

Model Based Engine Map Adaptation Using EKF

Erik Höckerdal ^{*,**}, Erik Frisk ^{*}, and Lars Eriksson ^{*}

^{*} *Department of Electrical Engineering, Linköpings universitet,
Sweden, {hockerdal,frisk,larer}@isy.liu.se*

^{**} *Scania CV AB, Södertälje, Sweden, erik.hockerdal@scania.com*

Abstract:

A method for on-line map adaptation is developed. The method utilizes the EKF as a parameter estimator and handles parameter aging, operating point dependent model and measurement quality. Map adaptation, by construction, gives marginally stable models with locally unobservable modes, that are handled. The method is also suitable for off-line calibration of maps where the only requirement of the data is that the entire operating region of the system is covered. The method is applied to a truck engine where an air mass-flow sensor adaptation map is estimated based on data from a diesel engine during an ETC. It is shown that an adaptation map can be found in a measurement sequence not specially designed for adaptation.

Keywords: bias compensation, EKF, non-linear, observer, engine map, adaptation

1. INTRODUCTION

Look-up tables and maps are frequently used to describe relations in modern control and diagnosis systems where physical models are unavailable, e.g. low level sensor and actuator characteristics.

In engine control systems, maps of different types are commonly used for example to compensate for changed ambient conditions and aging engine subsystems like cooler efficiency, injector characteristics, aftertreatment systems etc. These are typical examples of maps that need continuous on-line adaptation to avoid undesired system behavior, like e.g. biased air mass-flow estimates, causing increased emissions. A related topic of major concern in engine control system development is the calibration process of the complex control system with its variety of maps and parameters.

An attractive idea of how to handle these problems is to incorporate system models to aid in the development of sufficiently robust and fast adaptation algorithms that can be used both on-line to handle system aging and off-line to automatize the calibration process of engine control systems (Guzzella and Amstutz, 1998).

Basically map adaptation can be viewed as a desire to reduce and store the compensation of an operating point dependent estimation error, and a way to reduce stationary estimation errors in model based observers was developed in Höckerdal et al. (2008b). It utilizes an observable default state space model

$$x_{t+1} = f(x_t, u_t) \quad (1a)$$

$$y_t = h(x_t), \quad (1b)$$

and measurements, y, u , from the system and estimates a low order bias augmentation $A_q q$ that, when used in observer design, reduces the stationary estimation errors of the resulting observer. The result of this method is an observable augmented model

$$x_{t+1} = f(x_t - A_q q_t, u_t) \quad (2a)$$

$$q_{t+1} = q_t \quad (2b)$$

$$y_t = h(x_t), \quad (2c)$$

that can be used with any suitable observer design to construct an observer with reduced stationary estimation errors compared to using the default model directly.

However, since the observer designed in Höckerdal et al. (2008b) treats the bias as a random walk it is not able to keep track of the changes in bias between operating points over time. That is, as soon as the system changes operating point, all information about the bias in the previous operating point is discarded.

An observer that stores information about the bias could be useful in applications where the system is such that information with a certain quality only is present in some operating points. Then information collected at these operating points can be used to improve the quality of the estimated variables throughout the entire operating region of the system, e.g. air mass-flow adaptation in diesel engines with exhaust gas recirculation (EGR) and variable geometry turbine (VGT) (Höckerdal et al., 2008a). In engine map adaptation schemes it is also important that the algorithm is robust against outliers since occasional spurious measurements are inevitable.

Hence, the goal is to design a robust observer with memory that is able to handle old parameters, operating point dependent models, and varying measurement quality, and occasional spurious measurements.

2. METHOD OUTLINE

Based on the method developed in Höckerdal et al. (2008b), an information preserving observer, i.e. with memory, can be obtained by exchanging the assumption of a bias modeled as a random walk

$$q_{t+1} = q_t + v_t, \quad v_t \sim \mathcal{N}(0, Q), \quad (3)$$

driven by white noise, for a parameterized function, or map, describing the bias

$$\begin{aligned} q_t &= q_{\text{fcn}}(x_t, u_t, \theta_t) \\ \theta_{t+1} &= \theta_t + v_t, \quad v_t \sim \mathcal{N}(0, Q_{\text{fcn}}). \end{aligned}$$

Where $q_{\text{fcn}}(x_t, u_t, \theta_t)$ is a parameterized function or map with unknown parameters θ that describes the bias dependence on the system states x , and inputs u . The parameters are modeled as random walks in the same way as the bias state in (3).

If this is done for the general state space model (2), the following system is obtained,

$$x_{t+1} = f(x_t - A_q q_t, u_t) \quad (4a)$$

$$\theta_{t+1} = \theta_t \quad (4b)$$

$$q_t = q_{\text{fcn}}(x_t, u_t, \theta_t) \quad (4c)$$

$$y_t = h(x_t), \quad (4d)$$

where $A_q q$ adjusts the stationary points of the system. Here and henceforth system, measurement, and random walk noise terms are left out to increase readability. Note that this is similar to the formulation used in Höckerdal et al. (2008b), the difference is that the bias states have been exchanged for a parameterized function and the function parameters are introduced as new states. If (4c) is inserted into (4a), a standard state space form is obtained which means that any suitable observer design can be applied. For example, one way of estimating states while at the same time handling unknown parameters is to apply a joint parameter and state estimating extended Kalman filter (EKF) (Kopp and Orford, 1963). There the parameters are introduced as new states with constant time derivatives and augmented to the original states just as in the system described by (4).

To develop a model like (4) and use it for estimation and identification entails that some new issues have to be addressed compared to Höckerdal et al. (2008b). Of these the main concern in this paper is how to update the function parameters, θ_t , in a controlled manner. Another related issue is how to find a suitable parameterization (4c), with its structure and regressors. This is not treated here and the interested reader is referred to e.g. Lind and Ljung (2008).

Even though any suitable observer design can be applied to this system, the choice here is to use a stochastic filter. An advantage of stochastic filters, like for example EKF, compared to deterministic observers is that, not only the state estimate, but also an estimate of the estimation error distribution is computed. The estimation error statistics is used in the computation of the filter feedback gain, which gives the stochastic filters natural tuning parameters that allow filter tailoring to handle unknown state initialization, time dependent model and measurement quality, outlier rejection etc.

For these reasons and because of simplicity the joint parameter and state estimating EKF is used throughout this paper.

3. OBSERVABILITY

In all estimation problems observability or, at least, detectability of the system at hand is central. Since the system considered here can be viewed as an extension

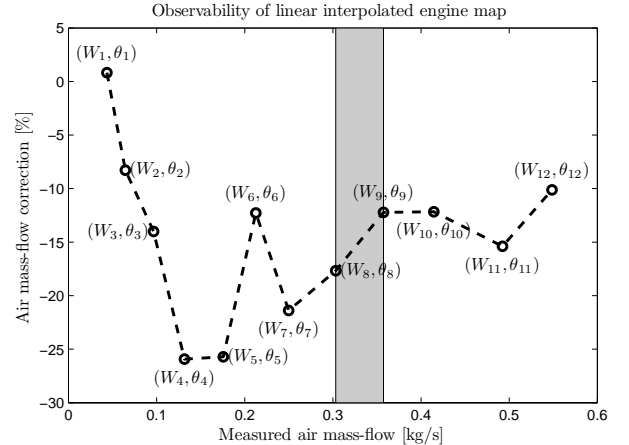


Fig. 1. Air mass-flow correction map with the grid points denoted with the pair (W_i, θ_i) corresponding to a correction factor of θ_i at a mass-flow of W_i .

of the augmented system addressed in Höckerdal et al. (2008b) it is natural to assume that the system

$$x_{t+1} = f(x_t - A_q q_t, u_t) \quad (5a)$$

$$q_{t+1} = q_t \quad (5b)$$

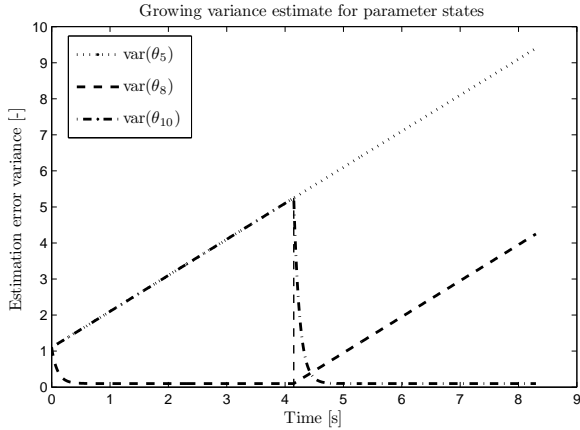
$$y_t = h(x_t) \quad (5c)$$

is observable and analyze how the replacement of (5b) by a parameterized function, (4b – 4c), affects the observability. The observability is therefore directly dependent on the properties of the parameterized function (4c) that describes the operating point dependence of the bias $q_t = q_{\text{fcn}}(x_t, u_t, \theta_t)$.

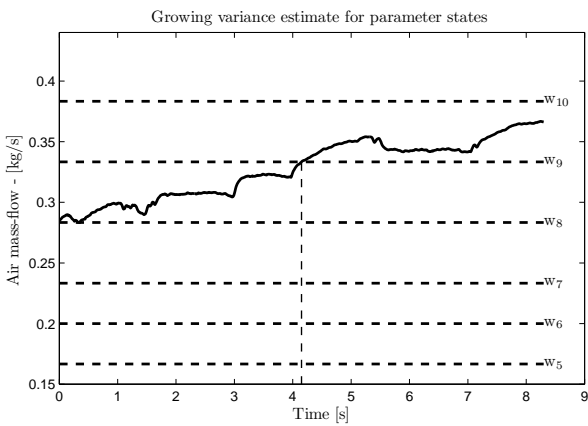
If for example (4c) is an engine map, implemented as a look-up table with the grid points as parameters, and an interpolation algorithm computing the output. Then the system will have locally unobservable states (Hermann and Krener, 1977), i.e. parameters that are not used in the interpolation in the current operating point are not observable. In Figure 1 the local observability of the parameter states in linear interpolation is illustrated using an air mass-flow correction map from an inline six cylinder Scania diesel with EGR and VGT (Höckerdal et al., 2008a). If the operating point, in this case defined by the air mass-flow measurement, lies in the shaded region of the figure then only the two grid points constituting the border of the region are observable.

3.1 Growing Estimation Error Covariance

At any given time there are generally, for systems like (4), some parameters θ_i that are locally unobservable. A property of systems with locally unobservable modes is that the estimation error covariance matrix in an EKF framework grows linearly for the unobservable modes if system noise for these modes is present. That is, in regions where the system seldom operates the estimation error covariance matrix coefficients, corresponding to locally unobservable modes or parameter states, will grow linearly without bound. This linear growth in covariance matrix elements have two sides. 1) It offers a way to achieve fast update of old parameters while protecting often updated parameters from spurious measurements. 2) It may cause numerical problems affecting the system stability when considering the life-time of the system, which has to be handled.



(a) Parameter variance.



(b) Air mass-flow trajectory.

Fig. 2. The figure shows the evolution of the estimation error variance for three parameter states.

This effect is illustrated in Figure 2, where the variance of three parameter states, θ_5 , θ_8 , and θ_{10} from Figure 1 are plotted versus time. In Figure 2(a) θ_5 corresponds to a parameter that is not observed at all for the studied trajectory, while the parameter θ_8 is observable during the first half of the trajectory and unobservable for the second half. For the parameter θ_{10} the case is reversed, that is the parameter is first unobservable and then observable.

Experiences from adaptive maps in engine applications, not using the EKF and joint parameter and state estimation, indicate problems concerning parameter aging and occasional spurious measurements. For example an engine that, during normal operation, does not cover the entire parameter space and only occasionally enters some areas, may suffer from undesired system behavior caused by old parameters. Many of today's adaptation schemes apply the same adaptation algorithm in each update step and do not adjust the update procedure with respect to when the parameters were last updated (Wu, 2006; Peyton Jones and Muske, 2007). In these cases, a linearly growing uncertainty for seldom updated parameters enables a fast parameter update rate of old parameters without risking large errors in the state estimates. This can in some sense be thought of as a dynamic forgetting factor similar to recursive least square (RLS) techniques and is a highly desirable property in engine adaptation algorithms not only

to handle aging parameters but also to protect updated parameters from occasional spurious measurements, that are fairly common in engine applications.

A direct and intuitive way of handling the linear growth of estimation covariance of locally unobservable parameter states, i.e. 2), is to introduce an upper limit for the corresponding estimation error covariance matrix elements. A possible upper limit is the initializing error covariance matrix, P_0 . Since it is desirable to limit the estimation error covariance of only the locally unobservable parameters it is appropriate to perform the limitation element wise, i.e. compare $P_{i,i}$ to $P_{0,i,i}$, and limiting $P_{i,i}$ by setting $P_{i,i} = P_{0,i,i}$ when $P_{i,i} \geq P_{0,i,i}$. It is straightforward to show (Jaynes, 1996) that the off-diagonal elements in P do not affect the estimation error covariance for a single parameter, and by using P_0 as an upper limit, the introduction of yet another tuning parameter is avoided.

4. FILTER TUNING – Q VS. Q_{FCN}

Even though the method developed here is quite similar to the one developed in Höckerdal et al. (2008b) there are some important differences. One difference is the rate at which the bias and parameter states are updated.

For the observer designed in Höckerdal et al. (2008b) it is necessary for the bias state to change approximately as fast as the system dynamics, otherwise it will not be able to track a change in system operating point. However, a rapidly changing bias state captures also high frequency disturbances, and is thereby sensitive to outliers, which makes the bias state in this method unsuitable for engine map adaptation.

In an observer utilizing a parameterized function to describe the bias, the parameter states operate with an update rate determined by system aging, which is substantially slower than for a bias state that has to track changes in system operating point. This makes the observer based on a model containing a parameterized function or map less sensitive to temporary disturbances, compared to an observer using only one state to describe the bias. However, both methods can be used to find an adaptation map, the first estimates it directly and the latter after some post processing like mean value computations, which makes it less suitable for on-line applications.

Another issue that is, to some extent, straightforwardly handled by stochastic filters is initialization of the unknown bias or function parameters. By proper tuning of the corresponding elements in the estimation error covariance matrix, P_0 , a temporary faster update of unknown bias or function parameter states is achieved. That is, due to an initially faster update rate of unknown parameters, in the same way as old parameters are allowed a faster update rate, a rapid convergence of the otherwise quite slow parameter states is achieved.

5. METHOD EVALUATION

To evaluate the method two studies, a simulation study and a study utilizing experimental data, are performed where the aim is to adapt the air mass-flow sensor in diesel engine, characterized by a 1-D adaptation map. The

simulation study shows the convergence of the method and includes a minor analysis of model error, and noise sensitivity while the experimental part shows the result of the method applied to experimental data.

In both studies a non-linear model of a heavy duty truck engine developed in Wahlström and Eriksson (2006) is used together with measurements from an engine in an engine test cell. The model has three states, intake and exhaust manifold pressures, and turbine speed which all are present in the model output together with the air mass-flow through the compressor. The data used are collected during a European transient cycle (ETC).

5.1 Observers

Three observers are designed and evaluated both in simulation and on experimental data. The observer designs are: An EKF based on the default model developed in Wahlström and Eriksson (2006) directly,

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ y_t &= h_{W_{\text{cmp}}}(x_t), \end{aligned}$$

referred to as Def. An EKF with an extra bias state introduced in the measurement equation to reduce the estimation error from the method developed in Höckerdal et al. (2008b),

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ q_{t+1} &= q_t \\ y_t &= h_{W_{\text{cmp}}}(x_t) + q_t, \end{aligned}$$

referred to as Aug. A joint state and parameter estimating EKF based on the default model and a parameterized bias,

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ \theta_{t+1} &= \theta_t \\ y_t &= (1 + q_{\text{fcn}}(h_{W_{\text{cmp}}}(x_t), \theta_t))h_{W_{\text{cmp}}}(x_t), \end{aligned}$$

referred to as Map. where q_{fcn} is presented in Figure 1.

In the simulation study all observers use only feedback from the air mass-flow sensor, $h_{W_{\text{cmp}}}(x_t)$, whilst in the experimental evaluation, feedback from all sensors except the exhaust manifold pressure sensor, x_2 , is used. Even though the model is observable from any of the outputs the model errors are such that an augmented feedback is needed in the experimental evaluation. The estimation performance evaluation is with respect to all states and outputs, i.e. intake and exhaust manifold pressures, turbine speed, and air mass-flow through the compressor.

For the study utilizing experimental data, all observers are augmented with an additional bias state, in addition to those introduced to handle the measurement error, with the purpose of reducing estimation errors due to model errors in the compressor model causing incorrect prediction of the compressor mass-flow, i.e. for Def.

$$\begin{aligned} x_{t+1} &= f(x_t, \Delta W_{\text{cmp},t}, u_t) \\ \Delta W_{\text{cmp},t+1} &= \Delta W_{\text{cmp},t} \\ y_t &= (x_{1,t} \ x_{3,t} \ h_{W_{\text{cmp}}}(x_t, \Delta W_{\text{cmp},t}))^T. \end{aligned}$$

5.2 Simulation data

The simulation study serves two purposes, 1) to compare modeling errors as stationary biases (Höckerdal et al.,

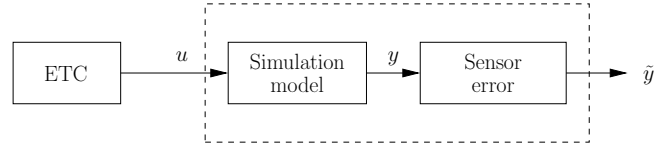


Fig. 3. Simulation set-up with sensor error from Höckerdal et al. (2008a).

2008b), or as parameterized functions, and 2) to analyze the effect of incorrect noise description and model error sensitivity when using the EKF as a parameter estimator.

Both these utilize the simulation set-up presented in Figure 3 for creating the data. The data is created by simulating the model with input data from an ETC segment. The segment is chosen to contain a wide range of air mass-flows such that a trajectory for which the system is observable is created. To simulate incorrect air mass-flow measurement a 1-D sensor error map is used in the simulation that adjusts the air mass-flow according to

$$W_{\text{meas}} = (1 + q(W_{\text{true}}))W_{\text{true}}, \quad (6)$$

where $q(W_{\text{true}})$ is the engine map presented in Figure 1. The distorted air mass-flow, W_{meas} , is then used for feedback in the observers.

Convergence One property of the estimation bias correction method developed in Höckerdal et al. (2008b) is that since no information about the bias in each operating point is saved the observer convergence speed depends on the speed of the bias states. While the extension presented here use a parameterized function and the convergence speed of the filter, when the parameters are adapted, will therefore not be dependent on the speed of the parameter states.

To analyze and compare the convergence speed of filters utilizing parameterized bias to filters utilizing bias states is a difficult task. A reason for this is that the convergence speed is highly dependent on the filter tuning. Since the maximum errors occur in transients, see for example the transient at the time 5s in Figure 4, they give an indication on how well the estimator is able to track transient behavior. Figure 4 presents the estimation errors from Aug., where the bias state has been tuned to be too slow and does not manage to track the change in bias. As references the estimation errors from Map. and Def. are also presented. Since a bias state has to be approximately as fast as the system dynamics and the parameters as fast as the system aging, filters tuned with these aspects in mind will have about the same performance with respect to estimation quality during normal operation, which is confirmed in Table 1. Nevertheless, since the bias state is allowed to change much faster than the parameter states, a filter utilizing that method will be more sensitive to disturbances, i.e. outliers as mentioned in Section 4. A filter with a parameterized function with slow parameters does not allow disturbances to affect the estimation of neither model nor parameter states to the same extent as a filter with a bias state, i.e. have a stronger smoothing effect.

Figure 5 shows the true and estimated map from Map., and the correction made by the slowly varying bias from Aug. computed according to

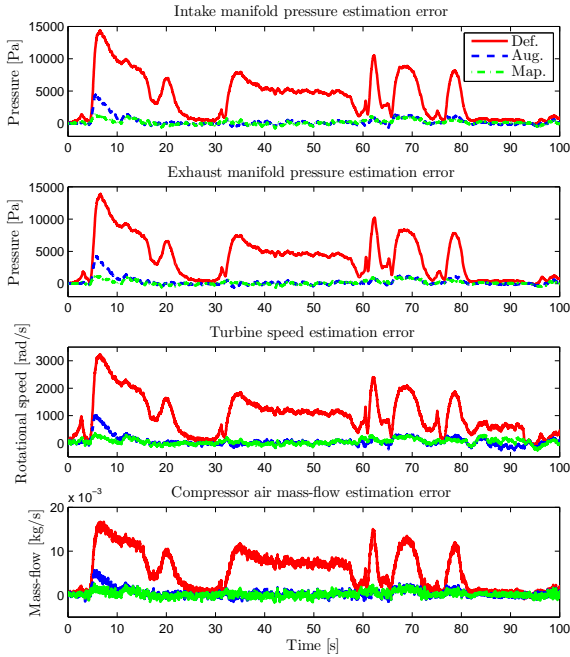


Fig. 4. Estimation error of Aug., where the bias is too slow which gives large estimation errors in transients, together with estimation errors of Map. and Def.

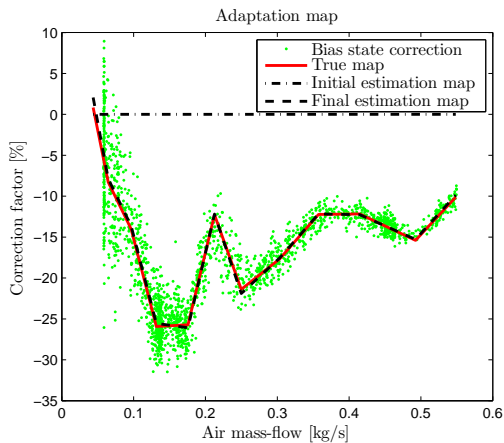


Fig. 5. Mass-flow correction estimated by Aug. and Map.

$$r = \frac{\hat{q}_t}{\hat{y}_t - \hat{q}_t}.$$

From this figure it is obvious that Map. manages to estimate a correction map out of a cycle, not specially designed for map adaptation, without any post processing. Also, the correction made by Aug. captures the true map but some post processing, like for example mean value computations, is needed to get a map that can be used for interpolation etc.

Evolution of adaptation map In an application where the method is used for engine map adaptation it is important that the method converges. However, since the model description probably never will be entirely correct, it is impossible to converge to something that can be called the true map. However, for simulated data this can be achieved. Figure 6 shows the evolution of the adaptation map over time. The parameters are all initiated to zero,

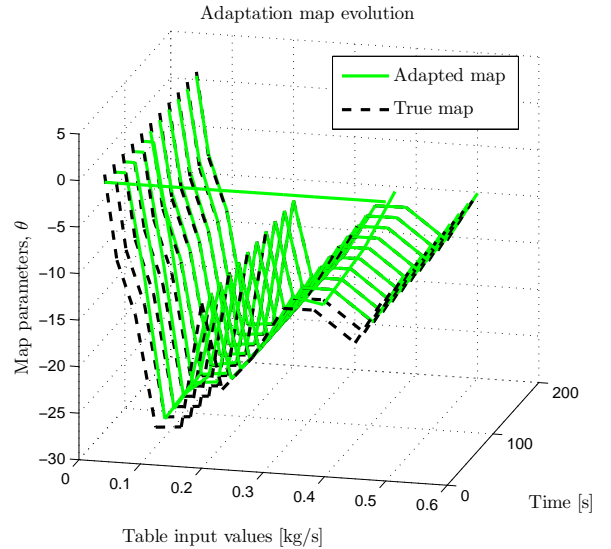


Fig. 6. Adaptation map evolution showing fast convergence.

Table 1. Def., Aug. and Map. estimation error.

Meas.	Max abs. error			Mean error		
	Def.	Aug.	Map.	Def.	Aug.	Map.
p_{im} [Pa]	13901	1156	1075	2521	47	46
p_{em} [Pa]	13482	1071	999	2383	43	42
n_{trb} [rpm]	3118	446	432	769	23	22
W_{air} [kg/s]	0.02	0.003	0.003	0.003	0.00007	0.00006

indicated by the straight line at time $t = 0$, and converges to the true map as the operating region is covered.

In Table 1 it is seen that all measures, both maximum and mean, are approximately the same for Aug. and Map. for all system states and the system output, while Def. has significantly larger errors. From this it can be concluded that the estimation performance with respect to the default states and outputs are similar for the two observers Aug. and Map. Though, Map. also automatically estimates a map that describes the air mass-flow sensor error, that can be used by other algorithms or functions in the engine control unit (ECU).

Robustness The robustness of the proposed algorithm is only briefly analyzed by i) introducing a variety of model errors in minor sub models known to have inaccuracies, e.g. the EGR system, ii) tampering with model and measurement noise structures and intensity.

The first to study the properties of filter divergence due to modeling errors and the second to analyze the effect due to incorrect noise properties. None of the experiments showed any tendencies of divergence or lack of convergence consistency.

5.3 Experimental data

To see how the method can be expected to work in a real application an evaluation using measurements from an engine in an engine test cell is conducted. Since in this case there are other model errors present besides the air mass-flow sensor error an extra bias state is introduced that compensates for incorrect mass-flow through the compressor, see Section 5.1. With the introduction of an

Table 2. Mean estimation error using experimental data for Def., Aug and Map.

Meas.	With $\Delta W_{\text{cmp},t}$			Without $\Delta W_{\text{cmp},t}$		
	Def.	Aug.	Map.	Def.	Aug.	Map.
p_{im} [Pa]	4208	169	-210	13875	13176	13019
p_{em} [Pa]	-10987	-13746	-14420	-2339	-3454	-3450
n_{trb} [rpm]	69	16	29	-57	-19	5
W_{air} [kg/s]	0.008	-0.020	-0.020	0.031	0.035	0.035

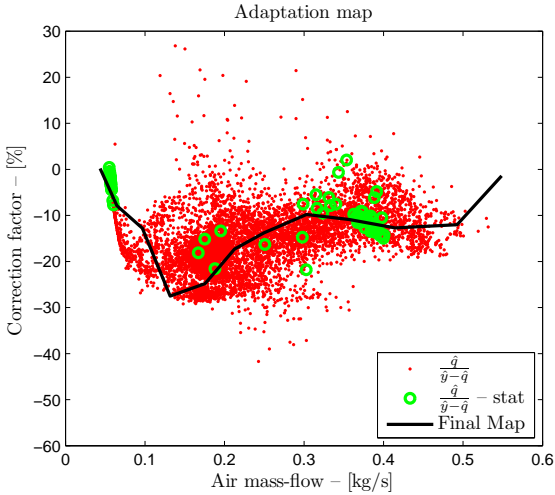


Fig. 7. By Aug. and Map estimated mass-flow correction.

extra state compensating for the compressor mass-flow, known to be biased, the estimate of the intake manifold pressure becomes significantly better at the expense of the exhaust manifold pressure estimate, while the estimates of turbine speed, and air mass-flow is almost unaffected, see Table 2.

In Section 4 the tuning of Aug. and Map. are discussed, especially the different philosophies of the bias describing states – the relatively fast bias state in Aug. and the slow map states in Map. In Figure 7, and Table 2 the similarity in estimation performance between Aug. and Map. is striking, which is an expected result. The mean estimation errors for p_{im} and n_{trb} are reduced while the mean estimation errors for p_{em} and W_{cmp} are slightly increased. That is, in absence of outliers Aug. and Map. are comparable with respect to output estimation performance. The benefit with Map. is that it also estimates an adaptation map.

From Figure 7 it is seen that, even though there are unknown model errors present, besides the compressor mass-flow, the method manages to estimate a map that describes the difference between modeled and measured air mass-flow through the compressor well. Figure 7 shows the correction factor between modeled and measured air mass-flow, similar to what was observed in Figure 5.

Finally, all this shows that an adaptation map can be estimated even though the data used is from the highly transient ETC, not specially designed for air mass-flow sensor adaptation.

6. CONCLUSIONS

A method for storing bias information from different operating points is developed. With this method it is possible to achieve simultaneous estimation of original

model states and parameters, like for example adaptation of engine maps.

Stochastic filters together with a parameterized bias that has locally unobservable states is in fact an asset that handles seldom updated parameters and gives robustness against occasional spurious measurements in ordinary map adaptation algorithms. The linear growth of estimation error covariance, that comes as a result of local unobservability of the parameters, also form a potential numerical problem for the filter and a way to limit this growth without extra filter parameters is provided.

The method shows promising results in a simulation study, where it manages to estimate the engine states while at the same time estimating a parameterized air mass-flow adaptation map. In an evaluation with experimental data it is shown that while maintaining the same estimation quality with respect to mean and maximum absolute error, as the method developed in Höckerdal et al. (2008b), an engine adaptation map can be estimated as well. That is, simultaneous state estimation and map adaption is achieved without, for adaptation, specially designed cycles.

REFERENCES

- Guzzella, L. and Amstutz, A. (1998). Control of diesel engines. *Control Systems Magazine, IEEE*. 01/11/199811/1998, 18(5), 53–71.
- Hermann, R. and Krener, A.J. (1977). Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, AC-22(5), 728–740.
- Höckerdal, E., Eriksson, L., and Frisk, E. (2008a). Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. *SAE Int. J. Passeng. Cars – Electron. Electr. Syst.*, 1(1), 393–402.
- Höckerdal, E., Frisk, E., and Eriksson, L. (2008b). Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*. doi:10.1016/j.conengprac.2008.09.004.
- Jaynes, E.T. (1996). *Probability Theory: The Logic of Science*. Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, UK.
- Kopp, R.E. and Orford, R.J. (1963). Linear regression applied to system identification for adaptive control systems. *AIAA*, 1(10), 2300–2306.
- Lind, I. and Ljung, L. (2008). Regressor and structure selection in NARX models using a structured ANOVA approach. *Automatica*, 44(2), 383–395.
- Peyton Jones, J.C. and Muske, K.R. (2007). Automatic calibration of 1 and 2-d look-up tables using recursive least-squares identification techniques. In *Electronic Engine Controls*, number 2007-01-1343 in SAE Technical paper series SP-2087, 205–213. SAE World Congress, Detroit, USA.
- Wahlström, J. and Eriksson, L. (2006). Modeling of a diesel engine with VGT and EGR including oxygen mass fraction. Technical Report LiTH-R-2747, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden.
- Wu, G. (2006). A table update method for adaptive knock control. In *Electronic Engine Controls*, number 2006-01-0607 in SAE Technical paper series SP-2003. SAE World Congress, Detroit, USA.