

Linköping Studies in Science and Technology
Dissertations, No 1366

Model Error Compensation in ODE and DAE Estimators

with
Automotive Engine Applications

Erik Höckerdal



Linköping University
INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering
Linköping 2011

Linköping Studies in Science and Technology
Dissertations, No 1366

Model Error Compensation in ODE and DAE Estimators with Automotive Engine Applications

Erik Höckerdal



Linköping University
INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering
Linköping 2011

Linköping Studies in Science and Technology
Dissertations, No 1366

Erik Höckerdal
hockerdal@isy.liu.se
www.vehicular.isy.liu.se
Division of Vehicular Systems
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

Copyright © 2011 Erik Höckerdal, unless otherwise noted.

All rights reserved.

Paper A reprinted with permission from Control Engineering Practice ©2009 Elsevier.

Paper B reprinted with permission from Control Engineering Practice ©2011 Elsevier.

Höckerdal, Erik

Model Error Compensation in ODE and DAE Estimators
with Automotive Engine Applications

ISBN 978-91-7393-209-7

ISSN 0345-7524

Cover illustration based on a photo by Dan Boman.

Typeset with L^AT_EX 2_ε

Printed by LiU-Tryck, Linköping, Sweden 2011

To My Family

ABSTRACT

Control and diagnosis of complex systems demand accurate information of the system state to enable efficient control and to detect system malfunction. Physical sensors are expensive and some quantities are hard or even impossible to measure with physical sensors. This has made model-based estimation an attractive alternative.

Model based observers are sensitive to errors in the model and since the model complexity has to be kept low to enable use in real-time applications, the accuracy of the models becomes limited. Further, modeling is difficult and expensive with large efforts on model parametrization, calibration, and validation, and it is desirable to design robust observers based on existing models. An experimental investigation of an engine application shows that the model have stationary errors while the dynamics of the engine is well described by the model equations. This together with frequent appearance of sensor offsets have led to a demand for systematic ways of handling operating point dependent stationary errors, also called biases, in both models and sensors.

Systematic design methods for reducing bias in model based observers are developed. The methods utilize a default model, described by systems of ordinary differential equations (ODE) or differential algebraic equations (DAE), and measurement data. A low order description of the model deficiencies is estimated from the default model and measurement data, which results in an automatic model augmentation. The idea is then to use the augmented model in observer design, yielding reduced stationary estimation errors compared to an observer based on the default model. Three main results are: a characterization of possible model augmentations from observability perspectives, a characterization of augmentations possible to estimate from measurement data, and a robustness analysis with respect to noise and model uncertainty.

An important step is how the bias is modeled, and two ways of describing the bias are analyzed. The first is a random walk and the second is a parameterization of the bias. The latter can be viewed as an extension of the first and utilizes a parameterized function that describes the bias as a function of the operating point of the system. By utilizing a parameterized function, a memory is introduced that enables separate tracking of aging and operating point dependence. This eliminates the trade-off between noise suppression in the parameter convergence and rapid change of the offset in transients. Direct applications for the parameterized bias are online adaptation and offline calibration of maps commonly used in engine control systems.

The methods are evaluated on measurement data from heavy duty diesel engines. A first order model augmentation is found for an ODE of an engine with EGR and VGT. By modeling the bias as a random walk, the estimation error is reduced by 50 % for a certification cycle. By instead letting a parameterized function describe the bias, better estimation accuracy and increased robustness is achieved. For an engine with intake manifold throttle, EGR, and VGT and a corresponding stiff ODE, experiments show that it is computationally beneficial to approximate the fast dynamics with instantaneous relations, transforming the ODE into a DAE. A main advantage is the possibility to use more than 10 times longer step lengths for the DAE based observer, without loss of estimation accuracy. By augmenting the DAE, an observer that achieves a 55 % reduction of the estimation error during a certification cycle is designed.

POPULÄRVETENSKAPLIG SAMMANFATTNING

I dagens samhälle har transporter av olika slag en betydande roll och på land står den tunga lastbilen för en majoritet av dessa. Samtidigt som transportbehovet ständigt ökar ställer både emissionslagstiftning och kunder allt högre krav på minskade utsläpp och minskad bränsleförbrukning. För dieselmotorer är det utsläpp av partiklar, det vill säga oförbränt bränsle och smörjoljerester, samt utsläpp av kväveoxider och koldioxid som omfattas. Kraven innebär både att hålla förbränningsemissionerna nere under normal drift och att fel som medför risk för förhöjda emissioner måste kunna upptäckas, vilket driver den tekniska utvecklingen framåt. Med introduktionen av nya tekniska lösningar samt hårdare emissionskrav följer behovet av tillförlitlig information om motorns interna tillstånd för att möjliggöra robust och säker drift. Till exempel behöver information om tryck, temperatur och syre/bränsle-förhållande tas fram.

Dock är det inte ekonomiskt eller praktiskt möjligt att använda fysiska sensorer för att mäta alla dessa parametrar. Det här har medfört introduktionen av matematiska modeller över motorn, vilka tillsammans med tillgängliga sensorer används för att ta fram information om motorns tillstånd. Modellerna baseras ofta på fysikaliska samband för exempelvis energi- och massbevarande. De är dyra att utveckla då det tar tid att ta fram de matematiska samband som krävs. Dessutom tillkommer aktiviteter såsom parametrering, kalibrering och validering. Oavsett hur mycket tid som läggs på att ta fram modellen kommer den aldrig att bli perfekt. I de fall där kraven på modellens beräkningskomplexitet är höga blir detta extra tydligt, vilket är fallet i de flesta realtidsapplikationer. Resultatet från modellen kommer alltså att avvika från de verkliga värdena, och det blir viktigt att reducera fel i skattningar som uppkommit till följd av fel i modellen.

Det har därför vuxit fram ett intresse för metoder som möjliggör användning av modeller behäftade med fel för att beräkna motorers interna tillstånd med hög noggrannhet. Syftet med forskningen som presenteras i avhandlingen är därför att utveckla systematiska metoder som, utan att involvera extra modellering, höjer noggrannheten i skattningar baserade på modeller som innehåller fel. Metoderna hjälper ingenjören, som har god kännedom om systemet, modellen och dess brister, att svara på frågan om kompensation för ett visst fel är möjlig, samtidigt som metoderna kan peka ut andra potentiella felkällor. Ur metoderna fås en felbeskrivning som används för att utöka modellen. Genom att nyttja denna modell, utökad med felbeskrivning, kan information om motorns tillstånd beräknas med högre noggrannhet. I motorstyrenheter är dessutom uppslagstabeller för att beskriva komplicerade fenomen där fysikaliska modeller saknas vanligt förekommande. Dessa är ofta i behov av kontinuerlig anpassning för att kompensera för drift, åldrande och slitage av motorns fysiska komponenter och de framtagna metoderna lämpar sig väl även för detta ändamål.

Sammanfattningsvis förenar metoderna teori, som garanterar tillförlitliga och stabila skattningar, med industriella tillämpningar såsom anpassning av uppslagstabeller. Metoderna är utvärderade med hjälp av mätdata från standardiserade certifieringscykler insamlade i motorprovceller på Scania i Södertälje. I dessa cykler uppvisas minskningar av skattningsfel på i medel omkring 50 %. Reduktionen av skattningsfel möjliggör noggrann reglering, med minskade emissioner och bränsleförbrukning, samt förbättrar möjligheterna att upptäcka små fel.

ACKNOWLEDGMENTS

This work has been carried out at the division of Vehicular Systems, department of Electrical Engineering at Linköping University and the division of Engine Performance Software at Scania CV AB.

First, I would like to thank my supervisors Erik Frisk and Lars Eriksson for their guidance and many interesting, and often very fruitful, discussions. I would also like to thank professor Lars Nielsen for letting me join his research group.

All colleagues at Vehicular systems also deserve a place in this acknowledgment for creating such a nice research atmosphere and a special thanks goes to Jan Åslund for his guidance in the theory of matrices. Some of the characters in the world of enlightenment that made the weeks go scary fast are; my roommates Andreas Myklebust and Peter Nyberg, the “baljan” crew composed by Oskar “toe-kick/proof-reader” Leufvén, Emil “pony-tail” Larsson, and Christoffer “orienteer” Sundström, all of whom participated in several intriguing, seldom research related, discussions.

I would also like to thank my colleagues at Scania for showing great interest in the project and for all valuable inputs. Special thanks goes to my managers Peter Madsen and Mats Jennische for letting me be a part of this project, and Mats Jennische, David Elfvik, Anders Larsson, and Ola Stenlås for providing valuable insights and measurement data. During my short detour to the engine calibration department, made possible by Christer Eriksson, I met Martin Jonsson and Anders Gau who guided me in the world of engine maps, for which I am deeply grateful. While in the Scania department, I would also like to acknowledge my former supervisors Björn Völcker and Erik Geijer Lundin, and my project manager Lars Dahlén, whom contributed gratefully to the project. A special thanks goes to Carl Svärd, Tommy Sahi, and Dan Hallgren, for making my visits to Scania joyful and for introducing me to the Gröndal facility, which made the evenings go fast.

This work has been supported by SCANIA CV AB and Swedish Governmental Agency for Innovation Systems VINNOVA through the research program GRÖNA BILEN 2.

Finally, I would like to thank my family and friends for their encouragement and support. Evermore gratitude goes to my parents, Suzanne and Karl Johan for endless encouragement and my brothers Gunnar and Henrik for great support.

Erik Höckerdal
Linköping, April 2011

Contents

1	Introduction	1
1.1	Problem Statement	3
1.2	Thesis Outline	3
1.3	Contributions	5
1.4	Publications	5
2	Model Error Compensation	7
2.1	Application Example	7
2.2	Gas Flow Measurement	9
2.2.1	Air Mass-Flow Sensor Variations	11
2.3	Gas Flow Estimation	12
2.3.1	Methods for Improving Sensor Signals	12
2.4	Publications and Contributions	15
2.4.1	Paper A – Model Augmentation for ODE:s	16
2.4.2	Papers B and C – Map Adaptation	16
2.4.3	Paper D – ODE vs. DAE in Estimation	16
2.4.4	Paper E – Model Augmentation for DAE:s	17
	References	18
A	Model Details	23
B	Experimental Setup and Data	29
	Publications	31
A	Observer Design and Model Augmentation for Bias Compensation with a Truck Engine Application	33
1	Introduction	36
2	Problem Formulation	36
2.1	Problem and Paper Outline	38
3	Discretization and Linearization	38
4	Possible Augmentations	39

5	Augmentation Estimation	41
5.1	Bias Estimation	41
5.2	Augmentation Computation	43
5.3	Properties of the Estimated Augmentation	44
5.4	Approach Evaluation	45
5.5	Method Summary	46
6	Experimental Evaluation	47
6.1	Evaluation Using Simulated Data	47
6.2	Two Experimental Evaluations	48
7	Conclusions	53
	References	55
A	Engine Model and Data	56
B	Proofs of Theorems 4.2 and 5.1	56
 B EKF-Based Adaptation of Look-Up Tables with an Air Mass-Flow Sensor Ap- plication		59
1	Introduction	62
2	Method Outline	63
3	Observability	65
3.1	Unobservable Modes and Covariance Growth	68
4	Parameter and Bias State Convergence Rates	72
5	Method Evaluation	73
5.1	Observers	73
5.2	Observer Tuning	74
5.3	Study 1: Simulation	74
5.4	Study 2: Experimental data	81
6	Conclusions	84
	References	87
A	Proofs of Theorems 3.1 and 3.3	89
 C Off- and On-Line Identification of Maps Applied to the Gas Path in Diesel Engines		93
1	Introduction	96
2	Method Outline	96
3	Observability	99
3.1	Unobservable Modes and Covariance Growth	100
3.2	Method for Bias Compensation	101
4	Method Evaluation	102
4.1	Study 1: Simulation	102
4.2	Study 2: Experimental data	104
5	Conclusions	109
	References	110

D DAE and ODE Based EKF:s and their Real-Time Performance Evaluated on a Diesel Engine 111

- 1 Introduction 114
- 2 Background and Problem Motivation 115
 - 2.1 Stiffness in Engine Models 115
- 3 DAE Observer 117
 - 3.1 Observer Design – EKF for DAE:s 117
 - 3.2 EKF Algorithm 120
 - 3.3 Observability of Engine Model 120
- 4 Evaluation with Respect to Step Length 122
 - 4.1 Effects of Stiff ODE Dynamics 123
 - 4.2 Influence of Discretization Method 124
 - 4.3 Step Length Analysis 124
- 5 Conclusions 128
- References 130
- A Engine Model and Data 132

E Bias Reduction in DAE Estimators by Model Augmentation: Observability Analysis and Experimental Evaluation 135

- 1 Introduction 138
- 2 Problem Formulation & Solution Outline 138
- 3 EKF for DAE Systems 140
- 4 Observability of the Augmented Model 141
 - 4.1 DAE Observability 141
 - 4.2 Possible Augmentations 143
- 5 Augmentation Estimation 145
 - 5.1 Augmentation Properties 146
- 6 Experimental Evaluation 146
 - 6.1 Augmentation Estimation 146
 - 6.2 Estimation Performance 147
- 7 Conclusions 148
- References 150

Chapter 1

Introduction

Transportation is of vital importance in the modern economy and a major part of these transportations are carried out by trucks, e.g., in Europe and United States road vehicles account for more than 70 % of the inland freight transport (Noreland, 2008; Bradley, 2000). As a consequence, a major part of the emissions from the vehicular traffic is from trucks. It is therefore necessary to reduce the emissions and fuel consumption.

Stricter emission legislations and customer demands on low fuel consumption drive the technical development of engines and force new solutions to be introduced. To cope with reduced emission limits on diesel engines, for example intake manifold throttle, exhaust gas recirculation (EGR), and variable geometry turbine (VGT) are introduced, see Figure 1.1. This technical development, with increased system complexity and tightened requirements from customers and legislators, increase the demands on the control and diagnosis systems. Two examples of important quantities that significantly affect the emissions from diesel engines are: air to fuel ratio (λ) and EGR-fraction (x_{egr}). The increased demands on the control and diagnosis systems, increase the required information quality of λ and x_{egr} . At the same time it is desirable to have as few and cheap sensors in the system as possible to keep the cost down. This has made estimation an important and active research area, see e.g. Colin et al. (2009); Lino et al. (2008); García-Nieto et al. (2008); Andersson and Eriksson (2004).

Model based estimators are often used to achieve cost-effective estimation with high accuracy. This has driven the development of new models that are suitable for estimator design. These models have to be simple enough to be evaluated in real time, by for example an engine control unit (ECU), and at the same time describe the system behavior accurately enough for the estimation task. Development of these models is a delicate balance between computational complexity of the model and how well it manages to describe the true system. Typically, a large engineering effort is spent on modeling, often combining first law physics and system identification techniques.

In all model based control or diagnosis systems, the performance of the system is directly dependent on the accuracy of the model. In addition, as stated above, modeling

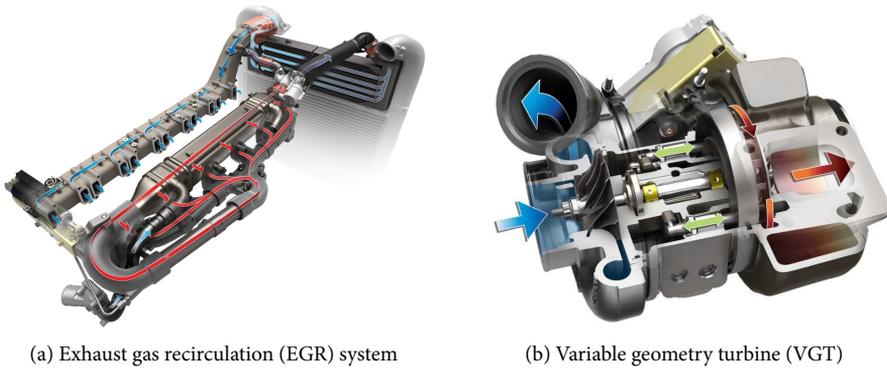


Figure 1.1: Technical solutions introduced on modern diesel engines to be able to fulfill the stricter emission legislations. *Courtesy Scania CV AB.*

is time consuming and even if much time is spent on physical modeling, there will always be errors in the model. The causes of these model errors can be quite varying; the model accuracy can depend on the operating point (Zimmerschied and Isermann, 2010), changes in ambient conditions (Won et al., 1998), the aging of components (Rupp and Guzzella, 2010), etc., all of which affect the system properties and hence the model errors. Model deficiencies are especially common if there are constraints on the model complexity, as is the case in most real time applications. Another scenario is that a model developed for some purpose, for example control, exists but needs improvement before it can be used for other purposes, for example diagnosis. That is, there exists a lot of models, on which much modeling time is spent, that needs improvement before they can be used in an estimation application. A common situation is that, while the dynamics is well captured by the model, there are stationary errors, possibly operating point dependent (Höckerdal et al., 2008). Hereafter, these already available models will be called default models. Since modeling is time consuming, and hence expensive, methods that enable use of these default models in estimation without involving extensive modeling efforts are needed.

In engine control and diagnosis, it is crucial to have good and unbiased estimates. In model based diagnosis (Ceccarelli et al., 2009) the true system is monitored using residuals, formed as the difference between estimated and measured signals. If the residual exceeds a threshold, it is concluded that something is wrong (Blanke et al., 2003; Isermann, 2011). In engine control (Stefanopoulou et al., 2000; Ortner and del Re, 2007; Plianos and Stobart, 2011), one objective is to control torque output while keeping the emissions below legislated levels and the fuel consumption as low as possible (Guzzella and Amstutz, 1998). Here, unbiased estimates are crucial since fuel consumption and emissions are often in conflict with each other. The hard constraints on the emissions force the engine operation away from the most fuel efficient operating point. With reduced stationary estimation errors the control system can balance closer to the fuel optimal operating point without the risk of violating the emission limits. For diesel

engines this is especially difficult since the control system normally does not have any feedback information from a λ or nitrogen oxides (NO_x) sensor and have to rely on estimated signals instead (Wang, 2008). In both cases, biased estimates impair the performance.

Finally, the development of engines and engine control systems involves extensive testing, both during the development of the control strategies and the engine calibration. Data is collected in engine test cells as well as in laboratory vehicles. This means that it is fairly easy to obtain system measurements. The sensors available are often both production sensors, that will be available on the commercial product, and high grade laboratory sensors added to enable extra monitoring. These laboratory sensors provide valuable information that can be used during the development phase, allowing estimation of model errors not possible to find with only production sensors.

1.1 PROBLEM STATEMENT

The objective is to develop systematic methods for reducing estimation errors given a default model and measurement data, without involving extensive modeling efforts.

The starting point is a default model and measurement data from the true system. From this it can be determined if the model describes the system sufficiently well or if it has to be modified to be applicable to the intended estimation application. The focus is on adjustments with respect to operating point dependent stationary estimation errors.

If it is concluded that the model suffers from too large stationary errors and cannot be used for estimation in its current state, then the methods developed for reducing stationary estimation errors can be applied. The ideas in the developed methods are to augment the default model with bias states that compensate for operating point dependent stationary errors. This augmented model can then be used in any suitable estimator design to get an adaptive estimator with reduced stationary errors compared to using the default model directly.

1.2 THESIS OUTLINE

The theme throughout the thesis is the successive development of methods for compensating operating point dependent stationary model errors in the design of estimators. The studied topics originate from estimation of gas flows in heavy duty diesel engines using existing mean value engine models (MVEM) (Hendricks, 1986; Jensen et al., 1991; Hendricks, 2001; Eriksson et al., 2002), referred to as default models.

Chapter 2 is based on Höckerdal et al. (2008) and describes an important estimation problem from the automotive industry. It gives an overview of the heavy duty diesel engine and model used for evaluation throughout the dissertation. This particular system is used to analyze how the quality of a sensor signal can be improved as well as how the quality can be assessed. The chapter illustrates the effect that a model with stationary errors has on the estimates when used in estimator design. Chapter 2 ends with a compilation of the contributions and their relation to other scientific work.

Papers A and E, based on Höckerdal et al. (2009) and Höckerdal et al. (Submitted), present systematic methods for bias compensation in model based estimator design for ordinary differential equation (ODE) and differential algebraic equation (DAE) models respectively. The methods apply the idea of introducing extra states, $q \in \mathbb{R}^{n_q}$, for adjusting the stationary operating point of the model, i.e. $x^\circ \rightarrow (x^\circ - A_q q)$, according to

$$\dot{x} = f(x - A_q q, u) \quad (1.1a)$$

$$\dot{q} = 0 \quad (1.1b)$$

$$y = h(x), \quad (1.1c)$$

where $x \in \mathbb{R}^{n_x}$ are the states, $u \in \mathbb{R}^{n_u}$ the inputs, and $y \in \mathbb{R}^{n_y}$ the outputs. In (1.1a), q represent the underlying cause of the bias, A_q its affection of the original states, x , and $A_q q$ shifts the stationary point of the model. Automatized methods for estimating low order augmentations, A_q , from measurement data are developed.

An operating point dependent bias can exhibit both fast and slow dynamics, arising from, for example, operating point dependent bias (Zimmerschied and Isermann, 2010) and aging (Rupp and Guzzella, 2010). Papers B and C address this problem in an integrated way by modeling the bias as a parameterized function,

$$q_{\text{fcn}}(x, u, \theta), \quad (1.2)$$

of known states and/or inputs instead of as an extra state

$$\begin{aligned} \dot{x} &= f(x, u, q_{\text{fcn}}(x, u, \theta)) \\ \dot{\theta} &= 0 \\ y &= h(x). \end{aligned} \quad (1.3)$$

The idea with a construction like (1.3) is to capture the operating point dependence of the bias by the parametrization (1.2), and use the parameters, $\theta \in \mathbb{R}^{n_\theta}$, introduced as new states, to track the aging. Paper B presents a solution and establishes necessary conditions for observability in the case the parameterized function is described by 1-D linear interpolation and an interpolation variable that is measured. Paper C extends the results with a simulation example using a 2-D cubic spline interpolation.

Paper D analyzes computational issues that arise when designing an observer for a stiff ODE system, containing both slow and fast dynamics, and especially what can be gained by approximating the fast dynamics with instantaneous relations resulting in a DAE system, i.e.

$$\begin{aligned} \dot{x}_{\text{slow}} &= f(x_{\text{slow}}, x_{\text{fast}}, u) \\ \dot{x}_{\text{fast}} &= g(x_{\text{slow}}, x_{\text{fast}}, u) \end{aligned} \quad \Rightarrow \quad \begin{aligned} \dot{x}_{\text{slow}} &= f(x_{\text{slow}}, x_{\text{fast}}, u) \\ 0 &= g(x_{\text{slow}}, x_{\text{fast}}, u). \end{aligned}$$

In an observer, efficient and accurate solution of these continuous-time models is necessary and has to be done in discrete-time. The properties of forward and backward Euler for discretization of the continuous-time model are also analyzed.

1.3 CONTRIBUTIONS

The main contributions are:

- * The experimental analysis of model and sensor errors of heavy duty diesel engines [Chapter 2].
- * Methods for estimating a low order bias compensating model augmentation using a default ODE or DAE model and measurements from the true system [Papers A and E].
- * Necessary and sufficient conditions for model augmentations that maintain system observability for ODE:s and DAE:s [Paper A, Theorem 4.2, and Paper E, Theorem 4.2].
- * Parametrization of all model augmentations that are possible to obtain with the proposed estimation algorithms [Paper A, Theorem 5.1].
- * An algorithm for engine map adaptation with variable parameter update rate [Paper B], with an additional 2-D cubic spline application example [Paper C].
- * An analysis of the benefits of approximating fast dynamics with instantaneous relations, transforming an ODE model into a DAE model, for EKF:s [Paper D].

1.4 PUBLICATIONS

The dissertation is based on the work presented in the following publications.

JOURNAL PAPERS

Erik Höckerdal, Erik Frisk, and Lars Eriksson. EKF-based Adaptation of Look-Up Tables with an Air Mass-Flow Sensor Application. In: *Control Engineering Practice*, 19(5):442–453, 2011. [**Paper B**]

Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer Design and Model Augmentation for Bias Compensation With a Truck Engine Application. In: *Control Engineering Practice*, 17(3):408–417, 2009. [**Paper A**]

Erik Höckerdal, Lars Eriksson, and Erik Frisk. Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. In: *SAE Int. J. Passeng. Cars – Electron. Electr. Syst.*, 1(1):393–402, 2008.

SUBMITTED

Erik Höckerdal, Erik Frisk, and Lars Eriksson. DAE and ODE Based EKF:s and their Real-Time Performance Evaluated on a Diesel Engine. In: *IEEE Transactions on Industrial Electronics*, 2011. [**Paper D**]

BOOK CHAPTER

Erik Höckerdal, Lars Eriksson, and Erik Frisk. Off- and On-Line Identification of Maps Applied to the Gas Path in Diesel Engines. In: *Identification for Automotive Systems*, Linz, Accepted for Publication, 2010. [**Paper C**]

CONFERENCE PAPERS

Erik Höckerdal, Erik Frisk, and Lars Eriksson. Model Based Engine Map Adaptation Using EKF. In: *6th IFAC Symposium on Advances in Automotive Control*. Munich, Germany, 2010.

Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer Design and Model Augmentation for Bias Compensation Applied to an Engine. *IFAC World Congress*. Seoul, Korea, 2008.

Erik Höckerdal, Lars Eriksson, and Erik Frisk. Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. In: *Electronic Engine Controls*. SAE Technical Paper 2008-01-0992. SAE World Congress, Detroit, USA, 2008.

SUBMITTED

Erik Höckerdal, Erik Frisk, and Lars Eriksson. Bias Reduction in DAE Estimators by Model Augmentation: Observability Analysis and Experimental Evaluation. In: *50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, Florida, 2011. [**Paper E**]

Model Error Compensation

As a prelude to the publications, some additional background is given with the purpose of putting the contributions into context. Even though the developed methods are general and applies to non-linear ODE and DAE models they are evaluated on automotive engine examples. A main challenge in engine control and diagnosis is accurate estimation of the internal state of the engine and was briefly described in Chapter 1 together with the contributions. This chapter elaborates on this, pointing out the necessity of unbiased estimates in engine control, and presenting some common properties of ordinary engine models. An overview of the heavy duty diesel engine with intake manifold throttle, EGR, and VGT is given in Section 2.1. Section 2.2 presents important control variables, the necessity of unbiased estimates, and the need for continuous adaptation in engine control and diagnosis, while Section 2.3 briefly describes the effect of biased models in model based estimation. Section 2.4 presents the publications with focus on the contributions and their relation to other scientific work.

2.1 APPLICATION EXAMPLE

This section serves as an overview of the system and the default models that are used for evaluation of the developed methods throughout the thesis. It also introduces the nomenclature, and presents important control quantities used in the control of diesel engines. Even though the methods developed are not specially devoted to engine applications, they are all applied and evaluated on the gas flow system of a Scania heavy duty diesel engine, like the one presented in Figure 2.1.

The default models used in the evaluations of the methods are developed in Wahlström and Eriksson (Accepted for publication), and Wahlström and Eriksson (2010). The main difference between the models are that the latter includes an intake manifold throttle, accompanied by an extra state for the intercooler pressure, and a state for the exhaust manifold temperature.

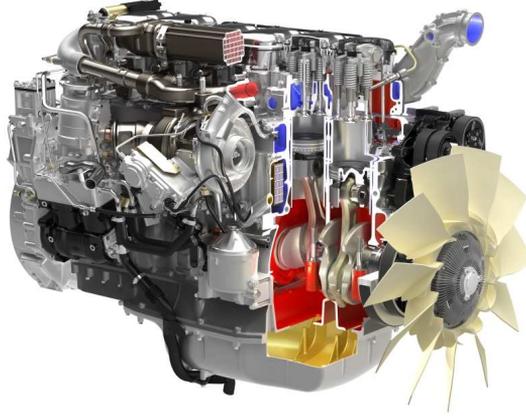


Figure 2.1: Cutaway view of the Scania inline six cylinder engine with VGT and EGR used for evaluation. *Courtesy Scania CV AB.*

Schematics of the more complex model from Wahlström and Eriksson (2010) is presented in Figure 2.2, where most of the modeled variables are presented. Control inputs to the model are injected amount of fuel u_δ and the positions of EGR, VGT, and throttle valves; u_{egr} , u_{vgt} , and u_{th} . The engine speed n_e is used as a parameterization input beside the control inputs, and thus the engine model can be expressed in state space form as

$$\begin{aligned}\dot{x} &= f(x, u, n_e) \\ y &= h(x).\end{aligned}$$

In these applications n_e is an input to the model which is due to the fact that the modeling is focused on the gas flows and does not include modeling of the produced torque and drive line. States are pressures in the intercooler, intake manifold and exhaust manifold, p_{ic} , p_{im} and p_{em} , turbine speed ω_t , and exhaust manifold temperature, T_{em} . Also presented are modeled signals for the, compressor mass-flow W_c , throttle mass-flow W_{th} , EGR mass-flow W_{egr} , mass-flow into the engine W_{ei} , mass-flow out of the engine W_{eo} , and turbine mass-flow W_t . Outputs from the model are the states, p_{im} , p_{em} , p_{ic} , and ω_t , and the compressor mass-flow W_c . Equations (2.1) and (2.2) presents a summary of the model and measurement equations and more details are presented in Appendix A.

$$\begin{aligned}\dot{p}_{im} &= f_{p_{im}}(p_{im}, p_{em}, p_{ic}, T_{em}, u_\delta, u_{egr}, u_{th}, n_e) \\ \dot{p}_{em} &= f_{p_{em}}(p_{im}, p_{em}, \omega_t, T_{em}, u_\delta, u_{egr}, u_{vgt}, n_e) \\ \dot{p}_{ic} &= f_{p_{ic}}(p_{im}, p_{ic}, \omega_t, u_{th}) \\ \dot{\omega}_t &= f_{\omega_t}(p_{em}, p_{ic}, \omega_t, T_{em}, u_{vgt}) \\ \dot{T}_{em} &= f_{T_{em}}(p_{im}, p_{em}, \omega_t, T_{em}, u_\delta, u_{egr}, u_{vgt}, n_e)\end{aligned}\tag{2.1}$$

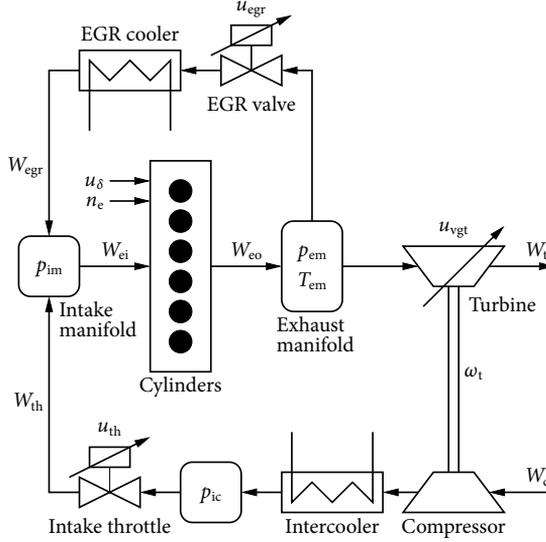


Figure 2.2: Schematic of the diesel engine model (Wahlström and Eriksson, 2010) with intake manifold throttle, EGR, and VGT, showing model states (p_{im} , p_{em} , p_{ic} , ω_t , and T_{em}), control inputs (u_{egr} , u_{vgt} , u_{δ} , and u_{th}), parametrization input (n_e), and flows between the different components (W_c , W_{th} , W_{egr} , W_{ei} , W_{eo} , and W_t). Rectangles with rounded corners represent control volumes.

$$y_1 = p_{im} \quad (2.2a)$$

$$y_2 = p_{em} \quad (2.2b)$$

$$y_3 = p_{ic} \quad (2.2c)$$

$$y_4 = \omega_t \quad (2.2d)$$

$$y_5 = W_c(p_{ic}, \omega_t) \quad (2.2e)$$

The data used is collected in engine test cells at Scania CV AB in Södertälje, Sweden, and a detailed sensor setup that includes accuracy and placement of the sensors used is presented in Appendix B.

2.2 GAS FLOW MEASUREMENT

The air mass-flow into the engine is a central quantity in the engine control systems and is hence often measured. It is used for many purposes and influences both the engine performance and emissions, and it is therefore essential to have an air mass-flow signal of good quality. One important issue with the air mass-flow sensor is its characteristics and

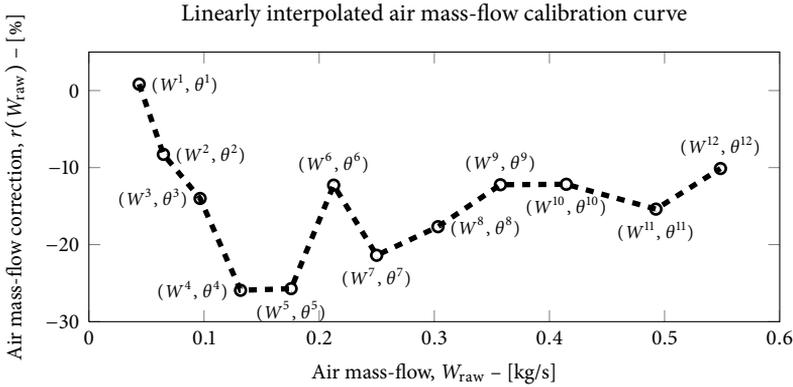


Figure 2.3: Air mass-flow sensor calibration curve with 12 grid points.

long term stability. To analyze this, two questions are addressed: how does the sensor characteristic evolve over time, and how does it vary between engine configurations?

To answer these questions, systematic engine test cell measurements have been conducted on a limited range of air mass-flow sensors over the span of several weeks. A central piece of information is a sensor calibration curve that has been recorded and stored for all days and tests. The data is analyzed with respect to day-to-day variations, aging, and changes between configurations. The calibration curve $r(W_{\text{raw}})$ is defined by

$$r(W_{\text{raw}}) = \frac{W_{\text{ref}}}{W_{\text{raw}}} - 1, \quad (2.3)$$

where W_{ref} is a reference sensor mass-flow sensor and W_{raw} is the raw engine air mass-flow measurement. The calibration curve is found by comparing the production air mass-flow sensor W_{raw} to a reference mass-flow sensor W_{ref} , for a long series of engine measurements. The reference sensor W_{ref} is available only in the engine test cell for the purpose of accurately being able to measure the air mass-flow into the engine, and has an uncertainty of less than 1 % and a response time of 12 ms. It is mounted on a straight pipe in the test cell, where the air mass-flow over the cross section of the pipe is orthogonal to the sensor and cylindrically symmetric, and is considered to give accurate measurements of the air mass-flow. The calibration curve is implemented as a lookup-table consisting of 12 grid points, see Figure 2.3 for an example. Using this calibration curve to adjust the raw sensor measurement an adapted sensor signal can be computed

$$W_{\text{adapt}} = (1 + r(W_{\text{raw}})) W_{\text{raw}}, \quad (2.4)$$

which gives a more accurate estimate of the true air mass-flow into the engine.

The air mass-flow signal is needed for computations of air to fuel ratio, λ , and EGR-fraction, x_{egr} . Both are important quantities that significantly affect the emissions. The air to fuel ratio is defined as

$$\lambda = \frac{W_{\text{air}}}{W_{\text{fuel}} (A/F)_s},$$

where W_{air} is the air mass-flow into the engine, W_{fuel} the fuel mass-flow, and $(A/F)_s$ the stoichiometric air to fuel ratio. In diesel engine control it is important to keep λ above a certain limit, $\lambda_{\text{smoke lim}}$, to avoid generating smoke. Normally, when λ is greater than $\lambda_{\text{smoke lim}}$, W_{fuel} is determined by the desired torque. However when the desired torque forces λ to $\lambda_{\text{smoke lim}}$, the control law enters a mode where W_{fuel} is proportional to W_{air} , (Wahlström, 2006). This is particularly important during transients where the torque demand is high, e.g. during acceleration. In these cases, an error in the air mass-flow signal results in either creation of smoke or reduced torque output. The other important quantity, is the EGR-fraction defined as

$$x_{\text{egr}} = \frac{W_{\text{tot}} - W_{\text{air}}}{W_{\text{tot}}},$$

where W_{tot} is the total gas mass-flow into the engine, i.e. $W_{\text{air}} + W_{\text{egr}}$. The x_{egr} is used in the engine control to reduce the NO_x emissions, and is governed by the EGR-valve and the VGT position. The following small example gives a rough estimate of the consequence of an incorrect air mass-flow measurement for the control of x_{egr} .

Example 1 Assume that the engine control system controls x_{egr} to 30 % based on the air mass-flow sensor and that the air mass-flow sensor signal is incorrect and reads $W_{\text{air}} \cdot 0.9$. That is,

$$x_{\text{egr}} = \frac{W_{\text{tot}} - W_{\text{air}} \cdot 0.9}{W_{\text{tot}}} = 30 \text{ \%}.$$

Then the true fresh air-fraction would become

$$(1 - 0.3) \cdot W_{\text{tot}} = W_{\text{air}} \cdot 0.9 \quad \Rightarrow \quad W_{\text{air}} = \frac{1}{0.9} \cdot (1 - 0.3) \cdot W_{\text{tot}} \approx 0.78 \cdot W_{\text{tot}},$$

and thereby the true $x_{\text{egr}} \approx 22 \text{ \%}$, which would have a significant effect on the NO_x emissions (Heywood, 1988). That is, in this example the control system controls the engine to run with less EGR than needed to fulfill the legislated NO_x levels. \diamond

An analogous analysis can be made for λ close to $\lambda_{\text{smoke lim}}$ which further supports the statement that an accurate estimate of the air mass-flow is important.

Both λ and x_{egr} are important for the emissions and the air mass-flow W_{air} is central in their control. Hence, it is important to have a high quality measurement or estimation of the air mass-flow. Note that the x_{egr} estimate also depends on W_{tot} , which is computed using the volumetric efficiency of the engine and is, by experience, considered to be accurate.

2.2.1 AIR MASS-FLOW SENSOR VARIATIONS

Calibration curves from two diesel engines, one inline 6 cylinder and one V8, are gathered from test runs in an engine test cell. 13 calibration curves are collected over a total time of about two weeks for the 6 cylinder engine and 21 calibration curves over four weeks for the V8 engine. Figure 2.4 presents the typical appearance of a calibration curve, the

upper for a 6 cylinder engine and the lower for a V8 engine. These calibration curves are used to analyze the quality of the air mass-flow sensor.

The difference between engine configurations can be seen by comparing the upper and lower plot in Figure 2.4 and Figure 2.5, where Figure 2.4 presents the day-to-day variations of the calibration curve and Figure 2.5 presents the trend of the four grid points, θ^2 , θ^5 , θ^8 , and θ^{11} , in the calibration curves, see Figure 2.3.

Figure 2.4 shows that the day-to-day variations are quite large, especially for the V8 engine where the standard deviation varies between 2 – 3 %-units. For the 6 cylinder engine the variations are smaller. Further, the difference between the minimum and maximum values for each parameter in the calibration curve varies between approximately 1.5 – 4 %-units for the inline six cylinder engine and 3 – 12 %-units for the V8 engine. Another difference between the two engine configurations is the appearance of the calibration curve. For the 6 cylinder engine the line starts at approximately 5 %, has a slightly positive slope, and ends at approximately 10 %, which corresponds to the computations in Example 1. For the V8 engine the line is quite different, it starts at about 1 %, varies quite a bit, and ends at -1 %. These investigations indicate that the air mass-flow sensor has to be continuously monitored and adapted, to ensure safe and clean engine operation over time.

The large spread among the calibration curves for the V8 engine plot, of about 10 %-units from min to max, indicates that an ad hoc approach for compensating the sensor signal using only a calibration curve (2.4) might not be enough, see Example 1. The quality has to be improved in a way that the spread is reduced as well. These observations together with the importance of the estimates of λ and x_{egr} necessitate an accurate estimate of the air mass-flow.

As Figure 2.5 shows there are no obvious trends in the data over time. However, due to the relatively short time span over which the data is collected, it is hard to draw any conclusions regarding long term aging of the air mass-flow sensors.

2.3 GAS FLOW ESTIMATION

In the previous section it was shown that a sensor is not sufficient for acquiring an accurate air mass-flow signal, and the main two reasons were; i) the sensor needs to be calibrated to compensate for its positioning in the intake system, ii) it needs continuous adaptation to compensate for system aging and different operating conditions caused by geographical location, for example pressure, temperature, and humidity of the surrounding air. This section presents some basic approaches to cope with sensor adaptation and include ad hoc mapping, according to (2.4), and Kalman filtering (Kalman, 1960). The investigation analyzes the effect model quality has on the estimates from a model based estimator, and is the topic of Section 2.3.1.

2.3.1 METHODS FOR IMPROVING SENSOR SIGNALS

There exist several ways of acquiring accurate estimates of these control and diagnosis variables, e.g., direct measurement via physical sensors and model based estimation, and

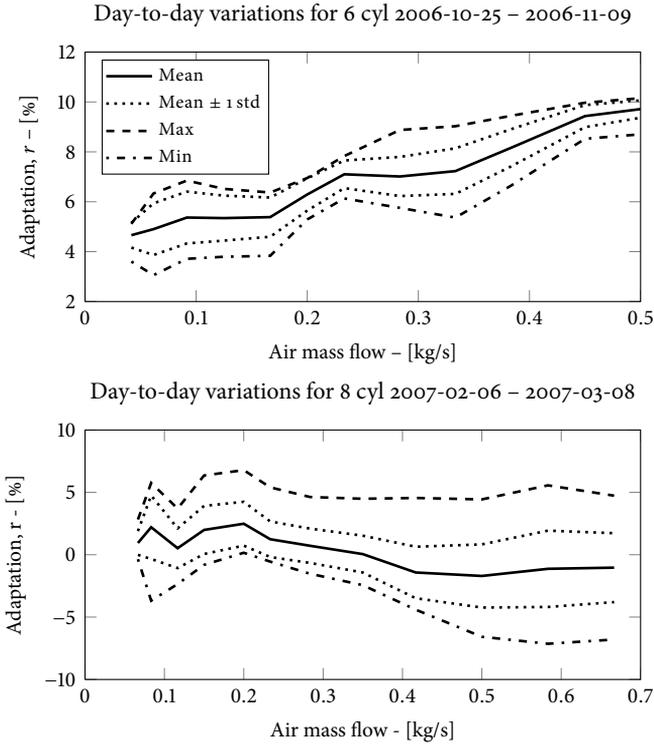


Figure 2.4: Min, max, mean, and standard deviation over all collected calibration curves are presented for a 6 cylinder engine (upper plot) and a V8 engine (lower plot). It can be seen that the variations are quite large for both engine configurations, especially for the V8 engine.

all model based estimators are highly dependent on the accuracy of the model used. This becomes especially apparent if the assumptions in the design method do not hold. If for example an EKF (Jazwinski, 1970) is used, the measurement and model errors are assumed to be described by zero mean white noise processes, i.e. biased measurements is not handled. Figure 2.6 presents estimates of the air mass-flow from five different sources; raw measurement from the production sensor, adapted production sensor (2.4), model output, EKF, i.e. combing the model and the adapted measurement, and a cell installed reference sensor. All representing means of acquiring estimates of the air mass-flow into the intake system of an engine.

One observation from Figure 2.6 is that the model output \hat{W}_{model} , computed using (2.1) and (2.2e), does not agree well with W_{ref} . It has an obvious offset that is different for low and high air mass-flows, but it manages to capture the system dynamics. From this observation it is clear that the model does not fully describe the engine and these model errors violate the assumptions made when utilizing the model to design an EKF,

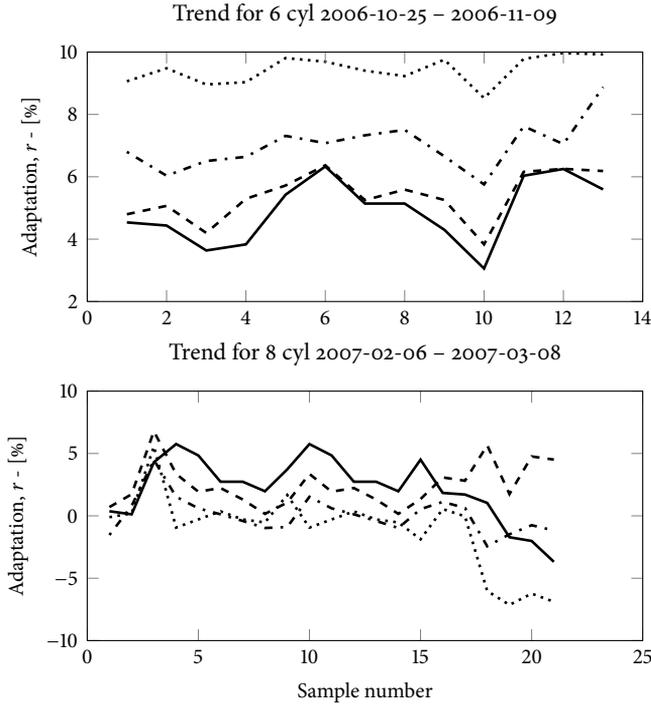


Figure 2.5: The trend of four support points for a 6 cylinder engine (upper plot) and a V8 engine (lower plot). It shows that there is no particular trend in either of the engine configurations. Note that the samples are not equidistant.

i.e. zero mean Gaussian system and measurement errors. Another observation is that also the raw measurement has an error that depends on the mass-flow. In this case a simple adaptation according to the calibration curve (2.3) in Section 2.2.1 significantly improves the estimation accuracy, see \hat{W}_{adapt} .

Obviously the model output and the raw measurement performs poorly, and by applying an adaptation scheme to the measurement much better estimates are acquired. Similarly, by combining the model with the adapted measurement in an EKF, even better estimates are achieved, see \hat{W}_{EKF} .

These estimators, the adapted mass-flow sensor and the EKF, compile the essence of the problems addressed in this thesis, i.e. the need for a systematic way of reducing operating point dependent stationary estimation errors in model based estimators, and the online adaptation of engine maps, or lookup-tables.

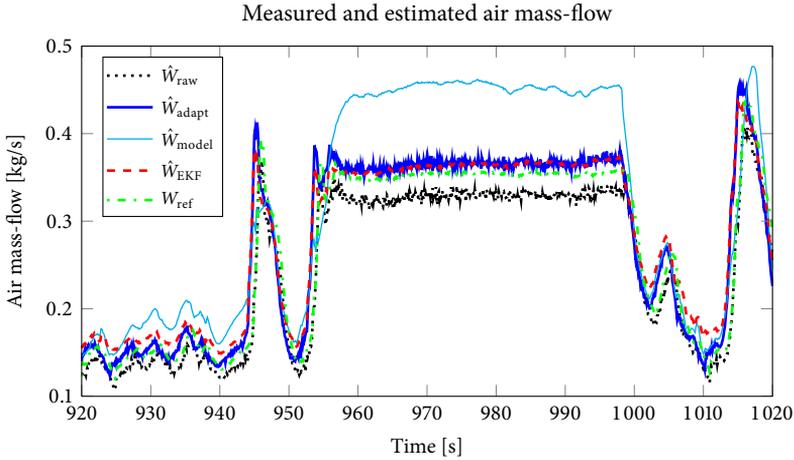


Figure 2.6: Typical example of model output from a biased model (Höckerdal et al., 2008), where W_{ref} is the air mass-flow measured by a reference sensor. As often is the case, the model captures the dynamics well but suffers from operating point dependent stationary errors. As comparison, the raw and adapted air mass-flow sensor measurements are presented, and an EKF using feedback from the adapted measurement is included as well.

2.4 PUBLICATIONS AND CONTRIBUTIONS

The overall goal with the work is the development of systematic methods that allow use of models with errors, referred to as default models, for estimator design. The focus is on models based on first principles physics and a primary condition on the methods is the preservation of the physical structure, or properties, of the models.

In system identification, model error modeling (MEM) is treated in for example Ljung et al. (1991); Stenman and Tjärnström (2000). However, since the focus here is on default models that have biases, or other stationary errors, and aims at preserving the physical structure of the model, the MEM path is not pursued. Methods that address the issue of biased default models for estimation exist in e.g. model augmentation using physical knowledge (Andersson and Eriksson, 2001) and proportional-integral (PI) observers (Söffker et al., 1995; Koenig and Mammar, 2002). The methods developed in this thesis unify these ideas with the idea of estimating a minimal description of the model bias.

In estimation, observability of the system is central to ensure consistent state and parameter estimates. This has made preservation of the default models' observability properties, in the developed methods, central. One method to check global observability is for example Ljung and Glad (1994) which is applied to an engine example in Sokolov and Glad (1999). However, this method applies to polynomial models and is not applicable to the models addressed in this dissertation. Hence, local analyses using model

linearizations, such as the Popov-Belevitch-Hautus (PBH)-test for ODE models (Kailath, 1980) and its DAE analogues (Dai, 1989), are used throughout the publications. An important observation is that the system (2.1) is coupled, meaning that several states have dependencies on both intake and exhaust states, which makes the default system locally observable from any output.

2.4.1 PAPER A – MODEL AUGMENTATION FOR ODE:S

The principal idea in the model augmentation is that local errors in the model may affect several model states. Consider for example an observer based on the engine model (2.1) consisting of three coupled volumes with one pressure state for each volume. Then an error in one of the mass-flow equations would, possibly, affect all three pressures. Some possibilities are then to, introduce a model augmentation using physical intuition (Andersson and Eriksson, 2001), or apply a PI-observer (Söffker et al., 1995). The first requires deep understanding of the modeled system while the latter only compensates for bias in measured states used for feedback and does not bother about the origin of the bias. The developed method applies a separate step in the observer design that estimates a low order model error description, which is used for model augmentation. The main contributions are a characterization of possible augmentations from observability perspectives, a parameterization of the augmentations from the method, and a robustness analysis of the proposed augmentation estimation method.

An advantage of the developed method, compared to e.g. PI observers, is its ability to incorporate information from extra sensors during the bias estimation. In this way compensation of states not available for feedback in the final application is made possible. It is also worth to note that both the model augmentation using physical knowledge and the PI-observer fits into the framework of the developed method.

2.4.2 PAPERS B AND C – MAP ADAPTATION

The ideas above address the bias compensation through model augmentation, by describing the bias as a random walk, and thus does not store any information about the bias in different operating points. A common technique to handle operating point dependencies in automotive applications is to introduce maps or look-up tables, (Guzzella and Amstutz, 1998; Peyton Jones and Muske, 2009). These maps are frequently used to describe relations when physical models are unavailable, e.g., sensor and actuator characteristics, cooler efficiency, injector characteristics, and aftertreatment systems. Common for these maps is that they benefit from continuous online adaptation to prevent undesired system behavior. Routines for online map adaptation have been considered in Wu (2006); Peyton Jones and Muske (2009), and a primary contribution in Paper B and Paper C is simultaneous bias compensation and online map adaptation.

2.4.3 PAPER D – ODE vs. DAE IN ESTIMATION

Using models with both fast and slow dynamics, i.e. stiff models, in real time estimation may be numerically problematic. The problem of stiff models, described by ordinary

differential equations (ODE), for engine control is closely connected to the embedded system in which it is implemented and its computational limitations. In engine control units (ECU), a main difficulty with stiff models is that the model execution is scheduled in loops with fixed frequencies, that limits the ECU:s capability of satisfactory solving the differential equations. A possible solution (Hairer and Wanner, 2000, Chapter 6), used in for example electrochemical and reactive distillation processes (Mandela et al., 2010), is to approximate fast dynamics with instantaneous relations, i.e. algebraic conditions. With this approach a stiff ODE would be transformed into a system of differential algebraic equations (DAE), while keeping the overall model structure. This is exploited for a diesel engine model with intake manifold throttle, EGR, and VGT in Paper D. The stability and estimation accuracy of an EKF based on the default stiff ODE is compared to that of an EKF based on the corresponding DAE. It is shown that even though the ODE, for each time-update, is less computational demanding than the resulting DAE, an EKF based on the DAE achieves better estimation performance with less computational effort. The main gain with the DAE based EKF is that it allows significantly increased step lengths without degrading the estimation performance compared to the ODE based EKF.

2.4.4 PAPER E – MODEL AUGMENTATION FOR DAE:S

The number of models described by DAE:s have increased, partly due to modern modeling tools such as DYMOLA, or similar tools using the Modelica® modeling language, and SIMSCAPE that often deliver DAE models and since DAE:s are a way of describing systems with both slow and fast dynamics. As more and more DAE models are available, it is natural to use them for observer or estimator design. Also these models may suffer from deficiencies that make them unsuitable for direct use in estimation and Paper E extends the model augmentation results for ODE:s from Paper A to DAE:s. The main contributions are necessary and sufficient conditions for the preservation of the observability properties of the default model during the augmentation.

REFERENCES

- Per Andersson and Lars Eriksson. Air-to-cylinder observer on a turbocharged SI-engine with wastegate. SAE Technical Paper 2001-01-0262, 2001. doi:10.4271/2001-01-0262.
- Per Andersson and Lars Eriksson. Cylinder air charge estimator in turbocharged SI-engines. In *Electronic Engine Controls*, number 2004-01-1366 in SAE Technical paper series SP-1822, 2004. doi:10.4271/2004-01-1366.
- Mogens Blanke, Michel Kinnaert, Jan Lunze, and Marcel Staroswiecki. *Diagnosis and Fault-tolerant Control*. Springer-Verlag Berlin and Heidelberg GmbH & Co.K, 2003.
- Ron Bradley. Technology roadmap for the 21st century truck program. Technical Report 21CT-001, U.S. Department of Energy, December 2000.
- Riccardo Ceccarelli, Carlos Canudas-de Wit, Philippe Moulin, and Antonio Sciarretta. Model-based adaptive observers for intake leakage detection in diesel engines. In *Proceedings of the 2009 conference on American Control Conference, ACC'09*, pages 1128–1133, Piscataway, NJ, USA, 2009. IEEE Press. ISBN 978-1-4244-4523-3. doi:10.1109/ACC.2009.5160133.
- Guillaume Colin, Gérard Bloch, Yann Chamailard, and Floriane Anstett. Two air path observers for turbocharged SI engines with VCT. *Control Engineering Practice*, 17(5): 571–578, 2009. ISSN 0967-0661. doi:10.1016/j.conengprac.2008.10.001.
- Council of European Parliament. Directive 2005/55/EC of the european parliament and of the council of 28 september 2005, 2005.
- Liyi Dai. *Singular Control Systems*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1989. ISBN 0387507248.
- Economic Commission for Europe – Inland Transport Committee. Regulation No 49 of the Economic Commission for Europe of the United Nations (UN/ECE). Official Journal of the European Union, August 2010.
- Lars Eriksson, Lars Nielsen, Jan Brugård, Johan Bergström, F. Pettersson, and Per Andersson. Modeling of a turbocharged SI engine. *Annual Reviews in Control*, 26(1): 129–137, October 2002. doi:10.1016/S1367-5788(02)80022-0.
- Sergio García-Nieto, Miguel Martínez, Xavier Blasco, and Javier Sanchis. Non-linear predictive control based on local model networks for air management in diesel engines. *Control Engineering Practice*, 16(12):1399–1413, December 2008. doi:10.1016/j.conengprac.2008.03.010.
- Lino Guzzella and Alois Amstutz. Control of diesel engines. *Control Systems Magazine, IEEE*, 18(5):53–71, October 1998. doi:10.1109/37.722253.
- E. Hairer and G. Wanner. *Solving Ordinary Differential Equations I*. Springer, 2000.

- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Bias reduction in DAE estimators by model augmentation: Observability analysis and experimental evaluation. Orlando, Florida, Submitted.
- Elbert Hendricks. A compact, comprehensive model of a large turbocharged, two-stroke diesel engine. SAE Technical Paper 861190, 1986. doi:10.4271/861190.
- Elbert Hendricks. Isothermal vs. adiabatic mean value SI engine models. In *3rd IFAC Workshop, Advances in Automotive Control, Preprints, Karlsruhe, Germany*, pages 373–378, March 2001.
- John B. Heywood. *Internal Combustion Engine Fundamentals*. McGraw-Hill, Inc., 1988.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. 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, 2008.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*, 17(3):408–417, 2009. doi:10.1016/j.conengprac.2008.09.004.
- Rolf Isermann. *Fault Diagnosis Applications*. Springer-Verlag Berlin and Heidelberg GmbH & Co.K, 2011.
- Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, April 1970. ISBN 0123815509.
- J.-P. Jensen, A.F. Kristensen, Spencer C Sorenson, Niels Houbak, and Elbert Hendricks. Mean value modeling of a small turbocharged diesel engine. SAE Technical Paper 910070, 1991. doi:10.4271/910070.
- Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc, Englewood Cliffs, New Jersey 07632, 1980.
- Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- Damien Koenig and Saïd Mammar. Design of proportional-integral observer for unknown input descriptor systems. *Automatic Control, IEEE Transactions on*, 47(12): 2057–2062, December 2002. ISSN 0018-9286. doi:10.1109/TAC.2002.805675.
- Paolo Lino, Bruno Maione, and Claudio Amorese. Modelling and predictive control of a new injection system for compressed natural gas engines. *Control Engineering Practice*, 16(10):1216–1230, October 2008. doi:10.1016/j.conengprac.2008.01.008.
- Lennart Ljung and Torkel Glad. On global identifiability of arbitrary model parameterizations. *Automatica*, 30(2):265–276, February 1994. doi:10.1016/0005-1098(94)90029-9.

Lennart Ljung, Bo Wahlberg, and Håkan Hjalmarsson. Model quality: The role of prior knowledge and data information. In *Proceedings of the 30th IEEE Conference on Decision and Control*, pages 273–278, Brighton, U.K., 1991. doi:10.1109/CDC.1991.261305.

Ravi Kumar Mandela, Raghunathan Rengaswamy, Shankar Narasimhan, and Lakshmi N. Sridhar. Recursive state estimation techniques for nonlinear differential algebraic systems. *Chemical Engineering Science*, 65(16):4548–4556, 2010. ISSN 0009-2509. doi:10.1016/j.ces.2010.04.020.

Jonas Noreland. Modal split in the inland transport of the EU – freight and passenger transport up to 2006. Technical Report 35, Eurostat, 2008.

Peter Ortner and Luigi del Re. Predictive control of a diesel engine air path. *Control Systems Technology, IEEE Transactions on*, 15(3):449–456, may 2007. ISSN 1063-6536. doi:10.1109/TCST.2007.894638.

James C. Peyton Jones and Kenneth R. Muske. Identification and adaptation of linear look-up table parameters using an efficient recursive least-squares technique. *ISA Transactions*, 48(4):476–483, October 2009. doi:10.1016/j.isatra.2009.04.007.

Alexandros Plianos and Richard K. Stobart. Nonlinear airpath control of modern diesel powertrains: a fuzzy systems approach. *International Journal of Systems Science*, 42: 263–275, 2011. doi:10.1080/00207721.2010.521864.

Daniel Rupp and Lino Guzzella. Adaptive internal model control with application to fueling control. *Control Engineering Practice*, 18(8):873–881, 2010. ISSN 0967-0661. doi:10.1016/j.conengprac.2010.03.011.

Dirk Söffker, Tie-Jun Yu, and Peter C. Müller. State estimation of dynamical systems with nonlinearities by using proportional-integral observer. *International Journal of Systems Science*, 26(9):1571–1582, 1995. doi:10.1080/00207729508929120.

Andrey Sokolov and Torkel Glad. Identifiability of turbocharged IC engine models. SAE Technical Paper 1999-01-0216, 1999. doi:10.4271/1999-01-0216.

Anna G. Stefanopoulou, Ilya Kolmanovsky, and James S. Freudenberg. Control of variable geometry turbocharged diesel engines for reduced emissions. *Control Systems Technology, IEEE Transactions on*, 8(4):733–745, jul 2000. ISSN 1063-6536. doi:10.1109/87.852917.

Anders Stenman and Fredrik Tjärnström. A nonparametric approach to model error modeling. In *Proceedings of the 12th IFAC Symposium on System Identification, Santa Barbara, USA*, pages 157–162, June 2000.

Johan Wahlström. Control of EGR and VGT for emission control and pumping work minimization in diesel engines. Technical report, 2006. LiU-TEK-LIC-2006:52, Thesis No. 1271.

Johan Wahlström and Lars Eriksson. Modeling of a diesel engine with intake throttle, VGT, and EGR. Technical Report LiTH-R-2976, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, 2010.

Johan Wahlström and Lars Eriksson. Modeling VGT EGR diesel engines with optimization of model parameters for capturing nonlinear system dynamics. *Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering*, Accepted for publication. doi:10.1177/0954407011398177.

Junmin Wang. Air fraction estimation for multiple combustion mode diesel engines with dual-loop EGR systems. *Control Engineering Practice*, 16(12):1479–1486, December 2008. doi:10.1016/j.conengprac.2008.04.007.

Mooncheol Won, S.B. Choi, and J.K. Hedrick. Air-to-fuel ratio control of spark ignition engines using gaussian network sliding control. *Control Systems Technology, IEEE Transactions on*, 6(5):678–687, September 1998. ISSN 1063-6536. doi:10.1109/87.709504.

Gang Wu. A table update method for adaptive knock control. SAE Technical Paper 2006-01-0607, 2006. doi:10.4271/2006-01-0607.

Ralf Zimmerschied and Rolf Isermann. Nonlinear time constant estimation and dynamic compensation of temperature sensors. *Control Engineering Practice*, 18(3):300–310, 2010. ISSN 0967-0661. doi:10.1016/j.conengprac.2009.11.008.

Appendix A

Model Details

Below follows a summary of the model equations using the symbols and indices presented in Tables A.1 and A.2. The model states, inputs, and outputs are presented in Table A.3 and more details about the model is found in Wahlström and Eriksson (2010).

MANIFOLDS

INTAKE MANIFOLD

$$\frac{d}{dt} p_{im} = \frac{R_a T_{im}}{V_{im}} (W_{th} + W_{egr} - W_{ei})$$

EXHAUST MANIFOLD

$$\begin{aligned} \frac{d}{dt} p_{em} &= \frac{R_e T_{em}}{V_{em}} (W_{eo} - W_{egr} - W_t) + \frac{p_{em}}{T_{em}} \frac{d}{dt} T_{em} \\ \frac{d}{dt} T_{em} &= \frac{R_e T_{em}}{p_{em} V_{em} c_{ve}} ((W_{eo} - W_{egr} - W_t) c_{ve} (T_{em,in} - T_{em}) + \\ &\quad R_e (T_{em,in} (W_{eo} - W_{egr} - W_t) - T_{em} (-W_{eo} + W_{egr} + W_t))) \end{aligned}$$

INTERCOOLER

$$\frac{d}{dt} p_{ic} = \frac{R_a T_{im}}{V_{ic}} (W_c - W_{th})$$

INTAKE THROTTLE

$$W_{th} = \frac{p_{ic} \Psi_{th}(\Pi_{th}) A_{th,max} f_{th}(u_{th})}{\sqrt{T_{im} R_a}}$$

$$\Psi_{th}(\Pi_{th}) = \begin{cases} \Psi_{th}^*(\Pi_{th}) & \text{if } \Pi_{th} \leq \Pi_{th,lin} \\ \Psi_{th}^*(\Pi_{th,lin}) \frac{1-\Pi_{th}}{1-\Pi_{th,lin}} & \text{if } \Pi_{th,lin} < \Pi_{th} \end{cases}$$

$$\Psi_{th}^*(\Pi_{th}) = \sqrt{\frac{2\gamma_{th}}{\gamma_{th}-1} \left(\Pi_{th}^{2/\gamma_{th}} - \Pi_{th}^{1+1/\gamma_{th}} \right)}$$

$$\Pi_{th} = \begin{cases} \left(\frac{2}{\gamma_{th}+1} \right)^{\frac{\gamma_{th}}{\gamma_{th}-1}} & \text{if } \frac{p_{im}}{p_{ic}} < \left(\frac{2}{\gamma_{th}+1} \right)^{\frac{\gamma_{th}}{\gamma_{th}-1}} \\ \frac{p_{im}}{p_{ic}} & \text{if } \left(\frac{2}{\gamma_{th}+1} \right)^{\frac{\gamma_{th}}{\gamma_{th}-1}} \leq \frac{p_{im}}{p_{ic}} \leq 1 \\ 1 & \text{if } 1 < \frac{p_{im}}{p_{ic}} \end{cases}$$

$$f_{th}(u_{th}) = b_{th1} (1 - \cos(\min(a_{th1} u_{th} + a_{th2}, \pi))) + b_{th2}$$

CYLINDER

CYLINDER FLOW

$$W_{ei} = \frac{\eta_{vol} p_{im} n_e V_d}{120 R_a T_{im}}$$

$$\eta_{vol} = c_{vol1} \frac{r_c - \left(\frac{p_{em}}{p_{im}} \right)^{1/\gamma_e}}{r_c - 1} + c_{vol2} W_f^2 + c_{vol3} W_f + c_{vol4}$$

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

$$W_{eo} = W_f + W_{ei}$$

EXHAUST MANIFOLD TEMPERATURE

CYLINDER OUT TEMPERATURE

$$T_e = T_{im} + \frac{q_{HV} f_{Te}(W_f, n_e)}{c_{pe} W_{eo}}$$

$$f_{Te}(W_f, n_e) = f_{TeWf}(W_f) \cdot f_{Tene}(n_e)$$

$$f_{TeWf}(W_f) = c_{fTeWf1} W_f^3 + c_{fTeWf2} W_{f,norm}^2 + c_{fTeWf3} W_{f,norm} + c_{fTeWf4}$$

$$f_{Tene}(n_e) = c_{fTene1} n_{e,norm}^2 + c_{fTene2} n_{e,norm} + 1$$

$$W_{f,norm} = W_f \cdot 100, \quad n_{e,norm} = \frac{n_e}{1000}$$

HEAT LOSSES IN THE EXHAUST PIPE

$$T_{em,in} = T_{amb} + (T_e - T_{amb}) e^{-\frac{h_{tot} \pi d_{pipe} l_{pipe} n_{pipe}}{W_{eo} c_{pe}}}$$

EGR-VALVE

$$W_{egr} = \begin{cases} \frac{A_{egr} p_{em} \Psi_{egr} \left(\frac{p_{im}}{p_{em}} \right)}{\sqrt{T_{em}} R_e} & \text{if } p_{em} \geq p_{im} \\ -\frac{A_{egr} p_{im} \Psi_{egr} \left(\frac{p_{em}}{p_{im}} \right)}{\sqrt{T_{egr,cool}} R_a} & \text{if } p_{em} < p_{im} \end{cases}$$

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

$$\Pi_{egrlim}(\Pi_{egr}) = \begin{cases} \Pi_{egropt} & \text{if } \Pi_{egr} < \Pi_{egropt} \\ \Pi_{egr} & \text{if } \Pi_{egr} \geq \Pi_{egropt} \end{cases}$$

$$A_{egr} = A_{egrmax} f_{egr}(u_{egr})$$

$$f_{egr}(u_{egr}) = b_{egr1}(1 - \cos(\min(a_{egr1}u_{egr} + a_{egr2}, \pi))) - b_{egr1}(1 - \cos(\min(a_{egr2}, \pi)))$$

TURBOCHARGER

TURBO INERTIA

$$\frac{d}{dt} \omega_t = \frac{P_t \eta_m - P_c}{J_t \omega_t}$$

TURBINE EFFICIENCY

$$P_t \eta_m = \eta_{tm} P_{t,s} = \eta_{tm} W_t c_{pe} T_{em} \left(1 - \Pi_t^{1-1/\gamma_e} \right)$$

$$\eta_{tm} = \eta_{tm,BSR}(BSR) \cdot \eta_{tm,\omega_t}(\omega_t) \cdot \eta_{tm,u_{vgt}}(u_{vgt})$$

$$\eta_{tm,BSR}(BSR) = 1 - b_{BSR} (BSR^2 - BSR_{opt}^2)^2$$

$$BSR = \frac{R_t \omega_t}{\sqrt{2 c_{pe} T_{em} \left(1 - \Pi_t^{1-1/\gamma_e} \right)}}$$

$$\eta_{tm,\omega_t}(\omega_t) = \begin{cases} 1 - b_{\omega t1} \omega_t & \text{if } \omega_t \leq \omega_{t,lim} \\ 1 - b_{\omega t1} \omega_{t,lim} - b_{\omega t2} (\omega_t - \omega_{t,lim}) & \text{if } \omega_t > \omega_{t,lim} \end{cases}$$

$$\eta_{tm,u_{vgt}}(u_{vgt}) = b_{vgt1} u_{vgt}^3 + b_{vgt2} u_{vgt}^2 + b_{vgt3} u_{vgt} + b_{vgt4}$$

$$\Pi_t = \frac{p_t}{p_{em}}$$

TURBINE MASS-FLOW

$$W_t = \frac{A_{\text{vgtmax}} p_{\text{em}} f_{\Pi t}(\Pi_t) f_{\omega t}(\omega_{t,\text{corr}}) f_{\text{vgt}}(u_{\text{vgt}})}{\sqrt{T_{\text{em}}} R_e}$$

$$f_{\Pi t}(\Pi_t) = \sqrt{1 - \Pi_t^{K_t}}$$

$$f_{\omega t}(\omega_{t,\text{corr}}) = 1 - c_{\omega t} (\omega_{t,\text{corr}} - \omega_{t,\text{corropt}})^2$$

$$\omega_{t,\text{corr}} = \frac{\omega_t}{100\sqrt{T_{\text{em}}}}$$

$$f_{\text{vgt}}(u_{\text{vgt}}) = c_{f_2} + c_{f_1} \sqrt{\max\left(0, 1 - \left(\frac{u_{\text{vgt}} - c_{\text{vgt}2}}{c_{\text{vgt}1}}\right)^2\right)}$$

COMPRESSOR EFFICIENCY

$$P_c = \frac{P_{c,s}}{\eta_c} = \frac{W_c c_{pa} T_{\text{amb}}}{\eta_c} \left(\Pi_c^{1-1/\gamma_a} - 1 \right)$$

$$\Pi_c = \frac{p_{ic}}{p_{\text{amb}}}$$

$$\eta_c(W_{c,\text{corr}}, \Pi_c) = \eta_{c,W}(W_{c,\text{corr}}, \Pi_c) \cdot \eta_{c,\Pi}(\Pi_c)$$

$$\eta_{c,W}(W_{c,\text{corr}}, \Pi_c) = 1 - a_{W3}(W_{c,\text{corr}} - (a_{W1} + a_{W2} \Pi_c))^2$$

$$\eta_{c,\Pi}(\Pi_c) = \begin{cases} a_{\Pi1} \Pi_c^2 + a_{\Pi2} \Pi_c + a_{\Pi3} & \text{if } \Pi_c < \Pi_{c,\text{lim}} \\ a_{\Pi4} \Pi_c^2 + a_{\Pi5} \Pi_c + a_{\Pi6} & \text{if } \Pi_c \geq \Pi_{c,\text{lim}} \end{cases}$$

$$a_{\Pi6} = \Pi_{c,\text{lim}}^2 (a_{\Pi1} - a_{\Pi4}) + \Pi_{c,\text{lim}} (a_{\Pi2} - a_{\Pi5}) + a_{\Pi3}$$

$$W_{c,\text{corr}} = \frac{W_c \sqrt{(T_{\text{amb}}/T_{\text{ref}})}}{(p_{\text{amb}}/p_{\text{ref}})}$$

COMPRESSOR MASS-FLOW

$$W_c = \frac{p_{\text{amb}} \pi R_c^3 \omega_t}{R_a T_{\text{amb}}} \Phi_c$$

$$\Psi_c = \frac{2 c_{pa} T_{\text{amb}} \left(\Pi_c^{1-1/\gamma_a} - 1 \right)}{R_c^2 \omega_t^2}$$

$$\Phi_c = \frac{k_{c1} - k_{c3} \Psi_c}{k_{c2} - \Psi_c}$$

$$k_{ci} = k_{ci1} (\min(Ma, Ma_{\text{max}}))^2 + k_{ci2} \min(Ma, Ma_{\text{max}}) + k_{ci3}, \quad i = 1, \dots, 3$$

$$Ma = \frac{R_c \omega_t}{\sqrt{\gamma_a R_a T_{\text{amb}}}}$$

Table A.1: Symbols used in the plant model.

Symbol	Description	Unit
A	Area	m^2
BSR	Blade speed ratio	–
c_p	Spec. heat capacity, constant pressure	$J/(kg \cdot K)$
c_v	Spec. heat capacity, constant volume	$J/(kg \cdot K)$
J	Inertia	$kg \cdot m^2$
n_{cyl}	Number of cylinders	–
n_e	Rotational engine speed	rpm
p	Pressure	Pa
P	Power	W
q_{HV}	Heating value of fuel	J/kg
r_c	Compression ratio	–
R	Gas constant	$J/(kg \cdot K)$
R	Radius	m
T	Temperature	K
u_{egr}	EGR control signal [†]	%
u_{th}	Throttle control signal [†]	%
u_{vgt}	VGT control signal [†]	%
u_δ	Injected amount of fuel	mg/cycle
V	Volume	m^3
W	Mass flow	kg/s
γ	Specific heat capacity ratio	–
η	Efficiency	–
Π	Pressure quotient	–
ρ	Density	kg/m^3
Φ_c	Volumetric flow coefficient	–
Ψ_c	Energy transfer coefficient	–
ω	Rotational speed	rad/s

[†] 0 – closed, 100 – open

Table A.2: Indices used in the plant model.

Index	Description
a	air
amb	ambient
c	compressor
d	displaced
e	exhaust
egr	EGR
ei	engine cylinder in
em	exhaust manifold
eo	engine cylinder out
ic	intercooler
f	fuel
im	intake manifold
t	turbine
th	throttle
vgt	VGT
vol	volumetric
δ	fuel injection

Table A.3: States, inputs, and outputs of the plant model.

States	Inputs	Outputs
p_{im}	u_{δ}	p_{im}
p_{em}	u_{th}	p_{em}
p_{ic}	u_{egr}	p_{ic}
ω_t	u_{vgt}	ω_t
T_{em}	n_e	W_c

Appendix B

Experimental Setup and Data

The data are collected in engine test cells at Scania CV AB in Södertälje, Sweden and are from two inline six cylinder Scania diesel engines with EGR and VGT. One of the engines was also equipped with an intake manifold throttle. The data were collected during a European transient cycle (ETC) (Council of European Parliament, 2005) for the engine without throttle, and a World harmonized transient cycle (Economic Commission for Europe – Inland Transport Committee, 2010) for the engine with throttle. The sensor signals used in all experimental evaluations are; intake and exhaust manifold pressures, turbine speed, and engine speed. Actuator signals used are; VGT and EGR positions, and injected amount of fuel. All these signals are available on a standard engine, i.e. no extra laboratory sensors were used, and collected at a sampling rate of 100 Hz.

An extra air mass-flow sensor, W_{ref} , is used as a reference for the experimental evaluation in Chapter 2. This signal is logged using a different measurement system at a sampling frequency of 10 Hz. The measurements from the different measurement systems are synchronized for the evaluation. The synchronization is made by comparing measurements of the engine speed which is logged with both systems, and performing a time shift.

SENSOR DYNAMICS

To justify that it is the system dynamics that is captured in the measurements, i.e. the sensors are fast enough to be able to track the system dynamics, a brief presentation of the sensor data is presented. The sensor specifications are provided by Scania.

The pressure sensors are capacitive pressure sensors and have a first order step response time constant of approximately 15 ms for the intake manifold pressure, and 20 ms for the exhaust manifold pressure. The intake manifold and intercooler pressure sensors are mounted directly in the intake manifold and right before the throttle, while the exhaust manifold sensor is mounted on an 0.4 m long pipe to avoid heat and soot.

The mass-flow sensor measuring the air mass-flow through the compressor is a hot-wire sensor also with a first order response and a time constant of 20 ms. The closed-loop control circuit maintains a constant temperature differential between the passing air stream and a platinum wire. The current required to heat the platinum wire provides an index of the air mass-flow.

The rotational speed sensors are inductive and measure the periodic variation in the magnetic flux generated by ferromagnetic ring gears passing induction coils. For the engine speed a cog passes the coil every sixth crankshaft degree and the signal used throughout this thesis is the mean value of 20 consecutive cog passes. This gives a time constant of approximately $20 \cdot (1 - e^{-1}) \approx 13$ samples. For the turbo speed there is only one cog on the ring gear and the signal used throughout the thesis is the median of three consecutive coil passes. That is, the maximum lag is roughly 13 times six crankshaft degrees and 2 times 360 turbo shaft degrees respectively. For the engine idle speed of 500 rpm, this gives a maximal time constant of approximately $13 \cdot (500/60 \cdot (360/6))^{-1} = 26$ ms, and for turbo speeds over 20 000 rpm, which is the minimum revolution speed for which the sensor works, this gives a maximum time delay of approximately $2 \cdot (20\,000/60)^{-1} = 6$ ms. Since these sensor responses are significantly faster than the dynamics seen in measurements they are neglected throughout the thesis.

REFERENCE SIGNAL – W_{REF}

The measured reference output W_{ref} is a cell sensor measuring the air mass flow into the engine. It is a Sensyflow P-Tube hot-wire sensor with type no. 14241-7962638 and a measuring range of 0.055-1.111 kg/s. The uncertainty is less than 1 % of reading and the sensor has a response time of 12 ms. This sensor is placed approximately 4 meters in front of the engine air mass-flow sensor on a straight pipe with a diameter of 0.28 m. The sensor reading is assumed to be without errors, due to the almost ideal sensor placement. The volume and distance between the two sensors give rise to unwanted dynamics. Calculations show that the filling and emptying dynamics from this volume has a time constant of approximately 10 ms and the effect from wave propagation has approximately the same traveling time, which is small in comparison to the time constants of the system.

PUBLICATIONS

Observer Design and Model Augmentation for
Bias Compensation with a Truck Engine
Application[☆]

A

[☆]Published in *Control Engineering Practice*, 17(3):408–417, 2009.

Observer Design and Model Augmentation for Bias Compensation with a Truck Engine Application

Erik Höckerdal, Erik Frisk, and Lars Eriksson

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, S-581 83 Linköping,
Sweden.*

ABSTRACT

A systematic design method for reducing bias in observers is developed. The method utilizes an observable default model of the system together with measurement data from the real system and estimates a model augmentation. The augmented model is then used to design an observer which reduces the estimation bias compared to an observer based on the default model. Three main results are a characterization of possible augmentations from observability perspectives, a parameterization of the augmentations from the method, and a robustness analysis of the proposed augmentation estimation method. The method is applied to a truck engine where the resulting augmented observer reduces the estimation bias by 50 % in a European Transient Cycle.

1 INTRODUCTION

In many application areas there are quantities that are important for control and diagnostics but that are not measured due to for example difficulties with the measurement methods or high costs of the sensors. This has made estimation an important and active research area, which is especially true in the automotive area where cost is important, see Lino et al. (2008); García-Nieto et al. (2008); Andersson and Eriksson (2004) for some examples.

In all model-based control or diagnosis systems, the performance of the system is directly dependent on the accuracy of the model. In addition, modeling is time consuming and, even if much time is spent on physical modeling, there will always be errors in the model. This is especially true if there are constraints on the model complexity, as is the case in most real time systems. Another scenario is that a model developed for some purpose, e.g. control, exists but needs improvements before it can be used for other purposes, for example diagnosis.

In many applications, like for example engine control and engine diagnosis, it is crucial to have unbiased estimates. In model based diagnosis, the true system is often monitored by comparing measured signals to estimated signals. If the magnitude of the difference, the residual, is above a certain limit a decision that something is wrong is made. In engine control, one objective is to maximize torque output while keeping the emissions below legislated levels and the fuel consumption as low as possible. For diesel engines this is especially hard since the control system does normally not have any feedback information from a λ - or NO_x -sensor and have to rely on estimated signals instead (Wang, 2008). In both cases, biased estimates impairs the performance.

The objective of this work is to develop a systematic method for reducing estimation bias in observers without involving further modeling efforts. This work is an extension of preliminary results in Höckerdal et al. (2008b) and the main extensions are a theoretical characterization of all solutions and additional method evaluations, including a robustness analysis with respect to measurement noise and model uncertainty.

The method utilizes an observable model and measurement data from the true system. The given model, referred to as the default model, and the measured inputs and outputs from the true system are used to estimate a suitable model augmentation. Then, the augmented model is used to design an observer that is shown to give estimates with reduced bias compared to an observer based on the default model. Three approaches for estimating a bias compensating augmentation are developed and evaluated with respect to measurement noise and model errors. Key results are a theoretical characterization of all possible augmentations from observability perspectives and a parametrization of the estimated augmentations. Finally the method is evaluated on a non-linear diesel engine model with experimental data from an engine test cell.

2 PROBLEM FORMULATION

Previous experience at Scania CV AB of state estimation based on an existing state-space model of a truck engine reveals that the model captures dynamic behavior reasonably

well but suffers from stationary errors (Höckerdal et al., 2008a). Designing an observer based on this model results in biased estimates. How to reduce the bias in a systematic manner is the topic of this paper.

The starting point is an existing model, referred to as the default model, that is provided in state-space form

$$\dot{x} = f(x, u) \quad (1a)$$

$$y = h(x), \quad (1b)$$

where x is the state-vector, u the known control inputs, y the measurement vector, and f and h are non-linear functions.

The objective is to find a systematic way to design an observer that gives an unbiased estimate of either the complete state x or a function of the state $z = g(x)$. This should be done even though the default model is subject to significant bias errors. A direct approach to compensate for constant, or slowly varying, biases is to augment the default model with bias variables q as

$$\dot{x} = \tilde{f}(x, u, q) \quad (2a)$$

$$\dot{q} = 0 \quad (2b)$$

$$y = \tilde{h}(x, q), \quad (2c)$$

and design the observer using this augmented model. If the augmentation captures the true modeling errors and the augmented system is observable, the observer estimates is made unbiased. An obvious question is then how to introduce the bias variable q in the model equations. One way could be through process knowledge, which have been successfully applied in Andersson and Eriksson (2004); Tseng and Cheng (1999). However, in this paper an estimation procedure based on available measurement data is proposed.

Besides the natural restriction, that the augmented model (2) is observable, it is also desirable not to introduce more extra bias states than necessary. It is therefore desirable to find a bias vector q with as low dimension as possible that manages to reduce the bias. Another reason for finding a low dimensional bias is that, since the model often is a first-principles physical model, bias in multiple states may be explained by one underlying bias affecting all these states. For example, bias in two pressures can originate from a bias in the mass flow between the two volumes or an incorrect modeling of energy conservation can give rise to bias in several states connected to the energy. However, the bias is necessarily not the same in the entire operating region of the system and may vary between operating points. This is part of the reason for introducing the bias as new states, rather than just a parameter, which allows some tracking ability of the bias.

In model (1) there are two natural ways to introduce biases, in the dynamic equation (1a) or in the measurement equation (1b). In the truck engine application the sensors, intake and exhaust manifold pressures and turbine speed, are considered more reliable than the model and the bias augmentation is therefore introduced in the dynamic equations

according to

$$\dot{x} = f(x - A_q q, u) \quad (3a)$$

$$\dot{q} = 0 \quad (3b)$$

$$y = h(x), \quad (3c)$$

where a stationary point of the system is moved by $A_q q$. The matrix A_q is thus a description of how the underlying bias variable q influences the stationary value of the state variable x . The model (3) will be referred to as the augmented model. It is worth mentioning that although the result in this paper focuses on biases in the dynamic equation, it is straightforward to modify the approach to also cover sensor biases.

2.1 PROBLEM AND PAPER OUTLINE

Based on the discussion above, the problem studied in the sections to follow can now be stated as: Given an observable default model (1) and available measurement data, find a low order bias augmented model (3) and design an observer that estimates x with reduced bias compared to using the default model. The observer should also be implementable in an Engine Control Unit (ECU).

To solve the problems, some issues need to be addressed. First, which matrices A_q are possible at all? All are not possible since it is required that the augmented system is observable and a characterization of possible augmentations is derived in Section 4. Among these possible bias augmentations, which should be used? Section 5 describes three approaches for how to estimate a, for bias compensation, suitable low order A_q based on measurement data.

Section 6 presents two examples of the proposed estimator design methodology applied to a Scania diesel engine using simulated and real measurement data respectively.

3 DISCRETIZATION AND LINEARIZATION

As a first step, the nonlinear augmented model (3) is transformed to a linearized time discrete model. A reason for the discretization is the demand on the implementation, which will be done in the ECU as a time discrete system. Here, a simple Euler forward discretization with step size T_s seconds is used. Note that observability does not depend on the choice of discretization method, since as long as T_s is chosen small enough the results are valid also for, e.g. zero-order-hold (Kalman et al., 1963).

One objective of the paper is to find a suitable A_q such that (3) is locally observable and to be able to use simple observability conditions, the observability analysis is here performed on a linearization of the non-linear model (3). Of course, non-linear observability is not guaranteed from observability of the linearization. Nevertheless, observability of a linearization in a stationary point is a sufficient condition for local observability of the non-linear system, see Theorem 6.4 in Lee and Markus (1968). Even though observability is not strictly guaranteed, e.g. in transient mode when moving between operating points, the referred result gives theoretical support for using the

linearized system in the observability analysis. Thus, when analyzing (3) the following model will be used

$$\begin{pmatrix} x_{t+1} \\ q_{t+1} \end{pmatrix} = \begin{pmatrix} I + T_s A & -T_s A A_q \\ 0 & I \end{pmatrix} \begin{pmatrix} x_t \\ q_t \end{pmatrix} + \begin{pmatrix} T_s B \\ 0 \end{pmatrix} u_t \quad (4a)$$

$$y_t = (C \quad 0) \begin{pmatrix} x_t \\ q_t \end{pmatrix}, \quad (4b)$$

where

$$A = \left. \frac{\partial f}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}}, \quad B = \left. \frac{\partial f}{\partial u} \right|_{\substack{x=x_0 \\ u=u_0}}, \quad \text{and} \quad C = \left. \frac{\partial h}{\partial x} \right|_{\substack{x=x_0 \\ u=u_0}}.$$

In the following, $I + T_s A$ is substituted for F to increase readability and (4) becomes

$$\begin{pmatrix} x_{t+1} \\ q_{t+1} \end{pmatrix} = \begin{pmatrix} F & -(F - I)A_q \\ 0 & I \end{pmatrix} \begin{pmatrix} x_t \\ q_t \end{pmatrix} + \begin{pmatrix} T_s B \\ 0 \end{pmatrix} u_t \quad (5a)$$

$$y_t = (C \quad 0) \begin{pmatrix} x_t \\ q_t \end{pmatrix}. \quad (5b)$$

4 POSSIBLE AUGMENTATIONS

Augmenting a model with more states may affect the observability of the model. Since the purpose of the augmented model is to use it for estimation, observability has to be maintained also after the augmentation. To find which augmentations that are possible an observability investigation of the augmented model is performed. The aim is to derive a necessary and sufficient condition on A_q such that the augmented model is observable. The observability criterion used in the analysis is known as the Popov-Belevitch-Hautus (PBH)-test (Kailath, 1980).

Similar results can be found in Bembenek et al. (1998), which also includes a discussion regarding the observability results, similar to the short discussion in the end of this section.

Theorem 4.1 *A pair (C, F) is observable if and only if*

$$\begin{pmatrix} C \\ \lambda I - F \end{pmatrix}$$

has full column rank $\forall \lambda \in \mathbb{C}$.

Now, using Theorem 4.1 and the assumption that the default model is observable the main result of this section can be formulated as

Theorem 4.2 *Assume that (C, F) in (5) is an observable pair then the augmented system (5) is observable if and only if*

$$\text{Ker}((F - I)(A_q \quad N_C)) = \{0\},$$

where the columns of N_C span $\text{Ker } C$.

Note that this is equivalent to

$$(F - I)(A_q \ N_C),$$

having full column rank.

Proof See Appendix B. □

This means that the space spanned by the columns in A_q can lie neither in $\text{Ker } C$ nor in $\text{Ker } (F - I)$ for the augmented model to be observable. These interpretations of the rank condition can be understood by analyzing the two requirements separately. First, the requirement that A_q can not lie in $\text{Ker } C$ is easily seen by studying a linear example.

Example 1 Starting with a linear model with a stationary bias

$$\begin{aligned} x_{t+1} &= Fx_t - (F - I)A_q q_t \\ q_{t+1} &= q_t \\ y_t &= Cx_t \end{aligned} \tag{6a}$$

and performing a change of variables, $z_t = x_t - A_q q_t$, gives

$$\begin{aligned} z_{t+1} &= Fx_t - (F - I)A_q q_t - A_q q_t = Fz_t \\ q_{t+1} &= q_t \\ y_t &= Cz_t + CA_q q_t, \end{aligned} \tag{6b}$$

which shows that columns of A_q in $\text{Ker } C$ are not observable. ◇

Second, a non-empty $\text{Ker } (F - I)$ implies that the system contains pure integrators, and a bias in $\text{Ker } (F - I)$ is not distinguishable from an unknown initialization of the integrator and is therefore not observable.

A closer look at the requirement that $(F - I)(A_q \ N_C)$ has to have full column rank conveys some other interesting results. First, assuming full column rank of $(F - I)$, it is easily seen that the number of augmented states n_q never can exceed the number of linearly independent measurement signals n_y since

$$\text{rank}(F - I)(A_q \ N_C) = \text{rank}(A_q \ N_C) \leq \text{rank } A_q + \text{rank } N_C = n_q + n_x - n_y \leq n_x, \tag{7}$$

i.e. $n_q \leq n_y$. Second, again imagine that $(F - I)$ has full rank which means that the model does not have any pure integrators, then the full column rank condition on $(F - I)(A_q \ N_C)$ reduces to requiring full column rank of $(A_q \ N_C)$ or, equivalently, full column rank of the product CA_q . Now if C has one or several zero columns, then CA_q will not contain any information from those rows in A_q corresponding to zero columns in C . That is, those rows in A_q that correspond to zero columns in C will not contribute to the observability, see the following example.

Example 2 Illustration of possible augmentations for a default model without pure integrators and

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let \times denote a non-zero element, then some possible augmentations are

$$A_q^1 = \begin{pmatrix} \times & 0 \\ 0 & \times \\ 0 & 0 \end{pmatrix}, \text{ and } A_q^2 = \begin{pmatrix} 0 & \times \\ \times & 0 \\ 0 & 0 \end{pmatrix}$$

since

$$CA_q^1 = \begin{pmatrix} \times & 0 \\ 0 & \times \end{pmatrix}, \text{ and } CA_q^2 = \begin{pmatrix} 0 & \times \\ \times & 0 \end{pmatrix},$$

which have full column rank. While an augmentation

$$A_q^3 = \begin{pmatrix} \times & 0 \\ 0 & 0 \\ 0 & \times \end{pmatrix} \text{ is not possible since } CA_q^3 = \begin{pmatrix} \times & 0 \\ 0 & 0 \end{pmatrix}$$

does not have full column rank. \diamond

5 AUGMENTATION ESTIMATION

Now that all possible model augmentations have been characterized, the next question is how to find a suitable augmentation, using measured data from the real system, that fulfills the requirements derived in Section 4. The proposed augmentation estimation procedure is divided into two steps, i) from measured data estimate samples of the bias and ii) compute a basis for the bias samples. Three approaches for how to conduct the first step are developed. In the second step a low order augmentation is computed by performing a Singular Value Decomposition (SVD) on selected samples of the bias found in step one.

5.1 BIAS ESTIMATION

The first step in the estimation of a low order model augmentation deals with estimating the bias, i.e. collect samples of the bias $\beta_t = A_q q_t$. The first approach is quite simple and its main purpose is to illustrate the basic ideas for the estimation of bias samples, whereas the second and third approach are applicable to more general systems. Since the method aims at reducing bias in stationary operating points only stationary behavior and data is studied.

APPROACH 1

The first approach utilizes the discretized linearization directly and the assumptions that all states are measured, i.e. C has full column rank, and that the system does not have any \diamond

pure integrators, i.e. $(I - F_t)$ has full column rank. The linearized and time discretized augmented model is

$$\begin{aligned}x_{t+1} &= F_t x_t + (I - F_t) A_q q_t + T_s B_t u_t \\ y_t &= C_t x_t.\end{aligned}$$

Due to the full column rank assumptions on C and $(I - F_t)$ it is possible to invert the measurement equation and insert the resulting x in the dynamic equation. This gives that

$$\beta_t = A_q q_t = (I - F_t)^{-1} (C_{t+1}^\dagger y_{t+1} - F_t C_t^\dagger y_t - T_s B_t u_t),$$

where \dagger denotes the pseudo inverse.

This approach has three evident flaws, it requires a full column rank C and $(I - F_t)$ and, since no filtering of the measurements is involved, it is sensitive to low Signal to Noise Ratio (SNR).

Therefore two other approaches are proposed for estimating bias samples. Common for both these approaches are that they utilize the residuals from an observer and the assumption that the true bias enters the model according to Equation (3). The fact that they are based on observers makes them less sensitive to low SNR and imply that they do not require full column rank C to work. The first employ an observer based on the default model and the bias samples are computed by inverting the observer system. The second employ a fully augmented model fulfilling the observability requirements developed in Section 4.

APPROACH 2

The second approach uses the residuals originating from an observer based on the default model. Here, the observer is an Extended Kalman Filter (EKF) (Kailath et al., 2000), where the noise covariance matrices Q and R are design parameters tuned by the user. Of course, other observer designs are equally possible but here an EKF is used. Let K_t be the EKF feedback gain then the estimation error becomes,

$$\begin{aligned}e_{t+1} &= x_{t+1} - \hat{x}_{t+1|t+1} \\ &= F_t x_t + (I - F_t) A_q q + T_s B_t u_t - \\ &\quad (F_t \hat{x}_{t|t} + T_s B_t u_t + K_t (y_{t+1} - C_t F_t \hat{x}_{t|t} - C_t T_s B_t u_t)) \\ &= \{y_{t+1} - C_t F_t x_t + C_t (I - F_t) A_q q + C_t T_s B_t u_t\} \\ &= (F_t - K_t C_t F_t) e_t + (I - K_t C_t) (I - F_t) A_q q.\end{aligned}\tag{9}$$

Equation (9) can not be used directly since the state estimation error is not known. Therefore, the output error

$$r_t = y_t - \hat{y}_{t|t} = C_t (x_t - \hat{x}_{t|t}) = C_t e_t,\tag{10}$$

is used for estimating the bias.

As previously stated, solely stationary parts of the residuals are involved in the bias estimation. It would be possible to use also dynamic parts of the residuals and a dynamic

inverse. The reason for not utilizing these here is to prevent dynamic estimation errors from affecting the estimation of the constant or slowly varying bias.

Now, utilizing that only stationary data is considered, (9) and (10) can be combined resulting in

$$\begin{aligned} r_{\text{stat}} &= C_{\text{stat}} e_{\text{stat}} \\ &= C_{\text{stat}}(I - F_{\text{stat}} + K_{\text{stat}} C_{\text{stat}} F_{\text{stat}})^{-1} \times (I - K_{\text{stat}} C_{\text{stat}})(I - F_{\text{stat}}) A_q q_{\text{stat}} \end{aligned}$$

and the bias can be estimated as

$$\beta_t = A_q q_t = (C_{\text{stat}}(I - F_{\text{stat}} + K_{\text{stat}} C_{\text{stat}} F_{\text{stat}})^{-1} \times (I - K_{\text{stat}} C_{\text{stat}})(I - F_{\text{stat}}))^{\dagger} r_t. \quad (11)$$

APPROACH 3

An alternative to Approach 2 for finding β_t is to augment the default model with as many extra states as possible. According to Theorem 4.2 the requirement on A_q is that, $(F - I)(A_q N_C)$ has to have full column rank. This means that A_q can have a maximum of n_y columns. These columns have to be linearly independent of the columns of N_C and can not lie in $\text{Ker}(F - I)$. One way to construct such an augmentation is to use C^{\dagger} and leave out those columns that become zero when multiplied by $(F - I)$ from the left. Then run the observer based on the augmented model, estimating both \hat{x} and \hat{q} , and assemble $\beta_t = C^{\dagger} \hat{q}_t$.

An advantage with this approach is that no inversions as those in (11) are needed. A disadvantage though is that since a fully augmented model is used the order of the observer might be unnecessarily high.

5.2 AUGMENTATION COMPUTATION

As stated in the problem formulation in Section 2, the bias is necessarily not the same in the entire operating region of the system. This makes it important to collect samples of the bias from stationary operating points selected such that the entire operating region is covered. From the first step of the proposed procedure, bias samples are collected according to this. Based on the discussion of only a few underlying biases affecting several states in Section 2, the task of step two is to find a low order basis spanning the space in which these bias samples are located.

To start with bias samples from N stationary operating points are assembled

$$\tilde{\beta}^{n_x \times N} = (\beta^1 \quad \dots \quad \beta^N),$$

Then the SVD of $\tilde{\beta}$ is computed,

$$\tilde{\beta} = U \Sigma V^*,$$

where U contains orthogonal vectors spanning the space in which the bias moves and Σ the corresponding singular values. The singular values in Σ are ordered in non-increasing order which means that the far left columns of U , corresponding to large singular values, represent the most dominating directions along which the bias moves. Therefore the

dimension of q can be found by comparing the singular values in Σ , and picking the most significant ones. Then the corresponding columns of U are used to assemble \hat{A}_q .

This way of computing an augmentation from bias samples is optimal with respect to the Frobenius norm.

5.3 PROPERTIES OF THE ESTIMATED AUGMENTATION

According to the discussion in the end of Section 4, the properties of C place restrictions on which A_q 's that are possible to find. The conclusion of that discussion is that rows in A_q corresponding to zero columns in C become zero in the estimation step. However, a more thorough analysis of the three bias estimation approaches shows that more can be said.

Theorem 5.1 *Assume that the observer gain, K , is chosen such that the observer is strictly stable and does not have any poles in the origin. Then, in absence of noise, the bias samples are spanned by the rows of C and can thereby be written as*

$$\beta_t = C^T \Gamma.$$

Proof See Appendix B. □

Note that Theorem 5.1 holds for the pseudo inverse and is not generally true for an arbitrary left inverse.

As a consequence, the observer based on an estimated augmentation may not be able to reduce the bias in the estimates to acceptable levels. This problem can be circumvented in, for example one of the two following ways. The first is for an engineer to design an A_q not possible to find through estimation, for example through knowledge of the underlying physics. The second is to, during the design phase, add extra sensors to the true system to acquire a full column rank C which enables estimation of all rows in A_q . When utilizing this possibility one must be cautious and check the observability of the augmented system that in the end will not rely on the additional sensors used for estimating \hat{A}_q . That is, check the column rank of $(F - I)(\hat{A}_q N_C)$, and in case of column rank deficiency remove those columns in \hat{A}_q causing rank deficiency. Since SVD is used, the columns in \hat{A}_q are arranged in non-increasing significance order which makes it appropriate to remove the columns in \hat{A}_q starting from the right to get an augmentation that is observable.

The example below illustrates the remarks regarding the effects that properties of C have on the augmentation estimation.

Example 3 Consider a true system with

$$F = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

and a true bias,

$$A_q = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Then the estimate of A_q , according to Theorem 5.1, will be

$$\hat{A}_q = \begin{pmatrix} 1 \\ 2 \times 7/5 \\ 1 \times 7/5 \end{pmatrix},$$

where the factor $7/5$ comes from minimizing

$$\left\| A_q - C^T \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix} \right\| = (1 - \Gamma_1)^2 + (2 - 2\Gamma_2)^2 + (3 - \Gamma_2)^2$$

with respect to Γ_1 and Γ_2 .

That is, because of the structure of C the top element in \hat{A}_q will be correct while the bottom two elements will not. \diamond

5.4 APPROACH EVALUATION

Two main approaches, approaches 2 and 3, for estimating the bias have been proposed. It is important to understand how these approaches perform under varying operating conditions and model uncertainty. Therefore, the approaches are evaluated with respect to robustness against model errors and robustness to changes in noise levels. This is done by introducing noise and modeling errors in a non-linear simulation model of a Scania diesel engine with exhaust gas recirculation (EGR) and variable geometry turbine (VGT), and performing Monte Carlo simulations. In the simulations, a one-dimensional q is also introduced, i.e. A_q is a vector with three elements.

Modeling errors can be introduced in many ways and it is difficult to obtain a comprehensive evaluation of robustness properties of a non-linear method. Therefore, a more pragmatic approach is adopted. First, model errors are introduced by manipulating physical constants in the simulation model and thus making the simulation model, that generates the observations, different from the default model used for designing the observer. Another way model errors are introduced is by pre-multiplying the vector field f in (1a) by a slowly varying sinusoid, i.e. the simulation is done with $\tilde{f}(x, u)$ defined as $\tilde{f}(x, u) = (1 + \gamma \sin(\Lambda t))f(x, u)$, where Λ is the model error frequency, and γ is a small number varied between 0.1 and 0.5. Doing Monte Carlo simulations with such model errors introduced reveal that both approaches react similarly to the model errors with respect to degraded performance in bias estimation and variance in the estimation. No certain conclusion can be drawn concerning which approach is more robust against modeling errors and the overall picture is that both approaches have similar graceful performance degradation with increased modeling errors.

Examining the effect of measurement noise is done by introducing white Gaussian noise with different noise levels in the simulation model and estimating the effect on the

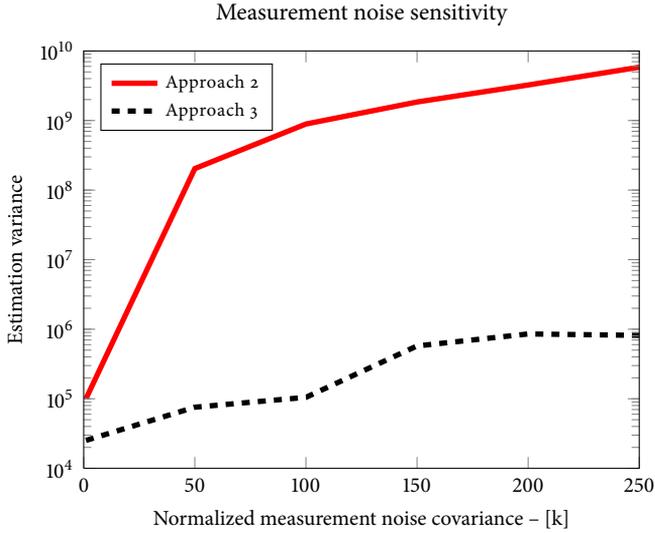


Figure 1: Estimation variance for bias estimation approaches 2 and 3 with different measurement noise – $k \cdot \mathcal{N}(0, R)$.

augmentation estimation by computing the variance in the β^i :s. In Figure 1 the effect of increased measurement-noise level on the variance in the estimated β^i :s is shown. It is seen that Approach 3 is significantly less sensitive to measurement noise and therefore preferable when estimating an augmentation.

5.5 METHOD SUMMARY

The procedure can be summarized in three steps.

Step 1 - Linearize and discretize the model if necessary. Normally, the default model is a non-linear time continuous model such as (1) and has to be linearized and discretized.

Step 2 - Find an appropriate augmentation, A_q , and compile an augmented model (4). Here the designer has a choice, either to estimate an augmentation from measured data, introduce an augmentation found in some other way, or to combine an estimated augmentation with one found through system knowledge.

The estimation procedure contains two steps, i) estimation of bias samples utilizing one of the three approaches presented in Section 5.1, ii) compute a basis for the bias samples using SVD.

With good knowledge of the system, the designer might have some idea of what is causing the bias in the estimates and can choose an appropriate A_q .

To combine an augmentation found through process knowledge with one found through estimation can be desirable if some model deficiencies are known but does not manage to achieve satisfactory bias reduction. In this case the estimation approach can be applied to the, by the engineer, partly augmented model to find an additional augmentation that captures the remaining dominating bias.

Step 3 - Design an observer based on the augmented model (3) and the A_q found in Step 2. In this paper, an Extended Kalman Filter is used but any non-linear observer design methodology is possible.

6 EXPERIMENTAL EVALUATION

To evaluate the method experiments are performed using a non-linear model of a heavy-duty truck engine. The experiments consist of a simulation study of the non-linear model, and evaluation of the method on measurement data from an engine test cell.

The non-linear model of the Diesel engine has three states: p_{im} , p_{em} , and n_{trb} , that represent intake and exhaust manifold pressures, and turbine speed respectively. See Appendix A for more information about the engine and model. In the second experiment, real data from the engine is used together with the engine model to illustrate the properties of the proposed approach in a real application. In both experiments the stationary parts of the data, used in the augmentation estimation, are separated out through visual inspection and estimation Approach 3 is chosen to estimate the bias.

6.1 EVALUATION USING SIMULATED DATA

The objective of the first experiment is to illustrate how the approach, which is based on linearization procedures, performs when fed with data from a non-linear simulation model. Thus, synthetic data is created where known biases are introduced in the simulation. The method is then applied to show how biases in non-linear systems can be estimated.

The introduced bias is represented by a matrix

$$A_q = \begin{pmatrix} 1 & -2 \\ 2 & 1 \\ 0 & 0.2 \end{pmatrix},$$

and two slowly varying biases q_1 and q_2 . This A_q means that there are two independent biases affecting the model states which varies between approximately 0 and 10 % of the state values. The default system has linear measurement equations where $y_1 = p_{im}$ and $y_2 = n_{trb}$. However, according to the discussion in Section 5.3, an augmentation as the one introduced in this example can not be estimated without a direct connection between p_{em} and y . Therefore the measurement equation is extended with an extra sensor for p_{em} for the augmentation estimation. Note that this extra sensor is not used for feedback neither in the observer based on the default model nor in the observer based on the

augmented observer. This reflects the situation that a lab environment or development system may be equipped with extra sensors to achieve a better augmentation estimation.

The observer based on the default model is referred to as the default observer while the observer based on the augmented model is referred to as the augmented observer. Both observers only use the p_{im} and p_{em} measurements. To make the simulation more realistic, white system and measurement noise are added in the creation of the synthetic data.

Using the simulated data and the default model, the augmentation estimation results in

$$\Sigma \approx \begin{pmatrix} 5.0259 & 0 & 0 \\ 0 & 4.8669 & 0 \\ 0 & 0 & 0.0024 \end{pmatrix} 10^5,$$

and

$$U \approx \begin{pmatrix} -0.8295 & -0.5527 & 0.0800 \\ 0.5515 & -0.8233 & -0.0388 \\ 0.0881 & 0.0123 & 0.9960 \end{pmatrix},$$

where Σ indicates that there are two slowly varying biases present. Hence, \hat{A}_q is estimated using the first two columns of U .

At a first look \hat{A}_q does not appear similar to A_q . However, the crucial fact is that the columns of \hat{A}_q and A_q span, approximately, the same space. A closer look reveals that the elements in the bottom row is significantly smaller than the other elements, and that the factor between row one and two are approximately 2. That is, the only thing that differs between A_q and \hat{A}_q is a scaling.

The objective was not only to estimate the bias, but rather to obtain an observer that compensated for the model bias. Thus, an observer is created using EKF methodology for a model augmented according to the estimated \hat{A}_q . The performance is compared to the default observer. The state estimates are presented in Figure 2 together with the true states. It is easily seen that the augmented observer estimates p_{im} and n_{trb} better than the default observer. To obtain a better view on observer performance, the estimation errors are plotted in Figure 3. Here it is clear that all three state estimates become better with the augmented observer than with the default observer.

The conclusion of this small simulation example is that the approach managed to get a good enough estimate of a bias in a non-linear model to improve the state estimates.

6.2 TWO EXPERIMENTAL EVALUATIONS

The experimental data described in Appendix A is used to evaluate the augmentation estimation and observer performance. The true states are approximated by non-causal, zero-phase, low-pass filtered measurements, where the filter has a cut off frequency of 2 Hz, see Figure 4. Note that parts of the turbine speed data is missing, which is due to the fact that the measuring range of the turbine speed sensor is limited, speeds below 20 000 rpm, or approximately 2 100 rad/s can not be measured.

Based on the measurement data, an augmentation is estimated using data from two stationary operating points in the European transient cycle (ETC) of about 1 000 samples

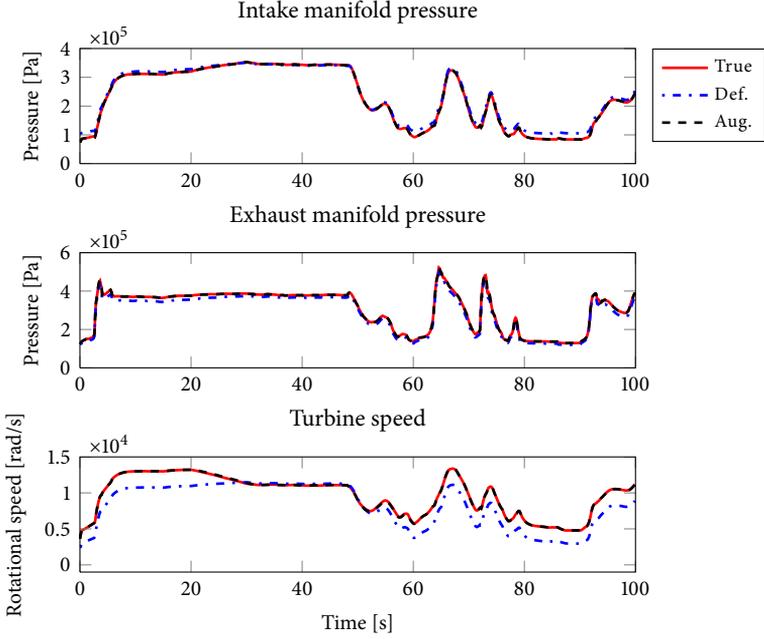


Figure 2: True states and estimated states using default and augmented observer in the simulation study.

each. All states are measured and the augmentation estimation results in

$$\Sigma \approx 10^5 \begin{pmatrix} 5.3230 & 0 & 0 \\ 0 & 0.3739 & 0 \\ 0 & 0 & 0.0044 \end{pmatrix},$$

and

$$U \approx \begin{pmatrix} -0.2610 & 0.9650 & -0.0249 \\ -0.9648 & -0.2671 & -0.0274 \\ -0.0329 & 0.0169 & 0.9993 \end{pmatrix}, \quad (12)$$

where Σ indicates that there is one dominant slowly varying bias present. Hence, \hat{A}_q is selected to be only the first column of U .

REDUCED AUGMENTATION ORDER

In this system it is possible to augment the system with three extra states and still have an observable system if all states are measured. One interesting question is if the proposed method that estimates a lower dimension augmentation can still capture most of the bias. Therefore, three observers are designed: the default observer, a fully augmented observer, and a one dimensional augmentation observer.

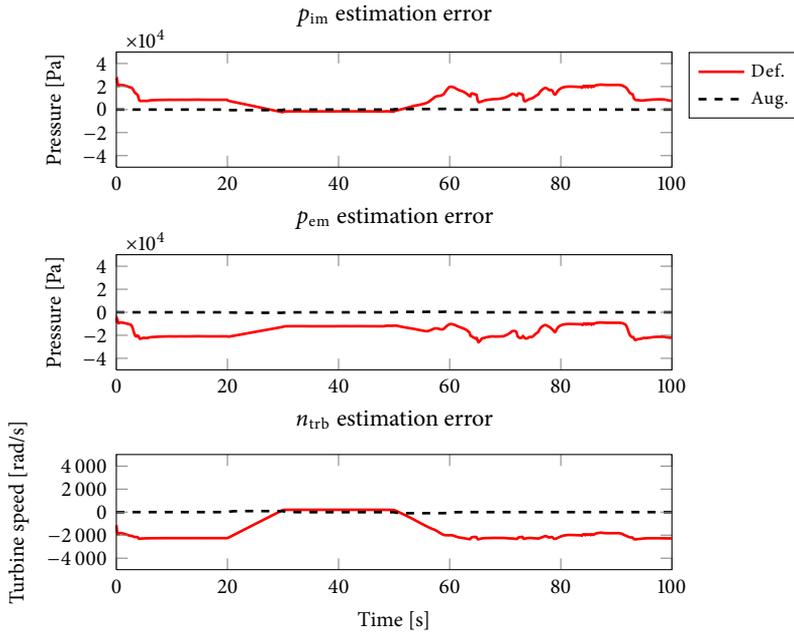


Figure 3: Estimation errors using default and augmented observer in the simulation study.

The aim of this comparison is thus to conclude whether the proposed method works, and is performed by analyzing the estimation errors from the three observers. The resulting Probability Density Functions (PDF) of the estimation errors are shown in Figure 5 and mean and maximum absolute errors for the entire ETC are presented in Table 1. From the data it is clear that the default observer has a bias and that the augmented observers reduce the bias. Now comparing the two augmented observers it is seen that the observer with only a one dimensional augmentation delivers close to the same reduction in bias as the fully augmented observer. This is a clear illustration that the method succeeds in finding the dominant bias in the model.

BENEFITS OF ADDITIONAL SENSOR DURING DESIGN

Another interesting question is what can be achieved by allowing extra sensors, compared to what is used in the final observer, while estimating an augmentation. The application chosen is to estimate the exhaust manifold pressure with reduced bias compared to a default observer without having a sensor measuring it. That is, design an observer for p_{em} using feedback from p_{im} and n_{trb} . The analysis is performed by comparing the estimates from two observers; one based on the augmentation

$$\hat{A}_q = (-0.2610 \quad -0.9648 \quad -0.0329)^T$$

Table 1: Data from observers None – default observer, H^\dagger – fully augmented observer, and \hat{A}_q – observer using reduced dimension augmentation found using augmentation estimation approach 3. All observers use feedback from all states.

States	Max abs. error			Mean error		
	None	H^\dagger	\hat{A}_q	None	H^\dagger	\hat{A}_q
p_{im} [Pa]	5459	6840	6599	-985	11	37
p_{em} [Pa]	14411	14277	14278	443	86	132
n_{trb} [rad/s]	0.8	0.7	0.6	0.005	-0.003	-0.007

Table 2: Static data from observers None – default observer, \bar{A}_q – observer using augmentation estimated using only measurements of p_{im} , and n_{trb} , and \hat{A}_q – observer using augmentation estimated using measurements of all states. All observers use feedback from p_{im} , and p_{em} only.

States	Max abs. error			Mean error		
	None	\bar{A}_q	\hat{A}_q	None	\bar{A}_q	\hat{A}_q
p_{im} [Pa]	4191	3650	3641	-622	-80	-176
p_{em} [Pa]	58758	58197	51322	6810	6328	-678
n_{trb} [rad/s]	0.1	0.1	0.1	0.02	0.004	0.006

Table 3: Dynamic data from observers None – default observer, \bar{A}_q – observer using augmentation estimated using only measurements of p_{im} , and n_{trb} , and \hat{A}_q – observer using augmentation estimated using measurements of all states. All observers use feedback from p_{im} , and p_{em} only.

States	Max abs. error			Mean error		
	None	\bar{A}_q	\hat{A}_q	None	\bar{A}_q	\hat{A}_q
p_{im} [Pa]	5748	6828	6316	-533	2	-34
p_{em} [Pa]	180279	177982	174486	16604	16479	8922
n_{trb} [rad/s]	0.9	0.6	0.5	0.02	0.0007	-0.001

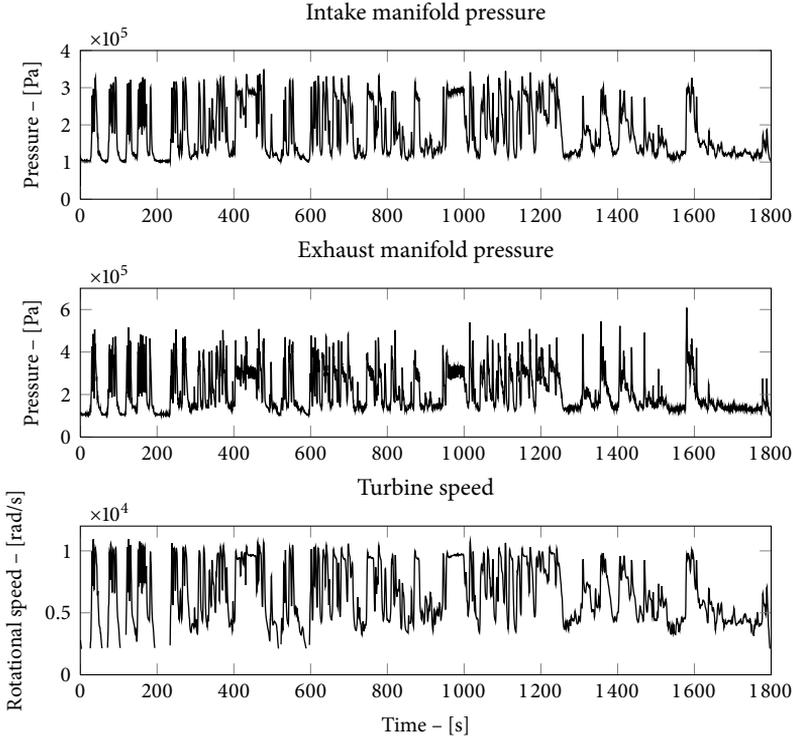


Figure 4: Measurements of p_{im} , p_{em} , and n_{trb} from the ETC used in the experimental evaluation. Note that turbine speeds below approximately 2100 rad/s are missing. This is due to the limited measurement range of the turbine speed sensor.

estimated using measurements of p_{im} , p_{em} , and n_{trb} , i.e. column one in (12), and another based on an augmentation

$$\bar{A}_q = (-0.9864 \quad 0 \quad -0.1644)^T$$

estimated using measurements of p_{im} , and n_{trb} only.

The two augmented observers are compared to the default observer and the results are shown in Figure 6 and Table 2 and 3. Figure 6 shows the probability density function for the estimation errors for the default observer, the observer based on the model augmented with \hat{A}_q , and the observer based on the model augmented with \bar{A}_q . It is seen that both augmented observers reduce the mean of the bias for p_{im} and n_{trb} compared to the default observer and that the observer based on the model augmented with \hat{A}_q significantly reduces also the bias in p_{em} . Table 2 and 3 show the mean and maximum absolute estimation errors for selected stationary parts of an ETC and for the entire ETC respectively. In both tables it is obvious that the observer based on a model augmented with \hat{A}_q significantly reduces the estimation bias. The mean error is

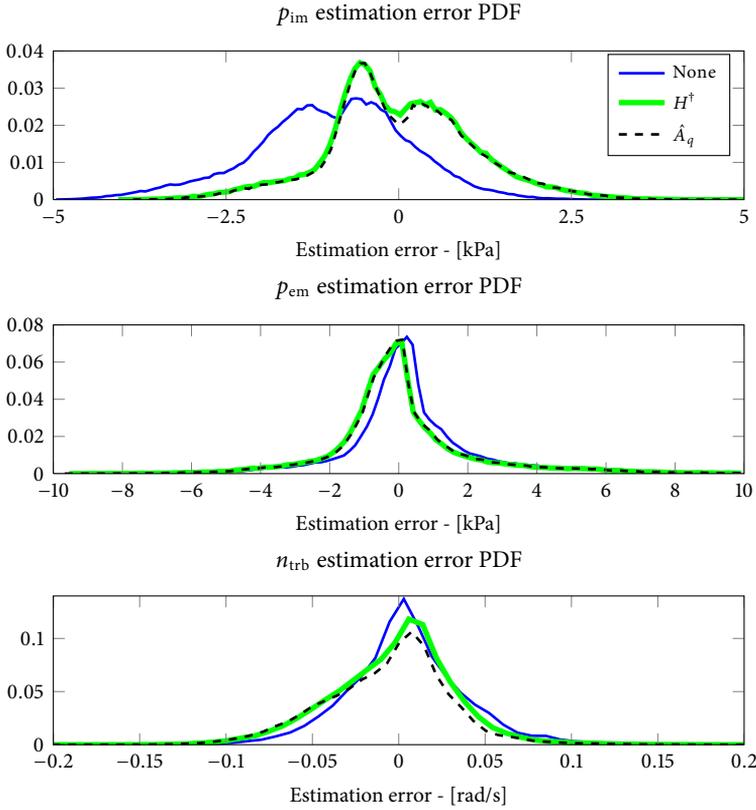


Figure 5: Probability density functions for three observers: None – default observer, H^\dagger – observer augmented with three states, and \hat{A}_q – observer augmented with one state and the estimated \hat{A}_q .

reduced by approximately 50 % during an entire ETC and by approximately 90 % for selected stationary parts, while the maximum absolute errors are almost unaffected. These, quite large, differences in the different measures are all explained by the fact that the suggested method reduces stationary bias and, as can be seen in Figure 4 the ETC is a rather dynamic sequence and the maximum absolute errors occur in transients.

7 CONCLUSIONS

A method for bias compensation in observers is developed. The idea is to, based on measurement data, compute a low dimension augmentation of the model that describes the most significant model biases. This augmented model is used to design an augmented observer that results in a state estimate with reduced bias. Three main results are a characterization of possible augmentations from observability perspectives, a parameter-

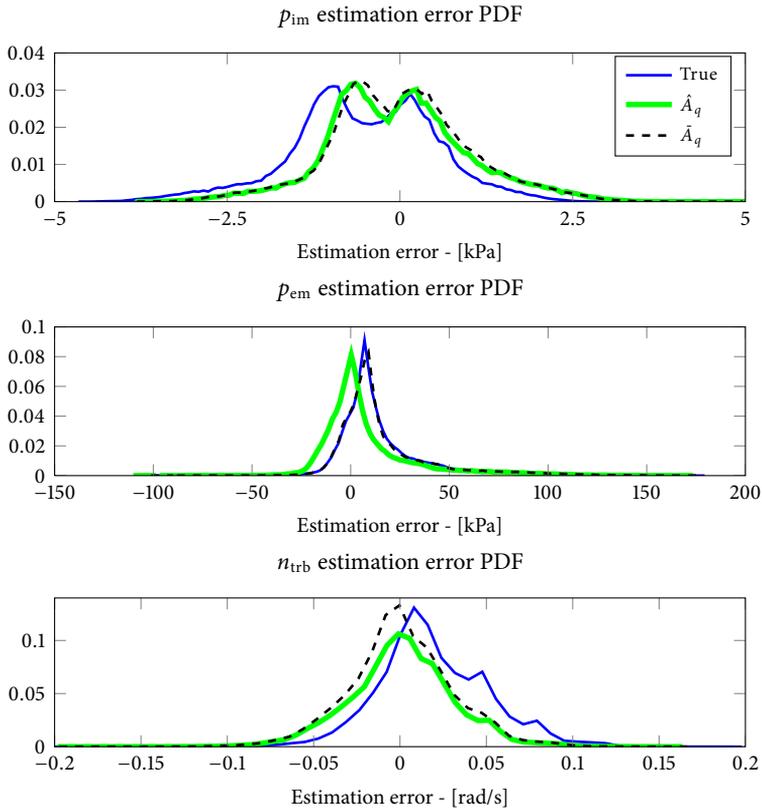


Figure 6: Probability density functions for default and augmented observers applied to real measurement data using feedback from p_{im} and n_{trb} . The two augmented observers are Red. – augmentation estimated measuring p_{im} and n_{trb} and Full – augmentation estimated measuring p_{im} , p_{em} , and n_{trb} respectively.

ization of the augmentations from the method, and a robustness analysis of the proposed augmentation estimation method.

The method is successfully applied to a diesel engine with VGT and EGR, using a non-linear default model and measurement data from an engine in a test cell. It is shown that an augmentation according to the suggested augmentation procedure reduces the mean estimation error, that is the bias, by approximately 50 % in an ETC.

REFERENCES

- Per Andersson and Lars Eriksson. Cylinder air charge estimator in turbocharged SI engines. In *Electronic Engine Controls*, number 2004-01-1366 in SAE Technical paper series SP-1822, 2004. doi:10.4271/2004-01-1366.
- Charlene Bembenek, Tom A. Jr. Chmielewski, and Paul R. Kalata. Observability conditions for biased linear time invariant systems. In *Proceedings of the American Control Conference*, pages 1180–1184, Philadelphia, Pennsylvania, June 1998. doi:10.1109/ACC.1998.703599.
- Sergio García-Nieto, Miguel Martínez, Xavier Blasco, and Javier Sanchis. Non-linear predictive control based on local model networks for air management in diesel engines. *Control Engineering Practice*, 16(12):1399–1413, December 2008. doi:10.1016/j.conengprac.2008.03.010.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. In *Electronic Engine Controls*, number 2008-01-0992 in SAE Technical paper series SP-2159, pages 137–145, SAE World Congress, Detroit, USA, 2008a. doi:10.4271/2008-01-0992.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation applied to an engine. IFAC World Congress, Seoul, Korea, 2008b. doi:10.3182/20080706-5-KR-1001.00347.
- Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc, Englewood Cliffs, New Jersey 07632, 1980.
- Thomas Kailath, Ali H. Sayed, and Babak Hassibi. *Linear Estimation*. Prentice-Hall, Inc, Upper Saddle River, New Jersey 07458, 2 edition, 2000.
- Rudolf Emil Kalman, Yu-Chi Ho, and Kumpati S. Narendra. Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1, 1963.
- Ernest Bruce Lee and Lawrence Markus. *Foundations of Optimal Control Theory*. John Wiley & Sons, Inc, New York, 1968.
- Paolo Lino, Bruno Maione, and Claudio Amorese. Modelling and predictive control of a new injection system for compressed natural gas engines. *Control Engineering Practice*, 16(10):1216–1230, October 2008. doi:10.1016/j.conengprac.2008.01.008.
- Tung-Ching Tseng and Wai K. Cheng. An adaptive air/fuel ratio controller for SI engine throttle transients. In *Electronic Engine Controls*, number 1999-01-0552 in SAE Technical paper series SP-1419, SAE International Congress and Exposition, Detroit, USA, 1999. doi:10.4271/1999-01-0552.
- Johan Wahlström and Lars Eriksson. Modeling of a diesel engine with VGT and EGR including oxygen mass fraction. Technical Report LiTH-R-2747, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, 2006.

Junmin Wang. Air fraction estimation for multiple combustion mode diesel engines with dual-loop EGR systems. *Control Engineering Practice*, 16(12):1479–1486, December 2008. doi:10.1016/j.conengprac.2008.04.007.

A ENGINE MODEL AND DATA

The model, on which the method is applied, is a third order non-linear state space model of a six cylinder Scania diesel engine with VGT and EGR. The model states are intake manifold pressure, p_{im} , and exhaust manifold pressure, p_{em} , and turbine speed, n_{trb} . The inputs are injected amount of fuel, engine speed, VGT and EGR positions. It is based on a model developed in Wahlström and Eriksson (2006) but slightly simplified. The simplifications are that the states for the EGR mass fraction and actuator dynamics are removed.

The data is collected in an engine test cell at Scania CV AB in Södertälje, Sweden. The data is from a six cylinder Scania diesel engine with VGT and EGR and was collected during an ETC. The sensor signals used are; intake and exhaust manifold pressures, turbine speed, and engine speed, and actuator signals used are; VGT and EGR positions, and injected amount of fuel. All these signals are available on a standard engine, i.e. no extra lab sensors were used, and collected with a sampling rate of 100 Hz.

B PROOFS OF THEOREMS 4.2 AND 5.1

Theorem 4.2 *Assume that (C, F) in (5) is an observable pair then the augmented system (5) is observable if and only if*

$$\text{Ker}((F - I)(A_q \ N_C)) = \{0\},$$

where the columns of N_C span $\text{Ker } C$.

Proof From Theorem 4.1 it follows that the augmented model (5) is observable if and only if $x = 0, q = 0$ is the only solution to

$$Cx = 0 \tag{13a}$$

$$(\lambda I - F)x + (F - I)A_q q = 0 \tag{13b}$$

$$(\lambda I - I)q = 0 \tag{13c}$$

for all $\lambda \in \mathbb{C}$. For $\lambda \neq 1$ it is immediate from (13c) that $q = 0$. Then the assumption that (C, F) is an observable pair together with (13a), (13b), and Theorem 4.1 gives that $x = 0$. Thus, only $\lambda = 1$ needs to be investigated further.

For $\lambda = 1$ in (13) the augmented model is observable if and only if $x = 0, q = 0$ is the only solution to

$$(F - I)(x - A_q q) = 0$$

$$Cx = 0.$$

Let the columns of N_C be a basis for $\text{Ker } C$, then $x = N_C \xi$ for some ξ and observability is equivalent to $q = 0$, $\xi = 0$ being the only solution to the equation

$$(F - I)(N_C \xi - A_q q) = 0.$$

This is equivalent to that the matrix

$$(F - I) \begin{pmatrix} A_q & N_C \end{pmatrix}$$

has full column rank which ends the proof. \square

Theorem 5.1 *Assume that the observer gain, K , is chosen such that the observer is strictly stable and does not have any poles in the origin. Then, in absence of noise, the bias samples are spanned by the rows of C and can thereby be written as*

$$\beta_t = C^T \Gamma.$$

Proof Since Approach 1 only is applicable if C has full column rank and due to the augmentation, C^\dagger , used in Approach 3 the theorem automatically holds for these cases. It is therefore sufficient to prove the result for Approach 2.

Now, starting with the output error and rewriting it

$$\begin{aligned} r_t &= C \overbrace{(I - F + KCF)}^W^{-1} (I - KC) (I - F) \beta_t \\ &= CW^{-1} (W - KC) A_q q = (I - CW^{-1}K) C \beta_t, \end{aligned} \quad (14)$$

where the assumption that K is chosen such that the observer system, $(I - F + KCF)$, is strictly stable and does not have any eigenvalues equal to zero which assures that W^{-1} exists, is used. Then, using the pseudo inverse, (14) can be written as

$$C \beta_t = (I - CW^{-1}K)^\dagger r_t = \bar{r}_t. \quad (15)$$

A unique solution to (15) is received by computing the minimum square solution with least Euclidean norm. Writing

$$\beta_t = \beta_t^o + \beta_t^\perp, \quad (16)$$

where

$$\beta_t^o \in (\text{Ker } C)^\perp = \text{span}\{C^T\} \quad (17)$$

and

$$\beta_t^\perp \in \text{Ker } C, \quad (18)$$

the solution with least Euclidean norm is the solution with $\beta_t^\perp = 0$, i.e.

$$\beta_t = \beta_t^o = C^T \Gamma \quad (19)$$

which concludes the proof. \square

EKF-Based Adaptation of Look-Up Tables with
an Air Mass-Flow Sensor Application[☆]

B

[☆]Published in *Control Engineering Practice*, 19(5):442–453, 2011.

EKF-Based Adaptation of Look-Up Tables with an Air Mass-Flow Sensor Application

Erik Höckerdal, Erik Frisk, and Lars Eriksson

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, S-581 83 Linköping,
Sweden.*

ABSTRACT

A method for bias compensation and online map adaptation using extended Kalman filters is developed. Key properties of the approach include the methods of handling component aging, varying measurement quality including operating-point-dependent reliability and occasional outliers, and operating-point-dependent model quality. Theoretical results about local and global observability, specifically adapted to the map adaptation problem, are proven. In addition, a method is presented to handle covariance growth of locally unobservable modes, which is inherent in the map adaptation problem. The approach is also applicable to the offline calibration of maps, in which case the only requirement of the data is that the entire operating region of the system is covered, i.e., no special calibration cycles are required. The approach is applied to a truck engine in which an air mass-flow sensor adaptation map is estimated during a European transient cycle. It is demonstrated that the method manages to find a map describing the sensor error in the presence of model errors on a measurement sequence not specifically designed for adaptation. It is also demonstrated that the method integrates well with traditional engineering tools, allowing prior knowledge about specific model errors to be incorporated and handled.

1 INTRODUCTION

Modern control and diagnosis systems rely upon accurate information about the system state. Determining the full state with direct measurements is not always possible as sensors can be unavailable, impossible to install, or too expensive for the intended application. Estimation schemes are therefore used, as they can often provide a cost-efficient means of obtaining the state (Colin et al., 2009). Such estimation schemes are often based on a model of the process, but a fundamental disadvantage is that models have errors. The causes of these model errors can be quite varied; the errors can vary with system's operating point (Zimmerschied and Isermann, 2010), changes in ambient conditions, the age of components (Rupp and Guzzella, 2010), etc., all of which affect the error dynamics. The variation in the system's operating point has rapid dynamics, approximately on the same order as the system itself during normal operation, while the second and third causes above can present time constants ranging from hours up to years, depending on the system and its location.

A common situation is that while the dynamics is well-captured by the model, stationary errors exist. This situation will be referred to as a biased model and has been observed in truck engines, for example (Höckerdal et al., 2008). There exist estimation algorithms for operating-point-dependent biases, see for example Höckerdal et al. (2009) and Lin and Bar-Shalom (2006), but they can not separately handle both the short-term variations induced by normal changes in operating point and the long-term drift caused by component aging or changes in the environment. Instead, the engineer has to choose which of these effects is most important and tune the observer accordingly.

A common technique to handle operating point dependencies in industrial applications is to introduce maps or look-up tables, see e.g., Guzzella and Amstutz (1998); Peyton Jones and Muske (2009) for automotive applications and Horton (1997); Larsson et al. (2009) for aerodynamic applications. These look-up tables provide information about operating point dependencies for certain operating conditions, e.g., default ambient conditions and the current state of system degradation, which changes over time.

For automotive applications in particular, look-up tables or maps are frequently used to describe relations when physical models are unavailable, e.g., sensor and actuator characteristics, cooler efficiency, injector characteristics, and aftertreatment systems. These are examples in which maps benefit from continuous online adaptation to prevent undesired system behavior. Routines for online map adaptation have been considered in Wu (2006); Peyton Jones and Muske (2009), and a primary contribution in the present work is a systematic way to handle aging and other slowly-varying uncertainties. Simultaneous bias compensation and online map adaptation are key properties of the proposed method, which has industrial value because the method integrates well with existing map-based solutions. Another interesting application is the offline calibration of the engine control system with its variety of maps and parameters. In this process, special system trajectories are often used when tuning the parameters. With the proposed method, special system trajectories are not needed, and the only requirement on the trajectory is that it spans the operating region of the system.

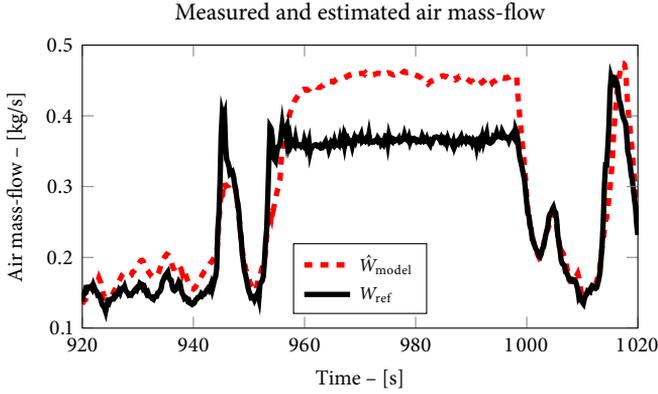


Figure 1: Typical example of model output from a biased model (Höckerdal et al., 2008), where W_{ref} is the air mass-flow measured by an additional, highly accurate, reference sensor available in the engine test cell.

2 METHOD OUTLINE

A systematic approach to designing an observer that adapts parameterized functions (maps) online while reducing stationary errors in a model is developed. The starting point is a default model, in discrete-time state-space form,

$$\begin{aligned} x_{t+1} &= f_{\text{def}}(x_t, u_t) \\ y_t &= h(x_t), \end{aligned} \quad (1)$$

where $x \in \mathbb{R}^{n_x}$ are the states, $u \in \mathbb{R}^{n_u}$ the inputs, and $y \in \mathbb{R}^{n_y}$ the outputs, that suffers from stationary errors, see Figure 1 for an example. These model errors can exhibit both fast and slow dynamics, arising from, for example, operating-point-dependent bias and aging.

The goal is to handle these errors in an integrated way, and this is achieved by extending (1) with a parameterized function,

$$q_{\text{fcn}}(x_t, u_t, \theta_t), \quad (2)$$

and augmenting the state vector with the parameter vector, $\theta \in \mathbb{R}^{n_\theta}$, in the following way

$$\begin{aligned} \theta_{t+1} &= \theta_t \\ x_{t+1} &= f(x_t, u_t, q_{\text{fcn}}(x_t, u_t, \theta_t)) \\ y_t &= h(x_t). \end{aligned} \quad (3)$$

The idea with a construction like this is to capture the operating point dependence of the bias by the parametrization (2), and use the parameters, θ , introduced as new states, to track the aging. With (3) as a basis, any non-linear observer design technique can be applied, to design an observer that estimates the augmented state vector, (x, θ) .

The following example of a parameterized function q_{fcn} , from an automotive application (Höckerdal et al., 2008), serves as motivation for the work and will be used later in the evaluation.

Example 1 In a heavy duty diesel engine with exhaust gas recirculation (EGR) and a variable geometry turbocharger (VGT), the air mass-flow through the compressor is vital information for safe and clean engine control. Therefore, the engine is typically equipped with a mass-flow sensor. However, the sensor signal is subject to an operating-point-dependent error, due to the sensor installation and local flow fields, and this measurement error has to be compensated for before use by the engine control unit (ECU). The relationship between the measured air mass-flow W_{meas} and the true air mass-flow W can be stated as

$$W_{\text{meas}} = (1 + q_{\text{fcn}}(W, \theta)) W$$

where the function q_{fcn} is presented graphically in Figure 2. W^i are the grid points, and θ^i are the corresponding correction parameters. Note that there exist several ways of computing the true air mass-flow and three of them are; i) to compute the mass-flow into the engine using intake air density and engine speed, which during stationary operation is equal to the measured air mass-flow, ii) to assume that the air mass-flow sensor give accurate measurements, or, iii) to install a reference sensor. Which approach is chosen depends on the application and is determined by which is most reliable among those that are available. Here, where a reference sensor is not available, the modeled air mass-flow into the engine is considered to be the most accurate (Höckerdal et al., 2008) and is used as a reference air mass-flow. In this case, the sensor model is represented by a 1-D linear interpolation map, where W lies between the grid points $W^i \leq W \leq W^{i+1}$ and the interpolation is defined by

$$q_{\text{fcn}}(W, \theta) = \theta^i + \frac{(\theta^{i+1} - \theta^i)}{W^{i+1} - W^i} \cdot (W - W^i). \quad (4)$$

Due to aging, it is necessary to adapt this map, see Höckerdal et al. (2008) for a longer discussion. \diamond

Note that with construction (3), using parameters to track aging, the observer stores information about the operating point dependent errors, i.e., the parameters act as a memory.

The development of a model like (3) for estimation and identification entails that some new issues have to be addressed. The main concern in this paper is how to update the function parameters θ_t in a controlled manner. Sections 3 and 4 treat this issue with respect to observability and observer tuning, respectively. Here, the parameterization of q_{fcn} is given, and the interested reader is referred to, for example Haber and Unbehauen (1990) or Lind and Ljung (2008) for a discussion on how to find a structure and suitable regressors.

The system (3) is in standard state-space form, which means that any suitable observer design can be applied, and a stochastic filter is chosen in this work. Example approaches

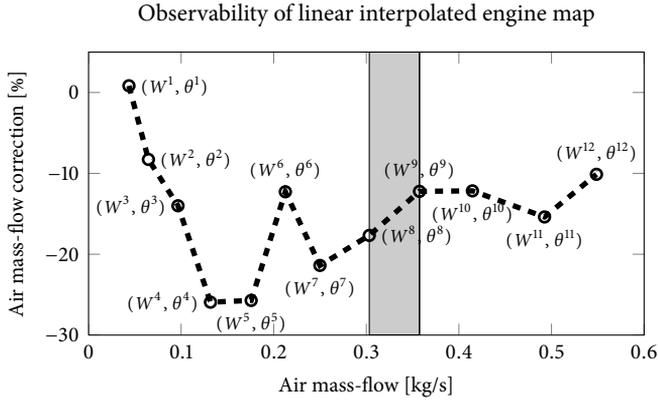


Figure 2: Air mass-flow sensor error map with the grid points denoted by the pair (W^i, θ^i) corresponding to a correction factor θ^i at a mass-flow of W^i .

to estimating states while at the same time handling unknown parameters are to apply a joint parameter-and state-estimating extended Kalman filter (EKF) (Kopp and Orford, 1963), or an unscented Kalman filter (UKF) (Wan et al., 2000). An advantage of stochastic filters compared to deterministic observers is that an estimate of the error statistics is computed in addition to the state estimate. The estimation error statistics are used in the computation of the filter feedback gain, which gives the stochastic filters natural tuning parameters that allow the filter to be tailored to handle system aging, unknown state initialization, a time-dependent model and measurement quality, outlier rejection etc.; see Section 4 for the discussion. Either the EKF or the UKF can be used, and since the EKF is a widely used standard method, which proved to be sufficient for achieving the results, the joint state and parameter estimating EKF is used.

3 OBSERVABILITY

In estimation, the observability or detectability of the system at hand is central to ensure correct and consistent estimates. The observability criterion states whether it is at all possible to reconstruct the model states from measurements, and this section is devoted to the observability of (3), where q_{fcn} represents linear interpolation. The observability analysis is conducted on the continuous-time system,

$$\begin{aligned}
 \dot{\theta} &= 0 \\
 \dot{x} &= f^c(x, u, q_{\text{fcn}}(x, \theta)) \\
 y &= h(x),
 \end{aligned} \tag{5}$$

corresponding to (3). The results are valid also for the discrete-time system as long as the sampling time is small enough (Kalman et al., 1963). The analysis is based on the definitions of observability in Hermann and Krener (1977), and Besançon (2007).

Intuitively, a system of linear or higher order interpolations with parameter states as the interpolation grid points is not locally observable. This insight can be understood by considering the air mass-flow adaptation map in Example 1. First, considering the total system and assuming more than one interpolation interval, this point can be understood by the intrinsic nature of the interpolation. Because not all parameters are involved in the interpolation computation at each operating point, some parameters are always locally unobservable. However, even if only one interpolation interval is studied and the non-interacting parameters are disregarded, e.g., consider only the shaded region of Figure 2, the ordinary rank condition, stating sufficient conditions for observability, is insufficient to assess the local observability of the reduced system. This is illustrated in Example 2.

Example 2 Consider a model that contains linear interpolation, with a single interpolation interval, in the measurement equation, defined by the grid points $(0, \theta_0)$ and $(1, \theta_1)$,

$$\dot{\bar{x}} = \begin{pmatrix} \dot{x} \\ \dot{\theta}_0 \\ \dot{\theta}_1 \end{pmatrix} = \begin{pmatrix} ax \\ 0 \\ 0 \end{pmatrix} = \bar{f}^c(\bar{x}) \quad (6a)$$

$$y = h(\bar{x}) = cx + \theta_0 + x(\theta_1 - \theta_0). \quad (6b)$$

From (6b) it can be seen that, with sufficient level of excitation, the variables x , θ_0 and θ_1 can be determined from the measurements y . However, the observability rank condition for (6) becomes

$$\left. \frac{\partial}{\partial \bar{x}} \begin{pmatrix} h(\bar{x}) \\ L_{\bar{f}^c} h(\bar{x}) \\ \vdots \\ L_{\bar{f}^c}^p h(\bar{x}) \end{pmatrix} \right|_{\bar{x} = \begin{pmatrix} x \\ \theta_0 \\ \theta_1 \end{pmatrix}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a^p \end{pmatrix} \underbrace{\begin{pmatrix} c + \theta_1 - \theta_0 & 1 - x & x \\ c + \theta_1 - \theta_0 & -x & x \\ \vdots & \vdots & \vdots \\ c + \theta_1 - \theta_0 & -x & x \end{pmatrix}}_{\text{rank}=2}$$

which clearly is rank deficient. \diamond

Hence, direct use of the rank condition for assessing the local and global observability of (5) is not possible. To ensure global observability of (5), a trajectory that spans the interpolation grid, and fulfills the conditions in Theorem 3.1, has to be considered.

The analysis is performed by first establishing the local observability of (5) at an operating point within a single interpolation interval, i.e., where the interpolation variable, y_2 , is in, $I_i =]y_2^i, y_2^{i+1}[$. By disregarding the non-interacting parameters for the case in which a 1-D linear interpolation is used and the interpolation variable $h_2(x) = y_2$ is measured, i.e.,

$$q_{\text{fcn}}(y_2, \theta^i, \theta^{i+1}) = \theta^i + \frac{(\theta^{i+1} - \theta^i)}{y_2^{i+1} - y_2^i} \cdot (y_2 - y_2^i), \quad (7)$$

model (5) becomes

$$\begin{aligned}
 \dot{\theta}^i &= 0 \\
 \dot{\theta}^{i+1} &= 0 \\
 \dot{x} &= f^c(x, u, q_{\text{fcn}}(y_2, \theta^i, \theta^{i+1})) \\
 y &= \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix} = h(x).
 \end{aligned} \tag{8}$$

Then global observability is established by considering a trajectory that passes through all interpolation intervals. Even though the formal analysis only applies to this particular case, it gives insight into more general cases.

The global observability of (5) is assessed using the result below on the local observability of (8). To state the result, introduce the notation

$$A = \left. \frac{\partial f^c}{\partial x} \right|_{\bar{z}_0}, \quad B_q = \left. \frac{\partial f^c}{\partial q_{\text{fcn}}} \right|_{\bar{z}_0}, \quad C = \left. \frac{\partial h}{\partial x} \right|_{\bar{z}_0},$$

where

$$\bar{z}_0 = (x_0^T \quad u_0^T \quad q_{\text{fcn},0}^T)^T,$$

is a linearization point and denote the numerator polynomials of $(pI - A)^{-1}B_q$ by $b_q(p)$, where p is the differentiation operator.

Theorem 3.1 *Model (8) with q_{fcn} defined by (7) is locally observable in \bar{z}_0 on a smooth trajectory (y, u) in I_i , containing the stationary point \bar{z}_0 , if $\dot{y}_2 \neq 0$, $b_q(p)y_2 \neq 0$, the pair (C, A) is observable and the matrix*

$$\begin{pmatrix} -A & B_q \\ C & 0 \end{pmatrix}$$

has full column rank.

Proof See Appendix A. □

Remark 3.2 *Requiring $\dot{y}_2 \neq 0$ is natural since the trajectory must move in the interval I_i for the parameters to be visible. Requiring $b_q(p)y_2 \neq 0$ says that the trajectory must not be canceled by the zero dynamics, which is also natural but can be seen as technical and seldom limiting.*

Even though introducing q_{fcn} into the system model is rather general, the interpolation variable must still be measured. However, for a system of the form

$$\dot{\theta}^i = 0 \tag{9a}$$

$$\dot{\theta}^{i+1} = 0 \tag{9b}$$

$$\dot{x} = f^c(x, u) \tag{9c}$$

$$y_1 = h_1(x) \tag{9d}$$

$$y_2 = q_{\text{fcn}}(h_2(x), \theta^i, \theta^{i+1}) \tag{9e}$$

the sufficient conditions for observability can be stated without having a measured interpolation variable. The system considered in Example 1 falls into this model class. Let C_1 be $\partial h_1(x)/\partial x$ evaluated at x_0 . Then, a sufficient observability condition is given by the following result:

Theorem 3.3 *Model (9) where q_{fcn} is defined by (7) and is locally observable in \bar{z}_0 on a smooth trajectory (y, u) in I_i , containing the stationary point \bar{z}_0 , if $\frac{d}{dt}h_2(x) \neq 0$, and the pair (C_1, A) is observable.*

Proof See Appendix A. □

With the results regarding observability in a given interpolation interval in Theorems 3.1 and 3.3, it is straightforward to state global results. For Theorem 3.1, the global result is given in Corollary 3.4, and for Theorem 3.3, global results can be stated in a similar way.

Corollary 3.4 (Global observability) *If there exists a trajectory such that y_2 ranges over $\cup_i I_i$ and Theorem 3.1 is fulfilled in each interval, then (5) is globally observable with respect to that trajectory.*

Proof Given a trajectory fulfilling Corollary 3.4, Theorem 3.1 can be applied to each interval, I_i , which ensures the observability of $(x, \theta^i, \theta^{i+1})$ within each interval. As this is true for all intervals, x and $\theta = (\theta^1 \dots)$ are observable. □

Finally, observability does not depend on the choice of discretization method as long as the sampling time is chosen small enough (Kalman et al., 1963). This together with Theorem 3.1/3.3 and Corollary 3.4 for the continuous-time system (5) implies observability of the corresponding discrete-time system (3).

Even though the analytical results of this section only state what is required in the case of a 1-D map on which the output is computed using linear interpolation, it is reasonable to assume that similar results exist for the cases of non-measured but observable interpolation variables, bilinear interpolation or higher-order interpolation algorithms.

3.1 UNOBSERVABLE MODES AND COVARIANCE GROWTH

According to the previous paragraph, at any given time in systems like (3), there are generally some parameters θ^i that are locally unobservable. The estimation error covariance matrix of EKF/UKF observers based on systems with locally unobservable states grows linearly for the unobservable states if system noise is present in those modes. That is, in regions where the system seldom operates, the estimation error covariance matrix coefficients that correspond to the locally unobservable parameter states will grow linearly. The effects of this covariance growth is two-sided: i) it offers a way to rapidly update old parameters while protecting frequently updated parameters from spurious measurements, ii) it may cause numerical problems that affect the filter's stability over the lifetime of the system, i.e., covariance matrix divergence, which must be handled. The latter has previously been discussed in e.g. Gustafsson (1997) where the solution is to

stop the update of the covariance matrix elements if the system excitation is poor which also is the solution used here.

Example 3 A model that contains a parameterized function describing a linearly interpolated map in the measurement equation, similar to the system in Example 1, is used to illustrate the linear growth of the covariance estimate P in an EKF. An analysis of the equations describing the evolution of the feedback gain and estimation error covariance gives insight into the growth of the estimation error covariance

$$K_t = P_{p,t-1} C_t^T (C_t P_{p,t-1} C_t^T + R)^{-1} \quad (10a)$$

$$\begin{aligned} P_{u,t} &= P_{p,t-1} - P_{p,t-1} C_t^T (C_t P_{p,t-1} C_t^T + R)^{-1} C_t P_{p,t-1} \\ &= P_{p,t-1} - K_t C_t P_{p,t-1} \end{aligned} \quad (10b)$$

$$P_{p,t} = A_t P_{u,t} A_t^T + Q. \quad (10c)$$

Here K_t is the Kalman gain, $P_{u,t}$ and $P_{p,t}$ are the estimation error covariance estimates in the update and prediction steps of the EKF, respectively, and Q and R are the system and measurement noise covariances, respectively. The property of locally unobservable parameter states is demonstrated using diagonal P_0 , Q , R and a linearly interpolated 1-D map in the system measurement equation.

Consider the following partition of the state vector

$$\bar{x} = (x^T \quad \theta_o^T \quad \theta_u^T)^T,$$

where x is the state vector in the original model, θ_o are parameter states that have been observed, and θ_u are the parameter states that are not observed. Start in time step t with

$$\begin{aligned} P_{p,t-1} &= \begin{pmatrix} *^{n_x \times n_x} & *^{n_x \times n_{\theta_o}} & 0 \\ *^{n_{\theta_o} \times n_x} & *^{n_{\theta_o} \times n_{\theta_o}} & 0 \\ 0 & 0 & D_{t-1}^{n_{\theta_u} \times n_{\theta_u}} \end{pmatrix}, \\ C_t &= \begin{pmatrix} *^{n_y \times n_x} & *^{n_y \times n_{\theta_o}} & 0 \end{pmatrix} \\ A_t &= \begin{pmatrix} *^{n_x \times n_x} & 0 \\ 0 & I^{n_{\theta} \times n_{\theta}} \end{pmatrix}, \end{aligned}$$

where D_{t-1} on the diagonal of $P_{p,t-1}$ is a diagonal matrix corresponding to unobserved parameter states, and the zero in C_t corresponds to parameters not used in the interpolation in time step t , i.e., with

$$\begin{aligned} y &= h(x) + q_{\text{fcn}}(x, \theta), \\ C_t &= \left(\frac{\partial h}{\partial x} + \frac{\partial q_{\text{fcn}}}{\partial x} \quad \frac{\partial q_{\text{fcn}}}{\partial \theta_o} \quad \frac{\partial q_{\text{fcn}}}{\partial \theta_u} \right) \Big|_{\bar{x}=\bar{x}_t} = \begin{pmatrix} * & * & 0 \end{pmatrix} \end{aligned}$$

which can be compared to the bright and shaded regions of Figure 2. Now (10a) becomes

$$K_t = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & D_{t-1} \end{pmatrix} \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \end{pmatrix},$$

that is, due to the zero in K_t , the unobservable states θ_u are not updated. Further, (10b) becomes

$$P_{u,t} = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & D_{t-1} \end{pmatrix} - \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & D_{t-1} \end{pmatrix},$$

i.e., the estimation error covariance for the unobserved parameter states is not reduced. Nevertheless, due to (10c), a linear increase in the estimation error covariance is obtained

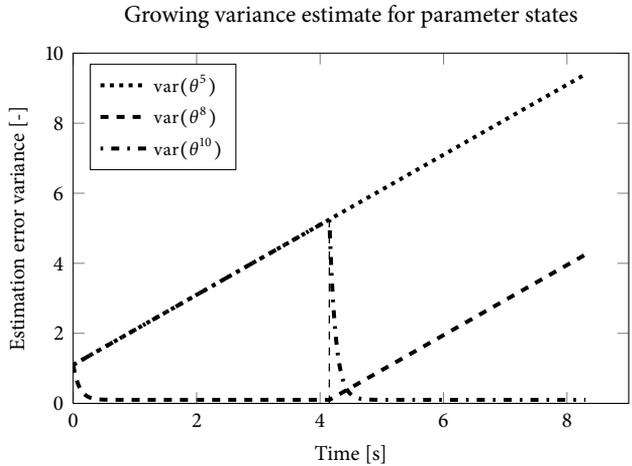
$$P_{p,t} = A_t P_{u,t} A_t^T + Q = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & D_{t-1} \end{pmatrix} + \begin{pmatrix} Q_x & 0 & 0 \\ 0 & Q_{\theta_o} & 0 \\ 0 & 0 & Q_{\theta_u} \end{pmatrix}.$$

This shows that $D_t = D_{t-1} + Q_{\theta_u}$, and it can be concluded that the covariance matrix coefficients corresponding to areas within the system operating region in which the system does not operate will increase linearly over time. \diamond

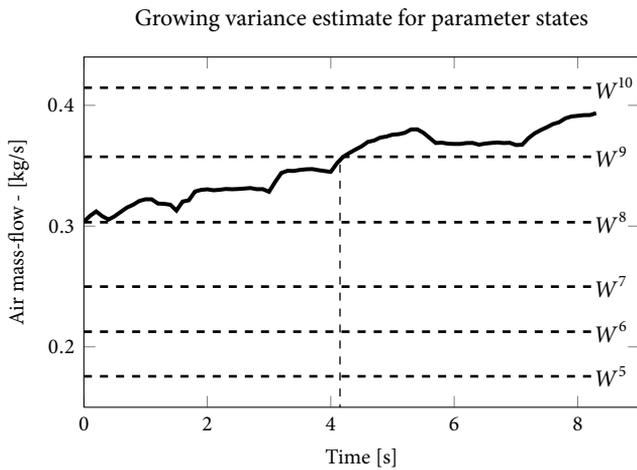
This effect is also illustrated in Figure 3, where the variance of three parameter states, θ^5 , θ^8 , and θ^{10} from Figure 2 is plotted versus time. In Figure 3a, parameter θ^5 corresponds to a parameter that is not observed at all during the studied trajectory, while the parameter θ^8 is observable during the first half of the trajectory and unobservable during the second half. This behavior is reversed for θ^{10} ; it is first unobservable and then observable.

Returning to the application of engine map adaptation, experiences indicate that adaptation algorithms, not using the EKF and joint parameter and state estimation, have problems concerning parameter aging and occasional spurious measurements. For example an engine, whose trajectories during normal operation, does not span the entire parameter space and only occasionally enters some areas, may suffer from undesired system behavior caused by old parameters corresponding to those seldom visited operating points. 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, 2009). 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.

After discussing the positive effects introduced by the locally unobservable parameter states it remains to discuss how to ensure stable observer operation even though some parameter states never may become observable, causing the corresponding covariance matrix elements to grow without bound. A direct and intuitive way of handling this linear growth of estimation error covariance of the locally unobservable parameter states is to introduce an upper limit for the corresponding estimation error covariance matrix elements. A possible such upper limit is the initial error covariance matrix, P_0 . Since it is desirable to limit the estimation error covariance of only the locally unobservable



(a) Parameter variance.



(b) Air mass-flow trajectory.

Figure 3: The figure shows the development of the estimation error variance for three parameter states. One that is unobservable during the entire trajectory – θ^5 , one that is observable for the first half of the trajectory – θ^8 , and one that is observable for the second half of the trajectory – θ^{10} .

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, Appendix E) that the off-diagonal elements in P do not affect the estimation error covariance for a single parameter, and with P_0 as an upper limit, the introduction of yet another tuning parameter is avoided. The effect of this is discussed in Section 5.3 and illustrated by simulation in Figure 10.

4 PARAMETER AND BIAS STATE CONVERGENCE RATES

This section briefly describes the difference between the proposed approach and existing (Bembenek et al., 1998; Höckerdal et al., 2009) approaches for reducing operating point dependent biases. Especially accentuated are the benefits of using stochastic observers for the locally unobservable models used in the new approach.

The developed approach differs from other approaches to reduce operating point dependent estimation errors such as the method presented in Höckerdal et al. (2009) in which the state vector is augmented with a bias state that is used directly to describe the model error,

$$\begin{aligned}x_{t+1} &= f(x_t, u_t, q_t) \\q_{t+1} &= q_t \\y_t &= h(x_t).\end{aligned}\tag{11}$$

A primary difference is that an observer based on (3) facilitates separate tracking of fast changes in operating point and slower changes due to aging. For an observer based on (11), this is not possible and it is necessary for the bias state to change approximately as fast as the system dynamics. Otherwise, the observer will not be able to track a change in the system's operating point. However, a rapidly changing bias state will also capture high-frequency disturbances, and will therefore not be able to withstand spurious measurements to the same extent as slower parameter states. This concern is one reason that makes the bias state in this method unsuitable for engine map adaptation.

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

Another issue, which is to some extent straightforwardly handled by stochastic filters, is initialization of the unknown bias or function parameters. By properly tuning the corresponding elements in the estimation error covariance matrix, P_0 , temporarily faster convergence of unknown bias or function parameter states can be achieved. That is, due to the initially faster convergence rate of unknown parameters, see Section 3.1, a rapid convergence of otherwise quite slow parameter states can be achieved in the same way that old parameters are allowed a faster convergence rate.

5 METHOD EVALUATION

Two studies are performed to evaluate the method, a simulation study and a study utilizing experimental data, in which the aim is to adapt the air mass-flow sensor in a diesel engine characterized by a 1-D adaptation map. The simulation study shows the convergence of the approach and the evolution of the adaptation map and includes an investigation of the effects of incorrect noise models when using the EKF as a parameter estimator (Ljung, 1979). The experimental study shows the results of applying the method to experimental data with the aim of analyzing the robustness and performance in the presence of model errors. Since the approach in some sense supplies a solution to the problem of biased default models solved in Höckerdal et al. (2009), the performance of that approach is presented as a comparison.

In both studies, the non-linear model of a heavy-duty diesel 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, all of which are present in the model output along with the air mass-flow through the compressor. The data used are collected during a European transient cycle (ETC) (Council of European Parliament, 2005).

5.1 OBSERVERS

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

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ y_t &= (x_{n_{\text{trb}},t} \quad h_W(x_t))^T, \end{aligned} \quad (12)$$

referred to as Def. An EKF with an extra bias state introduced in the measurement equation to describe the air mass-flow measurement error from the method developed in Höckerdal et al. (2009),

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ q_{t+1} &= q_t \\ y_t &= (x_{n_{\text{trb}},t} \quad h_W(x_t) + q_t)^T, \end{aligned} \quad (13)$$

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

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ \theta_{t+1} &= \theta_t \\ y_t &= (x_{n_{\text{trb}},t} \quad (1 + q_{\text{fcn}}(h_W(x_t), \theta_t))h_W(x_t))^T, \end{aligned} \quad (14)$$

referred to as Map. in which the parameterization of q_{fcn} is presented in Figure 2. The primary difference between the three observers are schematically presented in Figure 4, where the feedback part of each observer is highlighted.

The models (12)–(14) on which the three observers are based are observable, which is established using the ordinary observability rank condition for models (12) and (13), and Theorem 3.3 together with Corollary 3.4, for model (14).

In the simulation study, all observers only use feedback from the air mass-flow and turbine speed sensors, whilst in the experimental evaluation, feedback from all sensors except the exhaust manifold pressure sensor is used. Even though the Def. and Aug. models are observable from any of the outputs, and the trajectory is such that the interpolation grid is spanned for the Map. observer, the model errors are such that augmented feedback is needed in the experimental evaluation to improve the estimation performance. The performance is evaluated with respect to all states and outputs, i.e., intake and exhaust manifold pressures, turbine speed, and air mass-flow through the compressor.

5.2 OBSERVER TUNING

As in all observer applications, the tuning process is a delicate task. The main objective is to balance the rate of convergence and the smoothing effects, i.e., the stability and measurement noise sensitivity, of the observer.

In the case of an EKF, the tuning is performed in terms of the system, measurement, and initialization uncertainties. This is done by introducing noise terms to the system description

$$\begin{aligned}x_{t+1} &= f(x_t, u_t) + w_t \\ y_t &= h(x_t) + v_t,\end{aligned}$$

where w_t and v_t are independent white noise processes with covariance matrices

$$Q = T_Q Q_{\text{tune}} T_Q^T, \text{ and } R = T_R R_{\text{tune}} T_R^T,$$

where T_Q and T_R denote scaling matrices, and the initial state x_0 is assumed to be a random vector with covariance matrix P_0 (Kailath et al., 2000).

With this definition, all observers in both studies are tuned with equal respect to all parameters except the variance of the bias and parameter driving noises, which is done in accordance with Section 4. That is, all three observers have the same measurement noise R , and the same system noise corresponding to the default states x , i.e., the $n_x \times n_x$ upper left block of Q . The difference in bias modeling between Map. and Aug. produces an $n_q \times n_q$ lower right block in Q for Aug. that is a thousand times larger than the $n_\theta \times n_\theta$ lower right block of Q in Map. This difference is due to the different dynamics of the bias and parameter states. With this tuning, both Aug. and Map. are able to track the bias with approximately the same performance, while Map. also estimates a sensor error map.

5.3 STUDY 1: SIMULATION

The simulation study serves three purposes: i) to illustrate the convergence of the approach, ii) to study the adaptation map evolution, and iii) to evaluate the effects of incorrect noise assumptions (Ljung, 1979).

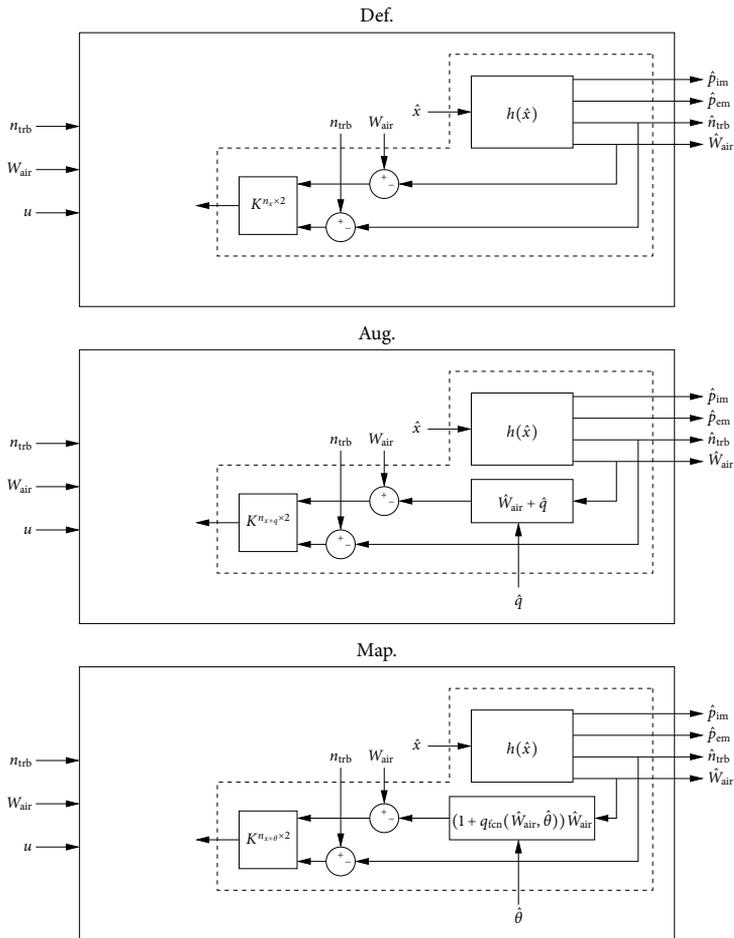


Figure 4: Schematics that highlight the feedback part of the three studied observers, Def., Aug., and Map. The main difference between the three configurations are shown in the block on the \hat{W}_{air} feedback signal, below the $h(\hat{x})$ block. To the left the inputs and feedback signals are shown, and to the right the four evaluation outputs are shown. Note that the dimension of the output signal of the Kalman gain block have different dimensions for the different observers.

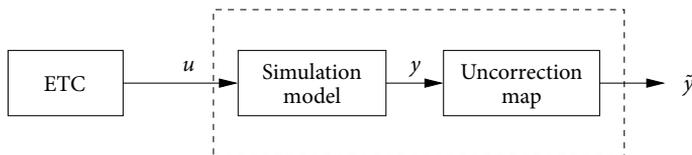


Figure 5: Simulation set-up with sensor error from Höckerdal et al. (2008).

All three utilize the simulation set-up presented in Figure 5 to create the data. The data are created by simulating the model with input data from the European transient cycle (ETC) segment, presented in Figure 6 and introducing artificial system and measurement noises with covariances of Q and R , respectively. The segment is chosen to contain a wide range of air mass-flows such that a trajectory for which the system is globally observable is created. To simulate incorrect air mass-flow measurement a 1-D sensor error map is used in the simulation and adjusts the air mass-flow according to

$$W_{\text{meas}} = (1 + q_{\text{fcn}}(W_{\text{true}}, \theta_{\text{true}})) W_{\text{true}},$$

where $q_{\text{fcn}}(W_{\text{true}}, \theta_{\text{true}})$ is the engine map presented in Figure 2. The distorted air mass-flow, W_{meas} , is then used for feedback to the observers.

CONVERGENCE RATE

One property of the estimation bias correction method developed in Höckerdal et al. (2009) is that, because no information about the bias at each operating point is saved, the observer convergence rate depends on the dynamics of the bias states. The extension presented here uses a parameterized function, and when the parameters are adapted, the convergence rate of the observer will not be dependent on the dynamics of the parameter states.

To analyze and compare the convergence rate of observers using parameterized bias to observers utilizing bias states is a difficult task because the convergence rate is highly dependent on the observer tuning. Since the maximum errors occur in transients, see the transient at the time 5 s in Figure 7 for an example, they give an indication on how well the estimator is able to track transient behavior. Figure 7 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, observers 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, because the bias state is allowed to change much faster than the parameter states, an Aug. observer will be more sensitive to disturbances, e.g., outliers, as mentioned in Section 4. An observer using a parameterized function with slow parameters does not allow disturbances to affect the estimation of the model or parameter states to the same extent as an observer with a bias state, i.e., the observer has a stronger smoothing effect.

Figure 8 shows the true and estimated map from Map. and the correction made by the slowly varying bias from Aug. computed according to

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

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

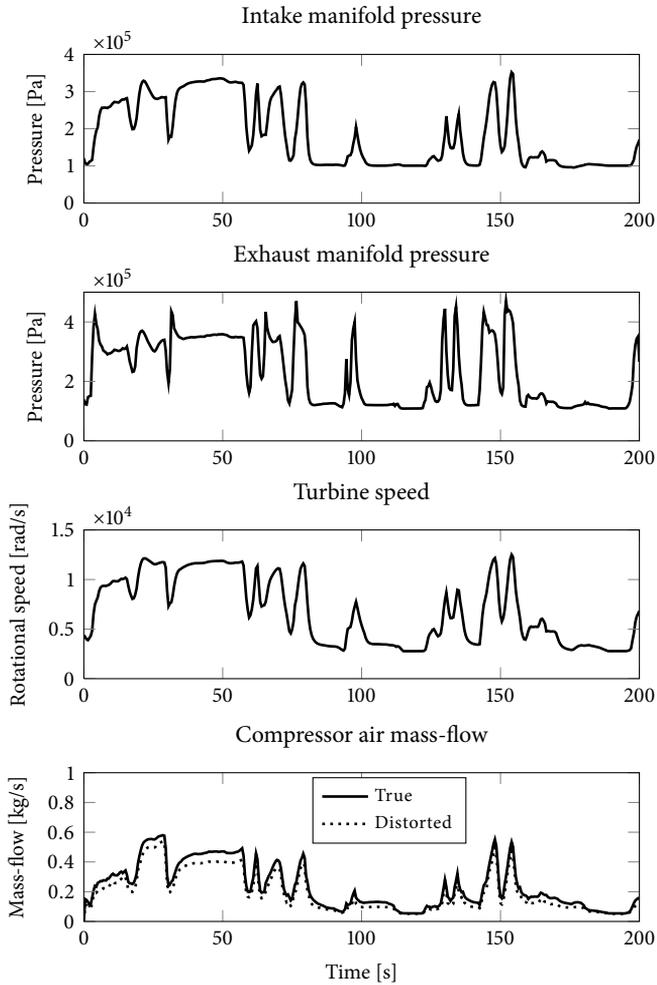


Figure 6: Simulated states and outputs from the ETC segment used in the experimental evaluation. The true states and outputs are plotted with solid lines and the distorted output used for feedback is plotted with a dotted line.

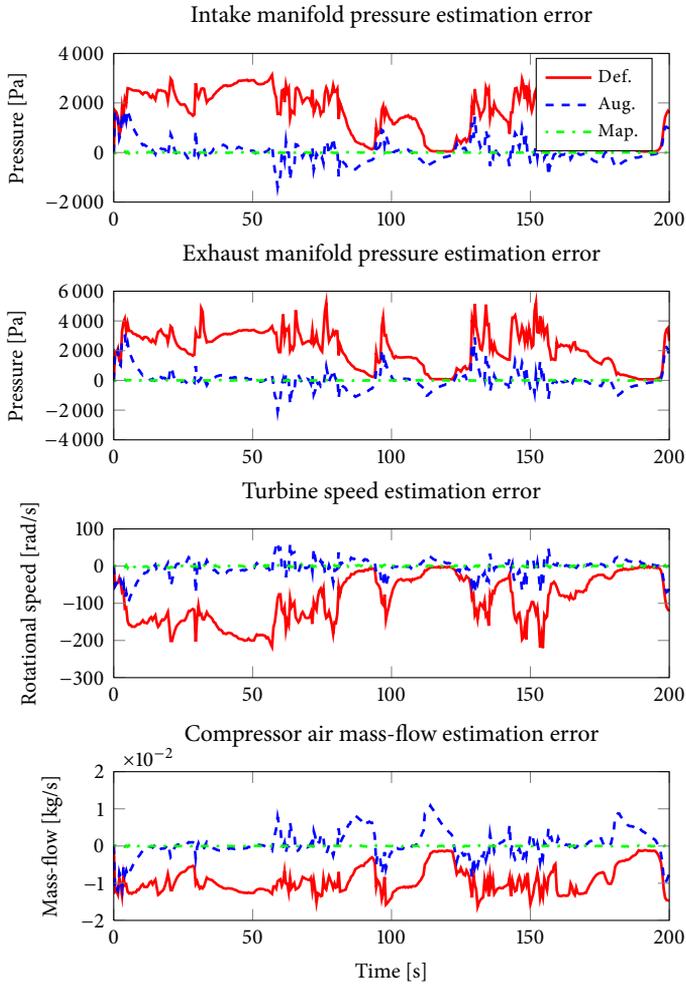


Figure 7: Estimation error of Aug., where the bias is too slow which gives large estimation errors in transients, together with estimation errors of Map. and Def.

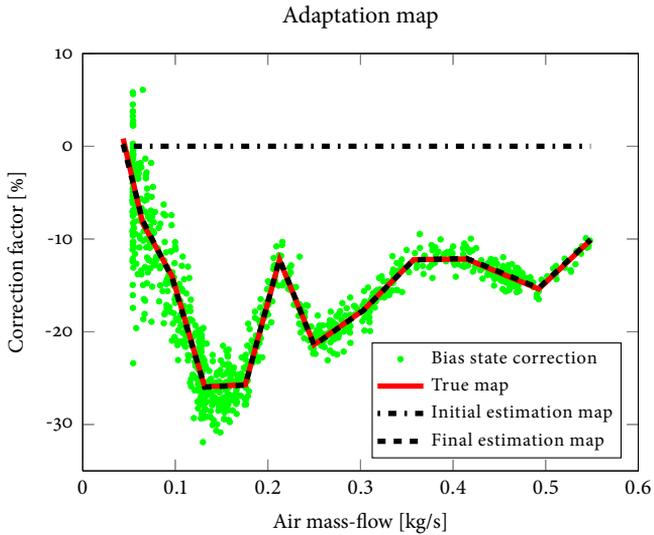


Figure 8: Mass-flow correction estimated by Aug. and Map.

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

ADAPTATION MAP EVOLUTION

In an application where the method is used for engine map adaptation it is important that the method converges. 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 9 shows the evolution of the adaptation map over time. The parameters are all initiated to zero, indicated by the straight line at time $t = 0$, and converges to the true map as the operating region is spanned.

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

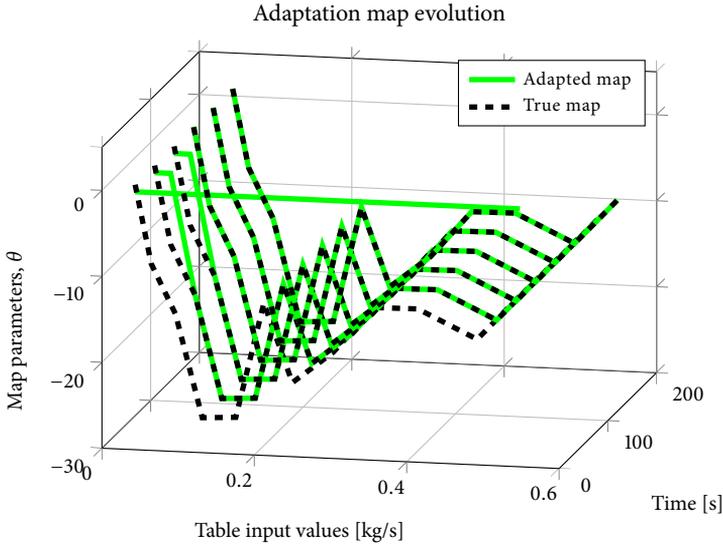


Figure 9: Adaptation map evolution showing fast convergence.

Figure 10 shows how the estimation error covariance is updated. In the figure the parameter noise covariance has been increased by a factor of 10^5 compared to simulation results presented in Figures 8 – 9, and Table 1 to clearly illustrate the growth in variance of the locally unobservable parameters. Notice that the estimation error covariance matrix element corresponding to a locally unobservable parameter grows linearly until it reaches the upper bound defined by P_0 , in agreement with the discussion in Section 3.1. However, for short time periods, like the ones simulated here, this linear growth of estimation error covariance is hardly a problem whilst it might become a problem when considering the entire lifetime of an engine.

NOISE MODEL SENSITIVITY

Ljung (1979) presents a theoretical analysis of the EKF as a parameter estimator and concludes that the parameter estimates will be biased if an incorrect noise description is used.

Inspired by this theoretical analysis and the fact that true nature of the noise seldom is completely known a simulation study is performed with the aim to analyze the influence of incorrect noise assumptions. The analysis is executed by applying the above designed observers to several sets of simulated data where the system and measurement noise variance and color had been adjusted. Neither of these simulations show of any problem of consistent convergence of the sensor error parameters.

Even though the analysis in the studied case does not indicate this to be an issue the potential problem cannot be discarded in general. Even if the main purpose of the parameters is, as in the engine applications addressed here, their function as a memory

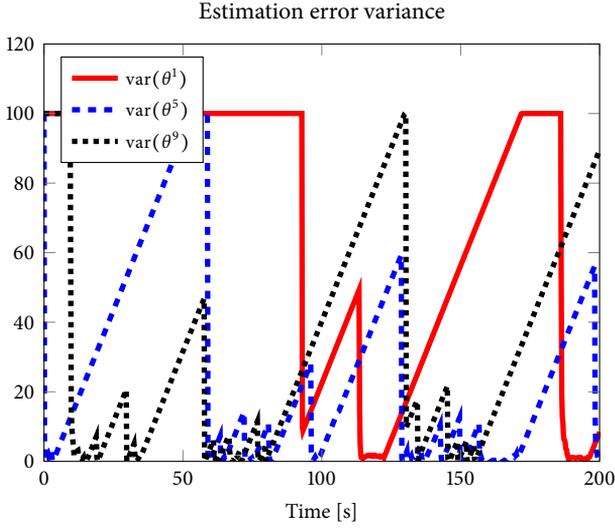


Figure 10: Evolution of the variance for three parameter states. Here the variance of the parameter noise is significantly increased compared to the observer which estimation error is presented in Table 1.

there exist applications where this is an issue. In such cases it is straightforward to use the in Ljung (1979) modified EKF algorithm for improved convergence properties.

5.4 STUDY 2: EXPERIMENTAL DATA

To assess the performance and robustness of the method an evaluation using measurements from an engine in an engine test cell is conducted. Since in this case there are other significant model errors present, besides the air mass-flow sensor error, an extra bias state is introduced to improve the state estimates. The additional bias state, complementing those for handling the measurement error, is introduced according to the ideas in Höckerdal et al. (2009). The purpose is to reduce the estimation errors due to a known error in the compressor model causing incorrect prediction of the compressor mass-flow, i.e. for Def.

$$\begin{aligned} x_{t+1} &= f(x_t, \Delta_{W,t}, u_t) \\ \Delta_{W,t+1} &= \Delta_{W,t} \\ y_t &= (x_{p_{im},t} \quad x_{n_{trb},t} \quad h_W(x_t, \Delta_{W,t}))^T. \end{aligned}$$

With the introduction of an extra state, compensating for the compressor mass-flow, 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 Figure 11 and Table 2.

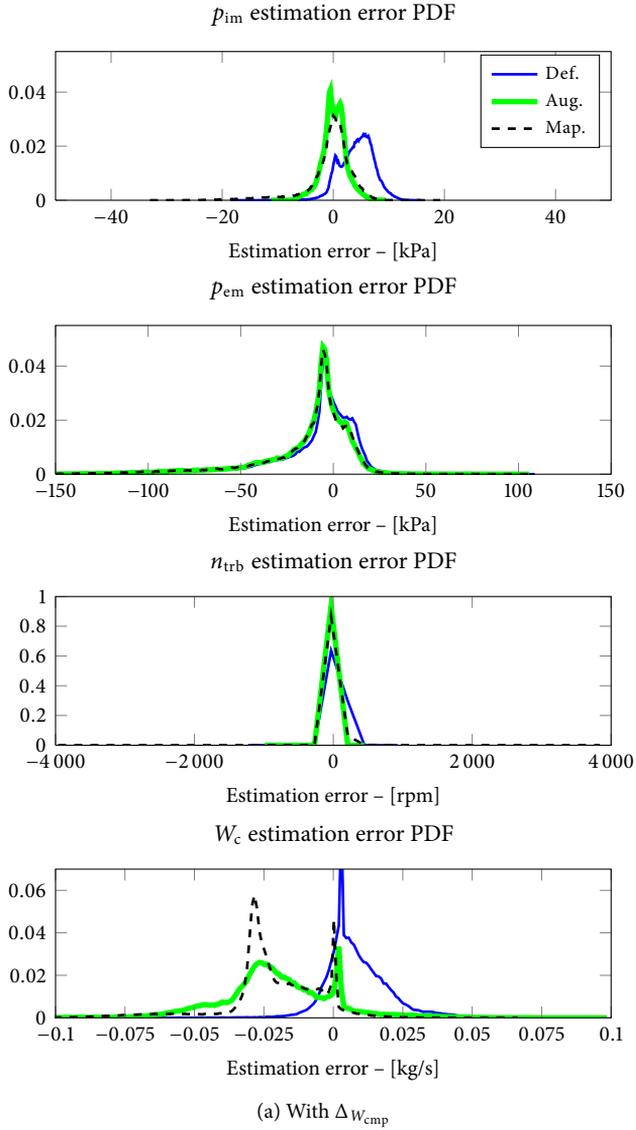


Figure 11: Estimation error PDF for estimators with and without ΔW_{cmp} . It is obvious that the estimate of p_{im} benefits from this extra state while p_{em} , n_{trb} , and W_{cmp} are moderately affected.

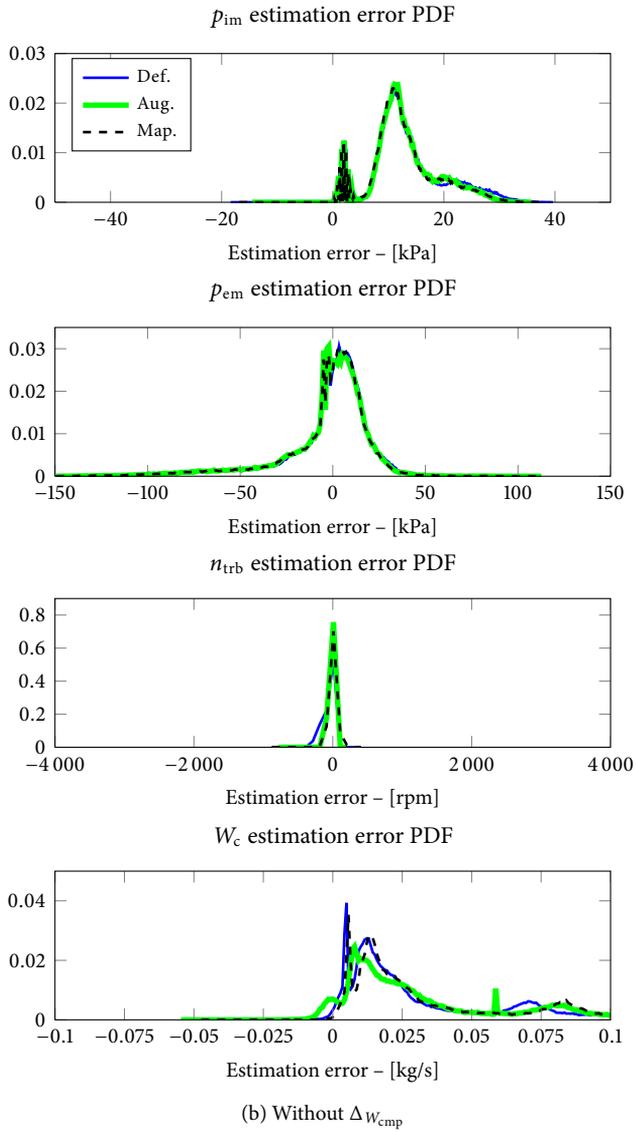


Table 2: Mean estimation error using experimental data for Def., Aug. and Map.

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

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 Figures 11 and 12, 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. The biggest benefit compared to Def. though is that the estimation errors become denser, see Figure 11a. That is, in the absence of outliers Aug. and Map. are comparable with respect to output estimation performance. The benefit with Map. is that it produces an adaptation map.

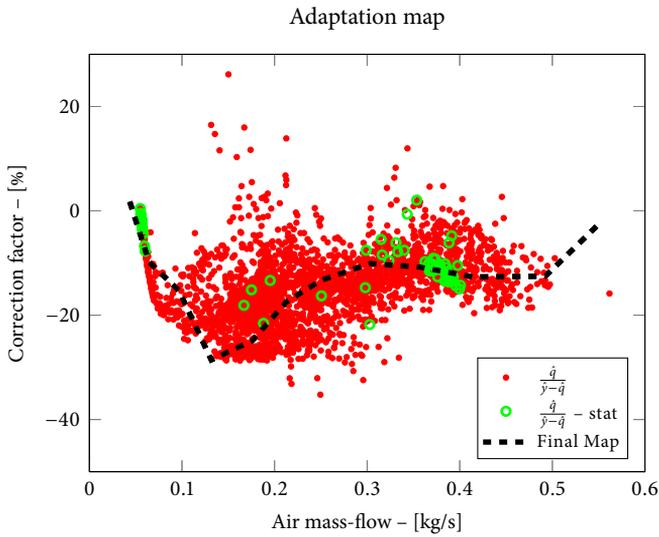
From Figure 12 it is seen that even though there are unknown model errors present, besides the compressor mass-flow, the method manages to estimate a map that describes the difference between modeled and measured air mass-flow through the compressor well. Figure 12a shows the correction factor between modeled and measured air mass-flow, similar to what was observed in Figure 8, and Figure 12b shows the adaptation map estimated by Map. and how the grid is spanned by the experimental data. Notice that there are a large number of samples at air mass-flows slightly below 0.2 kg/s and that very few of these samples represent stationary operation.

Finally, all of these results show that an adaptation map can be estimated even though the data are from a highly transient ETC not specifically designed for air mass-flow sensor adaptation.

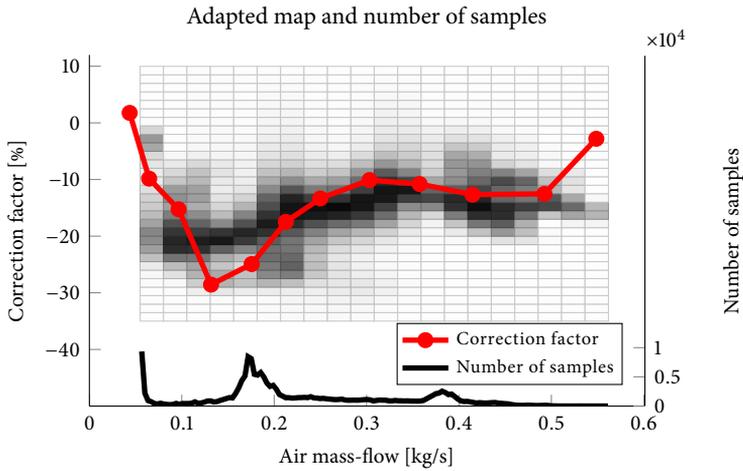
6 CONCLUSIONS

In modern control and diagnosis systems, model-based estimation has achieved increased attention. Especially interesting are techniques designed to handle the inevitable model errors. A technique common in engine applications is the use of adaptation maps that describe the model errors. A method for storing bias information from different operating points, based on the EKF as parameter estimator, is developed and investigated. This method achieves simultaneous estimation of original model states and parameters and also applies to adaptation of engine maps.

An inherent property of the adaptation of parameterized functions (maps) is that the system is locally unobservable. This has been analyzed, and a main contribution is a test for observability of systems with 1-D linearly interpolated map and measured or



(a) Aug. air mass-flow correction and final map estimated by Map.



(b) Final map plotted on a color map of the distribution of the air mass-flow correction from Aug.

Figure 12: By Aug. and Map estimated mass-flow correction.

observable interpolation variable. The observability analysis also gives insight into what applies to more general interpolation maps, such as bilinear interpolation, splines, etc.

Stochastic observers together with a parameterized bias that have 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, which comes as a result of local unobservability of the parameters, also forms a potential numerical problem for the observer, and a way to limit this growth without introducing extra tuning 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. An experimental evaluation shows that the method achieves the same estimation quality with respect to mean and maximum absolute error, as the method developed in Höckerdal et al. (2009) does, but adds value in that an engine adaptation map can be simultaneously estimated as well. Furthermore, the successful demonstration on experimental data, which inevitably introduces unknown model errors and noise, shows the robustness of the method. In the experimental evaluation it is also demonstrated that the method integrates well with traditional engineering tools, allowing prior knowledge of specific model errors, like the compressor mass-flow, to be incorporated and handled by the method. The method achieves simultaneous state estimation and map adaption, without using test cycles especially designed for adaptation, and it is therefore also suitable for offline calibration of maps.

REFERENCES

- Charlene Bembenek, Tom A. Jr. Chmielewski, and Paul R. Kalata. Observability conditions for biased linear time invariant systems. In *Proceedings of the American Control Conference*, pages 1180–1184, Philadelphia, Pennsylvania, June 1998. doi:10.1109/ACC.1998.703599.
- Gildas Besançon. *Nonlinear observers and applications*. Lecture notes in control and information sciences, 363. Springer, 2007.
- Guillaume Colin, Gérard Bloch, Yann Chamailard, and Floriane Anstett. Two air path observers for turbocharged SI engines with VCT. *Control Engineering Practice*, 17(5): 571–578, 2009. ISSN 0967-0661. doi:10.1016/j.conengprac.2008.10.001.
- Council of European Parliament. Directive 2005/55/EC of the European parliament and of the council of 28 September 2005, 2005.
- Fredrik Gustafsson. Slip-based tire–road friction estimation. *Automatica*, 33(6):1087–1099, 1997. ISSN 0005-1098. doi:10.1016/S0005-1098(97)00003-4.
- Lino Guzzella and Alois Amstutz. Control of diesel engines. *Control Systems Magazine, IEEE*, 18(5):53–71, October 1998. doi:10.1109/37.722253.
- Robert Haber and Heinz D. Unbehauen. Structure identification of nonlinear dynamic systems—a survey on input/output approaches. *Automatica*, 26(4):651–677, 1990. ISSN 0005-1098. doi:10.1016/0005-1098(90)90044-I.
- Robert Hermann and Arthur J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, AC-22(5):728–740, October 1977. doi:10.1109/TAC.1977.1101601.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. 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, 2008.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*, 17(3):408–417, 2009. doi:10.1016/j.conengprac.2008.09.004.
- M. P. Horton. Real-time identification of missile aerodynamics using a linearised Kalman filter aided by an artificial neural network. *IEE Proceedings. Control Theory and Applications*, 144(4):299–308, July 1997. doi:10.1049/ip-cta:19971125.
- Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, The Edinburgh Building, Cambridge CB2 2RU, UK, 1996.
- Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc, Englewood Cliffs, New Jersey 07632, 1980.

Thomas Kailath, Ali H. Sayed, and Babak Hassibi. *Linear Estimation*. Prentice-Hall, Inc, Upper Saddle River, New Jersey 07458, 2 edition, 2000.

Rudolf Emil Kalman, Yu-Chi Ho, and Kumpati S. Narendra. Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1, 1963.

Richard E. Kopp and Richard J. Orford. Linear regression applied to system identification for adaptive control systems. *AIAA*, 1(10):2300–2306, October 1963. doi:10.2514/3.2056.

Roger Larsson, Zoran Sjanic, Martin Enqvist, and Lennart Ljung. Direct prediction-error identification of unstable nonlinear systems applied to flight test data. In *Proceedings of the 15th IFAC Symposium on System Identification*, pages 144–149, Saint-Malo, France, July 2009. doi:10.3182/20090706-3-FR-2004.00024.

Ernest Bruce Lee and Lawrence Markus. *Foundations of Optimal Control Theory*. John Wiley & Sons, Inc, New York, 1968.

Xiangdon Lin and Yaakov Bar-Shalom. Multisensor target tracking performance with bias compensation. *IEEE Transactions on Aerospace and Electronic Systems*, 42(3): 1139–1149, 2006. ISSN 0018-9251. doi:10.1109/TAES.2006.248212.

Ingela Lind and Lennart Ljung. Regressor and structure selection in NARX models using a structured ANOVA approach. *Automatica*, 44(2):383–395, February 2008. doi:10.1016/j.automatica.2007.06.010.

Lennart Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24:36–50, February 1979. doi:10.1109/TAC.1979.1101943.

James C. Peyton Jones and Kenneth R. Muske. Identification and adaptation of linear look-up table parameters using an efficient recursive least-squares technique. *ISA Transactions*, 48(4):476–483, October 2009. doi:10.1016/j.isatra.2009.04.007.

Daniel Rupp and Lino Guzzella. Adaptive internal model control with application to fueling control. *Control Engineering Practice*, 18(8):873–881, 2010. ISSN 0967-0661. doi:10.1016/j.conengprac.2010.03.011.

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

Eric A. Wan, Rudolph Van Der Merwe, and Alex T. Nelson. Dual estimation and the unscented transformation. In *Neural Information Processing Systems*, pages 666–672. MIT Press, 2000.

Gang Wu. 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, 2006. doi:10.4271/2006-01-0607.

Ralf Zimmerschied and Rolf Isermann. Nonlinear time constant estimation and dynamic compensation of temperature sensors. *Control Engineering Practice*, 18(3):300–310, 2010. ISSN 0967-0661. doi:10.1016/j.conengprac.2009.11.008.

A PROOFS OF THEOREMS 3.1 AND 3.3

Theorem 3.1 *Model (8) with q_{fcn} defined by (7) is locally observable in \bar{z}_0 on a smooth trajectory (y, u) in I_i , containing the stationary point \bar{z}_0 , if $\dot{y}_2 \neq 0$, $b_q(p)y_2 \neq 0$, the pair (C, A) is observable and the matrix*

$$\begin{pmatrix} -A & B_q \\ C & 0 \end{pmatrix}$$

has full column rank.

Proof Since the interpolation variable, y_2 , is measured it is possible to prove the observability of θ_i and θ_{i+1} independently of x and the proof is split into two parts.

Consider (8) defined on $h_2(x) = y_2 \in I_i$, which since y_2 is measured can be written as

$$\dot{\theta}^i = 0 \tag{15a}$$

$$\dot{\theta}^{i+1} = 0$$

$$\begin{aligned} \dot{x} &= f^c(x, u, q_{\text{fcn}}(y_2, \theta^i, \theta^{i+1})) \\ y_1 &= h_1(x). \end{aligned} \tag{15b}$$

Linearizing (15b) at the stationary point \bar{z}_0 gives

$$\begin{aligned} \dot{x} &= f^c(\bar{z}_0) + \left(\frac{\partial f^c}{\partial x} \quad \frac{\partial f^c}{\partial u} \quad \frac{\partial f^c}{\partial q_{\text{fcn}}} \right) \Big|_{\bar{z}_0} (\bar{z} - \bar{z}_0) \\ &= f^c(\bar{z}_0) + A(x - x_0) + B_u(u - u_0) + B_q(q_{\text{fcn}} - q_{\text{fcn},0}) \\ y_1 &= h_1(x_0) + \frac{\partial h_1}{\partial x} \Big|_{\bar{z}_0} (x - x_0) = h_1(x_0) + C(x - x_0). \end{aligned} \tag{16}$$

Without loss of generality it can be assumed that $h_1(x_0) = 0$. The input-output form of (16) is then

$$a(p)y = b_u(p)u + b_q(p)q_{\text{fcn}}, \tag{17}$$

where it can be noted that x does not appear which allows observability analysis of θ^i and θ^{i+1} independently of x .

Combining (17) and (15) for each polynomial $b_q^l(p)$ in $b_q(p) = (b_q^1(p) \dots)^T$ gives

$$\begin{aligned} \dot{\theta}^i &= 0 \\ \dot{\theta}^{i+1} &= 0 \\ \zeta^l &= \frac{a(p)y - b_u(p)u}{b_q^l(p)} = q_{\text{fcn}} = \theta^i + (\theta^{i+1} - \theta^i)y_2 = \underbrace{(1 - y_2 \quad y_2)}_{=C(t)} \begin{pmatrix} \theta^i \\ \theta^{i+1} \end{pmatrix} \end{aligned} \tag{18}$$

on which observability of θ^i and θ^{i+1} can be established by a full rank condition on the observability gramian,

$$\Gamma(0, t) = \int_0^t \Phi^T(\tau, 0) C^T(\tau) C(\tau) \Phi(\tau, 0) d\tau,$$

under the condition that the influence of q_{fcn} is not cancelled by the zeros of $b_q(p)$. The transition matrix is here $\Phi(t, t_0) = I$ since $\dot{\theta}^i = \dot{\theta}^{i+1} = 0$. Without loss of generality it is assumed that (A, B_q, C) in (16) is a minimal realization. The zeros of $b_q(s)$ are then those s for which

$$\begin{pmatrix} sI - A & B_q \\ C & 0 \end{pmatrix} \quad (19)$$

loses rank (Kailath, 1980). The assumptions $b_q(p)y_2 \neq 0$ and the matrix

$$\begin{pmatrix} -A & B_q \\ C & 0 \end{pmatrix}$$

having full column rank assures that the influence of q_{fcn} is not cancelled.

Now it is possible to proceed analyzing the observability gramian for (18), which becomes

$$\Gamma(0, t) = \int_0^t I \begin{pmatrix} 1 - y_2 & y_2 \end{pmatrix}^T \begin{pmatrix} 1 - y_2 & y_2 \end{pmatrix} I d\tau = \int_0^t \begin{pmatrix} 1 - y_2(\tau) & y_2(\tau) \\ y_2(\tau) - y_2^2(\tau) & y_2^2(\tau) \end{pmatrix} d\tau,$$

and has full rank if and only if $\det \Gamma(0, t) \neq 0$. That is

$$t \int_0^t y_2^2(\tau) d\tau \neq \left(\int_0^t y_2(\tau) d\tau \right)^2, \quad (20)$$

which by algebraic computations give

$$\dot{y}_2(t) \neq 0$$

for any t such that the trajectory remains in I_i .

States θ^i and θ^{i+1} can now be uniquely determined from the trajectory if $\dot{y}_2 \neq 0$, and q_{fcn} in (16) can be computed and treated as an extra input allowing observability of x to be established using the usual observability rank condition. Finally, using the result in Lee and Markus (1968, Theorem 6.4) that a sufficient condition for local observability of a non-linear system is observability of a linearization in a stationary point, local observability of the system (15) is established. \square

Theorem 3.3 *Model (9) with q_{fcn} defined by (7) is locally observable in \bar{z}_0 on a smooth trajectory (y, u) in I_i , containing the stationary point \bar{z}_0 , if $\frac{d}{dt} h_2(x) \neq 0$, and the pair (C_1, A) is observable.*

Proof Equations (9c) and (9d) together with the assumption that (C_1, A) is an observable pair gives that state x is locally observable. Now, observability of θ^i and θ^{i+1} can be determined by rewriting (9a), (9b), and (9e) as

$$\zeta = \frac{y_2}{h_2(x)} - 1 = \begin{pmatrix} 1 - h_2(x) & h_2(x) \end{pmatrix} \begin{pmatrix} \theta^i \\ \theta^{i+1} \end{pmatrix}. \quad (21)$$

Using that $h_2(x)$ is known from the observability of x , (21) is in the form (18) and (θ^i, θ^{i+1}) can be uniquely determined if $\frac{d}{dt}h_2(x) \neq 0$, and Lee and Markus (1968) gives observability of (9). \square

Off- and On-Line Identification of Maps Applied
to the Gas Path in Diesel Engines[☆]

C

[☆] Chapter in book *Workshop on Identification in Automotive Systems*, 2010.

Off- and On-Line Identification of Maps Applied to the Gas Path in Diesel Engines

Erik Höckerdal, Erik Frisk, and Lars Eriksson

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, S-581 83 Linköping,
Sweden.*

ABSTRACT

Maps or look-up tables are frequently used in engine control systems, and can be of dimension one or higher. Their use is often to describe stationary phenomena such as sensor characteristics or engine performance parameters like volumetric efficiency. Aging can slowly change the behavior, which can be manifested as a bias, and it can be necessary to adapt the maps. Methods for bias compensation and on-line map adaptation using extended Kalman filters are investigated and discussed. Key properties of the approach are ways of handling component aging, varying measurement quality, as well as operating point dependent model quality. Handling covariance growth on locally unobservable modes, which is an inherent property of the map adaptation problem, is also important and this is solved for the Kalman filter. The method is applicable to off-line calibration of maps where the only requirement of the data is that the entire operating region of the system is covered, i.e. no special calibration cycles are required. Two truck engine applications are evaluated, one where a 1-D air mass-flow sensor adaptation map is estimated, and one where a 2-D volumetric efficiency map is adapted, both during a European transient cycle. An evaluation on experimental data shows that the method estimates a map, describing the sensor error, on a measurement sequence not specially designed for adaptation.

1 INTRODUCTION

There are high demands on the control and diagnosis functionality in engine control units and to achieve the desired performance they need accurate information about the system state. State information is acquired using measurements with sensors as well as through estimation algorithms that utilize models of the system. The latter is important since measurements are not always possible because sensors are either; unavailable, impossible to install, or too expensive for the intended application.

In control and diagnosis systems, models play an important role. These models can be represented in many ways either as equations or, as is very common, by utilizing maps (or look-up tables) that represent a function (Guzzella and Amstutz, 1998; Peyton Jones and Muske, 2009). Examples of maps can be found in sensor calibration data (1-D) or the volumetric efficiency of the engine (often 2-D). In particular, maps are frequently used to describe relations where physical models are unavailable or too complex for on-line implementation, e.g. sensor and actuator characteristics, cooler efficiency, injector characteristics, and aftertreatment systems.

A common situation is that the dynamics is well captured by the model but there are stationary errors (bias), which will be referred to as a biased model. This has been observed in for example truck engines (Höckerdal et al., 2008). Another important issue is that the system ages and there is need for adjusting the models accordingly to capture and account for such effects. There is thus a need for calibrating maps off-line at development time, and adapt the maps on-line while the system is running to capture aging.

For the off-line case it is worth noting that it is straight forward to include the map parameters in the total parameter vector θ and apply standard identification methods. Routines for on-line map adaptation have been considered in Peyton Jones and Muske (2009); Wu (2006); Höckerdal et al. (2011) and the approach discussed below builds upon the latter. In particular it includes a systematic way for handling aging and other slowly varying uncertainties. Simultaneous bias compensation and on-line map adaption is a key property which is of industrial value since the method integrates well with existing map-based solutions.

2 METHOD OUTLINE

The sections to come presents a systematic approach for designing an observer that adapt maps on-line at the same time as it reduces stationary errors in a model. The starting point is a default model, in discrete time state space form,

$$\begin{aligned}x_{t+1} &= f_{\text{def}}(x_t, u_t) \\ y_t &= h(x_t),\end{aligned}\tag{1}$$

with states $x \in \mathbb{R}^{n_x}$, inputs $u \in \mathbb{R}^{n_u}$, and outputs $y \in \mathbb{R}^{n_y}$, that suffers from stationary errors, see Figure 1 for an example. These model errors can exhibit both fast and slow dynamics, arising from for example operating point dependent bias and aging respectively.

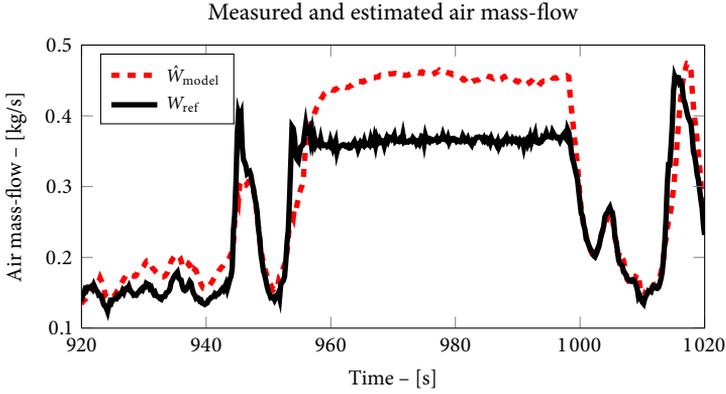


Figure 1: Typical example of model output from a biased model (Höckerdal et al., 2008), where W_{ref} is the air mass-flow measured by a reference sensor. As often is the case the model captures the dynamics well but suffers from operating point dependent stationary errors.

The objective is to handle these model errors in a systematic way by: using an on-line recursive algorithm, identifying model errors, and adapting the model. A direct way to achieve this is by introducing a parameterized function,

$$q_{\text{fcn}}(x_t, u_t, \theta_t), \quad (2)$$

in (1) and augmenting the state vector with the parameter vector θ

$$\theta_{t+1} = \theta_t \quad (3a)$$

$$x_{t+1} = f(x_t, u_t, q_{\text{fcn}}(x_t, u_t, \theta_t)) \quad (3b)$$

$$y_t = h(x_t). \quad (3c)$$

The idea with a construction like this is to let the parametrization (2) capture the operating point dependence, and use the parameters, θ , introduced as new states, to track the aging. With (3) as basis an observer that estimates the augmented state vector, (x, θ) , can be designed. Note that it is possible to incorporate prior information into the map identification process, such as limitations, smoothness, trends, etc., by altering the right hand side of (3a).

The following example of a parameterized function q_{fcn} , from an automotive application (Höckerdal et al., 2008), has served as motivation for the work and will later be used in the evaluation.

Example 1 In a heavy duty diesel engine with exhaust gas recirculation (EGR) and variable geometry turbocharger (VGT), the air mass-flow through the compressor is vital information for safe and clean engine control. Therefore the engine is typically equipped with a mass-flow sensor. However the sensor signal is subjected to an operating point

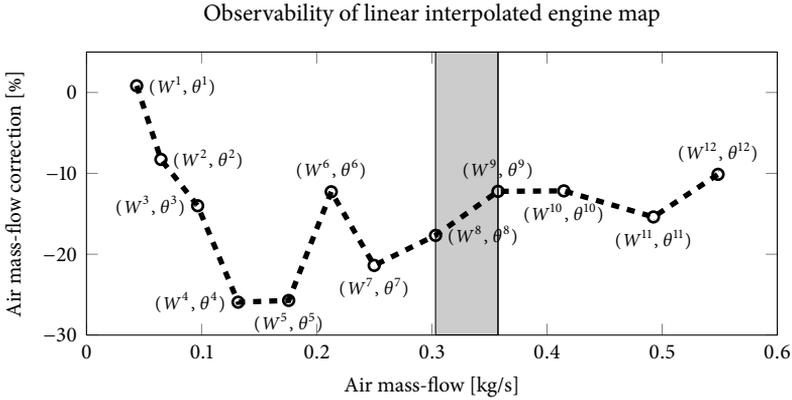


Figure 2: Air mass-flow sensor error map with the grid points denoted with the pair (W^i, θ^i) corresponding to a correction factor θ^i at a mass-flow of W^i .

dependent error, due to sensor installation and local flow fields, and this measurement error has to be compensated for when it is used in the engine control unit (ECU). The relation between the true air mass-flow W and measured air mass-flow W_{meas} can be stated as

$$W_{\text{meas}} = (1 + q_{\text{fcn}}(W, \theta)) W$$

where the function q_{fcn} is presented graphically in Figure 2. W^i are grid points, and θ^i are the corresponding correction parameters. In this case the sensor model is represented by a 1-D linear interpolation map, where W lies between grid points $W^i \leq W \leq W^{i+1}$ and the interpolation is defined by

$$q_{\text{fcn}}(W, \theta) = \theta^i + \frac{(\theta^{i+1} - \theta^i)}{W^{i+1} - W^i} \cdot (W - W^i). \quad (4)$$

Due to aging it is necessary to adapt this map, see Höckerdal et al. (2008) for a longer discussion. \diamond

Note that with the construction (3), using parameters to track aging, the observer stores information about the operating point dependent errors, i.e. the parameters act θ as a memory.

The development of a model like (3) for identification and its usage for estimation entails that some new issues have to be addressed. The main concern is how to update the parameters, θ , in a controlled manner, which is the topic of Section 3, with respect to observability and observer tuning. Here the parameterization of q_{fcn} is given and the interested reader is referred to for example Lind and Ljung (2008) for a discussion on how to find a structure and suitable regressors.

The system (3) is in standard state space form which means that any suitable observer design can be applied. The choice here is to use a stochastic filter which entails introduction of noise in (3) to describe model and measurement uncertainties. An approach for

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). An advantage of stochastic filters compared to deterministic observers is that, not only the state estimate, but also an estimate of the estimation error statistics 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 system aging, unknown state initialization, time dependent model and measurement quality, outlier rejection etc., see Section 3.2 for the discussion. For these reasons the joint parameter and state estimating EKF is used.

3 OBSERVABILITY

In estimation, observability or detectability of the system at hand is central in order to ensure correct and consistent estimates. This section is devoted to the observability of (3), where q represents linear interpolation. The observability discussion is conducted on the continuous time system,

$$\begin{aligned}\dot{\theta} &= 0 \\ \dot{x} &= f(x, u, q(x, \theta)) \\ y &= h(x),\end{aligned}\tag{5}$$

corresponding to (3). The results are valid also for the discrete time system as long as the sample time is chosen small enough (Kalman et al., 1963).

Intuitively, a system containing linear, or higher order, interpolation with the interpolation grid points as parameter states is not locally observable. This can be seen by considering the air mass-flow adaptation map in Example 1. Considering the total system, assuming more than one interpolation interval, this can be understood by the intrinsic nature of the linear interpolation (4). Since not all parameters are involved in the interpolation computation in each operating point, some parameters are always locally unobservable. However, even if only one interpolation interval is studied and the non-interacting parameters are disregarded, e.g. consider only the shaded region of Figure 2, the ordinary rank condition, stating sufficient conditions for observability, is insufficient for assessing local observability of the reduced system. To illustrate consider (5) in a single interval,

$$\begin{aligned}\dot{\theta}^i &= 0 \\ \dot{\theta}^{i+1} &= 0 \\ \dot{x} &= f(x, u, q(x_j, \theta^i, \theta^{i+1})) \\ y &= h(x), \quad y_k = x_j.\end{aligned}\tag{6}$$

For Example 1 this is realized by studying a stationary operating point where the interpolation variable, x_j , is in $I_i =]x_j^i, x_j^{i+1}[$, corresponding to the shaded region enclosed by W^8 and W^9 in Figure 2. Here the corresponding parameters θ^8 and θ^9 are unobservable

according to the ordinary rank condition, i.e., the Jacobian of

$$\begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \end{pmatrix},$$

where L_f denotes the Lie derivative along the vector field f , with respect to x , θ^i , θ^{i+1} will be rank deficient in all operating points. Hence direct use of the rank condition for assessing, both local and global, observability for this system is not possible. To ensure global observability of (5), a trajectory has to be considered and analytical results for establishing observability are derived in Höckerdal et al. (2011).

3.1 UNOBSERVABLE MODES AND COVARIANCE GROWTH

According to the previous paragraph, at any given time there are generally some parameters θ^i that are locally unobservable in (3). A property of EKF observers based on systems with locally unobservable states is that the estimation error covariance matrix grows linearly for the unobservable states if system noise for these modes is present, see Höckerdal et al. (2011). This effect is illustrated in Figure 3, where the variance of three parameter states, θ^5 , θ^8 , and θ^{10} from Figure 2 are plotted versus time. Parameter θ^5 corresponds to a parameter that is not observable 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. The effects of this covariance growth is two-sided: 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 filter stability when considering the life-time of the system, which has to be handled.

In engine map adaptation, experience indicates that adaptation algorithms, not using the EKF and joint parameter and state estimation, have problems concerning system aging and occasional spurious measurements. For example an engine, whose trajectories do not span the entire parameter space during normal operation and only occasionally enters some areas, may suffer from undesired system behavior caused by old parameters corresponding to seldom visited operating points. 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 (Ljung, 1999) and is a highly desirable property in engine adaptation algorithms not only to handle aging parameters but also to protect updated parameters.

After discussing the positive effects induced by the locally unobservable parameter states it remains how to ensure stable observer operation even though some parameter states never may become observable, causing the corresponding covariance matrix elements to grow without bound. A direct and intuitive way of handling this linear growth of estimation error covariance of the locally unobservable parameter states is to introduce an upper limit for the corresponding estimation error covariance matrix elements. A possible such upper limit is the initial error covariance matrix, P_0 , and with

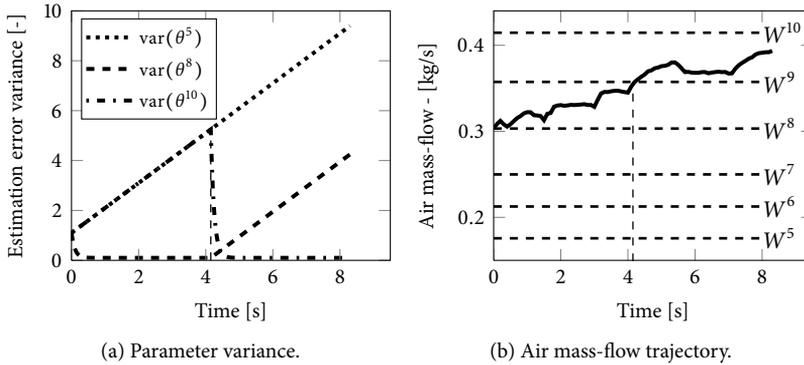


Figure 3: The figure shows the development of the estimation error variance for three parameter states. One that is unobservable during the entire trajectory – θ^5 , one that is observable for the first half of the trajectory – θ^8 , and one that is observable for the second half of the trajectory – θ^{10} .

this choice the introduction of yet another tuning parameter is avoided (Höckerdal et al., 2011).

3.2 METHOD FOR BIAS COMPENSATION

In the sense of map adaptation as a way of reducing bias, it is reasonable to compare the proposed map adaption method to other common ways of reducing operating point dependent estimation errors. One such method is presented in Höckerdal et al. (2009), where the state vector is augmented with a bias state that is used directly to describe the model error,

$$\begin{aligned}
 x_{t+1} &= f(x_t, u_t, q_t) \\
 q_{t+1} &= q_t \\
 y_t &= h(x_t).
 \end{aligned}
 \tag{7}$$

A main difference is that an observer based on (3) facilitates separate tracking of fast changes in operating point and slower changes due to aging. For an observer based on (7) this is not possible and 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 will also capture high frequency disturbances, and will thereby not be able to withstand spurious measurements to the same extent as the slower parameter states θ in (3).

4 METHOD EVALUATION

To evaluate the method two studies are performed, a simulation study and a study utilizing experimental data. The two evaluations address the problem of estimating and adapting the air mass-flow into an engine from two different perspectives that are mutually exclusive. That is either the mass-flow measurement, or the volumetric efficiency is considered to be most accurate. A third alternative is to install an accurate reference sensor, which is only possible in a test cell or development vehicle and not an option for a consumer product.

The simulation study demonstrates the capability to adapt a 2-D map for the volumetric efficiency while the experimental evaluation addresses the air mass-flow sensor adaptation problem presented in Example 1. The aim of the simulations are to show the convergence of the approach and the effect of a trajectory that only partly spans the map grid. As a part of the convergence analysis an investigation of the effect of incorrect noise models when using the EKF as a parameter estimator (Ljung, 1979) is included.

The experimental part shows the result of the method applied to experimental data with the aim to analyze robustness and performance in presence of model errors. Since the approach supplies a solution to the problem of biased default models solved in Höckerdal et al. (2009), the performance of that approach is presented as a comparison.

In both studies the non-linear model of a heavy duty diesel engine, developed in Wahlström and Eriksson (Accepted for publication), is used together with measurements from an engine in an engine test cell. An overview of the model schematics is presented in Figure 4. 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 is collected during a European Transient Cycle (ETC).

In the simulation study, feedback from the intake manifold pressure and turbine speed sensors are used whilst in the experimental evaluation, feedback from all sensors except the exhaust manifold pressure sensor are used. The performance is evaluated with respect to all states and outputs, i.e. intake and exhaust manifold pressures, turbine speed, and air mass-flow through the compressor.

4.1 STUDY 1: SIMULATION

The simulation study serves three purposes; illustrating convergence of the approach, show that the method applies also to higher map dimensions, and highlighting the effect of trajectories that does not span the entire map grid.

OBSERVER FOR THE SIMULATION EVALUATION

In the simulation evaluation the observer is constructed using the model developed in Wahlström and Eriksson (Accepted for publication) and replacing the volumetric efficiency sub model with a six-by-six parameterized cubical spline interpolation map, $q(n_{\text{eng}}, p_{\text{im}}, \theta)$, presented in Figure 5, which gives a model in the form (3). An alternative to using a spline for the interpolation would be to use bilinear interpolation. Here the spline is chosen to favor the smoothness of the volumetric efficiency.

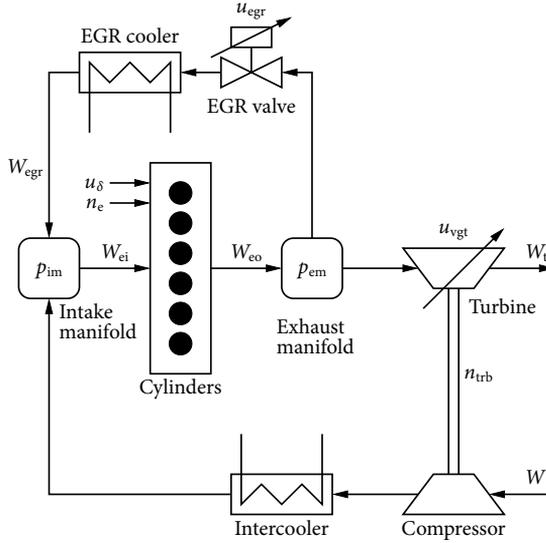


Figure 4: Schematic overview of the diesel engine model with EGR and VGT (Wahlström and Eriksson, Accepted for publication), showing model states (p_{im} , p_{em} , and n_{trb}), inputs (u_{δ} , u_{egr} , u_{vgt} , and n_e), and mass-flows between the different components (W , W_{ei} , W_{eo} , W_{egr} , and W_t). Rectangles with rounded corners denote control volumes.

SIMULATION EVALUATION

Figure 5 shows the true and adapted volumetric efficiency maps together with the ETC trajectory. The transparent upper surface corresponds to the true volumetric efficiency to be estimated and the lower, with a rough grid of 6×6 , the estimated map at the end of the ETC, which was initialized 10 percent units below the true map. At the bottom of Figure 5 level curves of the absolute estimation error are plotted together with the ETC trajectory to better illustrate the correlation between well adapted map areas and how the grid is spanned by the trajectory. As expected, map areas where the trajectory spends much time are well adapted while grid points in areas not spanned by the trajectory is unadapted and remains at their initial values.

NOISE MODEL SENSITIVITY

A theoretical analysis of the EKF as a parameter estimator is presented in Ljung (1979), which concludes that the parameter estimates will be biased if an incorrect noise description is used.

Inspired by this theoretical analysis and the fact that true nature of the noise seldom is completely known a simulation study is performed with the aim to analyze the influence of incorrect noise assumptions. The analysis is executed by applying the above designed observer to several sets of simulated data where the system and measurement noise

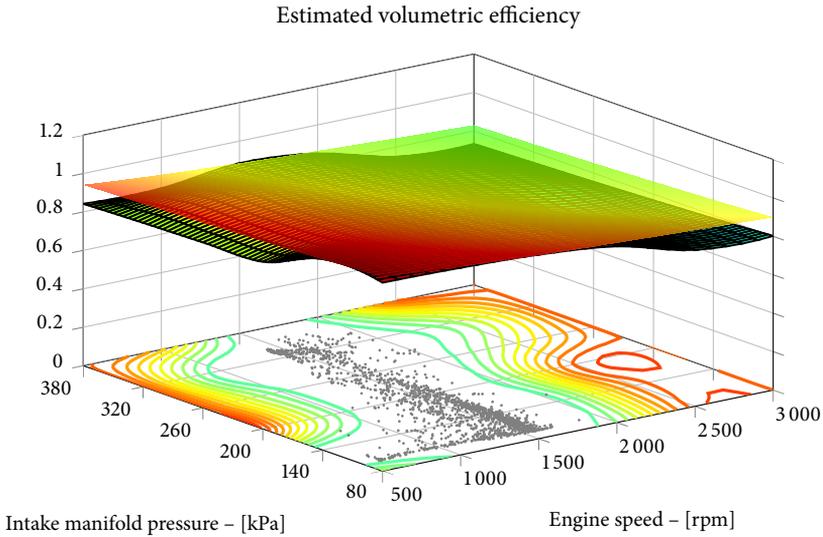


Figure 5: True and estimated volumetric efficiency together with level curves of the difference between the two. The level curves are also imposed by the ETC trajectory, used as adaptation cycle, represented by the gray dots.

variance and color had been adjusted. Neither of these simulations show of any problem of consistent convergence of the sensor error parameters.

Even though the analysis in the studied case does not indicate this to be an issue the potential problem can not be discarded in general. Even if the main purpose of the parameters is, as in the engine applications addressed here, their function as a memory there exist applications where this is an issue. In such cases it is straightforward to use the in Ljung (1979) modified EKF algorithm for improved convergence properties.

4.2 STUDY 2: EXPERIMENTAL DATA

To assess the performance and robustness of the method an evaluation using measurements from an engine in an engine test cell is conducted. Since in this case there are other significant model errors present, besides the air mass-flow sensor error, an extra bias state is introduced to improve the state estimates. The additional bias state is introduced according to (8) to compensate for a known model deficiency in the compressor mass-flow. The adaptation map, $q(W, \theta)$, is in this case defined by Example 1. Note that, in this case the adaptation map is situated in the measurement equation which is formally treated in Höckerdal et al. (2011).

OBSERVERS FOR THE EXPERIMENTAL EVALUATION

The observer designs for the experimental evaluation are, starting with those without an extra bias state: An EKF based on the default model developed in Wahlström and Eriksson (Accepted for publication) directly,

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ y_t &= \left(x_{p_{im},t} \quad x_{n_{trb},t} \quad h_W(x_t) \right)^T, \end{aligned}$$

referred to as Def. An EKF with an extra bias state introduced in the measurement equation to describe the air mass-flow measurement error from the method developed in Höckerdal et al. (2009),

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ q_{t+1} &= q_t \\ y_t &= \left(x_{p_{im},t} \quad x_{n_{trb},t} \quad h_W(x_t) + q_t \right)^T, \end{aligned}$$

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

$$\begin{aligned} x_{t+1} &= f(x_t, u_t) \\ \theta_{t+1} &= \theta_t \\ y_t &= \left(x_{p_{im},t} \quad x_{n_{trb},t} \quad (1 + q_{fcn}(h_W(x_t), \theta_t))h_W(x_t) \right)^T, \end{aligned}$$

referred to as Map. where the parameterization of q_{fcn} is presented in Figure 2. The primary difference between the three observers are presented schematically in Figure 6, where the feedback part of each observer is highlighted.

The additional bias state, complementing those for handling the measurement error, is introduced according to the ideas in Höckerdal et al. (2009). The purpose is to reduce the estimation errors due to a known error in the compressor model causing incorrect prediction of the compressor mass-flow, i.e. for Def.

$$\begin{aligned} x_{t+1} &= f(x_t, \Delta_{W,t}, u_t) \\ \Delta_{W,t+1} &= \Delta_{W,t} \\ y_t &= \left(x_{p_{im},t} \quad x_{n_{trb},t} \quad h_W(x_t, \Delta_{W,t}) \right)^T. \end{aligned} \tag{8}$$

EXPERIMENTAL EVALUATION

With the introduction of an extra state, compensating for the compressor mass-flow, 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 Figure 7 and Table 1. While Table 1 shows the

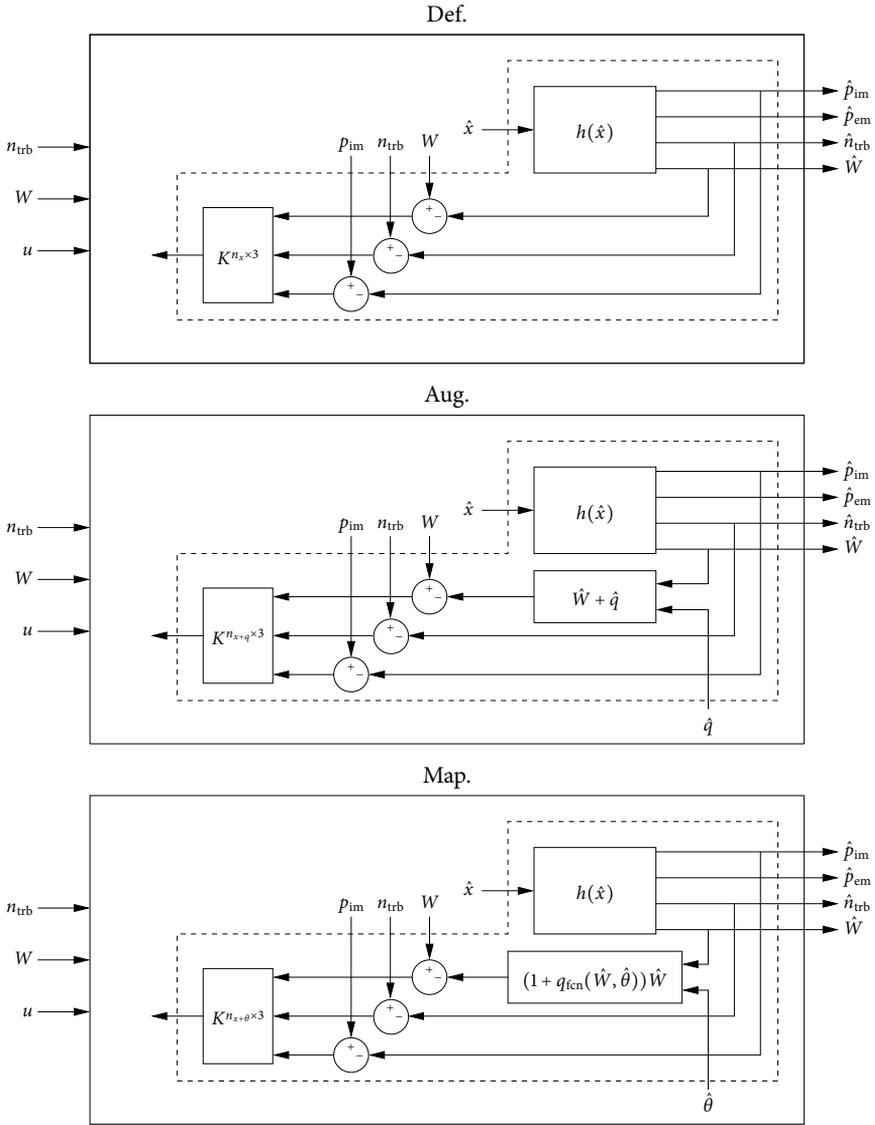


Figure 6: Schematics that highlight the feedback part of the three observers, Def., Aug., and Map, studied in the experimental evaluation. The main difference between the three configurations are shown in the block on the \hat{W} feedback signal, below the $h(\hat{x})$ block. To the left the inputs and feedback signals are shown, and to the right the four evaluation outputs are shown. Note that the dimension of the output signal of the Kalman gain block have different dimensions for the different observers.

Table 1: Mean estimation error using experimental data for Def., Aug. and Map.

Meas.	With $\Delta W_{cmp,t}$			Without $\Delta W_{cmp,t}$		
	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

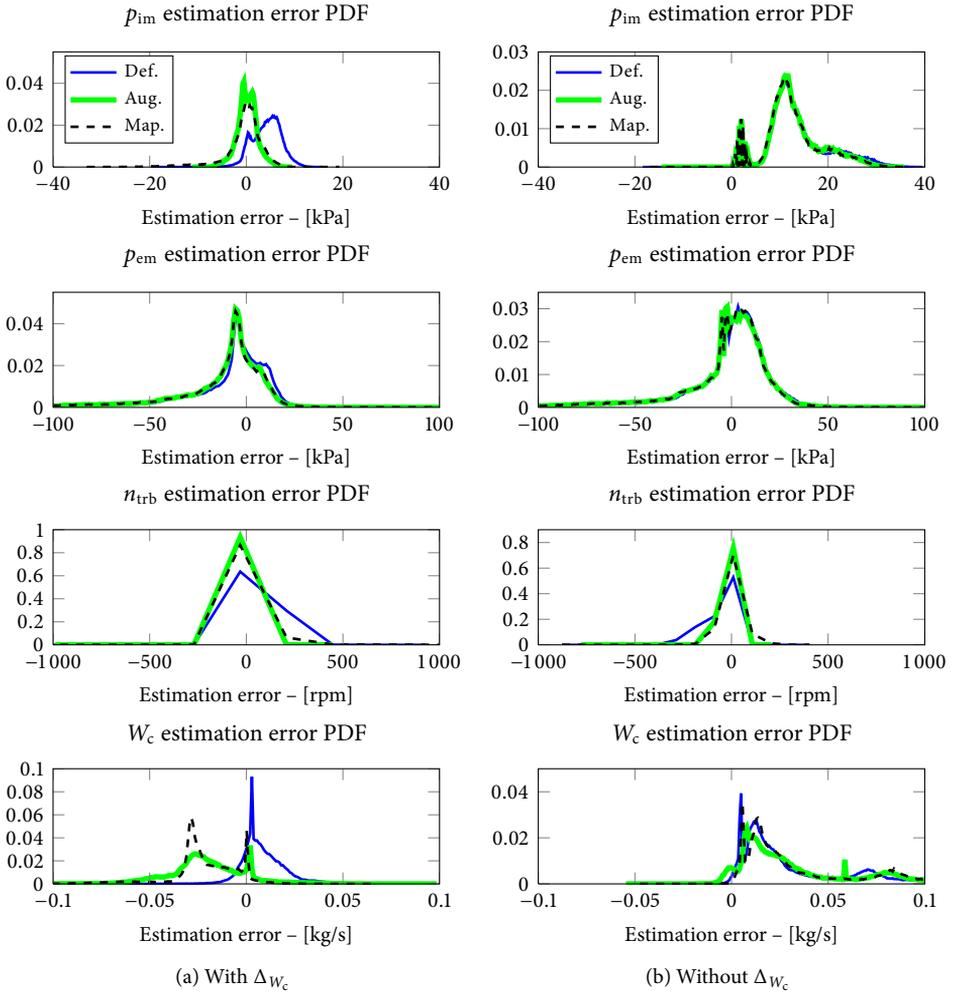


Figure 7: Probability density functions of the estimation errors showing bias and estimation error statistics.

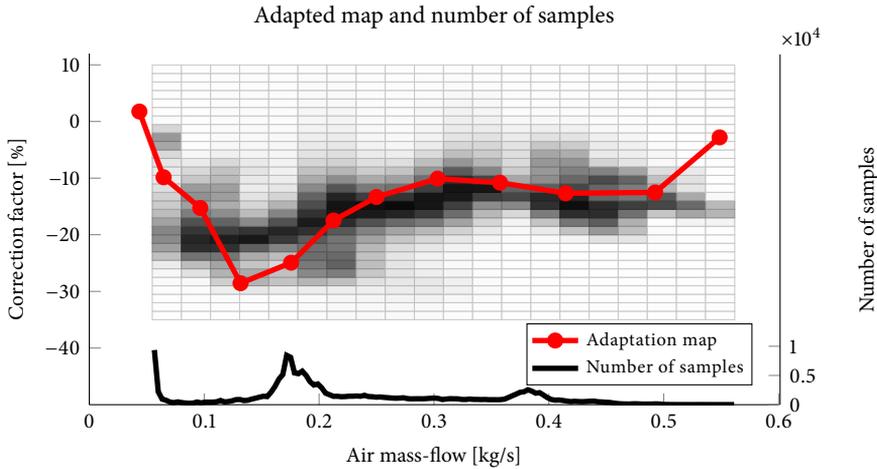


Figure 8: By Aug. and Map estimated mass-flow corrections showing the agreement in bias estimation of the two methods. The solid line with dots representing the final adaptation map of Map. and a color map of the distribution of the air mass-flow correction estimated by Aug.

mean estimation error, Figure 7 shows the estimation error probability density functions (PDF).

In Section 3.2 the tuning of Aug. and Map. were discussed, especially the different philosophies of the bias describing states, i.e. the relatively fast bias state in Aug. and the slow map states in Map. In Figures 7 and 8, and Table 1 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 produces an adaptation map, and enables the possibility to track aging and protection against corrupted measurements.

From Figure 8 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. Notice that there are large number of samples at air mass-flows slightly below 0.2 kg/s and that only few of these are from stationary operation.

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.

5 CONCLUSIONS

In modern control and diagnosis systems, model based estimation has achieved increased attention and especially interesting are techniques to handle the inevitable model errors. A method for on-line map identification, based on standard observer methodology, that handles the specific issues of aging parameters and occasional corrupted measurements is discussed. A common technique in engine applications is adaptation maps that describe the model errors. A method for storing bias information from different operating points, based on the EKF as parameter estimator, is investigated. This method achieves simultaneous estimation of original model states and parameters, and applies to adaptation of engine maps.

An inherent property of the adaptation of maps is that the system is locally unobservable. Stochastic observers together with a parameterized bias that has locally unobservable states is 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 forms a potential numerical problem for the observer and a way to limit this growth without introducing extra tuning parameters is provided.

The method is evaluated in a simulation study, where it is demonstrated that it estimates both the engine states and a parameterized 2-D adaptation map for the volumetric efficiency. An experimental evaluation shows that the method achieves the same estimation quality with respect to mean and maximum absolute error, as the method developed in Höckerdal et al. (2009), but adds value in that an engine adaptation map can be simultaneously estimated. Furthermore, the successful demonstration on experimental data, that inevitably introduces unknown model errors and noise, shows the robustness of the method. The method achieves simultaneous state estimation and map adaptation, without using test cycles especially designed for adaptation, and it is therefore suitable for both on- and off-line calibration of maps.

REFERENCES

- Lino Guzzella and Alois Amstutz. Control of diesel engines. *Control Systems Magazine, IEEE*, 18(5):53–71, October 1998. doi:10.1109/37.722253.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. SAE Technical Paper 2008-01-0992, 2008. doi:10.4271/2008-01-0992.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*, 17(3):408–417, 2009. doi:10.1016/j.conengprac.2008.09.004.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. EKF-based adaptation of look-up tables with an air mass-flow sensor application. *Control Engineering Practice*, 19(5):442–453, 2011. doi:10.1016/j.conengprac.2011.01.006.
- Rudolf Emil Kalman, Yu-Chi Ho, and Kumpati S. Narendra. Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1, 1963.
- Richard E. Kopp and Richard J. Orford. Linear regression applied to system identification for adaptive control systems. *AIAA*, 1(10):2300–2306, October 1963. doi:10.2514/3.2056.
- Ingela Lind and Lennart Ljung. Regressor and structure selection in NARX models using a structured ANOVA approach. *Automatica*, 44(2):383–395, February 2008. doi:10.1016/j.automatica.2007.06.010.
- Lennart Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24:36–50, February 1979. doi:10.1109/TAC.1979.1101943.
- Lennart Ljung. *System Identification - Theory for the user*. Prentice-Hall, Inc, Upper Saddle River, New Jersey 07458, 2 edition, 1999.
- James C. Peyton Jones and Kenneth R. Muske. Identification and adaptation of linear look-up table parameters using an efficient recursive least-squares technique. *ISA Transactions*, 48(4):476–483, October 2009. doi:10.1016/j.isatra.2009.04.007.
- Johan Wahlström and Lars Eriksson. Modeling VGT EGR diesel engines with optimization of model parameters for capturing nonlinear system dynamics. *Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering*, Accepted for publication. doi:10.1177/0954407011398177.
- Gang Wu. A table update method for adaptive knock control. SAE Technical Paper 2006-01-0607, 2006. doi:10.4271/2006-01-0607.

DAE and ODE Based EKF:s and their Real-Time
Performance Evaluated on a Diesel Engine[☆]

D

[☆]Submitted to *IEEE Transactions on Industrial Electronics*, 2010.

DAE and ODE Based EKF:s and their Real-Time Performance Evaluated on a Diesel Engine

Erik Höckerdal, Erik Frisk, and Lars Eriksson

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, S-581 83 Linköping,
Sweden.*

ABSTRACT

When estimating states in engine control systems there are limitations on the computational capabilities. This becomes especially apparent when designing observers for stiff systems since the implementation requires small step lengths. One way to reduce the computational burden, is to reduce the model stiffness by approximating the fast dynamics with instantaneous relations, transforming an ODE model into a DAE model.

Performance and sample frequency limitations for extended Kalman filters based on both the original ODE model and the reduced DAE model for a diesel engine is analyzed and compared. The effect of using backward Euler instead of forward Euler when discretizing the continuous time model is analyzed.

The ideas are evaluated using measurement data from a diesel engine. The engine is equipped with throttle, EGR, and VGT and the stiff model dynamics arise as a consequence of the throttle between two control volumes in the air intake system. It is shown that even though the ODE, for each time-update, is less computationally demanding than the resulting DAE, an EKF based on the DAE model achieves better estimation performance than one based on the ODE with less computational effort. The main gain with the DAE based EKF is that it allows increased step lengths without degrading the estimation performance compared to the ODE based EKF.

1 INTRODUCTION

Model based estimation for systems with both fast and slow dynamics is a wide and active research area, with different application areas such as combustion engines (Lino et al., 2008; García-Nieto et al., 2008), and electrochemical and reactive distillation processes (Mandela et al., 2010).

In the automotive industry there is a constant pursuit of designing cost effective systems that are able to meet stricter emission legislations and consumer demands on cheap and reliable operation. In particular this is true for the competitive market of heavy duty diesel engines where concepts like intake throttle, exhaust gas recirculation (EGR) and variable geometry turbines (VGT) are introduced to cope with the increased demands on lower emissions (Friedrich et al., 2009). To make accurate and robust control (Alt et al., 2009; Reichhartinger and Horn, 2009) and diagnosis (Morgan and Liu, 2009) of these systems feasible, reliable measurements or estimates of the internal states are required.

Due to installation limitations and sensor cost, the use of model based estimation has increased (Lino et al., 2008; García-Nieto et al., 2008; Reichhartinger and Horn, 2009). Often there exist models that have been developed for some purpose, e.g. system analysis, control, or diagnosis, which are ordinary differential equations (ODE) based on first principle physics (Müller et al., 1998; Eriksson, 2007). Much time are invested in these models, with large efforts on model parametrization, calibration and validation. It is therefore desirable to use the developed models in as many applications as possible. However, models like these can have properties, e.g., bias (Höckerdal et al., 2008) and/or fast dynamics (Wahlström and Eriksson, 2010; Friedrich et al., 2009), that limit their applicability in other areas, in this case estimation. Methods that can be used to attenuate such model deficiencies are desired. Biased models and their use in estimation are treated in for example Höckerdal et al. (2009, 2010) while the use of models that experience fast and slow dynamics, i.e. stiff models, is the topic of this work.

The problem of stiff models in engine control is closely connected to the embedded system in which the estimator is implemented and the computational limitations of the embedded system. In an engine control unit (ECU), a main difficulty with stiff models is that the model execution is scheduled in loops with fixed frequencies, which limits the possibility to satisfactorily solve the differential equations. A possible solution is to replace the fast dynamics in the ODE, with instantaneous dynamics, i.e. algebraic conditions, resulting in a differential algebraic equation (DAE), keeping the overall model structure (Brenan et al., 1989).

The objective is to analyze the potential in using, the popular and widely applied, Kalman filter (Kalman, 1960) and especially the extended Kalman filter (EKF) (Jazwinski, 1970) based on a DAE compared to using an EKF based on a stiff ODE. Focus is on the different approaches step length and stability requirements for different discretization methods, and thereby their computational demands, when running the corresponding ODE and DAE based EKF:s.

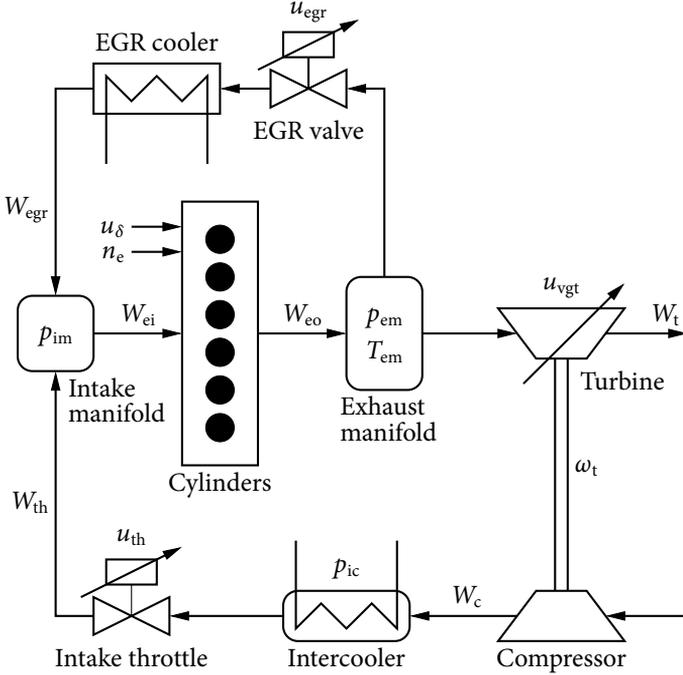


Figure 1: Schematic of the diesel engine model (Wahlström and Eriksson, 2010) with throttle, EGR and VGT, showing model states (p_{im} , p_{em} , p_{ic} , ω_t , and T_{em}), inputs (u_{egr} , u_{vgt} , u_{δ} , u_{th} , and n_e), and flows between the different components (W_c , W_{th} , W_{egr} , W_{ei} , W_{eo} , and W_t). Rectangles with rounded corners represent control volumes.

2 BACKGROUND AND PROBLEM MOTIVATION

Numerically efficient and stable observer solutions for embedded systems are desirable and an often limiting fact is the base frequencies in which the control system is scheduled, such as 20 or 100 Hz. This can cause problems when implementing EKF:s for models with fast dynamics, e.g. stiff models, due to their need for a small step length. One possible way around this is to simplify the model dynamics that cause the stiff behavior, by replacing it with algebraic constraints, and in this way introduce a DAE formulation.

2.1 STIFFNESS IN ENGINE MODELS

In the automotive industry, and especially in the field of engine control, a typical example where stiff models is a problem is when a throttle is used (Wahlström and Eriksson, 2010). A throttle in the air intake system, between the compressor and engine cylinders, connects the intake manifold and intercooler volumes with a variable restriction, see Figure 1. A common model for filling and emptying dynamics of volumes, is to consider

mass conservation of ideal gases in fixed control volumes (Heywood, 1988)

$$\dot{p}_{CV} = \frac{dp_{CV}}{dt} = \frac{R_{CV}T_{CV}}{V_{CV}} (W_{CV, in}(\cdot) - W_{CV, out}(\cdot)),$$

where p_{CV} and T_{CV} denotes the control volume pressure and temperature, R_{CV} the ideal gas constant for the gas in the control volume, V_{CV} the volume, and $W_{CV, in}/W_{CV, out}$ are the mass-flows in/out of the control volume. The flows depend on pressures and temperatures in the surrounding systems see Eriksson (2007), but these dependencies are omitted for the sake of clarity. An ODE for the intake system pressures presented in Figure 1 is

$$\dot{p}_{im} = \frac{R_a T_{im}}{V_{im}} (W_{th}(\cdot) + W_{egr}(\cdot) - W_{ei}(\cdot)) \quad (1a)$$

$$\dot{p}_{ic} = \frac{R_a T_c}{V_{ic}} (W_c(\cdot) - W_{th}(\cdot)), \quad (1b)$$

for the intake manifold and intercooler respectively. In operating points with fully open, or nearly open throttle, the filling and emptying dynamics become fast, resulting in oscillating estimates if the discretization method and/or step length is inadequate, i.e. too long, see the oscillating estimates in Figure 3b. For heavy duty diesel engines, these stiff operating conditions are common and need to be addressed.

The fast dynamics around the throttle and p_{ic} results in that the flows $W_c(\cdot)$ and $W_{th}(\cdot)$ follow each other closely. Therefore, one way to handle this is to remove the intercooler control volume and consider the flows in and out of that volume as an algebraic constraint. By doing this, the original ODE (1) is transformed into the following DAE,

$$\dot{p}_{im} = \frac{R_a T_{im}}{V_{im}^{DAE}} (W_{th}(p_{ic}, p_{im}, u_{th}) + W_{egr}(\cdot) - W_{ei}(\cdot)) \quad (2a)$$

$$0 = W_c(p_{ic}, \omega_t) - W_{th}(p_{ic}, p_{im}, u_{th}), \quad (2b)$$

with p_{ic} now becoming an algebraic variable. When a control volume is removed, the dynamics of the total system is affected, especially operating conditions with wide open throttle, where the dynamics becomes faster than that of the original model. Since most operation of diesel engines is with wide open throttle it is desirable to have a model, and an EKF, for these operating conditions. Figure 2 shows a simulation of the default ODE model, defined by

$$\dot{T}_{em} = f_{T_{em}}(p_{im}, p_{em}, \omega_t, T_{em}, u_{\delta}, u_{egr}, u_{vgt}, n_e) \quad (3a)$$

$$\dot{\omega}_t = f_{\omega_t}(p_{em}, p_{ic}, \omega_t, T_{em}, u_{vgt}) \quad (3b)$$

$$\dot{p}_{em} = f_{p_{em}}(p_{im}, p_{em}, \omega_t, T_{em}, u_{\delta}, u_{egr}, u_{vgt}, n_e) \quad (3c)$$

$$\dot{p}_{im} = f_{p_{im}}(p_{im}, p_{em}, p_{ic}, T_{em}, u_{\delta}, u_{egr}, u_{th}, n_e) \quad (3d)$$

$$\dot{p}_{ic} = f_{p_{ic}}(p_{im}, p_{ic}, \omega_t, u_{th}) \quad (3e)$$

and

$$y_{\text{fdb}} = \begin{pmatrix} p_{\text{im}} & p_{\text{em}} & \omega_t & W_c(p_{\text{ic}}, \omega_t) \end{pmatrix}^T \quad (4a)$$

$$y_{\text{eval}} = p_{\text{ic}}, \quad (4b)$$

together with simulations of the DAE model, with modifications of (3d) and (3e) according to (2). The measurement equation (4) is split in two, where (4a) is used for feedback in the EKF and (4b) is used as an evaluation output in Section 4 only. Two different control volume sizes, $V_{\text{im}}^{\text{DAE}} = V_{\text{im}}$ and $V_{\text{im}}^{\text{DAE}} = V_{\text{im}} + V_{\text{ic}}$, are evaluated. Especially it shows that the smaller volume gives undesired overshoots in the simulations of the pressures, turbine speed, and mass-flow for open throttle operation, see solid ellipses in Figure 2a, while the closed throttle operation is not that bad, see dotted ellipses in Figure 2b. Therefore the volume remaining in the DAE is increased to the sum of the original ODE's control volumes, i.e. $V_{\text{im}}^{\text{DAE}} = V_{\text{im}} + V_{\text{ic}}$ in (2a).

3 DAE OBSERVER

Before presenting the EKF algorithm used for the DAE observer in Section 3.2, a review of previous results and algorithms is conducted and presented in Section 3.1, while observability of the DAE model is the subject of Section 3.3.

3.1 OBSERVER DESIGN – EKF FOR DAE:S

For continuous time ODE:s, such as

$$\dot{x} = f(x, u) + w \quad (5a)$$

$$y = h(x) + v, \quad (5b)$$

where w and v are zero mean white noise processes with covariance matrices Q and R , EKF implementations are well known (Kailath et al., 2000). For general DAE:s

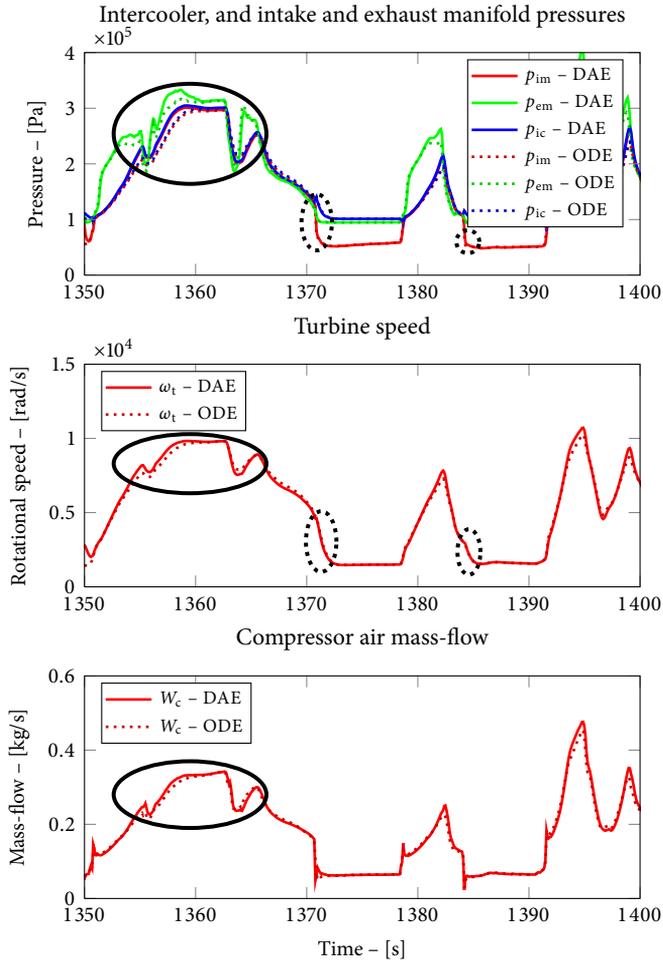
$$\dot{x} = f(x, z, u) + w \quad (6a)$$

$$0 = g(x, z, u) \quad (6b)$$

$$y = h(x, z, u) + v, \quad (6c)$$

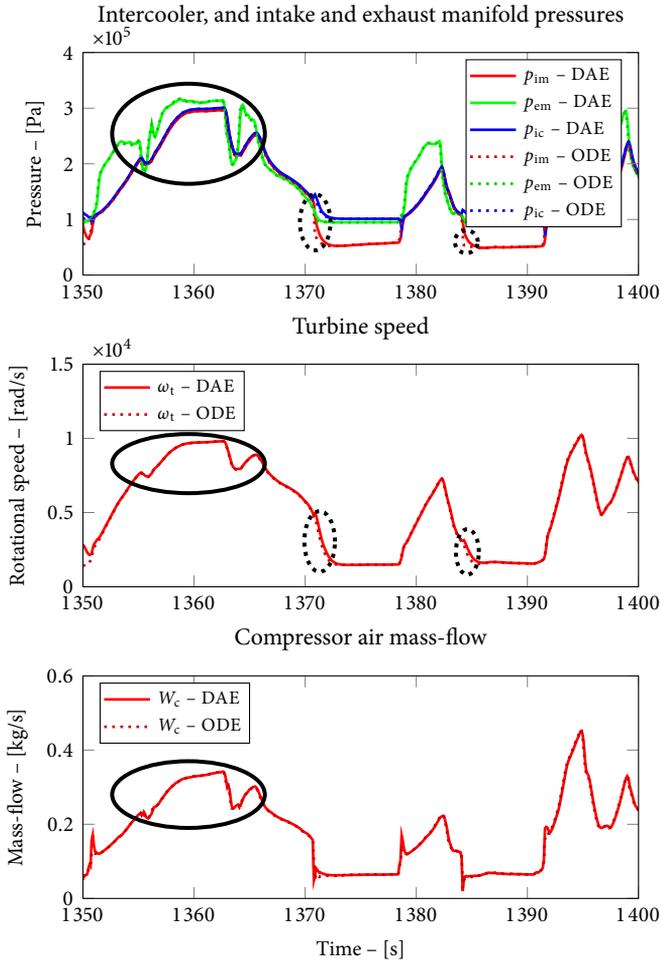
where x and z denotes differential and algebraic variables/states, respectively this is not the case. EKF:s for DAE:s are treated in for example Becerra et al. (2001); Mandela et al. (2010), that both consider semi-explicit DAE:s of index 1, i.e. $\frac{\partial g}{\partial z}$ has full column rank. The difference between the EKF in Becerra et al. (2001) and Mandela et al. (2010) is that the latter includes the algebraic subsystem (6b) in the standard EKF algorithm (Beccuti et al., 2009; Won et al., 2010). This is done by differentiating the linearized algebraic subsystem

$$\begin{aligned} \dot{x} &= A_t x + B_t z + w \\ 0 &= C_t x + D_t z \end{aligned} \quad \Rightarrow \quad \begin{aligned} \dot{x} &= A_t x + B_t z + w \\ \dot{z} &= -D_t^{-1} C_t \dot{x}, \end{aligned} \quad (7)$$



$$(a) \dot{p}_{im} = \frac{R_a T_{im}}{V_{im}} (W_{th}(\cdot) + W_{egr}(\cdot) - W_{ei}(\cdot))$$

Figure 2: DAE simulations with and without control volume adjustment for a step length of 50 ms using backward Euler, compared to high end simulation of the default ODE using the MATLAB variable step length solver, ode23t, with absolute and relative tolerances of 10^{-6} . The better dynamic behavior of the model with increased volume is seen in the areas highlighted with solid ellipses, representing open throttle operation. Also, there is only minor loss in simulation accuracy for the larger volume during closed throttle operation, indicated by dotted ellipses.



$$(b) \dot{p}_{im} = \frac{R_a T_{im}}{V_{im} + V_{ic}} (W_{th}(\cdot) + W_{egr}(\cdot) - W_{ei}(\cdot))$$

where

$$\begin{pmatrix} A_t & B_t \\ C_t & D_t \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial z} \end{pmatrix},$$

and the input, u , has been disregarded, which includes the estimation quality of the algebraic states in the Kalman gain computation. Note that inputs with discontinuous first-order derivatives are not allowed in the algebraic submodel (6b), due to the differentiation in (7). Normally this is not a problem since a low-pass filter on the inputs can be introduced to circumvent this.

Another issue with EKF:s for DAE:s is that, for the problem to be well posed, system noise must be introduced with care (Gerdin et al., 2007). In (6), as well as in Mandela et al. (2010), system noise is introduced only in the differential part of the DAE, which may be more restrictive than necessary but guarantees that the problem is well posed (Gerdin et al., 2007).

3.2 EKF ALGORITHM

The algorithm used has a lot in common with the one presented in Mandela et al. (2010). Among the differences are that the state prediction is performed with the explicit forward Euler, or implicit backward Euler, i.e. no dedicated DAE solver is used, and that the covariance matrix is predicted using the corresponding discrete-time linearization (8) instead of using the transition matrix $\phi = e^{\tilde{A}_{t|t}T_s}$. Note that the solution using forward Euler is supplemented with an equation solver to ensure consistency of the algebraic states after prediction.

The algorithm used is summarized in Algorithm 1. For the diesel engine f is composed by (3a–3d), g by (3e), and h by (4a), where modifications of (3d) and (3e) are made according to (2). In Algorithm 1, $\tilde{G}_{t|t}$ is a result of the differentiation in (7) and implies that system noise is present in the differential variables of the DAE only.

3.3 OBSERVABILITY OF ENGINE MODEL

To ensure that the observer estimates converge to the true states, observability of the underlying model along the studied trajectory is central. In contrast to standard state space descriptions, for DAE:s there exist several concepts of observability (Yip and Sincovec, 1981; Dai, 1989; Mehrmann and Stykel, 2005; Losse and Mehrmann, 2008). The concepts are \mathcal{C} -, \mathcal{R} - and \mathcal{I} -observability, denoting complete observability, observability within the reachable set and impulse observability, respectively.

Generally, descriptor systems, or DAE:s, are not \mathcal{C} -observable since they contain algebraic constraints that force the solution, and output, onto a specific manifold. For this reason the observability within the reachable set – \mathcal{R} -observability is introduced, which is the concept used here. This concept needs an appropriate projection of the dynamical part of the system, sometimes referred to as the slow sub-system, onto a manifold defined by the algebraic equations, or fast sub-system. For linear time-invariant systems

$$\begin{aligned} \tilde{E}\dot{\tilde{x}} &= \tilde{A}\tilde{x} \\ y &= \tilde{H}\tilde{x}, \end{aligned} \tag{9}$$

Algorithm 1 Extended Kalman Filter

1. Initialization:

$$\begin{pmatrix} \hat{x}_{0|0} \\ \hat{z}_{0|0} \end{pmatrix} = \begin{pmatrix} x_0 \\ z_0 \end{pmatrix} \quad \text{and} \quad P_{0|0} = \Pi_0,$$

where $\begin{pmatrix} x_0 \\ z_0 \end{pmatrix}$ is the initial state estimate and $\Pi_0 = \text{cov} \begin{pmatrix} x_0 \\ z_0 \end{pmatrix}$. Let $t = 0$.

2. Prediction:

$$\begin{aligned} \hat{x}_{t+1|t} &= \tilde{f}(\hat{x}_{t|t}, \hat{z}_{t|t}, u_t) && \Rightarrow && \begin{pmatrix} \hat{x}_{t+1|t} \\ \hat{z}_{t+1|t} \end{pmatrix} \\ 0 &= g(\hat{x}_{t+1|t}, \hat{z}_{t+1|t}) \\ \bar{P}_{t+1|t} &= \bar{A}_{t|t} \bar{P}_{t|t} \bar{A}_{t|t}^T + \bar{G}_{t|t} Q \bar{G}_{t|t}^T, \end{aligned}$$

where the implication indicates solving $\hat{x}_{t+1|t}$ and $\hat{z}_{t+1|t}$ from the system of equations,

$$\bar{A}_{t|t} = I + T_s \begin{pmatrix} A_{t|t} & B_{t|t} \\ -D_{t|t}^{-1} C_{t|t} A_{t|t} & -D_{t|t}^{-1} C_{t|t} B_{t|t} \end{pmatrix}, \quad (8)$$

T_s the sampling time, and

$$\bar{G}_{t|t} = \begin{pmatrix} I \\ -D_{t|t}^{-1} C_{t|t} \end{pmatrix},$$

from (7).

3. Measurement update:

$$\begin{aligned} S_{t+1} &= \bar{H}_{t+1|t} \bar{P}_{t+1|t} \bar{H}_{t+1|t}^T + R \\ \bar{K}_{t+1} &= \bar{P}_{t+1|t} \bar{H}_{t+1|t}^T S_{t+1}^{-1} \\ \hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + \bar{K}_{t+1} (y_{t+1} - h(\hat{x}_{t+1|t}, \hat{z}_{t+1|t}, u_{t+1})) \\ 0 &= g(\hat{x}_{t+1|t+1}, \hat{z}_{t+1|t+1}) && \Rightarrow && \hat{z}_{t+1|t+1} \\ \bar{P}_{t+1|t+1} &= \bar{P}_{t+1|t} - \bar{P}_{t+1|t} \bar{H}_{t+1|t}^T S_{t+1}^{-1} \bar{H}_{t+1|t} \bar{P}_{t+1|t} \end{aligned}$$

where

$$\bar{H}_{t+1|t} = \left(\frac{\partial h}{\partial x} \quad \frac{\partial h}{\partial z} \right) \Big|_{\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} \hat{x} \\ \hat{z} \end{pmatrix}_{t+1|t}}.$$

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

that are regular, i.e. that $\det(\alpha\bar{E} - \beta\bar{A}) \neq 0$ for some $(\alpha, \beta) \in \mathbb{C}^2$, this can be described in terms of \bar{E} , \bar{A} , and \bar{H} .

Definition 3.1 (\mathcal{R} -observable) *The system (9) is called observable within the reachable set if the zero output of the descriptor system with $u = 0$ implies that all solutions of this system satisfy $P_r \bar{x} = 0$, where P_r denotes the projection onto the right deflating sub-space corresponding to the finite eigenvalues of $\lambda\bar{E} - \bar{A}$.*

The concept of \mathcal{R} -observability can, in the linear time-invariant case (9), be characterized algebraically according to the following theorem (Dai, 1989).

Theorem 3.1 *The system (9) is \mathcal{R} -observable if and only if*

$$\begin{pmatrix} \lambda\bar{E} - \bar{A} \\ \bar{H} \end{pmatrix}$$

has full column rank for all $\lambda \in \mathbb{C}$.

\mathcal{R} -observability of the engine model is established in 3 steps. First, using Theorem 3.1 on linearizations in stationary operating points along the evaluation trajectory, local weak \mathcal{R} -observability of the linearized continuous time model is established for all t . Second, observability of a linearization in a stationary operating point is a sufficient condition for weak local observability of the non-linear system (Lee and Markus, 1968, Theorem 6.4). Third, observability does not depend on the choice of discretization method as long as T_s is chosen small enough (Kalman et al., 1963).

4 EVALUATION WITH RESPECT TO STEP LENGTH

The evaluation is performed by running the two EKF:s based on the default ODE model and the modified DAE model. Forward and backward Euler are used for the discretization of both observers, and the evaluation uses measurement data from an engine in a test cell, see Appendix A for more information. The EKF tuning is with respect to the system and measurement noise covariances, Q and R . For low dimension systems there exist methods for off-line computation of these tuning variables (Salvatore et al., 2010). However, normally at least Q is found by manual tuning while R could be found using sensor characteristics provided by the sensor manufacturer (Menegaldo et al., 2009). Here, the measurement noise is affected by both sensor characteristics as well as sensor installation, and is therefore treated in the same manner as Q , i.e. through manual tuning. The observers are tuned equally with respect to model and measurement uncertainty with,

$$Q = \text{diag}(10^{-4}, 0.1, 1, 1, 1)$$

and

$$R = \text{diag}(100, 100, 1, 10^{-8}),$$

corresponding to (3a–3d) and (4a) respectively. Stability limits for these settings are ~ 15 ms for the ODE observers and ~ 125 ms for the DAE observers which shows that a

Table 1: Normalized execution times for ODE and DAE based EKF:s with forward and backward Euler and different step lengths for a 220 s segment in the beginning of an WHTC. Table entries “–” indicate that the step length was too large for stable simulation with the current discretization method. “**Bold face**” entries denote the largest step length with non-oscillating and stable simulation results for each model and discretization method.

Model	Disc.	3 ms	10 ms	15 ms	50 ms	125 ms
ODE	FE	1	0.30 [†]	0.20 [†]	–	–
	BE	2.50	0.95	0.78	–	–
DAE	FE	6.23	1.99	1.35	0.45	0.19
	BE	4.93	1.63	1.11	0.39	0.16

[†] Oscillating estimates, see Figure 3b.

significant increase in step length can be achieved for the DAE based EKF compared to the ODE EKF, see Table 1.

The performance measure used in the evaluations is the root mean square error,

$$\text{RMSE} = \sqrt{\frac{\sum_t (\hat{y}_t - y_t)^2}{N}}, \quad (10)$$

where \hat{y}_t is the estimate and y_t the measurement. Since several of the model outputs are used for feedback (4a) in the EKF:s, an extra model output (4b) not used in the observer, is used for the evaluation, i.e. p_{ic} .

The evaluation highlights three aspects of the real-time performance of DAE and ODE based EKF:s. It starts by analyzing the effects of stiff models in estimation, then the forward and backward Euler for the discretization are analyzed, and it finishes up by studying the possibility of increased step lengths in the ODE and DAE based EKF:s with maintained estimation accuracy.

4.1 EFFECTS OF STIFF ODE DYNAMICS

Figure 3 presents the state estimates of both the ODE and DAE based EKF:s, using forward Euler for the prediction, for a 100 s segment of the WHTC (Economic Commission for Europe – Inland Transport Committee, 2010) where two different step lengths, 3 and 10 ms, have been used. The main difference between 3a, representing simulations with 3 ms step length which is the largest step size for which the ODE based EKF estimates are stable and without oscillations, and 3b, representing a step length of 10 ms, is the oscillating estimates of the ODE observer during high mass-flow operation. These oscillating properties, observed as a band of noise in Figure 3b of the ODE observer, can be explained by the stiffness of the ODE in operating points with wide open throttle, i.e. high compressor air mass-flow.

The estimated probability density functions (PDF) of the estimation errors for the two observers with two different step lengths are presented in Figure 4. In Figure 4a,

presenting the estimation error PDF:s for a step length of 3 ms, the appearance of the two observers are rather similar indicating that for a step length of 3 ms an EKF based on an ODE would be preferable due to its less demanding computational complexity. While for a step length of 10 ms the degradation of the ODE estimation error PDF:s for the outputs closely connected to the throttle, i.e. p_{im} , p_{ic} and W_c , is apparent in Figure 4b.

4.2 INFLUENCE OF DISCRETIZATION METHOD

A comparison of selected discretization step lengths for forward and backward Euler as discretization methods is presented in Table 1. It shows that, even though there is a gain in stability using backward Euler, e.g., non-oscillating estimates for the ODE model, the upper limit of possible step lengths for stable estimation is not affected for either the ODE or DAE based EKF:s. That is, the ODE and DAE based EKF:s lose stability for the same step length, independent of the discretization method, the ODE based EKF for 15 ms and the DAE based EKF for 125 ms. A key result here is that, while some gain in stability can be achieved by using a discretization method with better stability properties, the main gain is achieved by transforming the stiff ODE model into a DAE which results in a significant increase in step length without loss of stability.

It can be noted that the implementation of backward Euler utilizes a limited number of Newton iterations for finding the solution to the differential equation. The limited number of iterations ensures an upper execution time limit for each time step which is tuned for a 3 ms step length. From Table 1, studying the ODE columns for 3, 10, and 15 ms, it can be seen that the relative execution times between backward and forward Euler are 2.5, 3.2, and 3.9 respectively. The increasing trend comes from the fact that more Newton iterations are needed in backward Euler when the step length is increased. Another interesting remark is that, considering the total system with EKF and backward Euler, a benefit not utilized in the implementation is the fact that the EKF could provide the Jacobian for the Newton iterations, which would reduce the computational complexity further for backward Euler.

4.3 STEP LENGTH ANALYSIS

Figure 5 presents the normalized RMSE of the two observers as a function of step length for both forward and backward Euler. The normalization is with respect to the RMSE of the estimates of the ODE observer with 3 ms step length. The trend of the performance as a function of the step length is clear for all signals, i.e. the RMSE for the ODE observer increases more than the RMSE for the DAE observer as the step length increases. Even though the DAE observer estimates are not the best for all step lengths, for example W_c at 3 ms step length, the RMSE is equal or better than for the ODE for all step lengths from 10 ms and up, which represents a common timing loop in modern truck ECU:s. The reason for the better performance of the ODE observer for W_c and 3 ms step length originates from the change of model dynamics that was introduced when the intercooler pressure dynamics was removed, which also affected the throttle dependent dynamics, discussed in the last paragraph of Section 2.1.

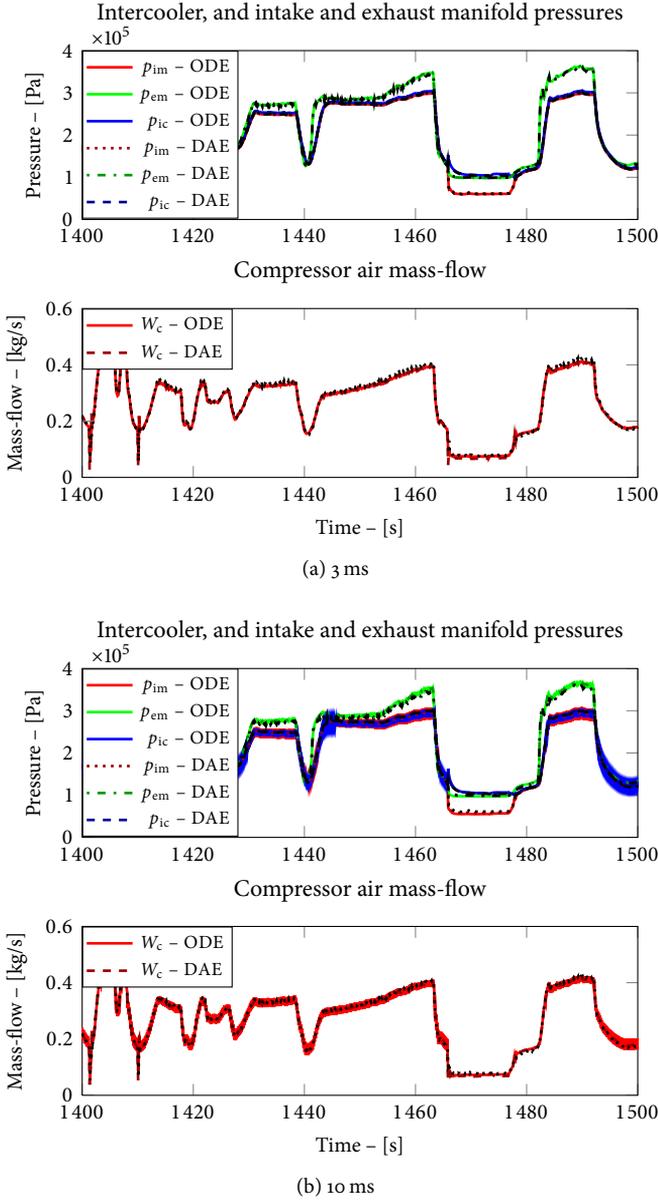


Figure 3: Estimates of EKF observers based on ODE and DAE models with 3 and 10 ms step lengths and using forward Euler for the prediction. Using the ODE based EKF and a step size of 10 ms, results in a high frequency estimation error, especially prominent for high pressures corresponding to open throttle operation. It can also be noted that also the ODE based EKF with 10 ms step length would predict the states well if a low pass filter was applied. Dotted black lines represent measurements.

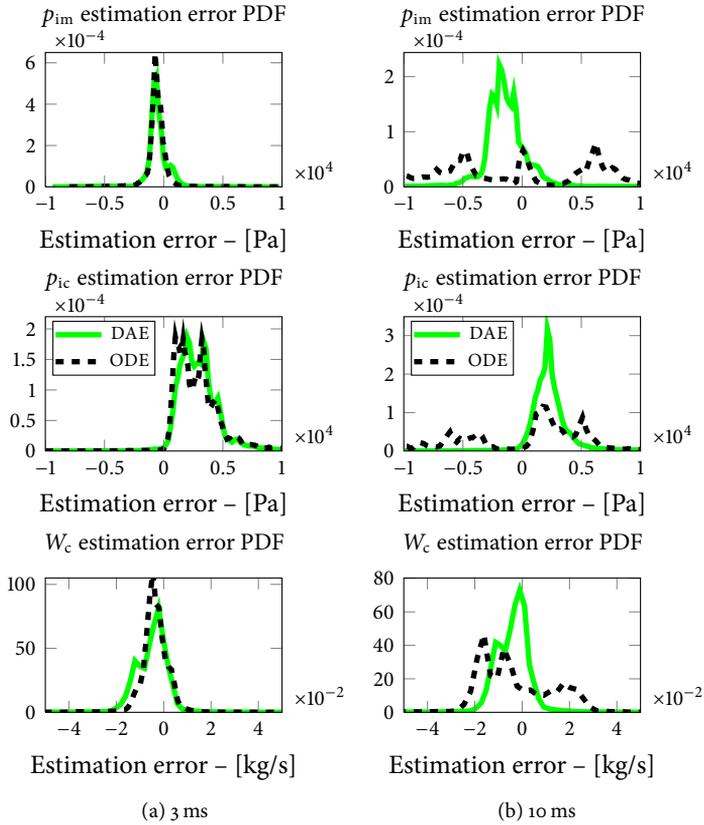
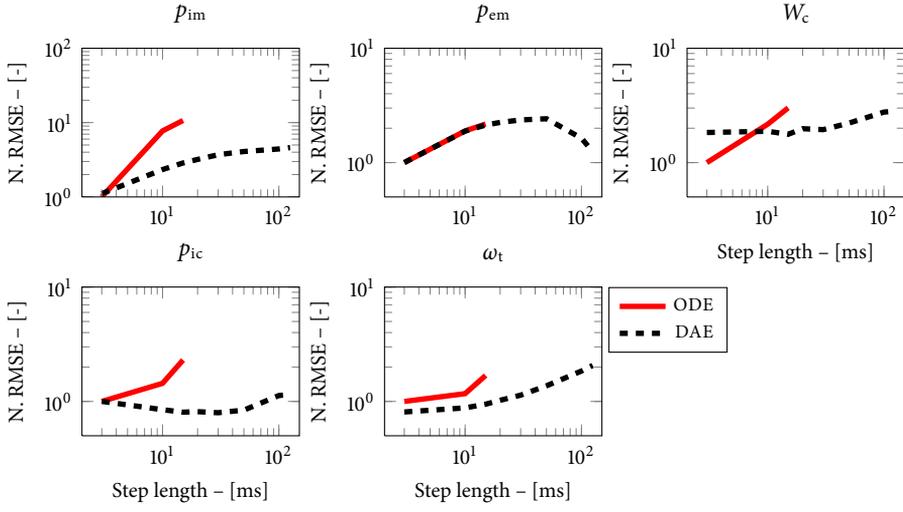
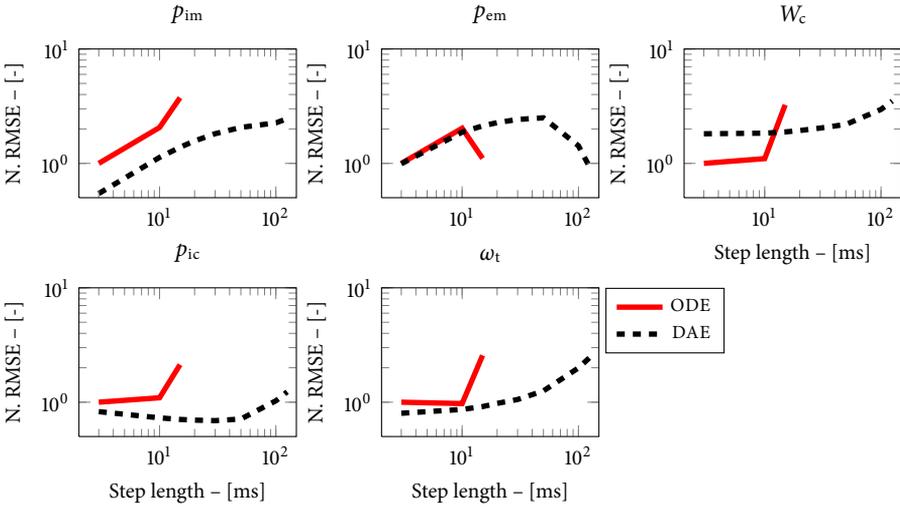


Figure 4: PDF of estimation errors for p_{im} , p_{ic} , and W_c from observers based on ODE and DAE models with different step lengths. The PDF:s of the estimation errors for simulation step lengths of 3 ms are similar for both the DAE and the ODE based EKF:s. For step lengths of 10 ms the ODE based EKF estimation error PDF:s are wider than those of the DAE based EKF which agrees with the noisy estimates of the ODE based EKF observed in Figure 3b.



(a) Forward Euler



(b) Backward Euler

Figure 5: Normalized RMSE as function of step length for DAE and ODE EKF:s. Both forward and backward Euler results are presented. The RMSE is normalized by the ODE RMSE with 3 ms step length for each variable. Presented discretization step lengths are $T_s = 3, 10, 15, 20, 30, 50, 100,$ and 125 ms, where the last five only applies to the DAE observers due to stability issues of the ODE observers. Note that the scale is logarithmic.

The performance of the DAE observer with a step length of 125 ms is comparable to the performance of the ODE observer with a step length of ~ 10 ms. So even if the DAE observer is computationally more demanding than the ODE observer in each time step the gain in step length results in a gain in total computational effort.

It can be noted that the solver used to compute the state prediction in the DAE based EKF is not optimized with respect to efficient evaluation and there exist several improvements that would reduce the computation time, see for example Brenan et al. (1989, Section 5.2). Table 2 contains the data of Figure 5a and shows the absolute values of the RMSE for the estimates. In particular it shows the relation of the RMSE between the estimated variables that is not shown in Figure 5. For example it shows, for the DAE observer, that no particular variable estimate distinguishes itself as better, or worse, than any other including p_{1c} . The data of Figure 5b is consistent with the observations made in Table 2 for Figure 5a and is omitted.

5 CONCLUSIONS

The benefit of using DAE based EKF:s instead of ODE:s in real-time applications with fixed discretization step lengths has been analyzed. The analysis is performed using an ODE model of a heavy duty Scania diesel engine with throttle, EGR and VGT, and a reduced DAE model obtained through removal of one of the volumes surrounding the throttle causing stiff model dynamics at wide open throttle operation. Criteria for observability for DAE:s are given and observability of both the DAE and ODE is established.

Using measurement data, the step length required to get equal performance of the two observers is investigated. It is shown that, even though the computational complexity for each time step of the DAE based EKF is higher than for the ODE based EKF, the possibility to use more than 10 times larger step lengths for the DAE, compared to the ODE, results in a more computationally efficient implementation with maintained estimation performance. The effect of discretization method for solving the differential equations is also studied. By comparing forward and backward Euler, representing an explicit and an implicit differential equation solver, it is shown that, even if there is a significant gain in increased discretization step length for backward Euler, the main gain is obtained by using a DAE based EKF.

Table 2: RMSE of DAE and ODE observers with forward Euler for different discretization step lengths.

T_s [ms]	P^{in}		P^{em}		P^{ic}		ω_t		W_c	
	DAE	ODE	DAE	ODE	DAE	ODE	DAE	ODE	DAE	ODE
3	$1.012 \cdot 10^3$	$9.143 \cdot 10^2$	$1.876 \cdot 10^3$	$1.873 \cdot 10^3$	$3.394 \cdot 10^3$	$3.389 \cdot 10^3$	$8.852 \cdot 10^1$	$1.094 \cdot 10^2$	$1.060 \cdot 10^{-2}$	$5.778 \cdot 10^{-3}$
10	$2.135 \cdot 10^3$	$7.070 \cdot 10^3$	$3.534 \cdot 10^3$	$3.546 \cdot 10^3$	$2.870 \cdot 10^3$	$4.880 \cdot 10^3$	$9.612 \cdot 10^1$	$1.280 \cdot 10^2$	$1.086 \cdot 10^{-2}$	$1.252 \cdot 10^{-2}$
15	$2.619 \cdot 10^3$	$9.774 \cdot 10^3$	$3.991 \cdot 10^3$	$4.080 \cdot 10^3$	$2.731 \cdot 10^3$	$7.800 \cdot 10^3$	$1.035 \cdot 10^2$	$1.845 \cdot 10^2$	$1.023 \cdot 10^{-2}$	$1.740 \cdot 10^{-2}$
20	$2.921 \cdot 10^3$	-	$4.209 \cdot 10^3$	-	$2.759 \cdot 10^3$	-	$1.120 \cdot 10^2$	-	$1.144 \cdot 10^{-2}$	-
30	$3.380 \cdot 10^3$	-	$4.425 \cdot 10^3$	-	$2.701 \cdot 10^3$	-	$1.241 \cdot 10^2$	-	$1.123 \cdot 10^{-2}$	-
50	$3.734 \cdot 10^3$	-	$4.531 \cdot 10^3$	-	$2.846 \cdot 10^3$	-	$1.498 \cdot 10^2$	-	$1.281 \cdot 10^{-2}$	-
100	$4.041 \cdot 10^3$	-	$3.063 \cdot 10^3$	-	$3.817 \cdot 10^3$	-	$2.018 \cdot 10^2$	-	$1.599 \cdot 10^{-2}$	-
125	$4.223 \cdot 10^3$	-	$2.336 \cdot 10^3$	-	$3.902 \cdot 10^3$	-	$2.258 \cdot 10^2$	-	$1.623 \cdot 10^{-2}$	-

REFERENCES

- Benedikt Alt, Jan Peter Blath, Ferdinand Svaricek, and Matthias Schultalbers. Multiple sliding surface control of idle engine speed and torque reserve with dead start assist control. *Industrial Electronics, IEEE Transactions on*, 56(9):3580–3592, 2009. ISSN 0278-0046. doi:10.1109/TIE.2009.2021593.
- Andrea Giovanni Beccuti, Sébastien Mariéthoz, Sébastien Cliquennois, Shu Wang, and Manfred Morari. Explicit model predictive control of DC–DC switched-mode power supplies with extended kalman filtering. *Industrial Electronics, IEEE Transactions on*, 56(6):1864–1874, 2009. ISSN 0278-0046. doi:10.1109/TIE.2009.2015748.
- Victor M. Becerra, Peter D. Roberts, and Graham W. Griffiths. Applying the extended kalman filter to systems described by nonlinear differential-algebraic equations. *Control Engineering Practice*, 9(3):267–281, 2001. ISSN 0967-0661. doi:10.1016/S0967-0661(00)00110-6.
- K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. siam, 1989.
- L. Dai. *Singular Control Systems*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1989. ISBN 0387507248.
- Economic Commission for Europe – Inland Transport Committee. Regulation no 49 of the economic commission for europe of the united nations (UN/ECE). Official Journal of the European Union, 2010.
- Lars Eriksson. Modeling and control of turbocharged SI and DI engines. *Oil & Gas Science and Technology - Rev. IFP*, 62(4):523–538, 2007. doi:10.2516/ogst:2007042.
- Ingo Friedrich, Chia-Shang Liu, and Dale Oehlerking. Coordinated EGR-rate model-based controls of turbocharged diesel engines via an intake throttle and an EGR valve. In *Vehicle Power and Propulsion Conference, 2009. VPPC '09. IEEE*, pages 340–347, 2009. doi:10.1109/VPPC.2009.5289828.
- Sergio García-Nieto, Miguel Martínez, Xavier Blasco, and Javier Sanchis. Non-linear predictive control based on local model networks for air management in diesel engines. *Control Engineering Practice*, 16(12):1399–1413, December 2008. doi:10.1016/j.conengprac.2008.03.010.
- Markus Gerdin, Thomas B. Schön, Torkel Glad, Fredrik Gustafsson, and Lennart Ljung. On parameter and state estimation for linear differential-algebraic equations. *Automatica*, 43(3):416–425, 2007. ISSN 0005-1098. doi:10.1016/j.automatica.2006.09.016.
- John B. Heywood. *Internal Combustion Engine Fundamentals*. McGraw-Hill, Inc., 1988.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. 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, 2008.

- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*, 17(3):408–417, 2009. doi:10.1016/j.conengprac.2008.09.004.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Model based engine map adaptation using EKF. In *6th IFAC Symposium on Advances in Automotive Control*, Munich, Germany, 2010. doi:10.3182/20100712-3-DE-2013.00051.
- Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, April 1970. ISBN 0123815509.
- Thomas Kailath, Ali H. Sayed, and Babak Hassibi. *Linear Estimation*. Prentice-Hall, Inc, Upper Saddle River, New Jersey 07458, 2 edition, 2000.
- Rudolf Emil Kalman, Yu-Chi Ho, and Kumpati S. Narendra. Controllability of linear dynamical systems. *Contributions to Differential Equations*, 1, 1963.
- Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- Ernest Bruce Lee and Lawrence Markus. *Foundations of Optimal Control Theory*. John Wiley & Sons, Inc, New York, 1968.
- Paolo Lino, Bruno Maione, and Claudio Amorese. Modelling and predictive control of a new injection system for compressed natural gas engines. *Control Engineering Practice*, 16(10):1216–1230, October 2008. doi:10.1016/j.conengprac.2008.01.008.
- Philip Losse and Volker Ludwig Mehrmann. Controllability and observability of second order descriptor systems, 2008.
- Ravi Kumar Mandela, Raghunathan Rengaswamy, Shankar Narasimhan, and Lakshmi N. Sridhar. Recursive state estimation techniques for nonlinear differential algebraic systems. *Chemical Engineering Science*, 65(16):4548–4556, 2010. ISSN 0009-2509. doi:10.1016/j.ces.2010.04.020.
- Volker Ludwig Mehrmann and Tatjana Stykel. Descriptor systems: a general mathematical framework for modeling, simulation and control. Technical Report Report 292-2005, DFG Research center Matheon, Berlin, 2005.
- Luciano Loporini Menegaldo, Gustavo Ferreira, Melquisedec Francisco Santos, and Rodrigo Siqueira Guerato. Development and navigation of a mobile robot for floating production storage and offloading ship hull inspection. *Industrial Electronics, IEEE Transactions on*, 56(9):3717–3722, 2009. ISSN 0278-0046. doi:10.1109/TIE.2009.2025716.
- Martin Müller, Elbert Hendricks, and Spencer C. Sorenson. Mean value modelling of turbocharged spark ignition engines. SAE Technical paper 980784, 1998. doi:10.4271/980784.

Ian Morgan and Honghai Liu. Predicting future states with n -dimensional markov chains for fault diagnosis. *Industrial Electronics, IEEE Transactions on*, 56(5):1774–1781, 2009. ISSN 0278-0046. doi:10.1109/TIE.2008.2011306.

Markus Reichhartinger and Martin Horn. Application of higher order sliding-mode concepts to a throttle actuator for gasoline engines. *Industrial Electronics, IEEE Transactions on*, 56(9):3322–3329, 2009. ISSN 0278-0046. doi:10.1109/TIE.2009.2026382.

Nadia Salvatore, Andrea Caponio, Ferrante Neri, Silvio Stasi, and Giuseppe Leonado Cascella. Optimization of delayed-state kalman-filter-based algorithm via differential evolution for sensorless control of induction motors. *Industrial Electronics, IEEE Transactions on*, 57(1):385–394, 2010. ISSN 0278-0046. doi:10.1109/TIE.2009.2033489.

Johan Wahlström and Lars Eriksson. Modeling of a diesel engine with intake throttle, VGT, and EGR. Technical Report LiTH-R-2976, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, 2010.

Johan Wahlström, Lars Eriksson, and Lars Nielsen. EGR-VGT control and tuning for pumping work minimization and emission control. *IEEE Transactions on Control Systems Technology*, 18(4):993–1003, 2010.

Seong-hoon Peter Won, Wael William Melek, and Farid Golnaraghi. A kalman/particle filter-based position and orientation estimation method using a position sensor/inertial measurement unit hybrid system. *Industrial Electronics, IEEE Transactions on*, 57(5):1787–1798, 2010. ISSN 0278-0046. doi:10.1109/TIE.2009.2032431.

Elizabeth L. Yip and Richard F. Sincovec. Solvability, controllability, and observability of continuous descriptor systems. *IEEE Transactions on Automatic Control*, 26(3):702–707, jun. 1981. ISSN 0018-9286. doi:10.1109/TAC.1981.1102699.

A ENGINE MODEL AND DATA

The default model, on which the method is applied, is a 5:th order non-linear state space model of a six cylinder Scania diesel engine with EGR, VGT, and intake throttle. The model states and inputs are presented in Tables 3 and 4, respectively. It is based on a model originally developed in Wahlström et al. (2010) and extended with intake throttle in Wahlström and Eriksson (2010). The modifications are that the states for the intake and exhaust manifold oxygen concentrations and actuator dynamics are removed.

The data is collected in an engine test cell at Scania CV AB in Södertälje, Sweden. The data is from a six cylinder Scania diesel engine with EGR, VGT, and throttle and was collected during a world harmonized transient cycle (WHTC) Economic Commission for Europe – Inland Transport Committee (2010). The sensor signals used are presented in Table 5. All signals are collected with a sampling rate of 100 Hz. It can be noted that the turbo speed sensor is unable to measure rotational speeds below 20 000 rpm ($\sim 2\,094$ rad/s) and those measurements are therefore excluded in the turbo speed measure.

Table 3: Model states.

State	Description	Unit
T_{em}	Exhaust manifold temperature	K
ω_t	Turbo speed	rad/s
p_{em}	Exhaust manifold pressure	Pa
p_{im}	Intake manifold pressure	Pa
p_{ic}	Intercooler pressure	Pa

Table 4: Model inputs.

Input	Description	Unit
n_e	Engine speed [†]	rpm
u_δ	Injected amount fuel	mg/cycle
u_{egr}	EGR valve position [‡]	%
u_{vgt}	VGT position [‡]	%
u_{th}	Intake throttle position [‡]	%
T_{amb}	Ambient temperature [†]	K

[†] Parametrization inputs.

[‡] 0 – closed, 100 – open

Table 5: Test cell measurements and model outputs

Output	Description	Unit
p_{im}	Intake manifold pressure	Pa
p_{em}	Exhaust manifold pressure	Pa
ω_t	Turbo speed	rad/s
p_{ic}	Intercooler pressure [†]	Pa
W_c	Compressor air mass-flow	kg/s

[†] Extra evaluation sensor not used for observer feedback.

Bias Reduction in DAE Estimators by Model
Augmentation: Observability Analysis and
Experimental Evaluation[☆]

E

[☆]Submitted to *Conference on Decision and Control*, Orlando, 2011.

Bias Reduction in DAE Estimators by Model Augmentation: Observability Analysis and Experimental Evaluation

Erik Höckerdal, Erik Frisk, and Lars Eriksson

*Vehicular Systems, Department of Electrical Engineering,
Linköping University, S-581 83 Linköping,
Sweden.*

ABSTRACT

A method for bias compensation in model based estimation utilizing model augmentation is developed. Based on a default model, that suffers from stationary errors, and measurements from the system a low order augmentation is estimated. The method handles models described by differential algebraic equations and the main contributions are necessary and sufficient conditions for the preservation of the observability properties of the default model during the augmentation.

A characterization of possible augmentations found through the estimation, showing the benefits of adding extra sensors during the design, is included. This enables reduction of estimation errors also in states not used for feedback, which is not possible with for example PI-observers. Beside the estimated augmentation the method handles user provided augmentations, found through e.g. physical knowledge of the system.

The method is evaluated on a nonlinear engine model where its ability to incorporate information from additional sensors during the augmentation estimation is clearly illustrated. By applying the method the mean relative estimation error for the exhaust manifold pressure is reduced by 55 %.

1 INTRODUCTION

Accurate information of the internal state of systems is important for fulfilling the increasing demands on control accuracy and fault detection. At the same time the overall cost has to be kept as low as possible, which often implies that it is insufficient to rely upon only physical sensors. As a consequence, model based estimation has attracted attention.

A common situation is that models based on first principles exist and it is desirable to be able to use them for observer design. However, these models often have undesired properties that prevent them from being directly applicable for estimation in embedded systems, such as engine control units (ECU). One such deficiency is that, even if the system dynamics is well described, the models can suffer from stationary errors, or biases (Höckerdal et al., 2008). Figure 1 illustrates this with experimental data from an engine where the model captures dynamics well while there is a bias in the estimate.

The number of models described by differential algebraic equations (DAE) are increasing, partly due to modern modeling tools such as DYMOLA and SIMSCAPE that often deliver DAE models and since DAE:s are a way of describing systems with both fast and slow dynamics. The latter arise when approximating fast dynamics with algebraic constraints, i.e. instantaneous dynamics. DAE applications range from electrochemical and reactive distillation processes (Mandela et al., 2010) to combustion engines (Lino et al., 2008; García-Nieto et al., 2008; Höckerdal et al., 2011).

The objective is to develop a method that enables usage of biased default models for estimation with reduced estimation bias where the reduction is achieved using model augmentation. Central in the method is preservation of the observability properties of the, biased, default model. The method is as an extension of the method developed for ordinary differential equations (ODE) in Höckerdal et al. (2009) to systems described by DAE:s.

2 PROBLEM FORMULATION & SOLUTION OUTLINE

Designing an observer based on a model that predicts the system dynamics well but suffers from stationary errors will result in biased estimates (Höckerdal et al., 2008). Common ways to reduce bias in observers are; i) to use so called PI-observers (Söffker et al., 1995; Koenig and Mammar, 2002) that introduce integrators for the feedback signals, ii) by physical knowledge introduce extra states to compensate for known model deficiencies (Andersson and Eriksson, 2001), or iii) to estimate a minimal augmentation that reduces the bias (Höckerdal et al., 2009).

The objective is to reduce bias in estimates for observers based on DAE:s in a systematic manner without involving extensive modeling efforts in a similar way as for ODE:s in Höckerdal et al. (2009). The starting point is a default, semi-explicit, DAE model

$$\dot{x} = f(\bar{x}, u) \quad (1a)$$

$$0 = g(\bar{x}, u) \quad (1b)$$

$$y = h(\bar{x}, u), \quad (1c)$$

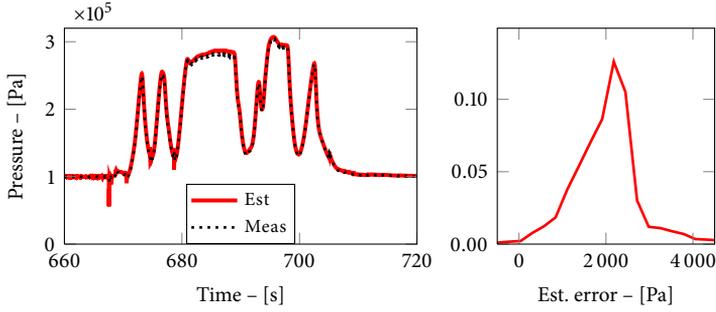


Figure 1: Biased estimates for intercooler pressure in a heavy duty diesel engine. **Left:** Measured and estimated intercooler pressure. **Right:** Normalized histogram of the corresponding estimation error which clearly is biased.

where $\bar{x} = \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^{n_x}$, x and z denote differential and algebraic variables respectively, and measurements, (u, y) , $y \in \mathbb{R}^{n_y}$, $u \in \mathbb{R}^{n_u}$. The method generally handles DAE:s of differential index 1, i.e. $\frac{\partial \begin{pmatrix} g \\ h \end{pmatrix}}{\partial z}$ has full column rank. Henceforth, and without loss of generality, it is assumed that $\frac{\partial g}{\partial z}$ has full column rank. It is assumed that the model described by (1) captures the system dynamics well but suffers from stationary errors, i.e. the model is biased. A possible solution is obtained by considering the bias as an offset error during stationary operation of the system.

A way to describe stationary errors is to adjust the default states of the system according to $(\bar{x} - A_q q)$, where q is a bias state and A_q is its effect on the default states. By introducing q as new states with constant derivatives, $\dot{q} = 0$, and driving noise, it is possible to describe a bias that varies with operating point. It is desirable to have as few bias states q as possible and the method describes a way to estimate a low order augmentation A_q from measurement data.

The idea is then to design an observer based on the augmented model,

$$\dot{x} = f(\bar{x} - A_q q, u) \quad (2a)$$

$$0 = g(\bar{x} - A_q q, u) \quad (2b)$$

$$\dot{q} = 0 \quad (2c)$$

$$y = h(\bar{x}, u), \quad (2d)$$

where $q \in \mathbb{R}^{n_q}$, denoted Aug., which will have better stationary estimation accuracy than an observer based on the default model. The observer design used is the discrete time extended Kalman filter (EKF) (Jazwinski, 1970) and specifically the for DAE:s modified version presented in Höckerdal et al. (2011).

3 EKF FOR DAE SYSTEMS

The EKF algorithm used originates from Mandela et al. (2010), and was slightly modified in Höckerdal et al. (2011). It is similar to the standard EKF; it consists of two steps, prediction and measurement update. The differential part of the DAE fits nicely into the standard EKF. To enable use of the standard EKF for both differential and algebraic states the linearized algebraic subsystem is differentiated once

$$\begin{aligned} \dot{x} &= A_t x + B_t z & \Rightarrow & \dot{x} = A_t x + B_t z \\ 0 &= C_t x + D_t z & \Rightarrow & \dot{z} = -D_t^{-1} C_t \dot{x}, \end{aligned} \quad (3)$$

where

$$\begin{pmatrix} A_t & B_t \\ C_t & D_t \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial z} \end{pmatrix}.$$

To save space the input u has been omitted in (3). Note that inputs with discontinuous first-order derivatives are not allowed in the algebraic submodel (1b), due to the differentiation in (3). Normally this is not a problem since a low-pass filter on the inputs can be introduced.

The algorithm used can be summarized as

Prediction:

State prediction is done as follows

$$\begin{aligned} \hat{x}_{t+1|t} &= \hat{x}_{t|t} + T_s f(\hat{x}_{t+1|t}, \hat{z}_{t+1|t}, u_{t+1}) \\ 0 &= g(\hat{x}_{t+1|t}, \hat{z}_{t+1|t}) \end{aligned} \Rightarrow \begin{pmatrix} \hat{x}_{t+1|t} \\ \hat{z}_{t+1|t} \end{pmatrix},$$

where the implication indicates solving $\hat{x}_{t+1|t}$ and $\hat{z}_{t+1|t}$ from the system of equations defined by (1a) and (1b) using backward Euler.

Covariance matrix prediction

$$\bar{P}_{t+1|t} = \bar{A}_{t|t} \bar{P}_{t|t} \bar{A}_{t|t}^T + \bar{G}_{t|t} Q_{t+1} \bar{G}_{t|t}^T,$$

where

$$\bar{A}_{t|t} = I + T_s \begin{pmatrix} A_{t|t} & B_{t|t} \\ -D_{t|t}^{-1} C_{t|t} A_{t|t} & -D_{t|t}^{-1} C_{t|t} B_{t|t} \end{pmatrix}, \quad (4)$$

T_s denotes discretization step length, and

$$\bar{G}_{t|t} = \begin{pmatrix} I \\ -D_{t|t}^{-1} C_{t|t} \end{pmatrix}.$$

$\bar{G}_{t|t}$ is a result of the differentiation in (3) and implies that system noise is present in the differential variables of the DAE only, i.e. x -variables.

Measurement update:

Innovation covariance

$$S_{t+1} = \tilde{H}_{t+1|t} \tilde{P}_{t+1|t} \tilde{H}_{t+1|t}^T + R_{t+1}$$

Kalman gain

$$\tilde{K}_{t+1} = \tilde{P}_{t+1|t} \tilde{H}_{t+1|t}^T S_{t+1}^{-1}$$

State

$$\begin{aligned} \hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + \tilde{K}_{t+1} (y_{t+1} - h(\hat{x}_{t+1|t}, u_{t+1})) \\ 0 &= g(\hat{x}_{t+1|t+1}, \hat{z}_{t+1|t+1}) \Rightarrow \hat{z}_{t+1|t+1} \end{aligned}$$

Covariance matrix

$$\tilde{P}_{t+1|t+1} = \tilde{P}_{t+1|t} - \tilde{P}_{t+1|t} \tilde{H}_{t+1|t}^T S_{t+1}^{-1} \tilde{H}_{t+1|t} \tilde{P}_{t+1|t}$$

4 OBSERVABILITY OF THE AUGMENTED MODEL

In estimation the concept of observability, or detectability, is used to analyze the estimators' ability to provide consistent estimates that asymptotically converge to the true states. This section addresses the observability of the augmented model (2) given the observability properties of the default model (1), i.e., which A_q 's are allowed, given that the observability of the default model must not be compromised. Since the observability of a linearization in a stationary operating point is a sufficient condition for local observability of the nonlinear system (Lee and Markus, 1968, Theorem 6.4), the observability analysis is conducted on model linearizations.

4.1 DAE OBSERVABILITY

Before presenting the main results an overview of definitions and theorems used to assess observability for systems described by differential algebraic equations is given.

For DAE:s there exist several concepts of observability, i.e. complete observability, observability within the reachable set, and impulse observability (Mehrmann and Stykel, 2005; Dai, 1989; Losse and Mehrmann, 2008).

Definition 4.1 (C-observable) *The system*

$$\begin{aligned} F(t, x, \dot{x}, u) &= 0 \\ y - H(t, x) &= 0, \end{aligned} \tag{5}$$

is completely observable if the zero output of the descriptor system with $u = 0$ implies that this system has the trivial solution $x = 0$ only.

Generally, descriptor systems are not \mathcal{C} -observable since they contain algebraic constraints that force the solution and output, onto a specific manifold. For this reason the observability within the reachable set, i.e. \mathcal{R} -observability, is introduced. This concept needs an appropriate projection of the dynamical part of the system, sometimes referred to as the slow sub-system, onto a manifold defined by the algebraic equations, or fast sub-system. For linear time-invariant systems

$$\begin{aligned} E\dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned} \tag{6}$$

that are regular, i.e. that $\det(\alpha E - \beta A) \neq 0$ for some $(\alpha, \beta) \in \mathbb{C}^2$, this can be expressed in terms of the system matrices E , A , and C .

Definition 4.2 (\mathcal{R} -observable) *The system (6) is called observable within the reachable set if the zero output of the descriptor system with $u = 0$ implies that all solutions of this system satisfy $P_r x = 0$, where P_r denotes the projection onto the right deflating sub-space corresponding to the finite eigenvalues of $\lambda E - A$.*

A last observability definition concerns the problem that arises when descriptor systems that are not strangeness-free (Kunkel and Mehrmann, 2006), i.e. have differentiation index larger than 1, are driven by inputs that are only piecewise continuous (Mehrmann and Stykel, 2005). Then, since the solution may depend on the derivative of the input, any classical solution may not exist.

Definition 4.3 (\mathcal{I} -observable) *The system (6) is called impulse observable if the output is continuous when a step is used as input.*

In the linear time-invariant case, (6), these concepts can be characterized algebraically in terms of E , A , and C , see for example Dai (1989) for theorems and proofs.

Theorem 4.1 *The system (6) is*

1. \mathcal{C} -observable if and only if

$$\begin{pmatrix} \alpha E - \beta A \\ C \end{pmatrix}$$

has full column rank for all $(\alpha, \beta) \in \mathbb{C}^2 \setminus \{(0, 0)\}$.

2. \mathcal{R} -observable if and only if

$$\begin{pmatrix} \lambda E - A \\ C \end{pmatrix}$$

has full column rank for all $\lambda \in \mathbb{C}$.

3. \mathcal{I} -observable if and only if

$$\begin{pmatrix} E \\ K_{E^T}^T A \\ C \end{pmatrix}$$

has full column rank, where the rows of K_{E^T} span $\text{Ker } E^T$.

Note that the system (6) is \mathcal{C} -observable if and only if it is \mathcal{R} -observable and $\begin{pmatrix} E \\ C \end{pmatrix}$ has full column rank. Furthermore, a descriptor system (6) with a regular pencil $\lambda E - A$ with differentiation index less than two is \mathcal{I} -observable (Mehrmann and Stykel, 2005; Dai, 1989).

4.2 POSSIBLE AUGMENTATIONS

Based on the theory presented in Section 4.1, necessary and sufficient conditions for preserving the observability properties of the default model throughout the augmentation, are now given and proven.

Using Theorem 4.1 it is possible to characterize the allowable model augmentations for a descriptor system,

$$\begin{aligned} E\dot{x} &= A(x - A_q q) + Bu \\ \dot{q} &= 0 \\ y &= Cx, \end{aligned} \quad (7)$$

i.e. augmentations A_q that preserve the observability of the default model (6).

Theorem 4.2 *The observability of (6) is preserved during model augmentation according to (7) if and only if*

$$A \begin{pmatrix} A_q & N_C \end{pmatrix},$$

has full column rank, where the columns of N_C span $\text{Ker } C$.

Worth noting is that even though there are three observability concepts, there is only one requirement. This due to the structure of the augmented system with $\tilde{E} = \begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix}$ which gives full column rank of the augmentation subsystem.

Proof The different observability properties of (7) are preserved if and only if $x = 0$, $q = 0$ is the only solution to the corresponding algebraic conditions in Theorem 4.1.

Rewriting the augmented system according to

$$\begin{aligned} \underbrace{\begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix}}_{\tilde{E}} \underbrace{\begin{pmatrix} \dot{x} \\ \dot{q} \end{pmatrix}}_{\dot{\tilde{x}}} &= \underbrace{\begin{pmatrix} A & -AA_q \\ 0 & 0 \end{pmatrix}}_{\tilde{A}} \underbrace{\begin{pmatrix} x \\ q \end{pmatrix}}_{\tilde{x}} + \underbrace{\begin{pmatrix} B \\ 0 \end{pmatrix}}_{\tilde{B}} \\ y &= \underbrace{\begin{pmatrix} C & 0 \end{pmatrix}}_{\tilde{C}} \begin{pmatrix} x \\ q \end{pmatrix} \end{aligned} \quad (8)$$

and applying Theorem 4.1 to the augmented system (8):

\mathcal{R} -observability is preserved if and only if $x = 0$, $q = 0$ is the only solution to

$$(\lambda E - A)x + AA_q q = 0 \quad (9a)$$

$$\lambda I q = 0 \quad (9b)$$

$$y = Cx \quad (9c)$$

for all $\lambda \in \mathbb{C}$. For $\lambda \neq 0$ it is immediate from (9b) that $q = 0$. Then the assumption that (6) is \mathcal{R} -observable together with (9a), (9c) and Theorem 4.1 gives that $x = 0$. Thus only $\lambda = 0$ needs further investigation.

For $\lambda = 0$ in (9) the augmented model is \mathcal{R} -observable if and only if $x = 0, q = 0$ is the only solution to

$$-Ax + AA_q q = 0, \quad (10a)$$

$$Cx = 0. \quad (10b)$$

Let the columns of N_C be a basis for $\text{Ker } C$, then, from (10b), $x = N_C \xi$ for some arbitrary ξ and \mathcal{R} -observability is equivalent to $q = 0, \xi = 0$ being the only solution to

$$-A(N_C \xi - A_q q) = 0,$$

which is equivalent to the matrix

$$A \begin{pmatrix} N_C & A_q \end{pmatrix}$$

having full column rank.

\mathcal{C} -observability is equivalent to \mathcal{R} -observability with the additional requirement that

$$\begin{pmatrix} \bar{E} \\ \bar{C} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & I \\ C & 0 \end{pmatrix} \quad (11)$$

has full column rank (Dai, 1989). The fact that the rank of a block diagonal matrix is equal to the sum of the ranks of the blocks (Lancaster and Tismenetsky, 1984) and the assumption that

$$\begin{pmatrix} E \\ C \end{pmatrix}$$

has full column rank now gives that also \mathcal{C} -observability is preserved if and only if

$$A \begin{pmatrix} N_C & A_q \end{pmatrix}$$

has full column rank.

\mathcal{I} -observability is preserved if and only if

$$\begin{pmatrix} \bar{E} \\ K_{\bar{E}^T}^T \bar{A} \\ \bar{C} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & I \\ K_{E^T}^T A & -K_{E^T}^T A A_q \\ C & 0 \end{pmatrix}$$

has full column rank. The Kernel of \bar{E} is equal to the Kernel of E , padded with zeros to the dimension of \bar{E} . Again using that the rank of a block diagonal matrix is the sum of the ranks of the blocks and the assumption that

$$\begin{pmatrix} E \\ K_{E^T}^T A \\ C \end{pmatrix}$$

has full column rank the preservation of the \mathcal{I} -observability follows. \square

5 AUGMENTATION ESTIMATION

After establishing a theoretical basis for all possible augmentations that preserve the default model observability properties, the method of estimating a low order model augmentation is presented.

The augmentation estimation procedure is divided into two steps; i) estimate the bias, and ii) compute a basis for the bias space. The bias estimation is performed by augmenting the default model fully, that is introduce as many bias states as possible without compromising the observability criteria,

$$\text{rank } A \begin{pmatrix} N_C & A_q \end{pmatrix} \leq n_{\hat{x}} - n_y + n_q \leq n_{\hat{x}},$$

found in Section 4.2. This means that the augmentation A_q can have at most as many columns as there are measurements, $n_q \leq n_y$. Also, from Theorem 4.2, these columns have to be linearly independent of the columns of N_C and can not lie in $\text{Ker } A$.

A simple way to construct such an augmentation is to use C^\dagger , where \dagger denotes the Moore-Penrose inverse (Lancaster and Tismenetsky, 1984, Exercise 5.1.7 and Proposition 12.8.2), and exclude the columns that become zero when multiplied by A from the left. Based on the fully augmented model, an observer that estimates both \hat{x} and \hat{q} , enabling the computation of bias estimates,

$$\beta_t = C^\dagger \hat{q}_t,$$

can be constructed.

Central in the bias estimation is that the entire operating region of the system is spanned, otherwise the estimated bias samples might not represent the actual bias for all operating points.

Given bias estimates, a basis for the bias is computed using a singular value decomposition (SVD) (Lancaster and Tismenetsky, 1984) of the bias estimates. To increase the computational efficiency and allow easier weighting of the biases from different stationary operating points, the average of the bias samples from each operating point is computed. These averaged bias samples from N operating points are collected

$$\tilde{\beta}^{n_{\hat{x}} \times N} = \begin{pmatrix} w^1 \beta^1 & \dots & w^N \beta^N \end{pmatrix}, \quad \sum_1^N w^i = 1,$$

for which the corresponding SVD is computed

$$\tilde{\beta} = U \Sigma V^*, \tag{12}$$

where β^i indicates the averaged bias in operating point i , and w^i is the corresponding importance weight. In (12) the columns of U contains orthonormal vectors spanning the bias space and Σ the corresponding singular values. The augmentation dimension can be found by analyzing the singular values and pick out the most significant ones. Then \hat{A}_q is constructed by assembling the corresponding columns of U .

5.1 AUGMENTATION PROPERTIES

In Section 4.2 the set of possible augmentations is analyzed, and it is apparent that the measurement equation plays a central role in which augmentations that are possible to find. This is also given by the bias estimation in Section 5, i.e. that $\beta = C^\dagger \hat{q}$.

From an engineering perspective this is interesting since it means that it is possible to, temporarily augment the measurement equation in for example a development environment, to increase the set of possible augmentations. Increasing the set of possible augmentations like this gives the possibility reduce bias also in states not used for feedback in the final application.

Also note that, even though the main idea with the method is to estimate an augmentation using system measurements, it is possible to provide an augmentation found through physical knowledge or engineering intuition, as long as it fulfills Theorem 4.2.

6 EXPERIMENTAL EVALUATION

The method is evaluated on a heavy duty Scania diesel engine with exhaust gas recirculation (EGR), variable geometry turbocharger (VGT), and intake manifold throttle. The evaluation is based on experimental data collected in an engine test cell. The model used, presented with states, mass-flows, inputs, and outputs in Figure 2, is the same as in Höckerdal et al. (2011), originally developed in Wahlström and Eriksson (2010). The modifications are twofold; i) removal of actuator dynamics, and ii) elimination of the intercooler pressure state with fast dynamics, both with the aim to improve the computational efficiency of the resulting EKF. By the latter modification the original model, described by ODE:s, is transformed into a system of DAE:s.

Designing a standard EKF on the default model directly gives biased estimates which is obvious from Figure 4, where the estimation error probability density functions (PDF) for all states are biased, i.e., the solid PDF:s are not centered at zero. The objective is to use the proposed method from Section 5 to reduce the estimation bias of all system states except the exhaust manifold temperature, using only measurements of intake manifold pressure, intercooler pressure, and turbine speed. The reason for not including T_{em} in the evaluation is that there is no reference measurement for that state available.

6.1 AUGMENTATION ESTIMATION

The model augmentation, A_q , is computed using a measurement sequence containing a large number of different stationary operating points, spanning the whole operating region of the engine. In the experimental environment of the engine test cell, the sensor setup is larger than in the intended customer application which is used to improve the augmentation. That is, extra sensors during the augmentation estimation allows bias compensation of also non-measured states, recall the discussion in Section 5.1. In the studied example the application includes measurements of the states ω_t , p_{im} , and p_{ic} while the experimental setup also allows measurement of p_{em} , which is utilized in the augmentation estimation.

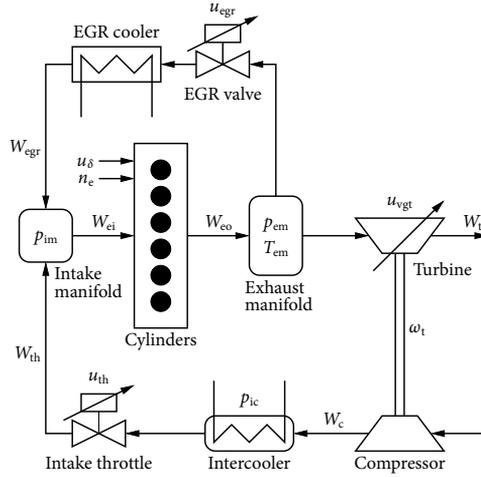


Figure 2: Schematics of the diesel engine used in the evaluation, showing the differential states (p_{im} , p_{em} , T_{em} , and ω_t), algebraic state (p_{ic}), inputs (u_{egr} , u_{vgt} , u_{δ} , u_{th} , and n_e), and flows between components (W_c , W_{th} , W_{egr} , W_{ei} , W_{eo} , and W_t).

Bias estimates, estimated according to Section 5, from 42 stationary operating points is used to find a low order augmentation A_q . In this case each operating point is weighted equally, i.e. $w^i = \frac{1}{N}$ since no information of a suitable distribution is available. The resulting singular values and vectors used to determine the augmentation are presented in Figure 3. The construction of $A_{q,i}$, where i indicates augmentation dimension, is done by assembling the i singular vectors that have the largest singular values, e.g.,

$$A_{q,3} = \begin{pmatrix} -0.013 & -0.023 & 0.071 \\ -0.555 & -0.485 & -0.676 \\ -0.431 & -0.528 & 0.729 \\ -0.712 & 0.697 & 0.084 \end{pmatrix}.$$

Three augmented observers, $A_{q,1}$, $A_{q,2}$, and $A_{q,3}$, are designed, where the maximal dimension, i.e. $\dim A_{q,3} = 3$, is limited by the three dimensional measurement equation in the EKF.

6.2 ESTIMATION PERFORMANCE

The estimation performance is evaluated by comparing the observers with different augmentation dimension and an observer based on the default model using data from a WHTC (Economic Commission for Europe – Inland Transport Committee, 2010). As performance measures the mean relative error, and the estimation error probability density function, are used. Together these measures capture both estimation bias and variance. The results are presented in Table 1 and Figure 4.

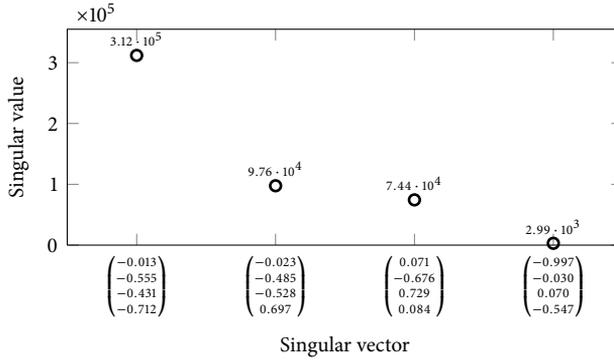


Figure 3: SVD singular values presented with the corresponding singular vector used for determining the augmentation, A_q . The vector components, from top to bottom, correspond to ω_t , p_{em} , p_{im} , and p_{ic} .

Table 1: Mean rel. error for the default and augmented observers. All with feedback from p_{im} , ω_t , and p_{ic} only.

Outputs	Mean relative error – [%]			
	Def.	$A_{q,1}$	$A_{q,2}$	$A_{q,3}$
ω_t	-2.1	0.1	0.3	0.5
p_{em}	6.8	5.6	8.5	3.0
p_{im}	-1.6	-1.5	-0.1	-0.2
p_{ic}	0.1	0.06	-0.09	0.07

Figure 4 shows that with a one dimensional augmentation, $A_{q,1}$, a significant reduction of the estimation error is achieved for the states where feedback is available, i.e. ω_t , p_{im} , and p_{ic} . A further indication of the performance of the augmented observer is given by the estimation error of the observer output that is not explicitly used for feedback, i.e. p_{em} . Table 1 clearly shows that in this case a mean estimation error reduction of approximately 55 % is achieved for p_{em} using a three dimensional augmentation, $A_{q,3}$. This shows a clear advantage of the proposed method compared to for example normal PI-observers (Söffker et al., 1995) that have integrators affecting the feedback variables only, i.e. the methods ability to incorporate extra sensors during the augmentation estimation.

7 CONCLUSIONS

A method for estimating a low dimension model augmentation for bias compensation given a default model and system measurements that is applicable to models described by DAE:s, is developed. A theorem that characterizes all possible augmentations that preserve the observability properties of the default model is given, and a characterization

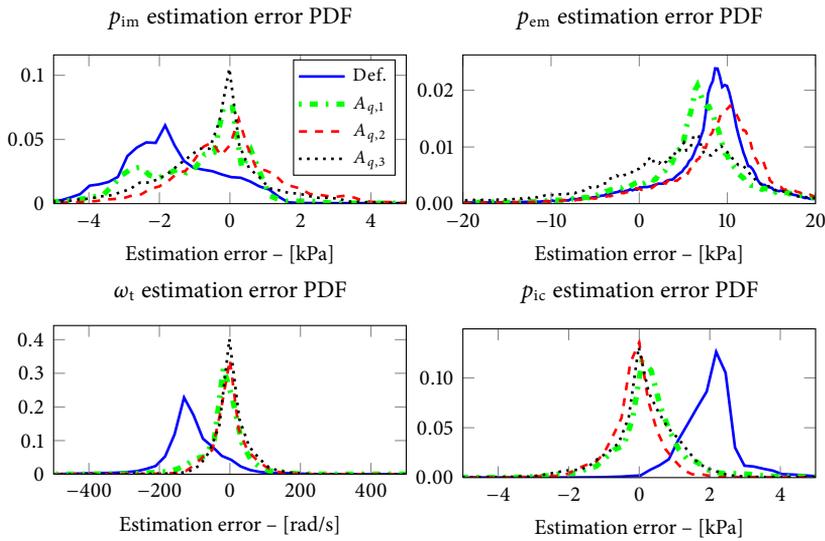


Figure 4: PDF for the estimation errors during a WHTC. It is obvious that all outputs that are compensated through feedback, i.e. p_{im} , p_{ic} , and ω_t , benefit from any choice of augmentation dimension. For p_{em} however, a three dimensional augmentation is required to achieve a significant bias reduction.

of the augmentations that are possible to estimate is presented. Common for both are that they are mainly limited by the available measurements. Beside the estimated augmentation the method allows user defined augmentations.

A main advantage of the proposed model augmentation method, compared to for example PI-observers, is its ability to incorporate information from additional sensors during the design to estimate an augmentation that can also reduce estimation errors in states not used for feedback.

The method is applied to a heavy-duty diesel engine with EGR, VGT, and intake throttle, using a nonlinear default DAE and measurements from an engine test cell. The data used is collected during a WHTC. It is shown that a one dimensional augmentation significantly reduces the mean estimation error for states where feedback is available. It is also shown that a 3 dimensional augmentation reduces the mean estimation error by as much as 55 % for the non-measured exhaust manifold pressure.

REFERENCES

- Per Andersson and Lars Eriksson. Air-to-cylinder observer on a turbocharged SI-engine with wastegate. SAE Technical Paper 2001-01-0262, 2001. doi:10.4271/2001-01-0262.
- Liyi Dai. *Singular Control Systems*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1989. ISBN 0387507248.
- Economic Commission for Europe – Inland Transport Committee. Regulation No 49 of the Economic Commission for Europe of the United Nations (UN/ECE). Official Journal of the European Union, August 2010.
- Sergio García-Nieto, Miguel Martínez, Xavier Blasco, and Javier Sanchis. Non-linear predictive control based on local model networks for air management in diesel engines. *Control Engineering Practice*, 16(12):1399–1413, December 2008. doi:10.1016/j.conengprac.2008.03.010.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. DAE and ODE based EKF:s and their real-time performance evaluated on a diesel engine. *Submitted to: IEEE Transactions on Industrial Electronics*, 2011.
- Erik Höckerdal, Lars Eriksson, and Erik Frisk. Air mass-flow measurement and estimation in diesel engines equipped with EGR and VGT. SAE Technical Paper 2008-01-0992, 2008. doi:10.4271/2008-01-0992.
- Erik Höckerdal, Erik Frisk, and Lars Eriksson. Observer design and model augmentation for bias compensation with a truck engine application. *Control Engineering Practice*, 17(3):408–417, 2009. doi:10.1016/j.conengprac.2008.09.004.
- Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, April 1970. ISBN 0123815509.
- Damien Koenig and Saïd Mammar. Design of proportional-integral observer for unknown input descriptor systems. *Automatic Control, IEEE Transactions on*, 47(12):2057–2062, December 2002. ISSN 0018-9286. doi:10.1109/TAC.2002.805675.
- Peter Kunkel and Volker Ludwig Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution*. European Mathematical Society, ETH-Zentrum FLI C4, CH-8092 Zürich, Switzerland, 2006. ISBN 3-03719-017-5.
- Peter Lancaster and Miron Tismenetsky. *The Theory of Matrices*. Academic Press, 2 edition, 1984.
- Ernest Bruce Lee and Lawrence Markus. *Foundations of Optimal Control Theory*. John Wiley & Sons, Inc, New York, 1968.
- Paolo Lino, Bruno Maione, and Claudio Amorese. Modelling and predictive control of a new injection system for compressed natural gas engines. *Control Engineering Practice*, 16(10):1216–1230, October 2008. doi:10.1016/j.conengprac.2008.01.008.

Philip Losse and Volker Ludwig Mehrmann. Controllability and observability of second order descriptor systems, 2008.

Ravi Kumar Mandela, Raghunathan Rengaswamy, Shankar Narasimhan, and Lakshmi N. Sridhar. Recursive state estimation techniques for nonlinear differential algebraic systems. *Chemical Engineering Science*, 65(16):4548–4556, 2010. ISSN 0009-2509. doi:10.1016/j.ces.2010.04.020.

Volker Ludwig Mehrmann and Tatjana Stykel. Descriptor systems: a general mathematical framework for modeling, simulation and control. Technical Report Report 292-2005, DFG Research center Matheon, Berlin, 2005.

Dirk Söffker, Tie-Jun Yu, and Peter C. Müller. State estimation of dynamical systems with nonlinearities by using proportional-integral observer. *International Journal of Systems Science*, 26(9):1571–1582, 1995. doi:10.1080/00207729508929120.

Johan Wahlström and Lars Eriksson. Modeling of a diesel engine with intake throttle, VGT, and EGR. Technical Report LiTH-R-2976, Department of Electrical Engineering, Linköpings Universitet, SE-581 83 Linköping, Sweden, 2010.

Linköping studies in science and technology, Dissertations
Division of Vehicular Systems
Department of Electrical Engineering
Linköping University

- No 1** Magnus Pettersson, *Driveline Modeling and Control*, 1997.
- No 2** Lars Eriksson, *Spark Advance Modeling and Control*, 1999.
- No 3** Mattias Nyberg, *Model Based Fault Diagnosis: Methods, Theory, and Automotive Engine Applications*, 1999.
- No 4** Erik Frisk, *Residual Generation for Fault Diagnosis*, 2001.
- No 5** Per Andersson, *Air Charge Estimation in Turbocharged Spark Ignition Engines*, 2005.
- No 6** Mattias Krysander, *Design and Analysis of Diagnosis Systems Using Structural Methods*, 2006.
- No 7** Jonas Biteus, *Fault Isolation in Distributed Embedded Systems*, 2007.
- No 8** Ylva Nilsson, *Modelling for Fuel Optimal Control of a Variable Compression Engine*, 2007.
- No 9** Markus Klein, *Single-Zone Cylinder Pressure Modeling and Estimation for Heat Release Analysis of SI Engines*, 2007.
- No 10** Anders Fröberg, *Efficient Simulation and Optimal Control for Vehicle Propulsion*, 2008.
- No 11** Per Öberg, *A DAE Formulation for Multi-Zone Thermodynamic Models and its Application to CVCP Engines*, 2009.
- No 12** Johan Wahlström, *Control of EGR and VGT for Emission Control and Pumping Work Minimization in Diesel Engines*, 2009.
- No 13** Anna Pernestål, *Probabilistic Fault Diagnosis with Automotive Applications*, 2009.
- No 14** Erik Hellström, *Look-ahead Control of Heavy Vehicles*, 2010.

