egr}(u_{egr}) p_{em} \Psi_{egr}}{\sqrt{R_e T_{em}}} \tag{C.11}$$

$$\Psi_{egr} = 1 - \left(\frac{1 - \Pi_{egr}}{1 - \Pi_{egropt}} - 1\right)^2 \tag{C.12}$$

$$\Pi_{egr} = \begin{cases} \Pi_{egropt} & \text{if } \frac{p_{im}}{p_{em}} < \Pi_{egropt} \\[2mm] \frac{p_{im}}{p_{em}} & \text{if } \Pi_{egropt} \le \frac{p_{im}}{p_{em}} \le 1 \\[2mm] 1 & \text{if } 1 < \frac{p_{im}}{p_{em}} \end{cases} \tag{C.13}$$

$$\Pi_{egropt} = \left(\frac{2}{\gamma_e + 1}\right)^{\frac{\gamma_e}{\gamma_e - 1}} \tag{C.14}$$

$$A_{egr}(u_{egr}) = A_{egrmax} f_{egr}(u_{egr}) \tag{C.15}$$

$$f_{egr}(u_{egr}) = \begin{cases} c_{egr1} u_{egr}^2 + c_{egr2} u_{egr} + c_{egr3} & \text{if } u_{egr} \le -\frac{c_{egr2}}{2 c_{egr1}} \\[2mm] c_{egr3} - \frac{c_{egr}^2}{4 c_{egr1}} & \text{if } u_{egr} > -\frac{c_{egr2}}{2 c_{egr1}} \end{cases} \tag{C.16}$$

Appendix D

Fault Tree Analysis

In the beginning of this thesis work a fault tree analysis was made to give a better understanding of what could be the causes of under-boost and over-boost respectively. The fault tree analysis is shown in Figures D.2 and D.1.
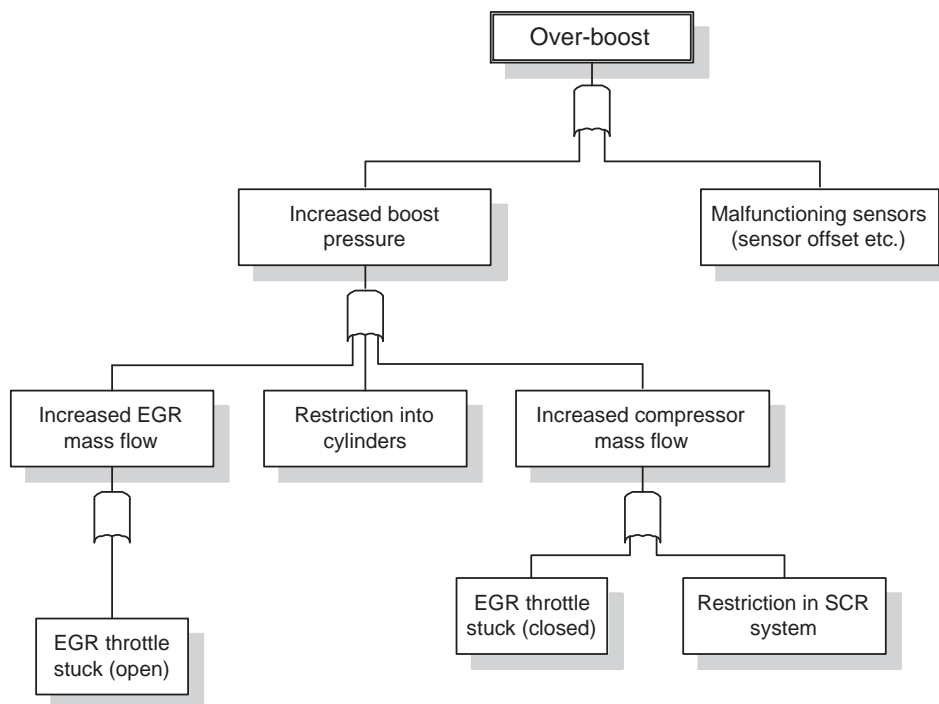

Figure D.1. This fault tree analysis describes what could be the causes of over-boost.
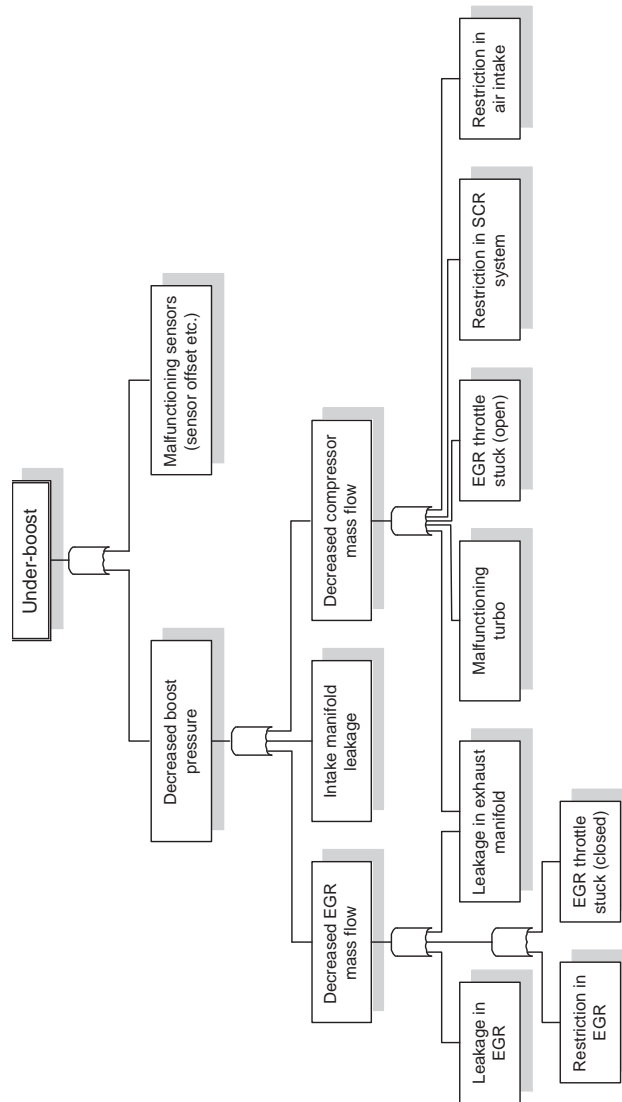
Figure D.2. This fault tree analysis describes what could be the causes of under-boost.