Design and Evaluation of an Automatically Generated Diagnosis System

Master's thesis
performed in **Vehicular Systems**
for **Scania CV AB**

by
Joakim Hansen & Jens Molin

Reg nr: LiTH-ISY-EX -- 06/3889 -- SE

December 14, 2006

Design and Evaluation of an Automatically Generated Diagnosis System

Master's thesis

performed in **Vehicular Systems**,
Dept. of Electrical Engineering
at **Linköpings universitet**

for **Scania CV AB**
by

Joakim Hansen & Jens Molin

Reg nr: LiTH-ISY-EX -- 06/3889 -- SE

Supervisor: **Mattias Nyberg PhD**
Scania CV AB
Anna Pernestål PhD student
Scania CV AB
Jonas Biteus PhD student
Linköpings unviversitet

Examiner: **Assistant Professor Erik Frisk**
Linköpings universitet

Linköping, December 14, 2006

Avdelning, Institution Division, Department Division of Vehicular Systems Department of Electrical Engineering Linköpings universitet SE-58183 Linköping, Sweden	**Datum** Date 2006-12-14

Språk Language ☐ Svenska/Swedish ☒ Engelska/English ☐ ________________	**Rapporttyp** Report category ☐ Licentiatavhandling ☒ Examensarbete ☐ C-uppsats ☐ D-uppsats ☐ Övrig rapport ☐ ________________	**ISBN** — **ISRN** LiTH-ISY-EX -- 06/3889 -- SE **Serietitel och serienummer** Title of series, numbering **ISSN** —

URL för elektronisk version

`http://www.vehicular.isy.liu.se`
`http://ep.liu.se`

Titel
Title

Konstruktion och utvärdering av ett automatgenererat diagnossystem

Design and Evaluation of an Automatically Generated Diagnosis System

Författare
Author

Joakim Hansen & Jens Molin

Sammanfattning
Abstract

Throughout recent years, legislations concerning emission levels for vehicles have become more restrictive and will be even more restrictive in the future. In the recent European environmental standards, EURO 4 (2006) and EURO 5 (2008), further requirements have been added on top of low emission demands. All heavy duty trucks have to be equipped with an OBD-system.
Scania CV AB has today an existing OBD-system that consists of several tests. Typically, a test is designed to check if a signal is inside specified limits or thresholds. To improve the system, Scania CV AB and Vehicular Systems at Linköping University have developed a method to design diagnosis systems in an automatic way, implemented in a toolbox called DSAME.
In this thesis, an automatic designed OBD-system has been created with DSAME and the corresponding parts in a manually designed OBD-system have been identified. The two systems have been compared. The result shows that both systems are equally at detecting faults but the automatic designed OBD-system is a lot better to isolate the faults than the existing OBD-system.

Nyckelord
Keywords

Model based diagnosis, Diagnosis performance, Evaluation

Abstract

Throughout recent years, legislations concerning emission levels for vehicles have become more restrictive and will be even more restrictive in the future. In the recent European environmental standards, EURO 4 (2006) and EURO 5 (2008), further requirements have been added on top of low emission demands. All heavy duty trucks have to be equipped with an OBD-system.
Scania CV AB has today an existing OBD-system that consists of several tests. Typically, a test is designed to check if a signal is inside specified limits or thresholds. To improve the system, Scania CV AB and Vehicular Systems at Linköping University have developed a method to design diagnosis systems in an automatic way, implemented in a toolbox called DSAME.
In this thesis, an automatic designed OBD-system has been created with DSAME and the corresponding parts in a manually designed OBD-system have been identified. The two systems have been compared. The result shows that both systems are equally at detecting faults but the automatic designed OBD-system is a lot better to isolate the faults than the existing OBD-system.

Keywords: Model based diagnosis, Diagnosis performance, Evaluation

Preface

This master´s thesis was performed during the summer and fall 2006 at Scania CV AB in Södertälje. Scania is a worldwide manufacturer of heavy duty trucks and engines for marine and industrial use. The work was carried out at the engine software development department (NED), which is responsible for the on board diagnostics (OBD) software.

Acknowledgment

We would like to express our gratitude to a number of people:

Our excellent supervisors, Mattias Nyberg and Anna Pernestål at Scania CV AB and Jonas Biteus at Linköping University, and our examiner, Erik Frisk, for suggestions and discussions. Gustaf Ahrrenius and Mats Jennische for the help with the truck. All other helpful people at NED, NEE and NEV. Rikard Falkeborn for all the nice discussions and all the help we got when we needed it. The other master thesis workers at our floor, Carl Svärd, Henrik Wassén, Klas Håkansson and Mikael Johansson for many nice coffee breaks.

Finally, we would like to thank our girlfriends, Karin and Cecilia, for helping us during the thesis.

Joakim Hansen
Södertälje, December 2006
Jens Molin
Venezuela, December 2006

Contents

Chapter 1

Introduction

Throughout recent years, legislations concerning emission levels for vehicles have become more restrictive and will be even more restrictive in the future. In the recent European environmental standards, EURO 4 (2006) and EURO 5 (2008), further requirements have been added on top of low emission demands. All heavy duty trucks have to be equipped with an on board diagnosis (OBD) system. The purpose of the OBD-system is to detect malfunctions leading to emissions above the permitted limits, and is motivated by the fact that the majority of vehicle emissions today are caused by malfunctioning emission control systems [8]. Another benefit by having an OBD-system is the possibility to discover small faults before they cause serious damage. It may also simplify troubleshooting at workshops and therefore improve repairability.

At Scania CV AB, the existing OBD-system consists of several tests. A test is a small system supervising a limited part of the engine process. Typically, a test is designed to check if a signal is inside specified limits, called thresholds. A more sophisticated test is to check if a measured signal deviates from an estimated value of the signal. To improve the OBD-system, Scania and Vehicular Systems at Linköping University have developed a method to design OBD-systems in an automatic way. Through several master theses ([2], [3] and [5]), a Matlab toolbox called DSAME (diagnostic structural analysis and modeling execution toolbox) has been developed for this purpose. Based on an existing engine model, DSAME finds candidates for tests, evaluates them and constructs a diagnosis system. In [3], a simple evaluation has been done of the automatically generated OBD-system designed by DSAME with the conclusion that the system seemed to work well. However, a full evaluation has not been done.

1.1 Problem Statement

The existing OBD-system used by Scania is constructed based on engineer knowledge. Studies have shown that it is able to detect large faults but many smaller faults may not be detected. The problem is that even small faults might cause too high emissions, for example by injection of too much fuel or that the control system does an incorrect change to a different control mode. Further, since the OBD-system is manually constructed, it has to be manually adapted to different engine types. If one small detail in the engine is changed, the OBD-system has to be changed manually. The problem is to find out if it may be possible in the future to enhance the development process of diagnosis systems using DSAME.

1.2 Objectives

The objective with this thesis is to compare an OBD-system automatically generated by DSAME with the existing Scania OBD-system and give a statement of the possibility to use DSAME in future OBD-system design. This objective can be parted into:

- Construct a diagnosis system for the engine using the DSAME-toolbox.
- Identify the corresponding part in a manually designed OBD-system.
- Create a method for evaluating the diagnostic performance of an OBD-system.
- Use the evaluation method to do a comparison between the two OBD-systems.
- Draw conclusions about advantages and disadvantages of the two systems and suggest how DSAME should be used in the future.

1.3 Thesis Outline

Chapter 1 gives an introduction to the thesis.

Chapter 2 gives a short introduction to principles of diagnosis.

Chapter 3 presents the diesel engine.

Chapter 4 describes a manually designed diagnosis system.

Chapter 5 describes the main principles of Scanias method for automatic design of an diagnosis system.

Chapter 6 presents the evaluation method that will be used in this thesis.

Chapter 7 describes how the measurements will be done.

Chapter 8 gives a comparison and analysis of the different diagnosis systems.

Chapter 9 concludes the thesis and discusses suggestions to possible future work.

1.4 Contributions

- Corresponding parts in the manually designed OBD-system have been identified.
- A method for evaluating OBD-systems concerning detectability, isolability and detection time has been created.
- A comparison between an automatic designed OBD-system and a manually designed OBD-system has been done.

Chapter 2

Principles of Diagnosis

In this chapter a short introduction to traditional diagnosis will be given and to a more recent model based approach, which is used in DSAME. It also treats how information from many tests can be used to isolate faults, i.e. to exactly point at one fault from other faults. This information is gathered from [8].

2.1 Traditional Diagnosis

Traditionally in heavy vehicles, diagnosis, i.e. to decide if there is a fault and if so identify the fault, was focused on safety critical processes but with restrictive laws the diagnosis has become more important. The principle of diagnosis can be divided into three parts that is shown in Figure 2.1. The first part is to observe the process, i.e. collect data from sensors and actuators. The observations will then be preprocessed in different tests. The final part is to calculate a diagnosis statement, i.e. which faults that can explain the observations, from the test results. This is done by an isolation unit.

This thesis will focus on the second part where the common diagnosis approaches that have been used are:

Duplication or hardware redundancy. A fault can for example be detected by having more than one sensor measuring the same value. If the values deviate then the conclusion that there is a fault can be drawn. This approach is common in safety critical systems like aeroplanes and power plants. The drawbacks with this approach is the cost of having extra sensors and the need of three sensors to isolate the faulty sensor.

Signal in range check tests. This type of tests checks if measured signals are in a specific range. If not, a fault has occurred. A drawback with this approach is that the signal might show similar values in the fault

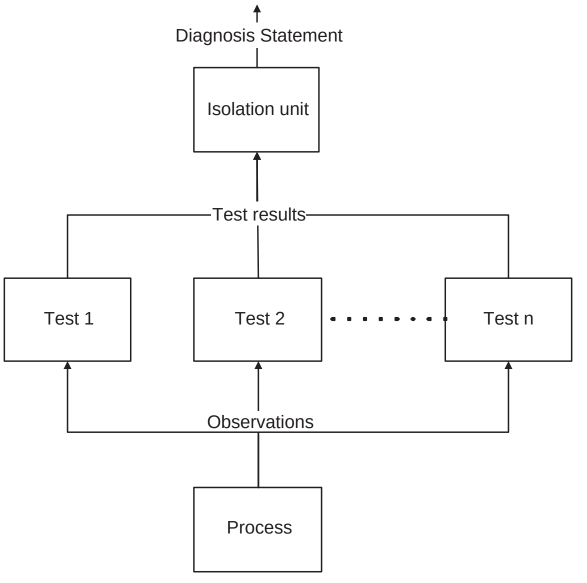


Figure 2.1: Overview of a general diagnosis system.

free case compared to the faulty case. The faulty case might therefore be hard to detect.

Active diagnosis. In this approach the process is controlled into working points where correct outputs are known. If the known value deviates from the measured value, a fault has occurred. A drawback is that it is often needed to interrupt the process which can be complicated and expensive.

2.2 Model Based Diagnosis

A recent approach in diagnosis is model based diagnosis. The idea with model based diagnosis is to create a model over the system. A measured value can then be compared to its corresponding estimated value and if they deviate, a fault has occurred. By, at time t, comparing a measured value , $y(t)$, with a estimated value, $\hat{y}(t)$, the residual, $r(t)$, can be formed

$$r(t) = y(t) - \hat{y}(t). \tag{2.1}$$

A residual, $r(t)$, is a function that is ideally zero in the fault free case. Good candidates for residuals are analytical redundancy relations[1], ARRs. An ARR is a relation that always holds in the fault free case. In reality residuals will

[1]Other terms often used in literature are consistency relations or parity relations

not be zero because of noise and model errors, instead a test quantity can be a better choice for diagnosis purpose.

2.3 Test Quantity

A test quantity is a function $TQ(z)$ that from the observations, z, calculates a scalar. The test quantity will be used to determine if there is a fault present in the process. If the test quantity is above (or below) the given thresholds (J_1 and J_2), the test will respond. Test quantities can be designed in many ways. In this thesis, test quantities are based on the mean value calculated for the residual $r(t)$ over $N+1$ samples where the observations are sensor and actuator values used to estimate $\hat{y}(t)$ and measure $y(t)$.

$$TQ(z) = \frac{1}{N+1} \sum_{t=0}^{N} r(t) \tag{2.2}$$

The test should respond if a fault has occurred but not respond if there are no fault. This means that the probability of false alarm should be low and the probability to respond, if there is a fault, should be high. By introducing the power function, β, the behavior of a test can be examined. Let the size of the fault in the process be θ, which is zero in the fault free case, then the power function can be formalized as

$$\beta(\theta) = P(TQ(z) > J_1 \text{ or } TQ(z) < J_2 | \theta) \tag{2.3}$$

The power function should be low for $\theta = 0$ and large for $\theta \neq 0$. Figure 2.2 shows a typical power function where it can be seen that the probability for the test to respond is low in the fault free case, $\theta = 0$, and higher when fault is present, $\theta \neq 0$.

2.4 Decision Structure

To investigate which faults each test might respond to, a decision structure can be used. A decision structure for a diagnosis system with three tests, δ_1 , δ_2 and δ_3, and three faults F_1, F_2 and F_3 can look like

	F_1	F_2	F_3
δ_1	X	X	
δ_2		X	X
δ_3			X

(2.4)

where an ’X’ on row i and column j means that test δ_i may respond to fault F_j. If there is no ’X’ the test will not respond to this fault.

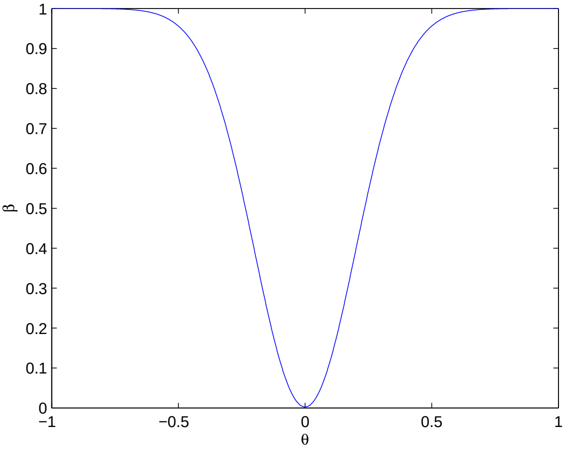


Figure 2.2: Typical power function.

2.5 Isolation

When a couple of tests have been constructed, each test will give a test result to an isolation unit, see Figure 2.1. By combining the different information from the test results, the isolation unit will give a diagnosis statement, a diagnosis. A diagnosis is a possible explanation of the observations preprocessed in the tests. The diagnosis statement can be calculated in different ways, for example as described in [9], and the basic idea to produce a diagnosis statement is that different tests will respond to different faults. Therefore conclusion about which faults that can explain the observations can be drawn.

By using the decision structure, a isolation structure, i.e. structure over which faults that can be isolated from each other, can be calculated. The isolation structure for the decision structure given in (2.4) becomes

$$\begin{array}{c|ccc} & F_1 & F_2 & F_3 \\ \hline F_1 & \mathrm{X} & \mathrm{X} & \\ F_2 & & \mathrm{X} & \\ F_3 & & & \mathrm{X} \\ & & & \end{array} \tag{2.5}$$

The isolation structure should be read as follows: an 'X' on row i and column j means that fault F_j can not be isolated from fault F_i. In (2.5) it can be seen that if fault F_1 occur, then it can be concluded from the isolation structure that F_1 can not be isolated from F_2. However, if fault F_2 occurs, then it can be isolated because there is only one 'X' in the second row.

Chapter 3

Background to Diesel Engines

In this chapter the principles of the diesel engine and the main solutions to decrease emissions will be presented.

3.1 Diesel Engine

The fundamental principle of a diesel engine is simple. Air is taken in through the inlet manifold into the cylinder. The working cylinder compresses the air, diesel is injected and immediately set on fire by the high pressure. Two things are controlled, the injection angle, α, i.e. the time when the fuel is injected, and the amount of injected fuel, δ. To have an efficient engine with high performance, it is necessary to pre-compress the inlet air by a turbo charger. The compressor is driven by the exhaust pressure.

One environmental problem with diesel engines is discharge of nitrogen oxides (NOx) which are created when the combustion temperature is too high. Today there are two main solutions to reduce this kind of emissions. One solution is to use an SCR (Selective catalyst reduction) system which aftertreats the exhaust gases in a catalyst. The other solution, which is used in this thesis, is the use of exhaust gas recirculation, EGR. With EGR, some exhaust gases are lead back to the inlet manifold and lowers the amount of oxygen in the combustion. The combustion temperature therefore decreases and the amount of NOx in the exhaust gases is reduced.

To get a satisfying combustion, the EGR-flow has to be controlled. The EGR-flow mainly depends on two things, the position of the EGR valve and the pressure difference between exhaust manifold and inlet manifold. To improve the control of the pressure difference it is possible to change the flow through the turbine by changing the geometry of the turbine, i.e. to use a variable geometry turbine, VGT. Control of EGR-flow is complicated and might be impossible if some sensor or actuator is malfunctioned. Diagnosis of gas-flow components is therefore necessary to guarantee low emissions. A figure of a modern diesel engine with EGR and VGT is shown in Figure 3.1. The figure shows the fundamental parts of the diesel engine. Air is taken in and compressed by the compressor. Then the compressed air is mixed up with EGR-gases. Finally, The turbine, which drives the compressor, is driven by the exhaust gases.

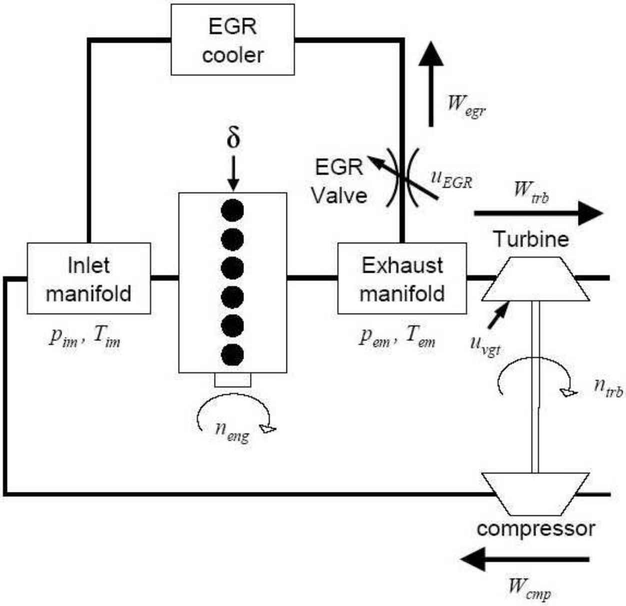


Figure 3.1: Schematic figure of the gasflow of a diesel engine with VGT and EGR.

3.2 Engine Control

The principle for an engine with control system is shown in Figure 3.2. In the figure, it is shown that the engine control unit (ECU) uses inputs from the driver, and by using measurements from the sensors, actuator values are calculated.

The control system optimizes the combustion to minimize the emissions while maintaining engine power. If a sensor or actuator is faulty, the control system will not work satisfactory and the emissions will increase. The observations, i.e. engine sensors and actuators, used in this thesis are listed in Table 3.1.

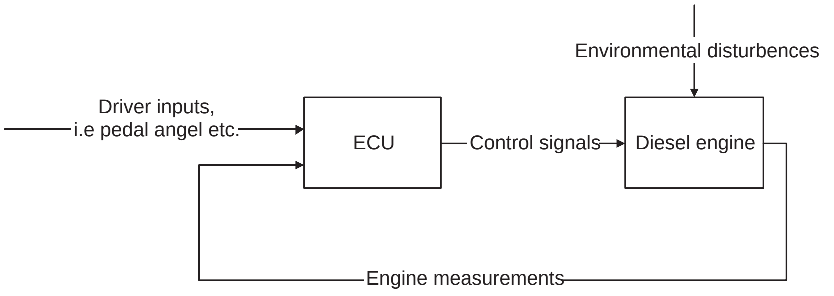


Figure 3.2: Principle of engine with control system.

Table 3.1: Sensors and actuators included in the comparison.

Sensor	Unit	Description
t_{amb}	[K]	Ambient temperature
t_{im}	[K]	Temperature in the intake manifold
p_{amb}	[Pa]	Ambient pressure
p_{im}	[Pa]	Pressure in the inlet manifold
p_{em}	[Pa]	Pressure in the exhaust manifold
w_{cmp}	[kg/s]	Air mass flow after the compressor
n_{trb}	[rps]	Turbine speed
n_{eng}	[rps]	Engine speed
u_{egr}	[volt]	The EGR actuator
u_{vgt}	[volt]	The VGT actuator
δ	[mg/stroke]	Injected fuel
α	degrees	Fuel injection angle

3.3 Engine Model

To be able to use model based diagnosis a model is needed. In this thesis, the engine model described in [1] and [10] is used. This model is used together with DSAME to create an automatic designed OBD-system, more about this in Chapter 5. The tests in the diagnosis system based on a model can not become better than the model it is derived from, i.e. not detect faults smaller than model errors and noise.

Chapter 4

Manually Designed OBD-system

In this chapter, different types of tests in the manually designed OBD-system are described. A deeper description of tests that are interesting for comparison between the manually designed OBD-system and the automatic designed OBD-system will be presented. The manually designed system consists of tests that have been handmade by engineers. Each test is usually designed with the purpose to discover only a few faults. The tests are classified into electrical tests and plausibility tests.

4.1 Electrical Tests

Electrical tests are designed to detect electrical faults like short and open circuit. These faults are easy to detect because of their special behavior. The common way to do these tests is to check if the electrical value has been equal to a maximum or a minimum value for a longer time. If so, the component is assumed to be broken. Electrical tests for open circuit and short circuit will not be investigated further in the thesis.

4.2 Plausability Tests

Plausibility tests check if a sensor takes reasonable value. If it does not, the sensor is assumed to be faulty. The plausibility tests can be divided into five different types:

- Duplication tests
- Signal in range check tests

- Adaption tests
- Control error tests
- Model based tests

Each type will be described in the following sections. In each section the specific tests of interest for the evaluation will be presented.

4.2.1 Duplication Tests

In some working points, some sensors placed at different locations are actually measuring the same value, for example all pressure sensors when the engine is shut off. Tests can be constructed by comparing these sensors to each other at these specific working points.

Another way to construct duplication tests is to add extra sensors in the system for duplication purpose. The main drawback with this is the extra cost. The duplication tests considered in this this thesis are static pressure sensors tests and a temperature sensor test.

Static Pressure Sensors Tests

Static behavior of the pressure sensors can for example be tested when the driver turn on the vehicle with the key. The three pressure sensors, inlet manifold pressure, p_{im}, ambient pressure, p_{amb}, and exhaust manifold pressure, p_{em}, can then be pairwise compared. When the engine is shut off, all the sensors should show the same value. In this thesis three static pressure sensors tests are considered. These tests are from now on denoted SPT_{ai}, SPT_{ei}, SPT_{ae} (static pressure test$_{sensor1sensor2}$). The faults affecting these tests will be faults in the pressure sensors.

Temperature Sensors Test

Another duplication test can be made by comparing the ambient temperature, T_{amb}, and the temperature at the mass flow sensor, T_w. If they deviate, the test responds. In this thesis, T_w is assumed to always be fault free. This test is from now on denoted ATT (ambient temperature test).

4.2.2 Signal in Range Check Tests

A common way to construct a plausibility test is to check if a signal is above (or below) a static threshold. If so the test will respond. A typical signal in range check test can be to check if the turbine speed is above a very high static threshold. As explained in Section 2.1, these kind of tests are assumed to have low performance compared to other tests and are not further investigated in the thesis.

4.2.3 Adaption Tests

When time goes by, many sensors can get small deviations without being faulty. Sensors might also be affected by different environments e.g different levels of humidity. Some sensors are therefore adjusted after a model or another sensor which is assumed to be more reliable. This thesis consider two relevant adaption procedures affecting the engine.

Adaption of Pressure Sensors

The pressure sensors p_{em} and p_{im} might have a small bias. To make the sensors usable anyway, the sensors can be adapted. The adaption is in this thesis made by comparing the sensors p_{em} and p_{im} with the p_{amb}-sensor. If they deviate, p_{em} and p_{im} are adapted to p_{amb}. Because of its close measure range and that it is not exposed to such a strenuous environment as the other pressure sensors, the p_{amb}-sensor can be considered more reliable than the others. Some tests can be constructed to alarm if the adaption becomes to large. These tests are from now on denoted $SPAT_{ai}$ and $SPAT_{ae}$ (static pressure adaption test$_{sensor1sensor2}$). Faults affecting the tests are faults on the pressure sensors.

Adaption of Mass Flow Sensor

The characteristics of the mass flow sensor normally changes with the environmental conditions. Therefore, measured mass flow is regularly compared to an estimated value, if they deviate too much, the mass flow sensor will be adapted to the estimation. In this thesis, a test is constructed to respond if the adaption become to large. This test is from now on denoted MAT (Mass flow adaption test). The signals affecting this test are the mass flow sensor and the signals used to estimate the mass flow, p_{im} and T_{im}.

4.2.4 Control Error Tests

Consider the system given in Figure 4.1. Control error tests compares the reference value $y_{ref}(t)$ with a measured feedback value $y(t)$. If they deviate the tests will respond. This thesis will handle two control error tests. These are EGR-damper test and EGR-flow test.

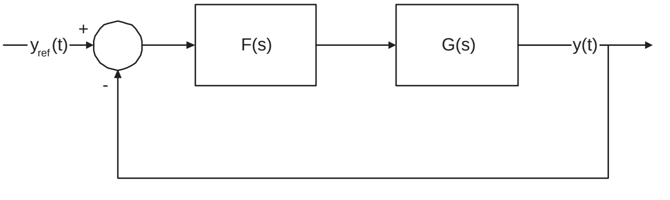


Figure 4.1: A system like the one in the control error tests.

EGR-damper Test

In this thesis, a test comparing the reference value of the EGR-damper position with a measured EGR-damper position will be considered. The controller used in this test will be an ordinary PID-controller. This test is from now called EDT (EGR-damper test) and will respond on faults in the EGR-damper.

EGR flow Test

In this thesis, one test comparing the EGR flow reference value with an estimated EGR-flow will be considered. The estimated flow is calculated by the model shown in (4.1). In this thesis, the controller that will be used is an ordinary PID-controller. A feed forward from the reference will also affect the system. This test will respond to the signals into the model i.e. w_{cmp}, p_{im} and T_{im} and to the control signal to the EGR-damper, u_{egr}. This test is from now on denoted EFT (EGR-flow test).

$$\hat{w}_{egr} = f(w_{cmp}, p_{im}, T_{im}) \tag{4.1}$$

4.2.5 Model Based Tests

There are some model based tests in the manually designed OBD-system. They are, with some modifications, constructed as described in Section 2.2, i.e. a measured value is compared to an estimated value. The following tests will be considered, two dynamic pressure sensors tests and one VGT test.

Dynamic Pressure Sensors Tests

This thesis considers two dynamic tests on the pressure sensors, p_{im} and p_{em}. These tests compare the dynamics of the signals with modeled dynamics If the dynamics deviate, the tests respond. These tests are constructed to detect dynamic faults in the pressure sensors like gain faults or sensors get stuck. A schematic figure is shown in Figure 4.2. These tests are from now on denoted DPT_i and DPT_e. (DynamicPressureTest$_{sensor}$). These tests are affected by the signals in the models for estimating the pressure values and the measured pressure values. The models used in this thesis are shown in (4.2) and (4.3).

$$\hat{p}_{im} = f(p_{em}, T_{amb}, T_{im}, w_{cmp}, u_{egr}) \tag{4.2}$$

$$\hat{p}_{em} = f(p_{im}, p_{amb}, T_{amb}, T_{im}, w_{cmp}, u_{egr}, u_{vgt}) \tag{4.3}$$

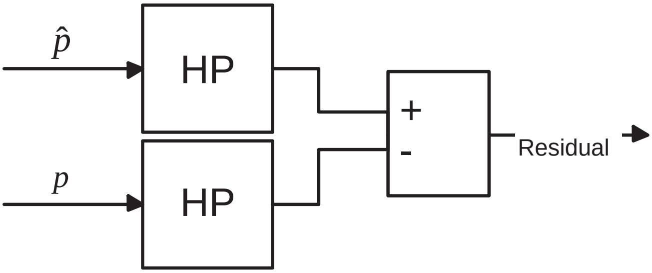


Figure 4.2: Schematic figure over the principle for the dynamic pressure tests.

VGT Test

This thesis consider a model based test for the VGT. This test compares a estimated and a measured value of the turbine speed. If they deviate, the test responds. The equation for estimating the turbine speed is seen in (4.4). This test is from now called VMT (VGTModelTest).

$$\hat{n}_{trb} = f(p_{amb}, p_{im}, T_{amb}, w_{cmp}) \tag{4.4}$$

4.3 Decision and Isolation Structure

As described earlier, it is sometimes hard to find out which faults affects a certain test. The tests should respond to faults at the signals used as input to each test, i.e sensors used for estimating signals and the measured signals. This is summarized in a decision structure, Table 4.1.

Table 4.1: Decision structure for manually designed OBD-system.

Testname	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
ATT				X					
SPT_{ae}	X	X							
SPT_{ai}	X		X						
SPT_{ei}		X	X						
VMT	X		X	X		X	X		
MAT			X		X	X		X	
DPT_i		X	X	X	X	X		X	
DPT_e	X	X	X	X	X	X		X	X
EFT			X		X	X		X	
EDT								X	
$SPAT_{ae}$	X	X							
$SPAT_{ai}$	X		X						

Table 4.1 shows which faults each test will respond to. Many of the tests will respond to several kinds of faults. This decision structure will lead on to a isolation structure as in Table 4.2. The isolation structure in the manually

designed OBD-system shows that it is not possible to isolate all faults. For example, w_{cmp} can not be isolated from p_{im}.

Table 4.2: Isolation structure for manually designed OBD-system.

	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
p_{amb}	X								
p_{em}		X							
p_{im}			X						
T_{amb}				X					
T_{im}			X		X	X		X	
w_{cmp}			X			X			
n_{trb}	X		X	X		X	X		
u_{egr}								X	
u_{vgt}	X	X	X	X	X	X		X	X

Chapter 5

Automatically Designed OBD-System

In this chapter a brief description is given of how the automatically generated diagnosis system is constructed from a Simulink model by DSAME. An OBD-system will be created and presented at the end of this chapter.

5.1 DSAME

Scania and Vehicular Systems at Linköping University have developed a method for designing OBD-systems in an automatic way. The process can be divided into five steps:

1. Structural analysis
2. Finding realizable residuals
3. Stability check of the residuals
4. Setting thresholds
5. Evaluation and selection of tests

These steps will be presented in the following sections and for a more detailed description see [2], [3] and [5].

5.1.1 Structural Analysis

The first step is to make a structural analysis of the model. The objective of structural analysis is to find ARRs that potentially can be used in diagnosis tests. This is done by finding relations between equations and variables. This is illustrated with Example 5.1.

Example 5.1
Consider a system M with two sensor signals y_1 and y_2, one actuator signal u, two states x_1 and x_2. The system is described by the model equations, e_{1-4}

$$M : \begin{cases} e_1 : & x_1 = u \\ e_2 : & x_2 = x_1 \\ e_3 : & y_1 = x_1 \\ e_4 : & y_2 = x_2 \end{cases}$$

This can be represented structually with a biadjacency matrix.

Equation	Unknown		Known		
	x_1	x_2	u	y_1	y_2
e_1	x		x		
e_2	x	x			
e_3	x			x	
e_4		x			x

(5.1)

The biadjacency matrix can now be used to find potential ARRs by using structural methods described in for example [6]. In this example, it can be seen that e_1 and e_3 can be used to form the ARR $y_1 - u = 0$, e_2, e_3 and e_4 can form $y_1 - y_2 = 0$. e_1, e_2 and e_4 can form the ARR $y_2 - u = 0$

However, it is not always as simple to find ARRs as in Example 5.1. Sometimes the equation system is to large that the calculations are impossible to do by hand. With structural analysis the ARRs can be found by some algorithms, see for example [6], but the calculation ways found with the algorithms are not always possible to do. The next step in DSAME is to find the ARRs in the ARR-set from structural analysis that can be realizable.

5.1.2 Finding Realizable Residuals

ARRs from the structural analysis are not necessary possible to realize. In fact many equations used in the engine model are not invertible. In DSAME, all derivatives are said to be not computable because of the problem to estimate derivatives in a noisy environment. Saturations, maps and other not strictly monotonous functions are also said to be not invertible. Residuals containing these kinds of expressions that must be inverted are removed, left is a set of realizable residuals.

5.1.3 Stability Check of Residuals

Realizability is a necessary but not a sufficient property for a residual to be usable. It must also be stable, at least for regions it is supposed to be used in. Investigation of stability for residuals can be made in several ways. Some are linearization or Lyaponov theory [7]. These both methods work for non-linear models but are difficult to implement in an automatic way. Because of this, in [3] an ad-hoc algorithm for stability check was implemented in DSAME. The algorithm regards following aspects:

Model drift. Check that the residual cross their own mean-value more than a certain number of times, depending on input data.

Feasibility. Check that the residual are inside feasible regions during the entire simulation.

Correlation. Check that cross-correlation between input and the residual is below a prespecified threshold.

Mean-error consistency. Check that the mean-value of the residual does not change too much with different data.

This method has been investigated in [3] and is assumed to work satisfactory. When realizable and stable residuals are found, it is possible to design tests by thresholding the residuals appropriately.

5.1.4 Thresholds

The test quantity is calculated by using (2.2) with N=2000. The thresholds for each test are then set by estimating the probability of false alarm for the test quantity that it is within a specific area when using fault-free measurements. In this thesis the probability for false alarm is 0.01% i.e.

$$P(J_2 < TQ(z) < J_1|\theta = 0) = 99.99\% \tag{5.2}$$

where θ is the fault level.

5.1.5 Evaluation and Selection of Tests

The final step is to select those tests which together give the best performance regarding detection and isolation. This selection is done by first making test candidate sets, i.e. pick those tests that can separate each fault from the other with the highest probability. The diagnosis system will be created by choosing the minimal set of tests that have an intersection with all the test candidate sets, i.e. the minimal set of tests which give maximal possible isolation. For further description, see [3].

5.2 Decision and Isolation Structure

When the design of the diagnosis system is done, the decision structure for the automatic designed OBD-system becomes as in Table 5.1 with corresponding isolation structure in Table 5.2. The isolation structure is decent but the system can e.g not isolate any other fault from p_{amb}. This can be explained by the decision structure, see Table 5.1. It is seen that p_{amb} is a signal in all tests and therefore can no other signals be isolated from p_{amb}.

Table 5.1: Decision structure of the automatic designed OBD-system.

Test	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
Test 1	X		X	X		X	X		
Test 2	X		X	X		X	X		
Test 3	X			X	X		X	X	X
Test 4	X				X	X	X	X	X
Test 5	X			X	X	X		X	X
Test 6	X	X	X	X	X		X	X	
Test 7	X	X	X		X	X	X	X	
Test 8	X	X		X	X		X	X	X
Test 9	X	X		X	X	X	X	X	
Test 10	X		X	X	X		X	X	X
Test 11	X		X		X	X	X	X	X
Test 12	X		X	X	X	X		X	X
Test 13	X	X		X	X			X	X
Test 14	X	X			X	X		X	X

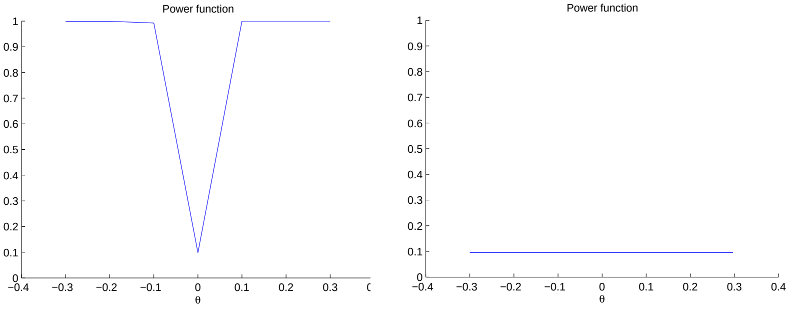
Table 5.2: Isolation structure of the automatic designed OBD-system.

	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
p_{amb}	X								
p_{em}	X	X			X			X	
p_{im}	X		X						
T_{amb}	X			X					
T_{im}	X				X			X	
w_{cmp}	X					X			
n_{trb}	X						X		
u_{egr}	X				X			X	
u_{vgt}	X				X			X	X

5.2.1 Correction of the Decision Structure

To improve the isolation the power functions for the residuals are investigated. Figure 5.1a shows a power function for a residual with high performance for a certain fault. The probability of false alarm is low compared to the probability to detect small faults. This can be seen in the figure because the probability for the test to respond is lowest for fault when $\theta = 0$. However, for the residuals of the automatic generated system, it was found that the power functions for some faults in several residuals appeared such as the power function in Figure

5.1b. The figure shows that the probability for the test to respond is equal even when a fault has occurred.



(a) Typical power function for a signal.

(b) A typical constant power function.

Figure 5.1: Two different power functions for residuals in the automatic designed OBD-system.

Thus the residual will not respond to the fault although it should due to Table 5.1. A deeper analysis shows that residuals with power functions as in Figure 5.1b will not be affected for larger faults either. To investigate this, gain-faults up to 10 times has been used, and it is not probable that faults in the sensors become larger. One explanation of the behavior is that the engine model has some equations where one of the input signals will not affect the output signal. By removing ’X’ in Table 5.1 where the power function is constant, the isolation structure can be improved. The adjusted decision structure is seen in Table 5.3 and the corresponding isolation structure is seen in Table 5.4. By comparing the new isolation structure in Table 5.4 with the old isolation structure given by Table 5.2, it can be seen that the isolabilty is significantly improved. For example, it is now possible to isolate some faults from faults in p_{amb}. This is possible because some tests now will not be affected by p_{amb}, see Table 5.3.

5.2.2 Model Complexity

When comparing the isolation structure of the automatic designed OBD-system given in Table 5.2 with the system generated in [3], the latter is better. The only difference between this thesis and the previous thesis is the engine model used as input to DSAME. The model used in this thesis is more complex since more details are modeled. Examples of these details are a model over the temperature exchange in the EGR-pipe, an extra pressure state has been used and that some maps has been extended to include more variables. The purpose of the extended complexity is to improve the model quality. A more exact model leads on to less model errors in the residuals and therefore

Table 5.3: Decision structure of the automatic generated system after correction.

Test	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
Test 1	X		X	X		X	X		
Test 2	X		X	X		X	X		
Test 3	X			X			X	X	X
Test 4	X				X	X	X		X
Test 5	X			X	X	X			X
Test 6	X	X	X	X	X		X	X	
Test 7		X	X		X	X		X	
Test 8	X	X			X		X	X	X
Test 9	X	X		X	X	X	X	X	
Test 10	X		X	X	X		X	X	X
Test 11	X		X		X	X	X		X
Test 12	X		X	X	X	X			X
Test 13	X	X		X	X				X
Test 14		X				X			X

Table 5.4: Isolability of the automatic generated system after correction.

	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
p_{amb}	X								
p_{em}		X							
p_{im}			X						
T_{amb}	X			X					
T_{im}					X				
w_{cmp}						X			
n_{trb}	X						X		
u_{egr}								X	
u_{vgt}									X

a better detectability. The drawback with a more complex model is that the isolability may decrease when using it with DSAME. More non-invertible equations have been introduced in the complex model, making it harder to find realizable residuals. An accurate model is not enough, the invertability of the equations has to be considered as well.

Chapter 6

Comparative Values

In this chapter the comparative values used when comparing the two diagnosis systems are described. From specific test cases, described in Chapter 7, the comparative values will be calculated. The result from the comparison can then be read in Chapter 8. In this thesis four comparative values will be used.

- Value of detectability level
- Value of isolability
- Value of detection time
- Residual performance

To make a comparison of detectability and isolability for two diagnosis systems possible, they must have equal probabilities of false alarm. The probability of false alarm has to be very low to avoid false alarm and to measure such low probability, a very large amount of process data is needed. Since these amounts are impossible to get for this thesis the probability will not be measured. Instead, the probability of false alarm for the two system are approximately set equal.

6.1 Value of Detectability Level

The value of detectability level describes the ability of the diagnosis system to detect a fault that has occurred. It can be measured in several different ways. In this thesis it is defined as the level for which a certain fault is detected by the diagnosis system. In practice, this means the fault level for which the diagnosis system responds.

6.2 Value of Isolability

Value of isolability describes the ability of the diagnosis system to correctly isolate a specific fault from other possible faults. The isolability, I_l, of a diagnosis system for a fault, F with a certain fault level, i, will in this thesis be measured by the number of faults it can not be isolated from. If all possible faults have equal probability to be chosen, the isolability can be calculated as

$$I_l(F_i) = \begin{cases} \frac{1}{(1+n)} & \text{If } F_i \text{ is in the diagnosis} \\ 0 & \text{Otherwise} \end{cases} \tag{6.1}$$

where n is the number of faults F_i can not be isolated from. The total isolability, I, for a fault F can then be calculated as

$$I(F) = \frac{1}{N} \sum_{i=1}^{N} I_l(F_i) \tag{6.2}$$

where $I_l(F_i)$ is the isolability for a fault with fault size i and N is the number of fault levels. The value of isolability is calculated for gain faults with the fault levels $i = 10\%$, $i = 20\%$ and $i = 30\%$. If the fault is not detectable for a given fault level, the isolability for this fault level is set to zero.

6.3 Value of Detection Time

Detection time describes the time from when a fault occurs to when it is detected for a detectable fault level. For non-detectable levels this time will go to infinity. Detection time is visualized in Figure 6.1. In the figure, a fault occurs after 240s. After 260s the test quantity is below the lower-threshold and the test responds. the detection time for this fault is then 20s. The detection time will then be normalized with the length of the time when the fault is present, In this thesis 180s, to get the value of detection time.

6.4 Residual Performance

Consider the residuals in Figure 6.2. These residuals are affected by faults in w_{cmp}.

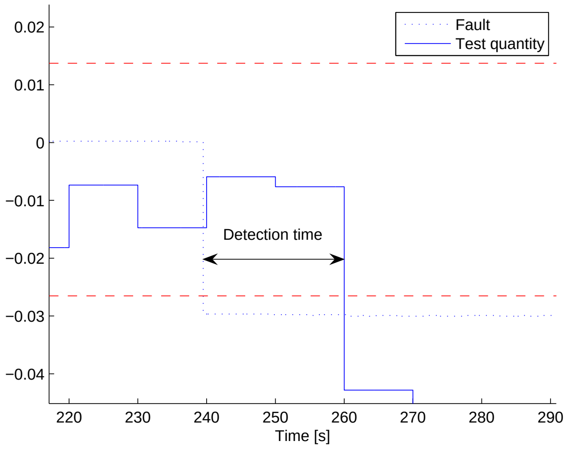
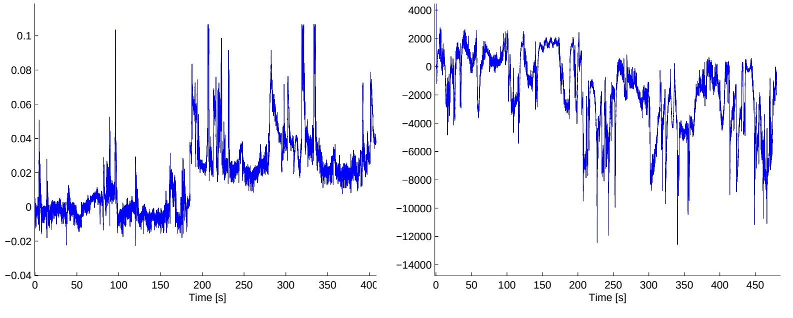


Figure 6.1: Detection time for a certain fault.



(a) Residual 1.

(b) Residual 2.

Figure 6.2: Two different residuals when a 30% gain fault in w_{cmp} occures at time 195s.

Without using any thresholds, residual 1 has intuitively a better performance than residual 2. The difference in mean value before and after the fault has occurred, is larger relative to the signal noise. Residual 1 has therefore higher probability of detection for a given probability of false alarm. To compare residual performances for certain faults, without setting a threshold, it would be convenient with a function describing the intuitive feeling of ”better performance”. Consider the distributions of a fault free process and a faulty process in Figure 6.3.

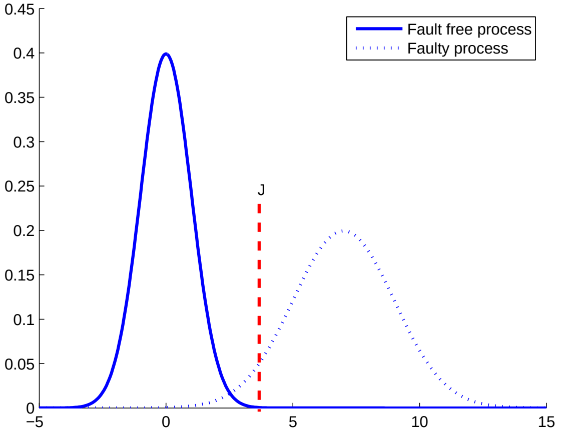


Figure 6.3: Distributions of residuals of a fault free process and a faulty process.

A measure of how a residual, $r(t)$, performs at a fault, F_i, is the probability to detect the fault given observations collected when F_i is actually present, $data(F_i)$ i.e.

$$P(r(t) > J|data(F_i)) \tag{6.3}$$

If two residuals have thresholds corresponding to the same probability of false alarm, the performances of the residuals can be compared. The probability to detect the fault can either be calculated numerically or analytically by assuming the residuals to be normal distributed. These calculations give curves which can be compared to each other.

When assuming the residuals to be normal distributed, $N(\mu_1, \sigma_1)$ for the fault free residual and $N(\mu_2, \sigma_2)$ for the faulty residual. For a certain probability of false alarm α, a threshold will be set as $J = k\sigma_1$. Using the mean value change $\mu = \mu_2 - \mu_1$, (6.3) can be calculated as

$$\begin{aligned} P(r(t) > J|data(F_i)) &= \frac{1}{\sqrt{2\pi}\sigma_2} \int_{k\sigma_1}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma_2^2}} dx = \\ &= \Big/ y = \frac{x-\mu}{\sqrt{2}\sigma_2}, dx = \sqrt{2}\sigma_2 dy \Big/ = \\ &= \frac{1}{2\sqrt{\pi}} \int_{\frac{k\sigma_1-\mu}{\sqrt{2}\sigma_2}}^{\infty} 2e^{-y^2} dy = \\ &= \frac{1}{2} erfc(T') \text{ where } T' = \frac{k\sigma_1 - \mu}{\sqrt{2}\sigma_2} \end{aligned} \tag{6.4}$$

Due to this, a residual performance function (RPF) will be used to evaluate residual preformance. RPF, for a given residual with distribution $N(\mu_1, \sigma_1)$ in the fault free case and $N(\mu_2, \sigma_2)$ when fault has occurred can be calculated as

$$RPF(k) = \frac{1}{2} erfc(T') \text{ where } \begin{cases} T' = \frac{k\sigma_1 - \mu}{\sqrt{2}\sigma_2} \\ \mu = \mu_2 - \mu_1 \end{cases} \tag{6.5}$$

This function is easy to calculate since the `erfc`-function is computable in Matlab. Expectation value, μ_i, and standard deviation, σ_i, are easy to estimate. However, in some cases the assumption of normal distribution is poor. Then (6.3) will be numerical calculated with different J. This will be called numerical RPF. Drawbacks with numerical RPF is that the threshold J must be estimated and the computational load will be higher.

If having two residuals, residual i is said to perform better than residual j if

$$RPF_i(k) > RPF_j(k) \text{ for all } k. \tag{6.6}$$

If 6.6 does not hold for all k, the detection probability slightly above 0,5 is the most interesting part. It is for that level a test quantity based on the mean value is above the threshold.

RPFs calculated for the residuals in Figure 6.2 are seen in Figure 6.4. The residual performance for residual 1 is better than for residual 2. This make sense and follows the intuitive feeling of which residual who performs the best.

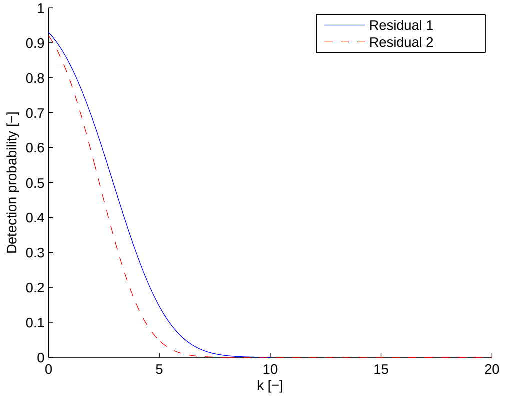


Figure 6.4: Residual performance for the residuals in Figure 6.2.

Alternative measures of residual performances

The power function defined earlier also measures residual performance. The difference between the two methods is that RPF is a function of the threshold and the power function is a function of the fault level. This makes it easy to use RPF for faults without a fault level, e.g. when sensors are stuck. RPF can also be used to investigate how sensitive the detectability is for changes in the thresholds. The derivative $\frac{dRPF(k)}{dk}$ can be used for this purpose.

Another possible measure of residual performance similar to RPF is ROC (Receive Operating Characteristics) which is used in detection theory. ROC and RPF measure almost the same but present the result in a different way. ROC has the false alarm probability at the x-axis while RPF has the threshold there. Since the probability of false alarm depend on the thresholds, the difference between the two methods might seem insignificant. But, the interesting area in RPF, where $3 \lesssim k \lesssim 10$ corresponds the area $99.5 \lesssim \alpha \lesssim 99.9999$. Thus, RPF has for the purpose a better scale at the x-axis. ROC is further described in [4].

Chapter 7

Test Cases

In this chapter, possible measurement methods to get test cases are described. To get relevant test cases faults have to be simulated. How the fault simulation will be done is described in this chapter. Also how the control system affects the diagnosis is discussed. An evaluation is done in the end of the chapter showing it necessary to have a control system in the evaluation. Therefore, all the measurements methods that have been investigated includes a control system.

7.1 Measurements

There are several ways to collect data to be able to evaluate the diagnosis system. In this thesis three methods have been investigated. Using a test bench, i.e. connecting an ECU to a computer with an engine model, using a real vehicle and using PC-simulation, i.e. simulating both the engine and the ECU on a computer. These methods will be described further in following sections.

7.1.1 Using Test Bench

A test bench is a simulation tool for the engine. A Simulink model is run by a computer in real time. To make the real time execution possible, the engine model is simplified. An ECU unit is electronically connected to the model. The big advantages using the test bench is the possibility to do all simulations and measurements with sensor and actuator faults automatically. A drawback is that many of the tests in the manually designed OBD-system are done when the vehicle starts and this is not simulated in the test bench, and can therefore not be evaluated. Another problem is that the engine model in the test bench is unexecutable for high torques.

7.1.2 Using a Real Vehicle

By driving a real vehicle on a test course, the manually designed OBD-system can be evaluated in a more complete way. Faults can be implemented by connecting a computer to the ECU. With a computer, signals and variables can be logged to evaluate the both systems off-line. Advantages with the method are that the measurements are realistic, and that disturbances and model errors have a realistic affect. Drawbacks are that it is time consuming and laborious, and that it is not possible to reiterate experiment. To test the OBD-systems on process faults like leakage, things has to be destroyed, they can not be simulated. A similar method of using a real vehicle is to use a test cell. Additional advantages with the test cell is that it is easy to reiterate experiment, many disturbances can be controlled and more extreme tests can be done. However, the test cells are not available for this thesis.

7.1.3 Using PC-Simulation

The point with using complete PC-simulation instead of using the test bench is that the control system is implemented in a PC, instead of in the ECU. Therefore real time execution is not needed, and the simulation can be done with exactly the same model as the one used to design the diagnosis system in DSAME. Testing a model based diagnosis system on a model from which the system is derived is equivalent with having an ideal model of the engine and measure at a real vehicle. This focus the evaluation to the design method, model errors should not affect the result. Another advantage is that process faults, e.g. leakage, can be modeled and tested. A drawback with this method is that it is very complex to connect the control system to an engine model.

7.1.4 The Choice in This Thesis

All methods has been investigated for use. The model simplifications in the test bench affected the model to much. The PC-simulation was not possible to implement inside the scope of this master thesis. The measurements has therefore been done using a real vehicle. It has been quite laborious but the resulting measurements are of high quality.

7.2 Fault Simulation

To make this thesis relevant the treated faults should be faults that really may occur. However, the faults that can really occur is not completely investigated and very difficult to find out. What is known is that the automatically generated system should perform well for sensor and actuator faults because of the fact that they are explicitly modeled in the engine model. Other faults affecting the engine can maybe be detectable but not isolable without constructing a model over the fault. This thesis treats sensor and actuator faults, it also tries to investigate some cases of leakage.

7.2.1 Simulation of Faults in a Vehicle

In a vehicle it is possible to connect to the ECU when driving. The ECU contains transformation curves from voltage to physical quantity for the sensors and vice versa for the actuators. These curves are described by some interpolation points which can be adjusted to manipulate the sensor values into the control system. Figure 7.1 shows the idea by simulating faults in the vehicle. In this thesis three different types of faults are simulated. These are gain faults, bias faults and stucked sensors.

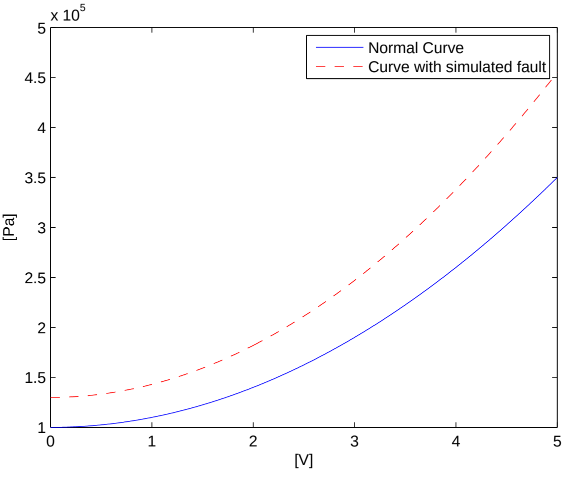


Figure 7.1: By changing the electrical curve a fault can be simulated.

7.2.2 Leakage Simulation

To evaluate the diagnosis system for leakage real holes have to be implemented in a vehicle. Since resources for these kinds of experiments are not available, it will be done in a model off-line. Therefore the affect of for example model error will not be seen.

Leakage Model in the Inlet Manifold

In the engine model, air mass flow into the cylinder, $\dot{m}_{cyl}$ is modeled as

$$\dot{m}_{cyl} = w_{cmp} + \dot{m}_{egr} \tag{7.1}$$

where w_{cmp} is the mass flow through the compressor and $\dot{m}_{egr}$ is the amount of EGR-gases. When there is a leakage in the inlet manifold, the air mass flow into the cylinder will be changed as

$$\dot{m}_{cyl} = w_{cmp} + \dot{m}_{egr} - \dot{m}_{leak} \tag{7.2}$$

where the leakage of air mass flow, $\dot{m}_{leak}$, can be modeled as

$$\dot{m}_{leak} = k(p_{im} - p_{amb}) \tag{7.3}$$

where k is a constant describing the properties of the hole. The signal p_{amb} is an input value to the engine model and p_{im} is an available state in the engine model. If k is introduced as a signal in the model used for diagnosis system generation the isolability of leakage can be investigated. By changing the level of k, different leakage flow can be simulated. The modeled leakage size is then available in the model during the simulation.

7.3 Affect of the Control System

In earlier evaluations, e.g [3], the affect of the control system has not been considered. As seen in Example 7.1 the control system should not affect the residuals. Instead, problems might occur when a fault makes so the control system controls the engine to other working points where the model is bad or to working points where the faults become undetectable. Also other control strategies like adaption (see section 4.2.3) might cause problems for the diagnosis system.

Example 7.1
Assume the system $G(s)$ with a control system $F(s)$ as in Figure 7.2. The system $G(s)$ has one input u and one output y.

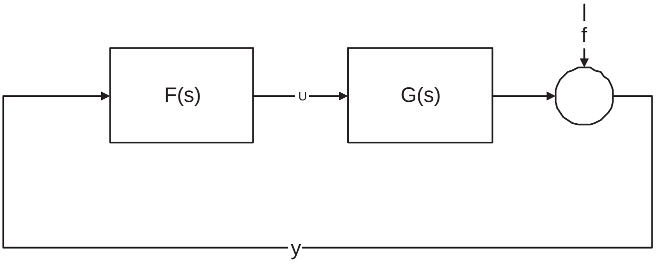


Figure 7.2: A general system.

An ARR can be constructed as

$$r = y - G(s)u$$

If having a control system, the fault f on the output-signal y would affect the residual such as

$$r = y - G(s)u = (G(s)u + f) - G(s)u = f$$

This means that the control system will not affect the residual. This result suggests the possibility to ignore the affect of the control system.

To evaluate this, residuals when faults are simulated in the truck are compared to residuals when faults are simulated off-line. Two aspects have been considered for the residuals when a fault occurs, difference in standard deviation and difference in mean value change.

7.3.1 Difference in Standard Deviation

Difference in standard deviation tends to become larger when faults are simulated off line for gain faults but tends to be equal for bias faults. This can be explained by that the measurement noise is amplified when the gain fault is added off line.

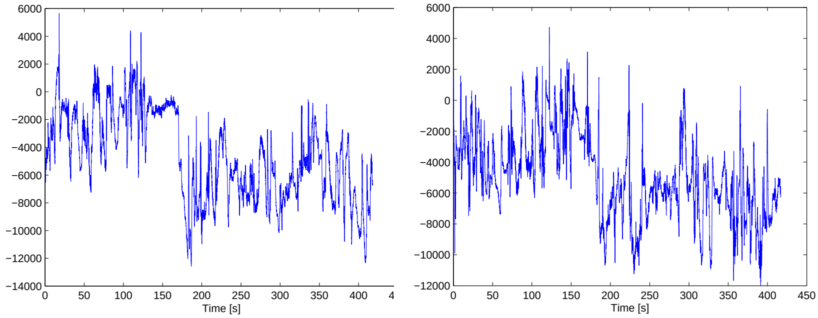
7.3.2 Difference in Mean Value Change

The difference in mean value changes in per cent are summarized in Table 7.1.

Table 7.1: Differences in mean value changes for different gain faults and fault levels.

	Fault level		
	10%	20%	30%
w_{cmp}	-20%	-4%	5%
p_{im}	-2%	8%	7%
p_{em}	1%	-3%	34%

For two different faults, 10% fault in w_{cmp} and 30% fault in p_{em}, there is a large difference that can not be explained by noisy mesurements. For the fault in w_{cmp} there are no differences for larger faults which indicates that the difference for 10% is caused by model error which affects more for small errors. This hypothesis makes even more sense by investigation of the residual in Figure 7.3, the residual is noisy and affected by model errors.



(a) Fault simulated off-line without a control system

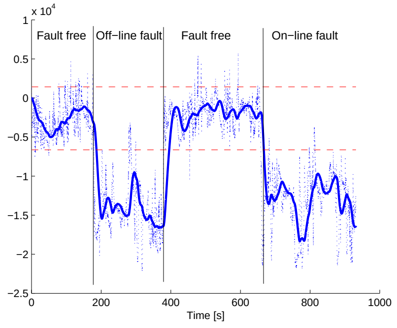
(b) Fault simulated on-line with control system

Figure 7.3: Residuals there a 10% gain fault in w_{cmp} are simulated.

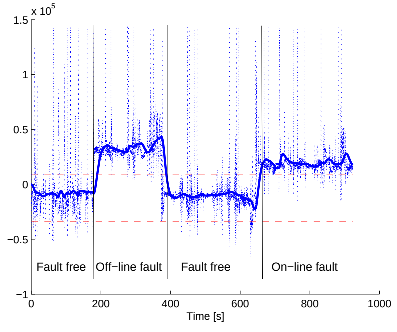
A residual with a large fault in p_{em} is seen in Figure 7.4. One explanation for the large mean value changes is that the measurement of p_{em} is very important for the EGR and VGT control. Therefore it makes sense that large faults in this sensor affects the control system the most and therefore affects the residuals the most. For the other faults in Table 7.1, the mean value changes can be explained by noise and model errors.

7.3.3 Conclusions

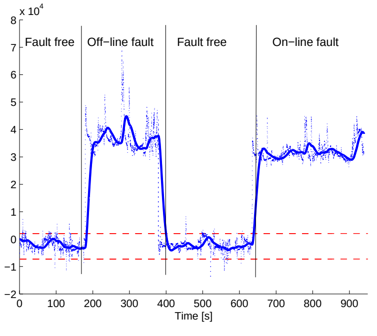
The affect of the control system must be considered when evaluating diagnosis systems. This is because the control system might control the engine to working points where the residuals will not respond equally as when simulating the fault off-line and without a control system. The affect tends to be larger for large faults in p_{em}.



(a) Residual with 30% gain fault in w_{cmp}.



(b) Residual with 30% gain fault in p_{em}.



(c) Residual with 30% gain fault in p_{im}.

Figure 7.4: Residuals where faults are simulated in different ways. When $200s < Time < 400s$ the fault is simulated off-line without a control system. When $670s < Time < 930s$ fault is simulated on-line with a control system. The thin dashed curves are raw residuals and the bold lines are the test quantities.

Chapter 8

Analysis and Results

In this chapter the main results from the comparison of the two diagnosis system and an analysis of the results are presented. The comparative values described in Chapter 6 are used and to get the diagnosis systems as equivalent as possible to be able to make a fair evaluation, the thresholds in the tests are set so the two diagnosis systems have the same probability for false alarm. The thresholds are calculated as described in Section 5.1.4.

In Section 8.1, the correctness of the assumption of residuals approximated as normal distributions is investigated. Section 8.2.1-8.2.4 will present the diagnosis performance for faults in pressure sensors, mass flow sensor, temperature sensors and EGR. These results are summarized in Table 8.1 on page 47. Leakage is evaluated for the automatic designed OBD-system and will be presented in Section 8.3. Due to the poor performance for the manually designed OBD-system compared to the automatic designed OBD-system shown in Section 8.2, an analysis of specific tests in the manually designed OBD-system will be presented in Section 8.4. The affect of control strategies like adaption may decrease the performances on an OBD-system, therefore consequences of using adapted values as inputs to an OBD-system are evaluated in Section 8.5.

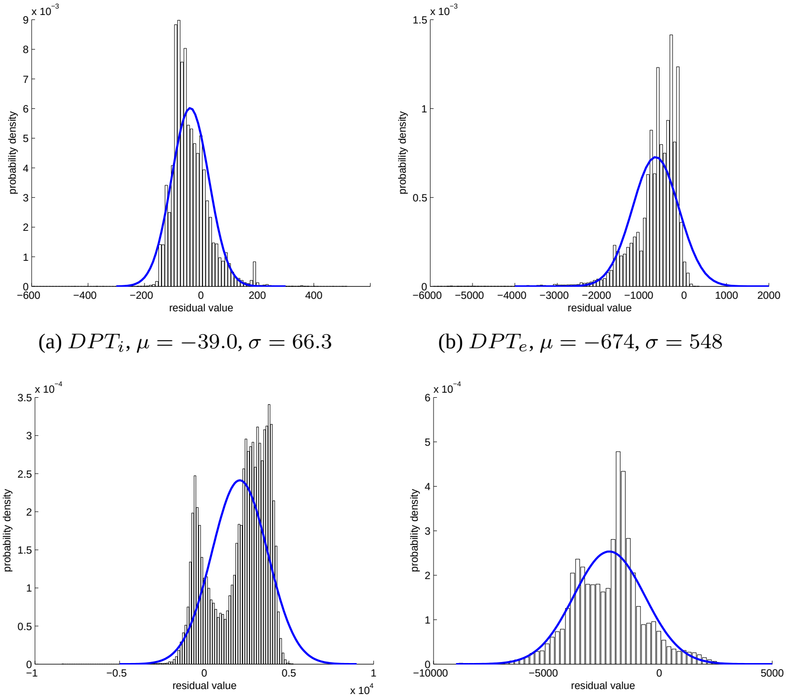
8.1 Assumption of Residuals Approximated as Normal Distribution

The correctness of the RPF, see Section 6.4, is dependent on how good the assumption of normal distribution of the residuals are. Therefore this must be evaluated.

To evaluate the correctness of assumption of normal distribution of a residual, a histogram of the collected data is plotted together with the probability density function (PDF) corresponding to the estimated expectation value and standard deviation of the residual. If the histogram follows the PDF, the approximation is good, if it does not, the approximation is bad.

The normal distribution will be evaluated for fault free residuals, for residuals with gain faults, and for residuals when a sensor is stuck. This investigation is important when using RPFs for evaluation of residual performance. The more normal distributed the residual is, the more correct are the RPFs. The evaluation of the normal distribution assumption is done for many faults, but only faults in p_{im} are presented, the results are similar for the other sensors. The residual performance will be evaluate for the residuals in DPT_i, DPT_e, VMT and the best residual from $DSAME$ for each fault.

Fault Free Residuals



(a) DPT_i, $\mu = -39.0$, $\sigma = 66.3$ (b) DPT_e, $\mu = -674$, $\sigma = 548$

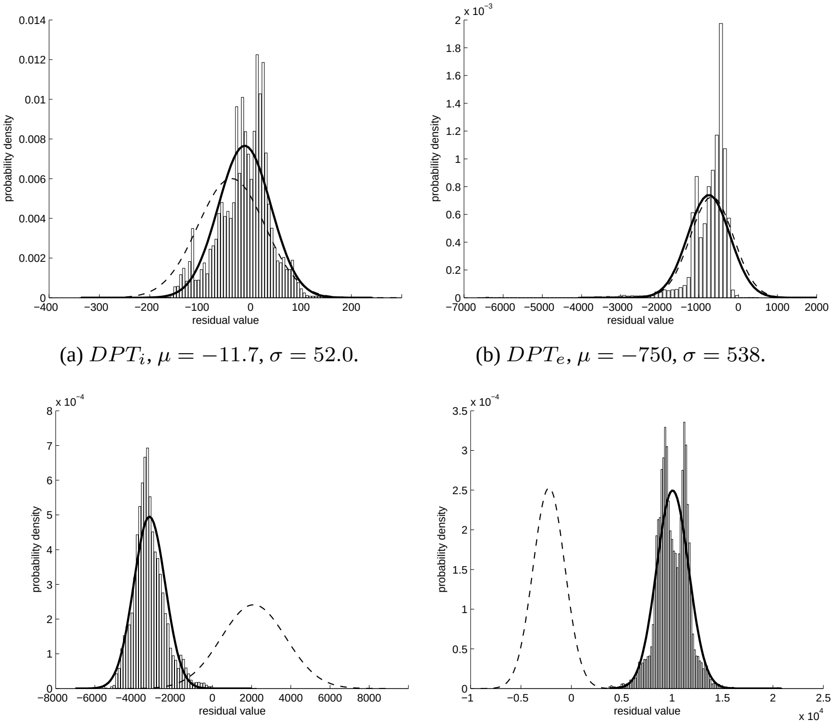
(c) VMT, $\mu = 2100$, $\sigma = 1650$ (d) DSAME-residual, $\mu = -2210$, $\sigma = 1570$

Figure 8.1: Comparison of measured histogram and the power density function from estimated $N(\mu, \sigma)$ for fault free residuals.

The distributions for fault free residuals are shown in Figure 8.1. For the DPT_i, and the $DSAME$-residuals the normal distribution approximation seems to be good, for the DPT_e-residual, the approximation is a bit worse.

The distributions all have one sided tails, e.g. DPT_i to the right and DPT_e to the left. The VMT-residual is bimodal distributed. After investigation of the residual and the EGR-damper position, this seems to depend on the fact that the residual changes expectation value when the damper closes, which probably depends on model errors.

Gain Fault



(a) DPT_i, $\mu = -11.7$, $\sigma = 52.0$.

(b) DPT_e, $\mu = -750$, $\sigma = 538$.

(c) VMT, $\mu = -3210$, $\sigma = 806$.

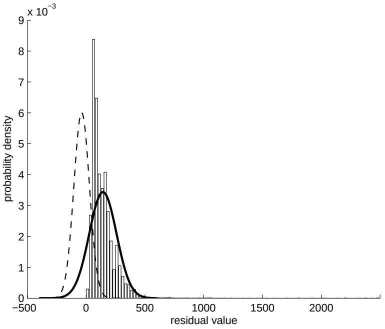
(d) DSAME-residual, $\mu = 10000$, $\sigma = 1600$

Figure 8.2: Comparison of measured histogram and the power density function from estimated $N(\mu, \sigma)$ for residuals with 30% gain fault in p_{im}. The dashed lines are the distribution from the fault free case for each residual.

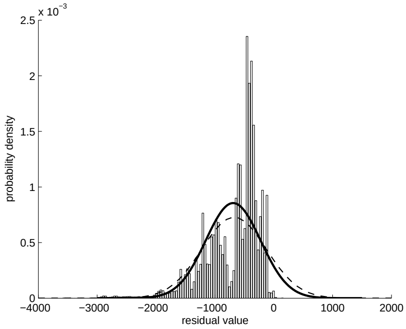
The distributions for residuals when a 30% gain fault in p_{im} is present are shown in Figure 8.2. The shape of the distributions of the DPT_i-residual and the DPT_e-residual have small changes compared to the fault free case. But, since also the changes in expectation value and standard deviation are small, the tests would not respond to the fault. The VMT-residual and the DSAME-residual have larger changes in expectation value and the tests will respond to the faults. The shape of the distributions of the residuals have changed.

The VMT-residual is more normal distributed than previous and the standard deviation has decreased. The gain fault seems to make the normal distribution approximation better. The DSAME-residual is less normal distributed and has became bimodal because of the fault.

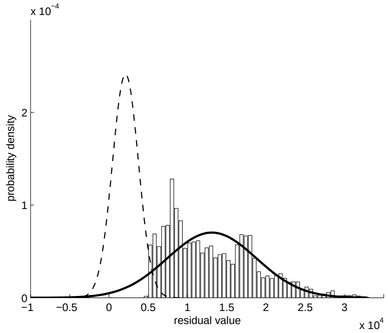
Sensor got Stuck



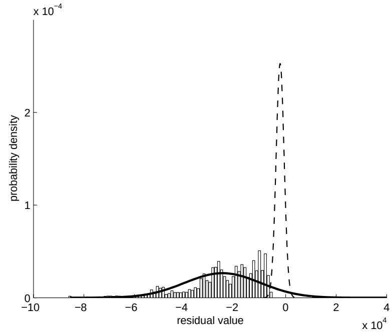
(a) DPT_i, $\mu = 143$, $\sigma = 46$.



(b) DPT_e, $\mu = -689$, $\sigma = 466$.



(c) VMT, $\mu = 13100$, $\sigma = 5670$



(d) DSAME-residual, $\mu = -24900$, $\sigma = 14800$.

Figure 8.3: Comparison of measured histogram and the power density function from estimated $N(\mu, \sigma)$ for residuals when p_{im} is stuck. The dashed lines are the distributions from the fault free case for each residual.

The distributions for residuals when p_{im} is stuck are shown in Figure 8.3. The expectation values of the DPT_i, VMT and DSAME-residuals have changed because of the fault. The changes are larger than for the gain faults shown in Figure 8.2. The expectation value of DPT_e has not changed significantly. The shape of the distributions of the residuals are all affected by the fault, more than for the gain fault. For the residuals with large changes in expectation values, the standard deviations have increased. This is because when a sensor got stuck it will in some points measure correct value and in some points not. The residuals will therefore depend on the working points. Thus, the normal distribution approximation is worse for faults when sensors got stuck than for other faults.

Conclusions about Normal Distribution

None of the residuals are perfect Gaussian processes, but, as wil be seen in the following sections, the conclusions will not change when using the assumption instead of doing a numerical calculation. Another important notice is that all residuals (more or less) have one sided tails. This means that the disturbances of the residuals are larger at one side and that, on these sides, the thresholds can be set closer to the mean value of the residual than the others. It also means that the performance will be different for positive and negative faults.

8.2 Comparative Values

8.2.1 Fault in Pressure Sensors

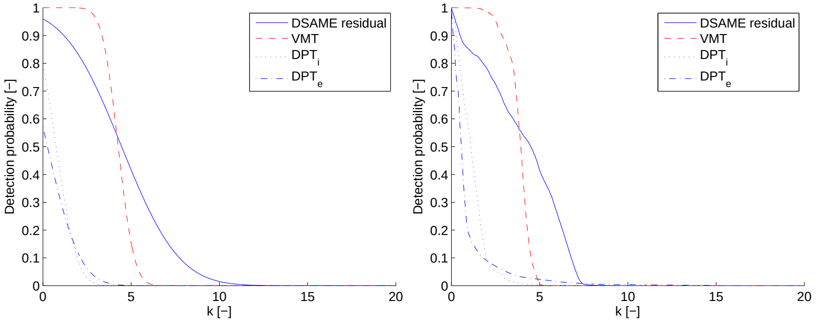
Results of detectability, isolability and detection time are summarized in Table 8.1.

Detectability The detectability levels for both systems are 0.1 for both p_{im} and p_{amb}. Faults in p_{em} can not be detected with the manually designed OBD-system if the fault level is lower than 0.5. The automatic designed OBD-system has a detectability level of 0.2 for this fault.

Isolability The isolability is a lot better for the automatically designed OBD-system. It is 1 for both positive and negative faults in p_{im} and p_{amb}, and 0.33 for negative and positive faults in p_{em}. The manually designed OBD-system has the isolability 0.22 for positive and negative faults in p_{im} and p_{amb}, and 0 for faults in p_{em}. This is poor and is because the lack of tests responding to small faults on the pressure sensors in the manually designed OBD-system.

Detection Time The detection time is small, 0.08, and equal between the two OBD-systems for almost all the detected faults. Only p_{em} for the manually designed OBD-system has worse, 1. This is because the manually designed OBD-system detects the fault when the engine is shut off.

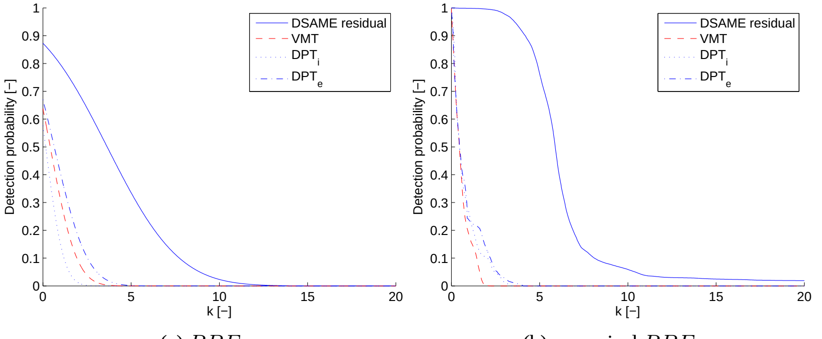
Residual Performance The RPFs for 30% gain fault in p_{im} is seen in Figure 8.4. The RPFs for the DSAME-residual and the VMT-residual crosses each other. The residuals perform about equally around the detection probability 0.5. The residuals can therefore be said to perform equal for the fault. If a high false alarm probability can be accepted, the VMT-residual performs the best. The residuals DPT_i and DPT_e both have very low performances for the fault.



(a) RPF. (b) Numerical RPF.

Figure 8.4: Residual performance for 30% gain fault in p_{im}.

The RPFs for 30% gain fault in p_{em} is shown in Figure 8.5. The DSAME-residual performs well for the fault since it has a high detection probabilitiy for high thresholds. It is also not very sensitive for changes of the threshold. The residuals from the exisiting OBD-system perform poorly and are not usable for detection of this fault. Altough the DPT_e-residual performs the best among the residuals from the manually designed OBD-system, its performance is still bad.

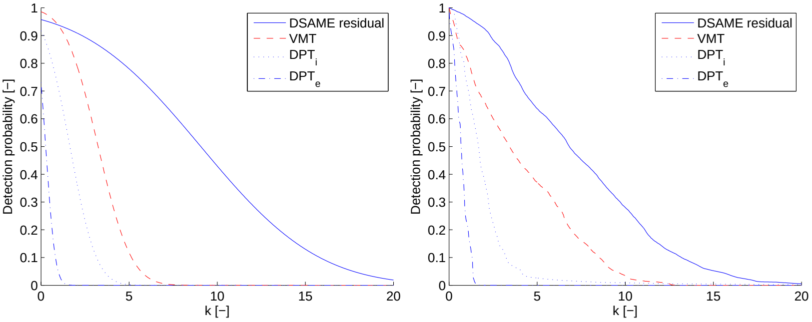


(a) RPF. (b) numerical RPF.

Figure 8.5: Residual performance for 30% gain fault in p_{em}.

The residual performance when p_{im} is stuck at atmosphere pressure is shown in Figure 8.6. The DSAME-residual performs the best. The VMT-

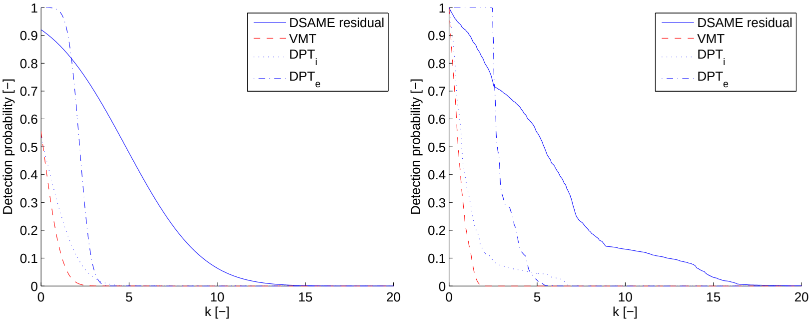
residual also performs well and is clearly the best among the residuals from the manually designed OBD-system.



(a) RPF. (b) numerical RPF.

Figure 8.6: Residual performance when p_{im} is stuck.

The residual performance when p_{em} is stuck is seen in Figure 8.7. The DSAME-residual performs the best, at least for high thresholds. The DPT_e-residual performs the best among the residuals from the manually designed OBD-system, which it was designed to do.



(a) RPF. (b) numerical RPF.

Figure 8.7: Residual performance when p_{em} is stuck.

8.2.2 Fault in Mass Flow Sensor

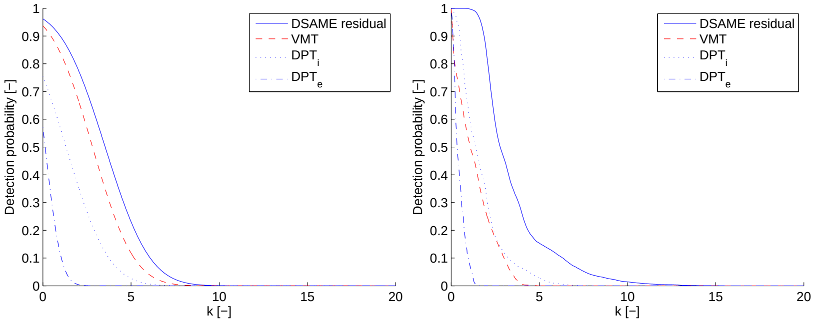
Results of detectability, isolability and detection time are summarized in Table 8.1.

Detectability Both systems will have the same detectability level for positive faults in the mass flow sensor, 0.1. The Automatic designed OBD-system is also able to detect negative faults on w_{cmp} at 0.1, the manually designed OBD-system only at 0.2.

Isolability Since there are only a few faults responding to small faults in w_{cmp} in the manually designed OBD-system, its isolability is poor, 0.22 for positive faults and 0.15 for negative faults. The automatic designed OBD-system has better isolability, 1 for both positive and negative faults. This means that the automatic designed OBD-system is able to completely isolate positive and negative faults from 0.1 and above in w_{cmp}.

Detection Time The detection time is small, 0.08, and equal between the two OBD-systems for the detected faults.

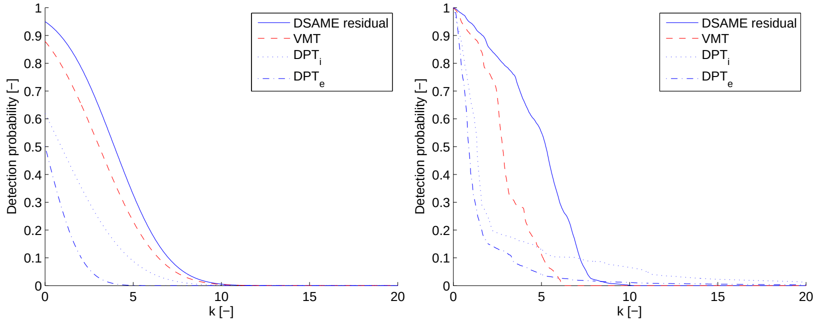
Residual Performance The RPFs for 30% gain fault in w_{cmp} is seen in Figure 8.8. It is clear that the automatic generated residual has the best performance. The VMT-residual is best among the existing residuals. Also DPT_i has high propability to detect the fault but has low performance. DPT_e has very low probability to detect the fault and the performance is bad. For this fault, there is a difference between the numerical and the analytical calculation for the VMT-residual. This is because the residual is not normal distributed when it is affected by fault in w_{cmp}.



(a) RPF. (b) Numerical RPF.

Figure 8.8: Residual performance for 30% gain fault in w_{cmp}.

The residual performance when w_{cmp} is stuck is seen in Figure 8.9. The DSAME-residual performs the best. The VMT_e-residual performs the best among the residuals from the manually designed OBD-system. Also for this fault, there is a difference between the numerical and the analytical calculation, at least for low probabilities of detection. The difference is even bigger than for the w_{cmp} gain fault. An explanation of this was presented in Section 8.1.



(a) RPF. (b) Numerical RPF.

Figure 8.9: Residual performance when w_{cmp} gis stuck

8.2.3 Fault in Temperature Sensors

Detectability Faults in temperature sensors, T_{amb} and T_{im} is hard to detect for the automatic designed OBD-system. This is because a fault in a temperature sensor has to be very large to affect the gas flow. Such large faults are obviously unreasonable before a test from the gas flow model is able to detect it. This problem is exemplified in Example 8.1 The manually designed OBD-system would be able to detect faults in T_{amb} if the duplication test had low thresholds.

Isolability Since the faults can not be detected, the isolability is zero.

Detection Time Since the faults can not be detected, the detection time goes to infinity.

Example 8.1
Imagine that the temperature are $20°C$ in the intake manifold and T_{im}got a positive gain-fault on 30%, $\theta = 1.3$, then the affect will be

$$T_{im_{fault}} = \theta T_{im} - T_{im} = 6° \tag{8.1}$$

The corresponding fault in Kelvin will be

$$\theta^* = \frac{T_{im_{fault}} + T_{im}}{T_{im}} = 1.0220 \tag{8.2}$$

This fault level is to small to detect for the automatic OBD-system.

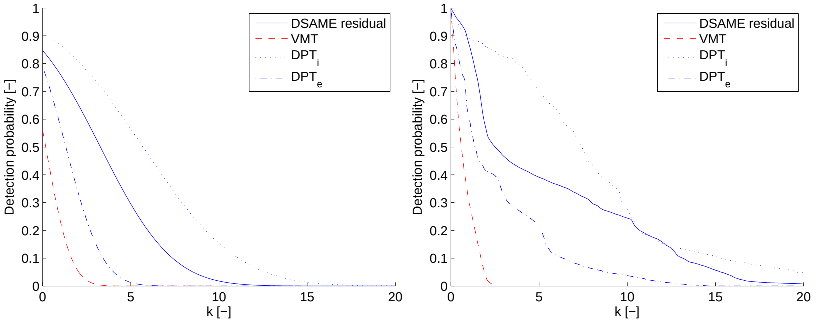
8.2.4 Fault in EGR

Detectability Faults in the EGR-actuator seems to be hard to detect. Both the automatic OBD-system and the manually designed OBD-system cannot find faults in the EGR-actuator up to 0.3. One reason to that the automatic designed OBD-system does not detect faults in EGR is because the signal does not affect the model behavior so much.

Isolability Since the faults can not be detected, the isolability is zero.

Detection Time Since the faults can not be detected, the detection time goes to infinity.

Residual Performance The residual performance when the EGR-valve is stuck open is seen in Figure 8.10. The DPT_i-residual performs the best. The DSAME residual performs decent. The DPT_i performs much better than the DSAME-residual. This is because the DPT_i-model depends more on u_{EGR}.



(a) RPF. (b) Numerical RPF.

Figure 8.10: Residual performance when u_{EGR} got stuck.

8.2.5 Conclusions from the Comparative Values

The comparative values in Section 8.2.1-8.2.4 can be used to decide which diagnosis system that is the best. By taking the mean value of the comparative values detectability, isolability and detection time and calculate a comparative

Table 8.1: Summary of the result from measurements. System 1 is the automatic designed OBD-system and system 2 is the manually designed OBD-system. Component describes which component has been investigated. Fault type describes which kind of fault the component has. Detectability level describes the value of detectability level for each fault. Isolability describes the value of isolability for each fault type. Detection time describes the value of detection time for each fault type.

Component	Fault type	Detectability level		Isolability		Detection Time	
		DSAME	man	DSAME	man	DSAME	man
p_{im}	positive gain	0.1	0.1	1	0.22	0.08	0.08
p_{im}	negative gain	0.1	0.1	1	0.22	0.08	0.08
p_{em}	positive gain	0.1	0.5	0.33	0	0.08	1
p_{em}	negative gain	0.1	0.5	0.33	0	0.08	1
p_{amb}	positive gain	0.1	0.1	1	0.22	0.08	0.08
p_{amb}	negative gain	0.1	0.1	1	0.22	0.08	0.08
w_{cmp}	positive gain	0.1	0.1	1	0.22	0.08	0.08
w_{cmp}	negative gain	0.1	0.2	1	0.15	0.08	0.08
T_{im}	positive gain	-	-	0	0	-	-
T_{im}	negative gain	-	-	0	0	-	-
T_{amb}	positive gain	-	-	0	0	-	-
T_{amb}	negative gain	-	-	0	0	-	-
u_{EGR}	positive gain	-	-	0	0	-	-
u_{EGR}	negative gain	-	-	0	0	-	-
Mean value[a]		0.1	0.2	0.83	0.16	0.08	0.31

[a] Mean value for the detected faults, i.e p_{im}, p_{em}, p_{amb} and w_{cmp}.

scalar, C. The system with lowest C will be the system that performs the best. This scalar, C can be calculated as

$$C = 1 + mean(\text{detectability level}) - mean(\text{Isolability}) + mean(\text{detection time}) \tag{8.3}$$

where the mean values will be taken over the detected faults, i.e p_{im}, p_{em}, p_{amb} and w_{cmp}. By using 8.3 the manually designed OBD-system got the value $C = 1.35$ and the automatic designed OBD-system got the value, $C = 0.35$. The values showing that the automatic designed OBD-system performs the best.

8.3 Leakage

Leakage was evaluated only for the automatic designed OBD-system. The simulated leakage was implemented as in section 7.2.2. Table 8.2 shows the result. The leakage is in the table described by k, mean leakage flow and maximum leakage flow during the simulation.

Table 8.2: Detection for leakage.

k	Mean leakage flow	Max leakage flow	Detection
$0.5 \cdot 10^{-6}$	11%	17%	No
$1 \cdot 10^{-6}$	17%	27%	Yes
$2 \cdot 10^{-6}$	25%	39%	Yes

Table 8.2 shows that holes leading to 17% mean leakage flow was possible to detect, at 25% it might be possible to isolate the leakage. Experiences shows that the manually designed OBD-system needs over 30% mean leakage for a long time for detection. As described in Section 7.3, it may also be possible to detect leakage on an engine with control system.

Diagnosis Performance for Leakage

The results showed that it was possible to detect leakage from a level about 17% of the air mass flow into the cylinder. To investigate possibilities to isolate leakage the fault is introduced in the original decision structure. The decision structure is seen in Table 8.3 and the theoretical isolability is seen in 8.4. The isolability for leakage is complete. The isolability of p_{em} has decreased. It can not be separated from leakage. By using the new decision structure, leakage will be isolated for a detection level about 25% mean leakage.

8.4 Analysis of Manually Designed OBD-system

Due to the poor performance shown by the manually designed OBD-system compared to the automatic designed OBD-system, an analysis of specific tests will be discussed in this section. The analysis will consider which faults that can be detected with the specific tests and how good they are. Duplication tests and signal in range check test

Table 8.3: Decision structure of the DSAME system when leakage has been modeled.

Test	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}	$leakage$
Test 1	X		X	X		X	X			
Test 2	X		X	X		X	X			
Test 3	X			X			X	X	X	X
Test 4	X				X	X	X		X	X
Test 5	X			X	X	X			X	X
Test 6	X	X	X	X	X		X	X		X
Test 7		X	X		X	X		X		X
Test 8	X	X			X		X	X	X	X
Test 9	X	X		X	X	X	X	X		X
Test 10	X		X	X	X		X	X	X	
Test 11	X		X		X	X	X		X	
Test 12	X		X	X	X	X			X	
Test 13	X	X		X	X				X	X
Test 14		X				X			X	X

Table 8.4: Isolability of the DSAME-system when leakage has been modeled.

	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}	$leakage$
p_{amb}	X									
p_{em}		X								X
p_{im}			X							
T_{amb}	X			X						
T_{im}					X					
w_{cmp}						X				
n_{trb}	X						X			
u_{egr}								X		
u_{vgt}									X	
leakage										X

will not be considered because these tests work as expected and does not need to be evaluated.

8.4.1 Static Pressure Tests

It can be concluded that it is not necessary to use both static pressure sensor tests, SPT (SPT_{ai}, SPT_{ae} and SPT_{ei}), and static pressure sensors adaption tests, $SPAT$ ($SPAT_e$ and $SPAT_i$), because these groups of tests are similar and will respond to the same faults. Therefore $SPAT$ should be removed because SPT consists of one more test and will isolate the specific pressure-sensor when a fault occurs.

8.4.2 VGT Test

The VGT model test-residual origins from a gas flow model through the compressor. This is a similar part of the engine model used to the residual in test 1 in the DSAME-system. The difference is that residual 1 compares measured and modeled w_{cmp} where the estimated w_{cmp} is calculated by n_{trb}. In VMT, n_{trb} has been modeled by w_{cmp} and necessary simplifications to make the residual executable in real time has been done. The lower residual performance for the VMT-residual seen in Figures 8.4-8.10 is probably caused by the simplifications made to make the residual executable in real time. If the residual from test 1 is executable in real time without simplifications it is better than the VMT-residual.

8.4.3 Dynamic Pressure Sensors Tests

The residuals in the dynamic pressure sensor tests are created by the difference between measured and modeled pressure values. They were designed to detect dynamic faults in the pressure sensors. As expected, among the tests in the manually designed OBD-system, DPT_e seems to perform the best for dynamic faults in p_{em} and DPT_i seems to perform the best for for dynamic faults in p_{im}. However, their performances are still low and the DSAME-residuals detects the faults better, but the DPT_i has excellent performance when the EGR-valve is stuck, even better than the DSAME-residuals.

8.4.4 Mass Flow Adaption Test

In this thesis the mass flow sensor, w_{cmp}, is adapted. The adaption was performed every 4 minutes to make the evaluation of the mass flow adaption test possible using limited amount of time. The modeled air mass flow ,which is used for adaption, in this thesis is linear dependent of p_{im} and T_{im}. Faults of 30% gain fault in p_{im} and T_{im} was implemented in the truck and the adaption behaved as expected. The adaption was affected equivalent for each of the faults and the adaption reached its threshold after 7-9 adaptions. Therefore the faults are detectable by the test but the detection time depends on how often the adaption is made. More about the mass flow adaption in Section 8.5.

8.4.5 EGR-damper Test

From the measurements of the residual created by the EGR-damper test, EDT, it is clear that the only fault affecting the residual is if u_{EGR} gets stuck. The residual is seen in Figure 8.11. The figure shows three different simulations with different faults. The residuals in Figure 8.11a and Figure 8.11b seem to be unaffected by any fault. If we instead look at Figure 8.11c, this residual seems to be affected by the fault. For an explaination, consider the following example for a linear system.

Example 8.2
Consider the system in Figure 8.12

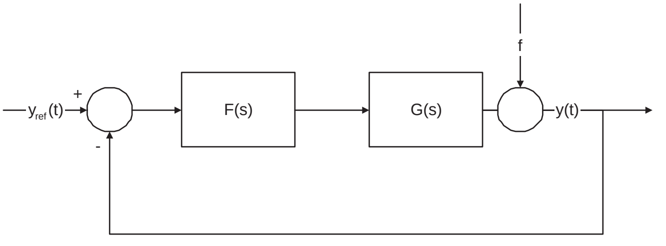


Figure 8.12: A system like the one in the EGR-damper test.

A residual is created by the control error, $r(t) = y_{ref}(t) - y(t)$ as in the EGR-damper test. A constant fault f is added to the output $y(t)$ simulating a bias fault in the sensor. The transfer function for the residual become

$$\begin{aligned}
& r(t) = y_{ref}(t) - y(t) \\
& \quad Y = F(s)G(s)(Y_{ref} - Y + f) + f = F(s)G(s)Y_{ref} - F(s)G(s)Y - f \\
\Rightarrow & \quad Y = \frac{F(s)G(s)}{1 + F(s)G(s)}Y_{ref} + \frac{1}{1 + F(s)G(s)}f \\
\Rightarrow & \quad R = -\frac{1}{1 + F(s)G(s)}Y_{ref} + \frac{1}{1 + F(s)G(s)}f
\end{aligned}$$

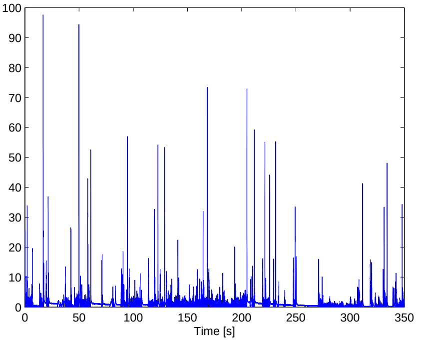
If F(s) is a PI-controller, $F(s) = K_P + \frac{K_I}{s}$, the transfer function from f to R is

$$G_{rf}(s) = \frac{1}{1 + F(s)G(s)} = \frac{s}{s + sK_P G(s) + K_I G(s)}$$

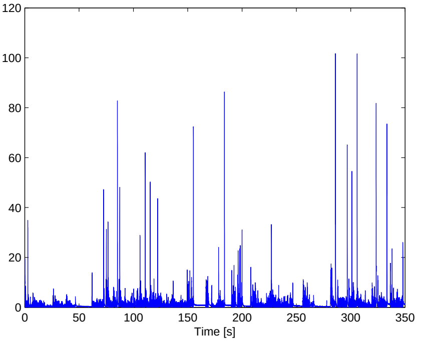
and when f is a constant error, the step response will go to

$$\lim_{s \to 0} sG_{rf}(s)\frac{f}{s} = \lim_{s \to 0} s\frac{f}{s}\frac{s}{s + sK_P G(s) + K_I G(s)} = 0$$

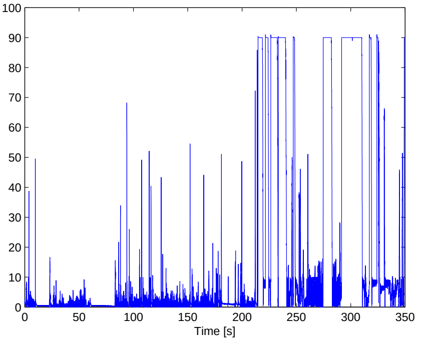
due to the final value theorem see e.g. [7]. A constant fault f is therefore not strongle detectable [8], i.e can not be detected static, in the residual based on the control error.



(a) Fault free residual.

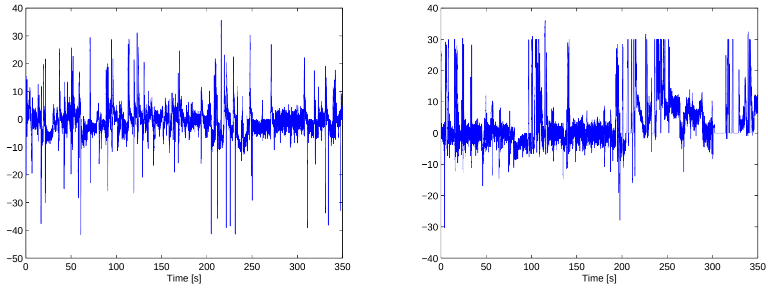


(b) Gain fault in u_{EGR}.



(c) u_{EGR} got stuck.

Figure 8.11: The residual for EGR-damper test with different fault-simulation. The faults are added at time=220s.



(a) fault free. (b) 30% gain fault in w_{cmp} occurs when t=195s.

Figure 8.13: The residual in EGR-flow test.

The EDT could not detect small gain faults in the EGR-damper, since the controller compensates for the fault as in Example 8.2. But, when the damper is stuck, the reference value is not reachable. Faults that are detectable in this test are those who make the reference not reachable, such as when the damper is stuck or some serious fault on u_{EGR}.

8.4.6 EGR-flow Test

From the logged data of the residual created by the EGR-flow test, EFT, it is clear that some of the faults significantly affects the residual. The faults that should affect the residual are w_{cmp}, p_{im} and T_{im}. The residual in the fault free case and a 30% gain fault in w_{cmp} are seen in Figure 8.13. It can be seen in the figure that the residual will be affected by a fault in w_{cmp} but the mean value of the residual seems to be zero. This can be explained in the same way as for EDT. The residual for EFT is similar as in Example 8.2 but the controller also has a feed forward from the reference value. When both the reference and actual values are modeled from other measured signals, faults in these signals will affect the residual and make it harder to reach the reference value. The feed forward seems to make the controller slower when faults have occurred. Slower means that the controller will take longer time to reach the reference-value because the feed forward is incorrect. It can be seen in Figure 8.13 that in both cases the residual will go to zero but in the case with fault in w_{cmp} it will take much longer time. EFT does not seem to detect all EGR-faults. Small faults in the EGR-damper are compensated by the controller and therefore not strongly detectable. On the other hand, small faults are probably not necessary to detect, as long as the EGR-flow can be controlled to the reference value. More serious faults as when the EGR-damper is stuck is not either detected. Probably because the ERG-flow reference is reachable by changing the exhaust pressure with the VGT.

8.5 Consequences When Using Adapted Values in OBD-systems

When having an OBD-system, a problem can be that control strategies like adaption can affect the tests. Therefore this will be investigated in this section. The adaption of some sensors described in Section 4.2.3 has some possible affects on the diagnosis that has been investigated. Assume that residuals in a diagnosis system using sensor values as input where two of them are $sensor_1$ and $sensor_2$. Assuming that $sensor_1$ is adapted to $sensor_2$, the following scenarios can then be considered:

- Residuals sensitive to faults in both $sensor_1$ and $sensor_2$ responds when $sensor_1$ gets faulty. It might be possible that the detectability decreases if $sensor_1$ is maladjusted by the adaption.
- Residuals sensitive only to fault in $sensor_1$ might respond to faults in $sensor_2$ because of the adaption. This might decrease the isolability.
- Residuals sensitive to fault in $sensor_2$ but not to $sensor_1$ should not be affected by the adaption of $sensor_1$.
- Residuals sensitive to neither fault in $sensor_1$ nor $sensor_2$ should not be affected by the adaption at all.

Mass Flow Adaption

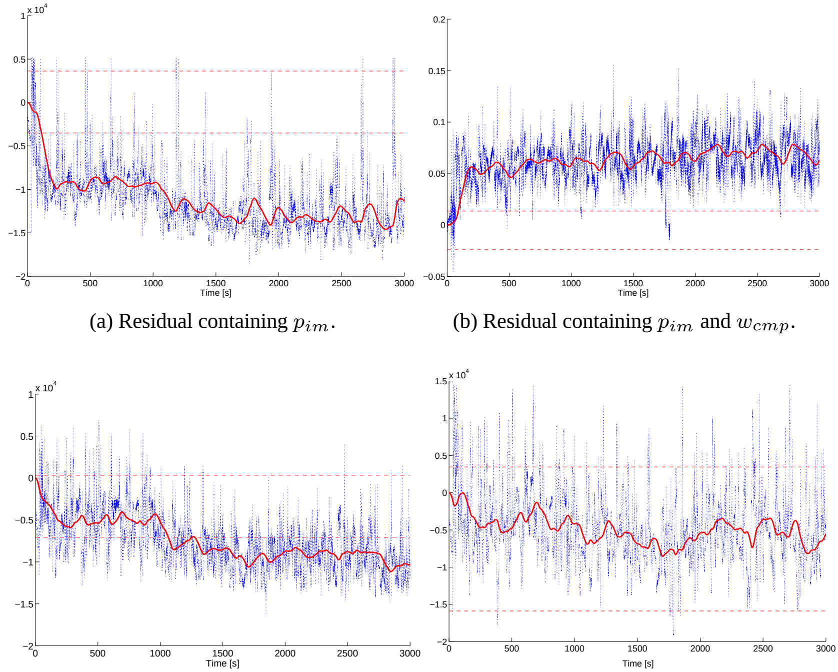
The hypothesis that adaption affects the diagnosis performance will be investigated for the adaption of the mass flow sensor. Consider the residuals in Figure 8.14. A 40% gain fault in p_{im}, which probably will affect the adaption, has occured when Time = 0, several adaptions has then been done during the measurement. The residuals are containing different combinations of p_{im} and w_{cmp} corresponding to the scenarios mentioned above.

- The residual in Figure 8.14a is containing p_{im} but not w_{cmp} and is responding to the fault. The change at time=1000s of the measurement may be caused by the adaption.
- The residual in Figure 8.14b is containing both p_{im} and w_{cmp}. This residual is responding to the fault and the fact that w_{cmp} is included in the residual does not disturb the detection.
- The residual in Figure 8.14c is containing w_{cmp} but not p_{im} and is not responding to the fault before the adaption. But after some adaptions, the residual changes and goes below the threshold.
- The residual in Figure 8.14d is containing none of the signals, w_{cmp} nor p_{im}. It does not respond to the fault in p_{im} and is not affected by the adaption.

A similar result can be found for the other signals affecting the adaption, T_{im}.

This result indicates two things.

- The detection is not affected so much by the adaption. Probably because the model used in the adaption is different to the model used in the residuals. The adaption of the mass flow sensor does therefore not compensate for the changes in the residuals.



(a) Residual containing p_{im}.

(b) Residual containing p_{im} and w_{cmp}.

(c) Residual containing w_{cmp} but not p_{im}.

(d) Residual containing none of w_{cmp} and p_{im}.

Figure 8.14: Residuals affected by adaption. A fault in p_{im} is present through the measurement. The bold lines in the middle of the residuals are the mean valued test quantities.

Table 8.5: Decision structure of an DSAME system after numerical correction, see Section 5.2.1, and correction for using of adapted value of w_{cmp}.

Test	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
Test 1	X		X	X	X	X	X		
Test 2	X		X	X	X	X	X		
Test 3	X			X			X	X	X
Test 4	X		X		X	X	X		X
Test 5	X		X	X	X	X			X
Test 6	X	X	X	X	X		X	X	
Test 7		X	X		X	X		X	
Test 8	X	X			X		X	X	X
Test 9	X	X	X	X	X	X	X	X	
Test 10	X		X	X	X		X	X	X
Test 11	X		X		X	X	X		X
Test 12	X		X	X	X	X			X
Test 13	X	X		X	X				X
Test 14		X	X		X	X			X

Table 8.6: Isolability table of an DSAME system after numerical correction, see Section 5.2.1, and correction for using of adapted value of w_{cmp}.

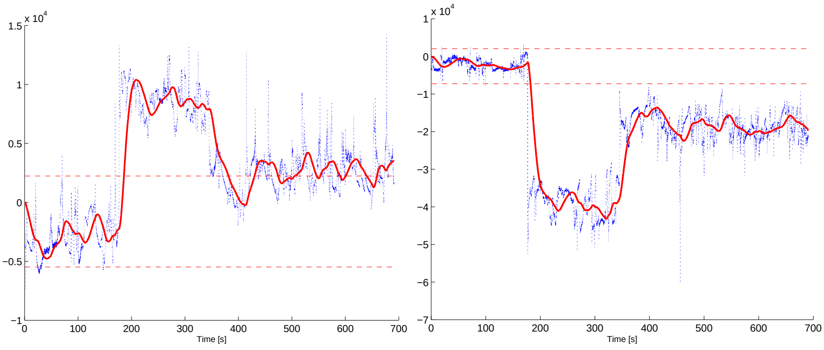
	p_{amb}	p_{em}	p_{im}	T_{amb}	T_{im}	w_{cmp}	n_{trb}	u_{egr}	u_{vgt}
p_{amb}	X								
p_{em}		X			X				
p_{im}			X		X				
T_{amb}	X			X					
T_{im}					X				
w_{cmp}		X			X	X			
n_{trb}	X						X		
u_{egr}								X	
u_{vgt}									X

- The isolability is affected by the adaption. The possible faults affecting the residual in Figure 8.14c is increased by w_{cmp} and the isolation structure does therefore change. The decision structure in Table 8.5 and the isolation structure in Table 8.6 shows decreased isolability compared to earlier.

However, if the fault has occurred and become detectable, it is possible to isolate the fault before the adaption is done. The problem is that the diagnosis statement can be changed when the adaption then is done. If the fault is an incipient fault, i.e. gradually developed from no fault to larger and larger, it might be more difficult to do a correct isolation. This problem could be solved by having some tests using unadapted values that are only run after a detection, to increase the isolability.

Pressure Sensors

The pressure sensors, p_{im} and p_{em}, are adapted after p_{atm} when the engine is shut off. Thus, as long as the p_{atm} measures correctly, the others will be measuring atmosphere pressure correct. Other faults as gain fault or changed sensor characteristics can not be eliminated by the adaption.



(a) Residual including p_{amb} and p_{em}.

(b) Residual including p_{amb}, p_{em} and p_{im}.

Figure 8.15: Residuals during pressure adaption. The fault in p_{amb} occurs after 180s. Then the adaption is simulated so after 370s, p_{im} and p_{em} are adapted.

The p_{atm}-sensor has no plausibility diagnosis at all. Therefore, if a bias fault occures in the sensor it will be propagated to the other pressure sensors and then to w_{cmp} through the adaption. This behavior is tested in the truck and some typical residual behavior from the DSAME-system is seen in Figure 8.15. The residuals first change mean values when the fault in p_{atm} occurs at time = 180s. When p_{im} and p_{em} are adapted to p_{amb} at time = 370s. For both residuals the mean value changes decrease which makes the faults harder to detect. For the residual in Figure 8.15a, the residual almost gets below the threshold after the adaption. Thus, the adaption might decrease the detectability but the faults seems to be still detectable. In some cases, the fault may become not detectable. The isolability decreases when residuals where p_{amb} is not included respond to the fault because of the adaption. As in the case with the adaption of the mass flow sensor, this could be solved by having some tests using unadapted values as input to increase the isolability. If adapted values are used as input to the test it seems to be a good idea to supervise the adaption (such as SPT and $SPAT$).

Chapter 9

Conclusions and Future Work

In this thesis an automatic designed OBD-system of the gas flow of a diesel engine has been designed using DSAME. The system was quite easy to design which makes it easy to redesign it if the engine model changes. The corresponding parts in a manually designed OBD-system have been identified and a comparison between the two systems has been done. The results show advantages and disadvantages with the two OBD-systems. The main conclusions are:

- The tests in the automatic designed OBD-system have higher residual performance than the manually designed OBD-system for almost all the evaluated faults. The comparative scalar which weighted together detectability, isolability and detection time, and should be as low as possible, become 0.35 for the automatic designed OBD-system and 1.35 for the manually designed OBD-system. This result together with the residual performance show that the automatic designed OBD-system is the best.

- Both OBD-systems have problems to detect faults in u_{egr} and the temperature sensors.

- With the ad hoc approach to design a diagnosis system used in the manually designed OBD-system, components might be tested several times and the faults for which a test responds to is often unknown. This makes it hard to evaluate the test and even harder to isolate the faults.

- OBD-systems can be used with adapted values as input. The adaption does not seem to affect the detectability. Whereas, the isolability is decreased a bit.

- It was shown in the thesis that leakage is possible to isolate with DSAME which indicates that also other faults can be detected if they are modeled. The manually designed OBD-system can not detect leakage.

9.1 Future Work

This thesis has shown that a diagnosis system with good performances can be designed by DSAME. Even better then the manually designed OBD-system. It has also discussed possible improvements that can be made. Work that may be needed is:

- It should be possible to improve the model to get a model better fitted for model based diagnosis instead of total engine simulation as it is today. This should lead to a model based diagnosis system even better than the one created in this thesis.
- The automatic generated OBD-system should be implemented in the real engine. The performance in real time is not investigated in this thesis and therefore needs to be evaluated.

References

[1] C. Eriksson. Mean value modelling of a poppet valve egr-system. Master's thesis LiTH-ISY-EX-3543, Vehicular Systems, Linköpings universitet, Linköping, Sweden, June 2004.

[2] L. Eriksson. Structural algorithms for diagnostic system design using simulink models. Master's thesis LiTH-ISY-EX-3601, Vehicular Systems, Department of Electrical Engineering, Linköpings universitet, Linköping, Sweden, January 2004.

[3] G. Arrhenius H. Einarsson. Automatic design of diagnosis systems, optimizing isolation and computational load. Master's thesis UPTEC F 1401-5757, Uppsala University, Uppsala, Sweden, Mars 2005.

[4] S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998.

[5] J. Nilsson K. Krigsman. Code generation for efficient realtime execution of diagnostic residual generators. Master's thesis UPTEC F 1401-5757, Royal Institute of Technology, Uppsala, Sweden, December 2004.

[6] M. Krysander. *Design and Analysis of Diagnosis Systems Using Structural Methods*. PhD thesis, Linköpings universitet, June 2006.

[7] L. Ljung and T. Glad. *Reglerteori -Flervariabla och olinjära metoder*. Number 6547 in ISBN 91-44-03003-7. Studentlitteratur, Linköpings universitet, 2 edition, 2003.

[8] M. Nyberg and E. Frisk. *Model Based Diagnosis of Technical Processes*. Studentlitteratur, Linköpings universitet, 2 edition, 2006.

[9] R. Reiter. A theory of diagnosis from first principles. *Artif. Intell.*, 32(1):57–95, 1987.

[10] J. Ritzén. Modelling and fixed step simulation of a turbo charged diesel engine. Master's thesis LiTH-ISY-EX-3442-2003, Linköpings Universitet, SE-581 83 Linköping, 2003.

Notation

Abbreviations

ARR	Analytical redundancy relation
ATT	Ambient temperature test
DPT	Dynamic pressure test
DSAME	Diagnostic structural analysis and modeling execution toolbox
ECU	Engine control unit
EDT	EGR-damper test
EFT	EGR-flow test
EGR	Exhaust gas recirculation
MAT	Massflow adaption test
MSO	Minimally structural overdetermined
OBD	On-board diagnosis
SCR	Selective catalyst reduction
SPAT	Static pressure adaption test
SPT	Static pressure sensor test
VGT	Variable geometry turbo
VMT	VGT model test
VOT	VGT overspeed test