

DIAGNOSIS OF EVAPORATIVE LEAKS AND SENSOR FAULTS IN A VEHICLE FUEL SYSTEM

Ingemar Andersson and Erik Frisk

*Dept. of Electrical Engineering, Linköping University
Linköping, Sweden
Email: ingemar@isy.liu.se, frisk@isy.liu.se*

Abstract: This paper describes a vacuum-decay based evaporative leak detection procedure for vehicle fuel systems. A physical model for an evaporative system is proposed containing parts for fuel evaporation, leakage flow and canister flow. Two methods for detecting evaporative leakages based on the model is presented. Both methods can detect a 0.5 mm diameter leakage in a laboratory environment.

Keywords: purge system, fault diagnosis, fault detection, model based diagnosis

1. INTRODUCTION

According to regulations for emissions from vehicles, fuel vapor leakage from the fuel tank must be detected. Fuel vapor is always generated in the fuel tank, the amount depends on ambient conditions like temperature and movement of the tank. Filling fuel also causes extra vapor to be generated. The fuel vapor may cause an over pressure that may push vapor out of the tank. Also, as fuel is consumed an under-pressure develops in the tank and it is required to level the fuel-tank gas pressure with ambient pressure at all times.

The most common way of handling fuel vapor from the fuel tank is to ventilate the fuel tank through a carbon canister. The canister is connected in one end to the fuel tank and the other end is open to the ambient. A purge valve connects the tank side of the canister to the intake manifold of the engine. The canister is purged from hydrocarbons when the purge valve is opened, causing a back-flow of air through the canister and into the engine.

There are two main principles for detecting leakages, vacuum decay and pressure decay principles. With the vacuum decay principle, an under-pressure is created in the fuel tank compared to the ambient pressure and the decay of the pressure difference is monitored and analyzed. The pressure decay principle creates an over-pressure to the fuel tank and the pressure difference is monitored and analyzed. A typical component setup for a vacuum decay principle is a pressure sensor and a closing valve (diagnostics valve), and for a pressure decay principle a separate pump. Currently the detection requirements move to smaller and smaller leakages, for

vehicles with model year 1996 and later, leakages as small as 0.040" (1mm)(*California's OBD-II Regulation, 1997*) in diameter must be detected. As of year 2000, the requirements is tightened and detection of leakages as small as 0.02" (0.5mm) is required (*California's OBD-II Regulation, 1997; Majkowski et al., 1999*).

This paper presents a model based approach for detecting leakage of fuel vapor and two methods of using the model. The system configuration is the commonly used vacuum decay type, based on a fuel tank pressure sensor, a diagnostics closing valve and the purge valve. The model developed and used in the detection algorithm is kept as simple as possible, while still maintaining its physical interpretation and good enough accuracy to meet required diagnosis performance.

2. MODEL

A principle sketch of the fuel evaporation system and its components is shown in Figure 1. The diagnosis system described in this work is model based and the only signal available is delivered by the pressure sensor located in the fuel tank. Since there is no hardware redundancy, a model describing the pressure signal is needed to detect and isolate the two faults considered, leakage in the fuel evaporative system and sensor bias.

Denote the absolute pressure in the tank with p_{ft} . The major processes that influences the fuel tank pressure is gas flow through the carbon canister into the engine and the environment, gas flows through leaks in the fuel

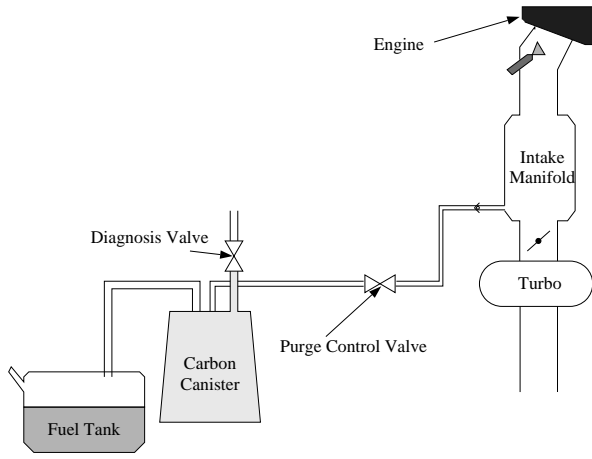


Fig. 1. The fuel evaporation system

tank, and fuel evaporation. Now, mathematical models of these processes will be derived:

The pressure change in the fuel tank is assumed to be proportional to the net gas flow into the tank, i.e. the main model equation is given by:

$$\dot{p}_{ft} = k(\dot{m}_b - \dot{m}_{tk} - \dot{m}_l) \quad (1)$$

where \dot{m}_b is the gas flow due to fuel evaporation, \dot{m}_{tk} gas flow through the carbon canister, and \dot{m}_l the gas flow due to leakages in the fuel tank.

Fuel evaporation

Gasoline evaporates from the liquid surface into the gas in the tank, resulting in a pressure rise. As long as there is a difference between the saturation pressure and the partial pressure of fuel vapor, the evaporation will continue. A simple model can be written as:

$$\dot{m}_b = k'_1(p_f^0 - p_f) \quad (2)$$

where p_f^0 is the fuel saturated pressure and p_f the fuel partial pressure. The constant k'_1 is the evaporation time constant which is dependent on e.g. the fuel tank temperature. Here, k'_1 is assumed to be constant but unknown.

Gas flow to the carbon canister

The gas flow through the carbon canister and the purge valve can be described by:

$$\dot{m}_{tk} = k'_2 D(p_{ft} - p_a) + k'_3 P f(p_{man}, p_{ft}) \quad (3)$$

where D and P are boolean variables indicating if the diagnosis or the purge valve is open (1) or closed (0), and p_a is the ambient pressure and p_{man} is the intake manifold pressure. An analytical expression for the flow into the engine $f(p_{man}, p_{ft})$ is not needed since the purge valve will be closed during the diagnosis procedure.

Leakage flow

For the gas flow rates that result from reasonable sized leakages, the leakage flow is assumed to be proportional

to the pressure difference over the leakage hole. The proportionality constant is then proportional to the leakage area. The model equation describing the leakage flow is then:

$$\dot{m}_l = k'_4(p_{ft} - p_a) \quad (4)$$

Complete Model

The process is equipped with one sensor measuring the difference pressure between tank and ambient pressure. Therefore, the measurement equation, including a model for the bias fault, is given by:

$$y = p_{ft} - p_a + b \quad (5)$$

where the variable b parameterizes the constant bias fault.

From now on, it will be assumed that the ambient pressure, leakage area, and bias fault are constant in time during the data collection, i.e. $\dot{k}'_4 = 0$, $\dot{b} = 0$, and $\dot{p}_a = 0$. Define k_i as $k_i = k k'_i$, and insert the flow models (2), (3), and (4) into the pressure equation (1):

$$\dot{y} = -(k_1 + k_2 D + k_4)(y - b) + k_1(p_f^0 + p_{air,t} - p_a) - k_3 P f(p_{man}, p_{ft}) \quad (6)$$

The term $k_1(p_f^0 + p_{air,t} - p_a)$ is still unknown since the partial air pressure in the tank $p_{air,t}$ is unknown. To use this model in the diagnosis algorithm described later, additional assumptions on this term is made.

If no leakage is present, $p_{air,t}$ will be constant when both valves are closed. If a leakage is present or the diagnostics valve is open, $p_{air,t}$ will adjust so that the total pressure inside the tank is leveled with the atmospheric pressure. During a pressure change $p_{air,t}$ will not be constant. However, the model will not consider the dynamics so $p_{air,t}$ denotes only the final value of the partial air pressure from this point.

The model is a first order differential equation with a simple solution. In this work it is only interesting to study the cases when the purge valve is closed ($P = 0$). The solution for the model with a closed purge valve is

$$y(t) = C_1 e^{-at} + C_2$$

where

$$C_2 = b(k_1 + k_2 D + k_4) + k_1(p_f^0 + p_{air,t} - p_a) \\ a = k_1 + k_2 D + k_4$$

The parameter C_2 is the final value of the exponential function $y(t)$. The final value is affected by both the pressure sensor offset and the partial air pressure in the fuel tank. For diagnostic purposes it is necessary to have the final value only dependent of the partial pressure of air. This can be accomplished by compensating the pressure reading for the sensor bias level. The parameter a represents the time constant of the exponential function and a larger leakage will cause a to be higher.

Finally, since the diagnosis algorithm is working with sampled data, a physical, time-discrete version of (1) is needed. Here, a forward Euler approximation of the time derivative is used for the transformation, i.e.:

$$\dot{y} \approx \frac{y(t+1) - y(t)}{T}$$

2.1 Model Validation and Noise modeling

As in any mathematical description of a physical process, there exists modeling errors. Here, a simple noise description of modeling errors is used. Also, with a noise description of model uncertainties, systematic threshold selection is possible based on false alarm rates or probability of missed detection. The model used here models unknown dynamics by white Gaussian noise $n(t)$ entering the model equation like this:

$$y(t+1) = (1 - T(k_1 + k_2D + k_4))(y(t) - b) + Tk_1(p_f^0 + p_{air,t} - p_a) - Tk_3Pf(p_{man}(t), p_{ft}(t)) + n(t) \quad (7)$$

This leads to a system model that can be written as a linear regression:

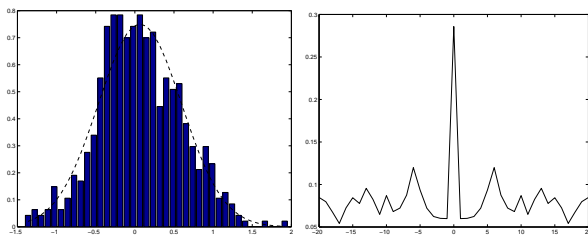
$$y(t+1) = \varphi(t)\theta + n(t)$$

Parameter estimation will show to be a core part of the diagnosis algorithm and ML-estimation of the θ parameter is particularly simple with this model structure. The estimate can be written analytically as:

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^TY \quad (8)$$

where Y is a vector of all collected measurements and Φ is a vector of all regression vectors. Also for this model structure, a recursive estimation procedure is easily derived, e.g. to get a more balanced processor load during calculations.

When comparing model performance against data collected from the real process, it becomes evident that for large leakage areas the whiteness assumption does not hold. However, large leakage holes is easy to detect and therefore is model performance only really important for really small holes. To evaluate model applicability, a noise realization has been estimated when running the model against data collected with a 0.5mm sized leakage hole. Figure 2 shows an estimate of amplitude distribution and also an estimate of the covariance function. In the figure it is seen that the amplitude distribution



(a) Estimate of the amplitude distribution of the noise. The dashed line is an estimated normal distribution and the bars the result of a histogram analysis of the noise estimation.

(b) Estimation of the covariance function. No strong correlation can be seen between $n(t)$ and $n(t + \tau)$ when $\tau \neq 0$, i.e. the whiteness assumption seems valid.

Fig. 2. Model validation by statistical analysis of model residuals.

resembles a Gaussian distribution reasonably well and also that there exists no major correlation between $n(t)$ and $n(t + \tau)$, $\tau \neq 0$, i.e. the whiteness assumption seems valid. This model performance is shown in Section 4 to be good enough for the desired diagnosis performance.

3. DIAGNOSIS ALGORITHM

The model handles two types of faults, sensor bias fault and leakage. The sensor bias is modeled as an additive offset to the actual pressure value as in (5). It is suitable to first estimate the sensor bias level since it can be used for offset correction of the pressure data when the leakage estimation is performed. The offset correction leads to a more simple model structure for leakage detection and good leakage detection performance. The sensor bias level and leakage are estimated on a separate data sets. The first data set is used for sensor bias level estimation and it is sampled with an open diagnostics valve to ensure that the fuel tank pressure is leveled with the ambient pressure. To detect a leakage the diagnostics valve has to be closed, and also a pressure difference between the fuel tank and the ambient is introduced by opening the purge valve.

A procedure that uses the model for detecting leakage is proposed:

1. In the initial state the system is configured as follows: purge valve closed, diagnostic valve open, engine running.
2. A bias level of the pressure reading is estimated based on an initial fuel tank pressure sampling session.
3. A pressure difference is created by closing the diagnostics valve and opening the purge valve. The purge valve is closed when a sufficient pressure difference has been created. It is clear, and also shown in Section 3.3, that the larger pressure difference that is created the better diagnosis performance (assuming equal amount of data can be collected).
4. The fuel tank pressure sensor value is sampled at a sufficient rate (10 Hz) for a suitable amount of time (20 s). The sampled data is corrected by the measured sensor bias.
5. Parameters in the leakage model are estimated either online or after the sampling session. Two methods are proposed.
6. A decision procedure with noise based thresholds evaluates the parameter estimates.

Sample rates and sampling duration were experimentally chosen. The leakage estimation can be performed in two ways depending on how the model is used. Both parameters *time constant* and *final value* contain information about a leakage presence. The model can be used in two ways, making each of the parameters alone sensible to a leakage. The two methods are described in detail in Section 3.3 and 3.4.

3.1 Formal Problem Formulation

Now follows a formal problem formulation based on classical hypothesis testing. As indicated in the algorithm outline above, the bias detection and leakage detection problem has been separated into two completely separated problems, thus this formal description of the diagnosis problem then consist of two hypothesis tests, one for each fault.

The bias fault is parameterized in the model (7), by the parameter b . The bias detection problem therefore consists of deciding between the hypotheses

$$H_b^0 : b = 0 \quad H_b^1 : b \neq 0$$

The parameter k_4 is modeled to be proportional to leakage area, thus rendering a hypothesis test:

$$H_l^0 : k_4 = 0 \quad H_b^1 : k_4 \neq 0$$

Now, for each test, define a test quantity and a rejection region to make the diagnosis decision. This means that for test i , define a function $T_i(y)$ and a rejection region \mathcal{S}_i such that the null hypothesis H_i^0 is rejected when $y \in \mathcal{S}_i$. Normally the rejection region is defined via a thresholding test, e.g. reject H_i^0 if

$$y \in \mathcal{S}_i \Leftrightarrow T_i(y) > J_i$$

Function $T_b(y)$ and threshold J_b is defined in Section 3.2, and two ways to define T_l and J_l is described in Sections 3.3 and 3.4.

3.2 Bias level estimation

The complete model was described in (7). The sensor bias level b is the target for the estimation. That is easiest accomplished by removing the uncertainty of the term containing $p_{air,t}$, by opening the diagnostics valve. The open valve will finally eliminate the pressure difference between the fuel tank and the ambient, and thus make the term $k_1(p_f^0 + p_{air,t} - p_a)$ zero.

The sensor bias level is estimated for two purposes:

1. to correct the data sequence in the leakage diagnosis
2. to detect pressure sensor failures. A large sensor bias may indicate a sensor problem.

For pressure sensor diagnosis the test quantity $T_b(y)$ is set to the bias estimate \hat{b} .

A sample sequence of tank pressure is taken when the diagnostics valve is open and the data is used for estimating the bias level b . The model can be written in a more general way as

$$y(t+1) = y(t)(1 - kT) + bTk$$

where k is the total time constant for the exponential function.

The parameter Θ is estimated from the data sequence $y(t)$ as described in Section 2.1 with

$$\Phi = [Y \ 1] \\ \Theta = \begin{bmatrix} 1 - kT \\ bkT \end{bmatrix}$$

and the bias level is calculated as

$$T_b(y) = \hat{b} = \frac{\hat{\Theta}_2}{1 - \hat{\Theta}_1}$$

The sensor bias level is used for correcting the data sequence for the leakage detection as

$$y^* = y(t) - \hat{b}$$

where $y^*(t)$ denotes the second sampling session.

3.3 Time constant method

The first method for detecting leakage estimates the time constant of the exponential function in the model, e.g the test quantity $T_l(y)$ is the estimate of the time

constant parameter. The model is first reduced using the assumption that a leakage is always present. Also the purge and diagnostics valves are closed and the data is corrected for sensor bias. Starting from equation 7 and removing zero-terms gives the following model:

$$y(t+1) = y(t)(1 - (k_1 + k_4)T) + n(t)$$

The equation can be written in a more general form as

$$y(t+1) = y(t)(1 - kT) + n(t) \quad (9)$$

Since the model does not distinguish between the two time constants k_1 and k_4 , they are replaced with a single time constant k .

The solution to (9) is an exponential function with a final value at zero and time constant k and the *time constant* procedure will fit this model to the sampled pressure data. In the case where a leakage is present this model is correct. In the case where no leakage is present the fitted exponential will have a large time constant since the sampled data will level out at a value well below zero.

The parameter estimation procedure described in Section 2.1 shall be used with

$$\Phi = Y \\ \Theta = 1 - kT$$

The leakage time constant is derived by

$$T_l(y) = \hat{k} = \frac{1 - \hat{\Theta}}{T}$$

The final step is to find a threshold J_l , i.e. decide how big $T_l(y)$ can be before H_l^0 is rejected and a leakage alarm is signalled. An alarm is signalled when $T_l(y) > J_l$.

For this threshold selection, a statistical analysis of $T_l(y)$ is necessary. Assume that the white Gaussian noise $n(t)$ in (7) has a variance σ_n^2 . Under the assumption that the model is correct, straightforward calculations give that $T_l(y)$ has a distribution

$$T_l(y) \sim N\left(k, \frac{\sigma_n^2}{\sum_{t=1}^N y^2(t)}\right)$$

This expression gives that the variance of the estimate varies with the size of the measurement signal $y(t)$ which is a natural situation, the more excitation, the more accurate estimate. This variation in variance however makes it difficult to set a fixed threshold, independent of system excitation and maintain a fixed significance level of the test. To remedy this, an *adaptive* threshold is selected as

$$J_l(y) = \frac{1}{\sqrt{\sum_{t=1}^N y^2(t)}} J$$

where J is selected such that $P(|X| > J) < \alpha$ where X is an $N(0, \sigma_n^2)$ distributed stochastic variable and α is the probability of false alarm.

It can be shown that, under the assumption that the model is correct, the test quantity $T_l(y)$ is Universally Most Powerful(UMP) (Casella and Berger, 1990), i.e. no other test quantity can be better than $T_l(y)$.

3.4 Final value method

The second method for detecting a leakage uses the model property that a leakage will level the fuel tank

pressure to the ambient pressure and the test quantity $T_l(y)$ becomes the estimate of the final value of the exponential function. The model is used in its original form as in equation 7, but with closed purge and diagnostics valves ($P = D = 0$) and sensor bias correction:

$$y(t+1) = y(t)(1 - (k_1 + k_4)T) + k_1 T(p_f^0 + p_{air,t} - p_a) + n(t) \quad (10)$$

or more generally:

$$y(t+1) = (1 - kT)y(t) + kTC + n(t)$$

where k is the time constant and C the final value of the exponential function $y(t)$. An estimate of C is used for leakage detection. When a leakage is present, C will reach zero or close to zero. In the case of no leakage C will stay at a level well below zero.

Estimating the parameter C is done as described in section 2.1 using

$$\Phi = [Y \ 1] \\ \Theta = \begin{bmatrix} 1 - kT \\ kTC \end{bmatrix}$$

The final value C is the test quantity and is derived as

$$T_l(y) = \hat{C} = \frac{\hat{\Theta}_2}{1 - \hat{\Theta}_1}$$

The decision procedure shall find a threshold where a final value C from evaporation can be separated from the final value for the smallest detectable leakage wanted.

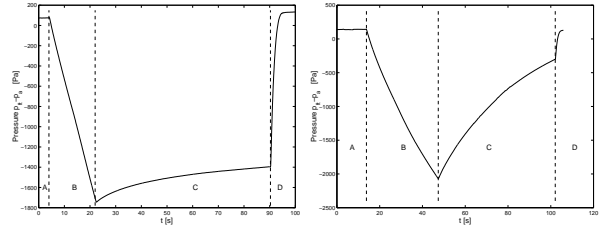
4. EVALUATION ON EXPERIMENTAL DATA

4.1 Experimental setup

Three sets of measurement data were taken in the engine laboratory at Vehicular Systems, Linköping University and in the laboratory of Mecel AB. The first set of data used a SAAB 2.3 liter engine for creating the vacuum and its control system Trionic 7 for controlling valves. The Trionic 7 standard fuel tank pressure sensor was used, mounted in a SAAB 9-5 fuel tank. The fuel tank was prepared with a hole where bolts with different sizes of drilled leakage holes could be mounted. Bolts with leakage diameters 0, 0.5, 1.0, 2.0, 3.5 and 5.0 mm in diameter were used in the experiments. A data collecting system sampled the signals fuel tank pressure, diagnostics valve command and purge valve command at a sampling rate of 10 Hz. The second and third sets used the same fuel tank and a vacuum pump for pressure control. Data were collected for no leak and 0.5 mm only, to further examine the detectability between the two.

Figure 3a shows fault free data. The time series is divided into four sections, A through D. During the A part, the diagnosis valve is open and the purge valve is closed. This is the part where the sensor bias estimation takes place. During the second part, the diagnosis valve is closed and the purge valve is opened, creating the desired pressure drop in the tank due to the low pressure in the intake manifold of the engine. In the C-section, both valves are closed and the pressure increases due to evaporation of the fuel in the tank. If given enough time, the pressure will stabilize when the fuel partial

pressure reaches its saturated level. Finally, in section D the diagnosis valve is opened. Figure 3b shows a similar



(a) Pressure trace when no leakage is present. Notice the pressure increase in section C of the data that is the result of fuel evaporation.

(b) Pressure plot when a 1mm in diameter leakage hole in the tank is present. The pressure characteristic in section C of the data changes significantly compared to the no leakage data

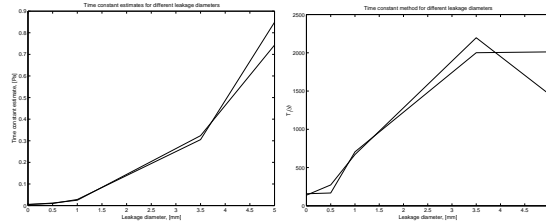
Fig. 3. Measured pressure traces

pressure plot where a 1mm leakage is present. There is a clear difference in pressure characteristic in the C section of the data and it is this portion of the data that is used for leakage detection.

Now follows an evaluation of the two leakage detection methods described in Section 3 on data collected with leakages ranging from 0.5mm to 5mm in diameter.

4.2 Evaluation of the time constant method

Figure 4a shows the value of the estimated time constant k for different leakage sizes. Figure 4b shows the value



(a) Time constant k for different leakage sizes.

(b) The test quantity $T_l(y)$ for different leakage sizes.

Fig. 4. Evaluation of all leakages in time constant method

of the normalized test quantity $T_l(y)$ for different leakage sizes. The data used in these two figures was sampled by 10 Hz during 20 seconds, two sequences for each leakage size. According to the modeling assumption, the parameter k should be proportional to the leakage area. This assumption seems valid since the test quantity resembles a x^2 curve, i.e. the area proportionality seems to hold. Larger leakages, ≥ 1 mm, are easy to detect but the distinction between no leak and 0.5 mm need some extra examination.

Data was collected at 5 Hz sampling rate during 20 seconds and eight sequences for leakage sizes 0 and 0.5 mm. The value of the test quantity is shown in Figure 5a. There is a separation between the two data sets, but the

separation is small compared to the variation within each data set. In an attempt to increase the separation, data sets with longer duration was sampled to investigate how much data is needed to securely detect 0.5mm sized leakages. Figure 5b show the result from 60 seconds sequence duration. The separation vs. variation is now large enough (≈ 10 standard deviations) for highly reliable detection of the smallest holes. Using this method

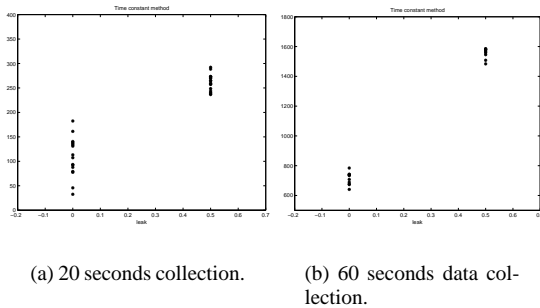


Fig. 5. The test quantity $T_l(y)$ for 0 and 0.5 mm leak.

it is possible to detect the smallest hole and separate all leakages from the no leakage case indicating that more than 20 and less than 60 seconds of data is needed to get reliable detection.

4.3 Evaluation of the final value method

Estimation of the final value was done for all five cases, 0 - 5 mm leakage diameter. Results are shown in Figure 6. There is an obvious separation between the ≥ 1 mm diameter leakage and the non-leakage cases. As the leakage gets larger, the final value estimate narrows around zero. A smaller leakage, 0.5 mm, show a sensitivity in the final value estimation and the worst value in this case is in the same order as the highest non-leakage value. Further investigations was done for the separation between 0 and 0.5 mm leaks. First a larger number

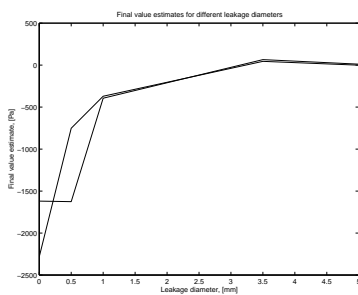
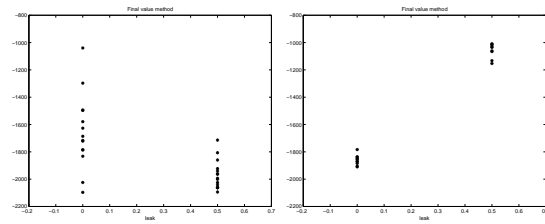


Fig. 6. The test quantity $T_l(y)$ for different leakage sizes.

of sequences was collected. Eight sequences, each 20 seconds duration at 5 Hz sampling rate, was collected for both 0 and 0.5 mm leaks. The result is shown in Figure 7a. The final value estimation is not good for 0.5 mm. There is no distinction between the sets. Increasing the sampling duration to 60 seconds gives a totally different result as seen in Figure 7b which is similar to the performance of the time-constant method.

The final value method can not be used to estimate the size of the leakage, only to separate no leakage from leakage. This because for all leakage sizes, the tank pressure will finally end up at ambient pressure. The



(a) 20 seconds collection. (b) 60 seconds data collection.

Fig. 7. The test quantity $T_l(y)$ for 0 and 0.5 mm leak.

estimation of final value is dependent on the pressure drop in phase B (see Figure 3).

The two methods attack the problem in somewhat different ways, using different properties of the model for diagnostic statements. In these tests the time constant method proved to be most efficient but the final value method may show advantages in other cases not covered by these tests.

5. CONCLUSIONS

Leakage detection and sensor bias detection in a fuel evaporative system for automotive vehicles has been considered. A physical model is derived, modeling fuel evaporation and flows through the purge valve into the engine, flow through the diagnosis valve into the environment, and leakage flows. The model is used in a systematic design procedure to derive two different diagnosis algorithms. The algorithms only utilize the model structure and no model parameters need to be stored, thus making the algorithm very stable against production variations between different individuals.

The diagnosis system is evaluated in a laboratory environment with a production car engine and evaporative system equipped with production sensors only. In these tests, the time constant method proved to be the most efficient and measured data indicates that it needs less data than the final-value method. Both algorithms presented in the paper successfully detect leakages as small as 1 mm in diameter using data collected during 20 seconds. For detecting a 0.5 mm leakage, longer data sequences is needed to perform reliable detection. In the experiments 60 seconds of data proved to provide very reliable detection of 0.5mm sized leakages. For more realistic performance, significantly less than 60 seconds of data is needed.

REFERENCES

- California's OBD-II Regulation (1997). section 1968.1, Title 13, California Code of Regulations, Resolution 93-40, September 25. URL:<http://www.arb.ca.gov>.
- Casella, G. and R.L. Berger (1990). *Statistical Inference*. Duxbury.
- Majkowski, S., K.M. Simpson and J. Steckler (1999). Development and validation of a 0.020" evaporative leak diagnostic system utilizing vacuum decay methods. In: *SAE Technical Paper Series*. number 1999-01-0861.