Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Meccanica e Scienze Avanzate dell'Ingegneria
INGEGNERIA DELLE MACCHINE E DEI SISTEMI ENERGETICI

Ciclo XXVI

Settore Concorsuale di afferenza: 09/C1

Settore Scientifico disciplinare: ING-IND/08

AUTOMOTIVE DIESEL ENGINE TRANSIENT OPERATION: MODELING, OPTIMIZATION AND CONTROL

Presentata da: **Giorgio Mancini**

Coordinatore Dottorato	**Relatore**
Prof. Vincenzo Parenti Castelli	**Prof. Nicolò Cavina**

Esame finale anno 2014

Ai miei genitori

Abstract

Traditionally, the study of internal combustion engines operation has focused on the steady-state performance. However, the daily driving schedule of automotive engines is inherently related to unsteady conditions. In fact, only a very small portion of a vehicle's operating pattern is true steady-state, e.g., when cruising on a motorway.

There are various operating conditions experienced by (diesel) engines that can be classified as *transient*; these may last from a few seconds up to several minutes. Besides the variation of the engine operating point, in terms of engine speed and torque, also the warm up phase can be considered as a transient condition. Chapter 2 has to do with this thermal transient condition; more precisely the main issue is the performance of a Selective Catalytic Reduction (SCR) system during cold start and warm up phases of the engine. The exhaust temperature is too low to allow thermal activation of the reactor and, consequently, to promote high conversion efficiency and significant NO_x concentration reduction. This is increasingly evident the smaller the engine displacement, because of its lower exhaust system temperature (reduced gross power while producing the same net power, i.e., higher efficiency). The proposal of the underlying work is to investigate and identify optimal exhaust line heating strategies, to provide a fast activation of the catalytic reactions on SCR. The main constraint is to limit the potential fuel consumption increase, and possibly to even increase global efficiency, and the chosen application is a small EU5-compliant diesel engine.

Chapters 3 and 4 focus the attention on the dynamic behavior of the engine, when considering typical driving conditions. To satisfy the increasingly stringent emission regulations and a demand for an ever lower fuel consumption, diesel engines have become complex systems with many interacting actuators. As a consequence, these requirements are pushing control and calibration to their limits. The calibration procedure nowadays is still based mainly on engineering experience, which results in a highly iterative process to derive a complete engine calibration. Moreover, automatic tools are available only for stationary operation, to obtain control maps that are optimal w.r.t. some predefined objective function. Therefore, the exploitation of any leftover potential during transient operation is crucial. Control trajectories resulting from the solution of an optimal-control problem pro-

vide a guideline to the calibration engineer, serve as a benchmark, and might be used for a partial automation of the calibration procedure. The common approach to dynamic optimization involves the solution of a single optimal-control problem. However, this approach requires the availability of models that are valid throughout the whole engine operating range and actuator ranges. In addition, the result of the optimization is meaningful only if the model is very accurate.

Chapter 3 proposes a methodology to circumvent those demanding requirements: an iteration between transient measurements to refine a purpose-built model and a dynamic optimization which is constrained to the model-validity region. Moreover all numerical methods required to implement this procedure are presented. The crucial steps are analyzed in detail, and the most important caveats are indicated. Finally, an experimental validation demonstrates the applicability of the method and reveals the components that require further development.

Chapter 4 proposes an approach to derive a transient feedforward control system in an automated way. It relies on optimal control theory to solve a dynamic optimization problem for fast transients. A partially physics-based model is thereby used to replace the engine. From the optimal solutions, the relevant information is extracted and stored in maps spanned by the engine speed and the torque gradient. These maps complement the static control maps by accounting for the dynamic behavior of the engine. The procedure is implemented on a real engine and experimental results are presented along with the development of the methodology.

Sommario

Lo studio del funzionamento dei motori a combustione interna è tradizionalmente focalizzato sulla prestazione in condizioni stazionarie. Tuttavia, il ciclo di guida giornaliero dei motori automobilistici è per sua stessa natura caratterizzato da condizioni non stazionarie. Infatti, solo una piccolissima porzione del ciclo operativo di un veicolo è realmente stazionaria, ad esempio quando si tiene una certa velocità di crociera in autostrada.

Le condizioni di funzionamento dei motori (diesel) che possono essere classificate come "transitorie" sono svariate; queste possono durare da pochi secondi fino a diversi minuti. Oltre alla variazione del punto di funzionamento del motore, in termini di velocità e coppia, anche la fase di riscaldamento può essere considerata come una condizione transitoria. Il Capitolo 2 tratta esattamente questa condizione di transitorio termico; più precisamente il problema principale è la prestazione di un sistema SCR (Selective Catalytic Reduction) durante le fasi di avviamento a freddo e riscaldamento del motore. La temperatura di scarico è troppo bassa per consentire l'attivazione termica del reattore e, di conseguenza, per promuovere un'elevata efficienza di conversione e una significativa riduzione della concentrazione di NO_x. Tale comportamento è sempre più evidente quanto più piccola è la cilindrata del motore, a causa della più bassa temperatura del suo sistema di scarico (ridotta potenza lorda a parità di potenza netta, ovvero maggiore efficienza). Il seguente lavoro si propone di individuare e identificare strategie di riscaldamento ottimo della linea di scarico, al fine di fornire una rapida attivazione delle reazioni catalitiche in seno al SCR. Il vincolo principale è quello di limitare il potenziale aumento del consumo di combustibile, e possibilmente di riuscire persino a incrementare l'efficienza globale. L'applicazione scelta è un piccolo motore diesel, che rispetta i limiti della normativa Euro 5.

I Capitoli 3 e 4 spostano l'attenzione sul comportamento dinamico del motore, considerato nelle tipiche condizioni di guida. Per soddisfare le sempre più stringenti normative sulle emissioni e la richiesta di un sempre più basso consumo di combustibile, i motori diesel sono diventati sistemi complessi con molti attuatori interagenti. Di conseguenza, tali requisiti stanno spingendo il controllo e la calibrazione ai loro limiti. La procedura di calibrazione è ancora oggi basata principalmente sull'esperienza ingegneristica, il quale richiede un processo altamente iterativo per poter derivare una calibrazione

completa del motore. Inoltre, sono disponibili diversi sistemi di calibrazione automatica applicabili al caso stazionario, al fine di ottenere mappe di controllo che siano ottime rispetto a funzioni obiettivo predefinite. Pertanto, lo sfruttamento di qualsiasi potenziale residuo durante il funzionamento transitorio è di cruciale importanza. Traiettorie di controllo risultanti dalla soluzione di un problema di controllo ottimo forniscono una linea guida per l'ingegnere di calibrazione, servono come punto di riferimento, e potrebbero essere usate per una parziale automazione della procedura di calibrazione. L'approccio comune all'ottimizzazione dinamica comporta la soluzione di un singolo problema di controllo ottimo. Tuttavia, questo approccio richiede la disponibilità di modelli validi su tutto il campo di funzionamento del motore e degli attuatori. Inoltre, il risultato dell'ottimizzazione è significativo solo se il modello è molto accurato.

Il Capitolo 3 propone una metodologia per aggirare tali requisiti esigenti: un'iterazione tra misure transitorie, per perfezionare un modello realizzato ad hoc, e un'ottimizzazione dinamica vincolata alla regione di validità del modello. In aggiunta, vengono presentati tutti i metodi numerici necessari per implementare tale procedura. I passi fondamentali sono analizzati nel dettaglio, soffermandosi in particolar modo sui punti più critici. Infine, una validazione sperimentale dimostra l'applicabilità del metodo e rivela gli aspetti che richiedono un ulteriore sviluppo.

Il Capitolo 4 propone un approccio per realizzare un sistema di controllo dei transitori in modo automatizzato. Esso si basa sulla teoria del controllo ottimo, applicata per risolvere un problema di ottimizzazione dinamica di transitori veloci. Un modello parzialmente basato sulla fisica viene quindi usato per rimpiazzare il motore. Dalle soluzioni ottime vengono estratte le informazioni rilevanti e archiviate in mappe dipendenti dalla velocità e dal gradiente di coppia. Queste mappe compensano le mappe di controllo statico tenendo conto del comportamento dinamico del motore. La procedura è implementata su un motore reale e i risultati sperimentali vengono presentati insieme allo sviluppo della metodologia.

Contents

Nomenclature

The time derivative of a variable x is denoted by $\dot{x}$, whereas $\overset{*}{x}$ represents a flow of mass, heat or energy, for instance. Bold symbols indicate vectors and matrices. The following list introduces the abbreviations and the symbols that are used consistently throughout the text. Indices and specific symbols that are used only in a narrow context are introduced and explained directly at the corresponding locations in the text. For each symbol that can assume different meanings, the respective context is indicated in brackets.

Symbols

σ_0	Stoichiometric AFR	[−]
λ	Normalized Air-to-fuel ratio $\lambda = {}^{\mathrm{AFR}}/_{\sigma_0}$	[−]
w	Speed	[rad/s]
m_{fcc}	Mass of fuel injected per cycle per cylinder	[kg/cc]
φ_{SOI}	Start of main injection	[°BTDC]
p_{IM}	Intake manifold pressure	[Pa]
p_{EM}	Exhaust manifold pressure	[Pa]
ϑ_{IM}	Intake manifold temperature	[K]
ϑ_{EM}	Exhaust manifold temperature	[K]
$x_{\mathrm{bg,IM}}$	Burnt-gas ratio in the intake manifold	[−]
$x_{\mathrm{bg,EM}}$	Burnt-gas ratio in the exhaust manifold	[−]
N_{eng}	Engine speed	[rpm]
T_{eng}	Engine coolant temperature	[K]

Acronyms and Abbreviations

AFM	Air-flow Meter
AFR	Air-to-fuel ratio
aTB	after Turbine

bTB	before Turbine
bmep	brake mean effective pressure
BSA	back-sweep angle of the blade
BTDC	Before Top Dead Center
CoV	Coefficient of Variation
DC	Dynamic Compensation
DOC	Diesel Oxidation Catalyst
DPF	Diesel Particulate Filter
ECU	Electronic Control Unit
EGC	Exhaust Gas Cooler
EGR	Exhaust Gas Recirculation
EM	Exhaust Manifold
FB	Feedback
FC	Fuel Consumption
FF	Feedforward
IM	Intake Manifold
IMEP	Indicated Mean Effective Pressure
LIP	Linear in Parameters
LSQ	Linear least-squares
NEDC	New European Driving Cycle
OCP	Optimal-control Problem
ODE	Ordinary Differential Equation
SCR	Selective Catalytic Reduction
SOI	Start Of Injection
VGT	Variable Geometry Turbine
TB	Time-based
TcatIN	Temperature catalyst inlet
TVA	Throttle Valve Actuator
TWC	Three Way Catalyst

Chapter 1

Introduction

1.1 Background and motivation

In recent years, the increasingly stringent exhaust emission regulations have dominated the automotive industry, and forced manufacturers to new developments. Automotive diesel engines have experienced a significant technological development, resulting in increased requirements on calibration and control [1]. For diesel engines, the emphasis is on reducing emissions of nitrogen oxides (NO_x) and particulate matter (PM), due to the toxicity of the inhaled nanoparticles and because these pollutants are typically higher than those from equivalent rated, port-injected gasoline engines equipped with three-way catalysts.

Unfortunately, there is a trade-off between NO_x and PM reduction, resulting in a complicated control strategy requiring complex after-treatment systems. Sophisticated, high-pressure common rail injection systems, exhaust gas recirculation (EGR), or selective catalytic reduction (SCR), multi-valve configurations with variable valve timing, variable geometry turbochargers (VGT), exhaust after-treatment systems with particulate traps or urea-based deNO_x are among the measures applied for pollutant emissions reduction and fuel consumption. For the research community, the introduction of these systems has led to a number of challenging problems as far as control and combustion research are concerned. In particular, the control of the EGR [2] and the conjoint control of the EGR and VGT [3, 4, 5, 6] have resulted in numerous publications.

Furthermore, carbon dioxide (CO_2) emissions are becoming increasingly important owing to their connection with global warming; limiting CO_2 production can be achieved, primarily, through improvements in fuel economy and use of biofuels. Today's diesel-engined automobiles not only demonstrate greater fuel efficiency than ever before, but they also achieve emission levels at least 50% lower than those of a few years ago.

Unsurprisingly, the various technological advances mentioned above have

also led to a significant increase in the complexity and cost of the engine and its control system, and this trend is sure to continue.

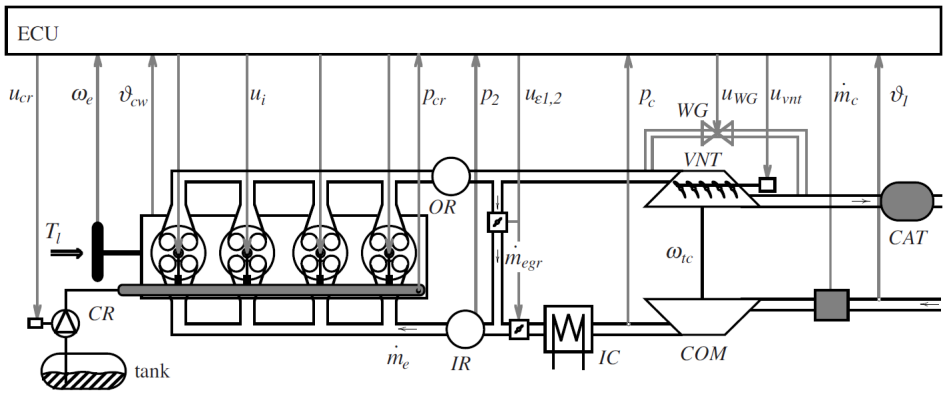
Figures 1.1 and 1.2 show an overview of a typical modern diesel engine as used in passenger cars. The comprehensive open- or closed-loop control systems employed gather the signals from the various sensors located on the engine, fuel pump and turbocharger (Figure 1.2), process them by means of look-up tables (steady-state maps) or, better still, model-based control theory with, for instance, feed-forward control, and eventually determine the optimum position of the various valves, vanes, etc.

The main objective for electronic diesel-engine control-systems is to provide the required engine torque while minimizing fuel consumption and complying with exhaust-gas emissions and noise level regulations. This requires an optimal coordination of injection, turbocharger and exhaust-gas recirculation systems in stationary and transient operating conditions. Traditionally, the control optimization is undertaken during the design stage for steady-state operations, with the calibrated parameters, e.g., injection strategy (pre-, main- and post-injection scheme, rate, timing and pressure of injection), VGT vanes position, boost pressure, EGR valve position etc., stored in 3-D maps spanned by engine speed and load/fueling.

From a control-engineering point of view, there are three important paths which have to be considered: fuel, air and EGR. Figure 1.3 shows a schematic overview of the basic structure of a typical diesel-engine control-system, clearly pointing out these three paths. Notice that a speed controller is standard in diesel engines: the top speed must be limited in order to prevent engine damage whereas the lower limit is imposed by the desired running smoothness when idling.

1.2 Transient operation fundamentals

Traditionally, the study of internal combustion engines operation has focused on the steady-state performance, with minor, if any, attention paid to the unsteady-state or more accurately termed *transient operation*. However, the majority of daily driving schedule involves transient conditions. In fact, only a very small portion of a vehicle's operating pattern is true steady-state, e.g., when cruising on a motorway. Historically, however, the research on transient diesel engine operation dates back to 1960s, when the engine manufacturers observed that when highly-rated, medium-speed diesel engines were employed in sudden 0-100% step load changes, severe difficulties were encountered, even leading to engine stall. In recent years, the global concern about environmental pollution has led the respective studies to be intensified; particulate, gaseous and noise emissions typically go way beyond their acceptable values following the extreme, non-linear and nonsteady-state conditions experienced during dynamic engine operation. A few representative



CAT	oxidation catalytic converter	$u_{\varepsilon 1,2}$	EGR valve(s) command	p_c	pressure after COM
COM	compressor	u_{cr}	CR pump command	p_2	intake pressure
CR	common-rail system	u_i	injection command	p_{cr}	CR injection pressure
IR	intake receiver	u_{vnt}	turbine nozzle command	$\dot{m}_c$	intake air mass flow
OR	outlet receiver	u_{WG}	WG command	ϑ_l	intake air temperature
IC	intercooler	T_l	load torque at the flywheel	ϑ_{cw}	cooling water temperature
VNT	variable nozzle turbine	ω_{tc}	turbocharger speed	ω_e	engine speed
WG	waste-gate (alternative to VNT)	$\dot{m}_e$	total engine-in mass flow	$\dot{m}_{egr}$	exhaust gas recirculation

Figure 1.1: Overview of a typical system structure of a Diesel engine ([7]).

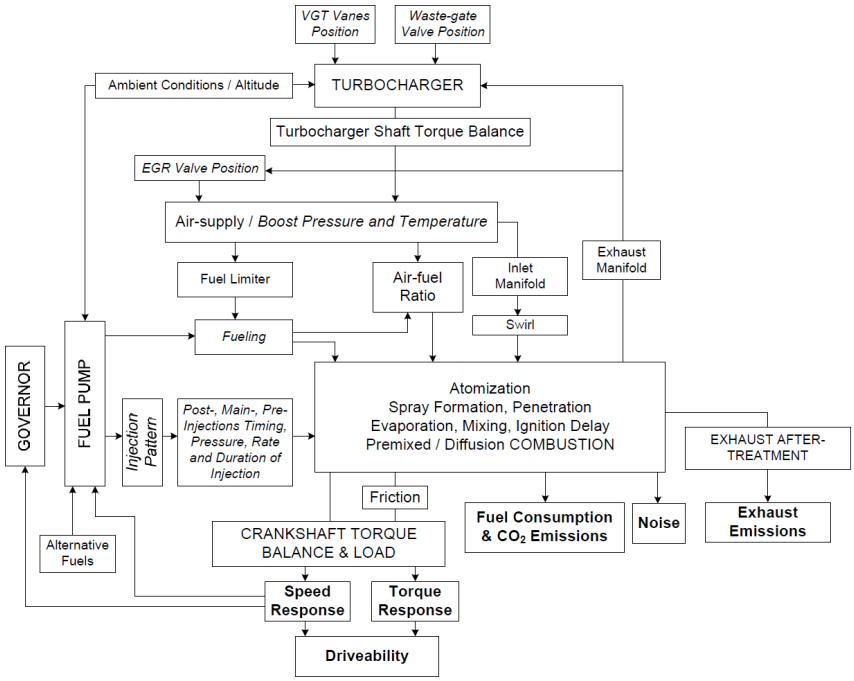


Figure 1.2: Simplified diagram showing some major air-supply and fueling controllable inputs (*italics*) and engine/vehicle outputs (**bold**), highlighting the complexity of a modern diesel engine powertrain ([8].)

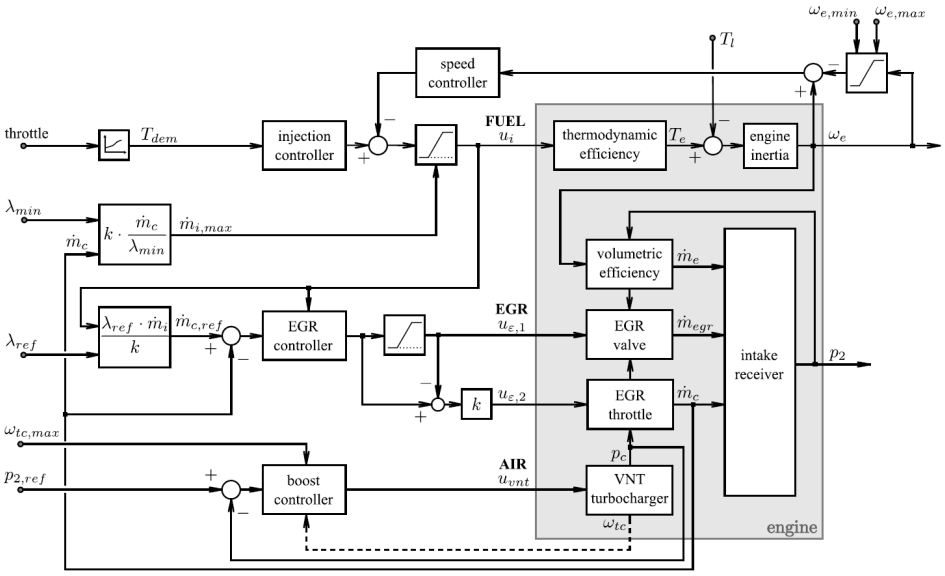


Figure 1.3: Basic diesel-engine control-system structure, variables as defined in 1.1 ([7]).

results follow: cold- or warm-start emissions from heavy-duty diesel engines have been found to exceed up to 15 times their steady-state values; 50% of NO_x emissions from automotive engines during the European Driving Cycle stem from periods of acceleration, whereas instantaneous particulate matter and NO_x emissions during load increase transients have been measured to be 1 to 2 orders of magnitude higher than their respective quasi-steady values.

Starting from the 1980s, (diesel-engined) vehicles have been tested for exhaust emissions, prior to type approval, using sophisticated standardized transient tests (Transient Cycles); these are usually characterized by long duration (up to 30 minutes) consisting of both speed and load changes under varying operating schedules. A Transient Test Cycle is a sequence of test points each with a defined vehicle speed to be followed by the vehicle under study, or with a defined rotational speed/torque to be followed by the engine under transient conditions; these test points are divided in time steps, mostly seconds, during which acceleration is assumed constant. Such standardization is necessary as it makes it possible to compare different vehicles/engines that fulfill the same operation.

In order for the exhaust emission measurements to be representative of real engine operation, Transient Test Cycles incorporate some or all of the following driving conditions:

- cold and hot starting
- frequent accelerations and decelerations

- changes of load
- idling conditions typical of urban driving
- sub-urban or rural driving schedule
- motorway driving

By applying a Transient Cycle for the testing of new vehicles, the complete engine operating range is tested and not just the maximum power or torque operating points. Moreover, the serious discrepancies that are experienced during abrupt transients are taken into account. However, it should be pointed out that the primary objective of a Transient Cycle procedure is to establish the total amount of exhaust emissions rather than indicate the specific parts or conditions under which these emissions are produced. Further, legislative Test Cycles assume straight roads with zero gradient, thus no account is taken of the respective road-dependent resistance torque. Transient Cycles require highly sophisticated experimental facilities (a fully automated testbench with electronically controlled motoring and dissipating (chassis) dynamometer, fast response exhaust gas analyzers, dilution tunnels, etc.) in order to be accurately reproduced. Many countries in the world have developed Transient Cycles for emission testing of their vehicles. These Transient Cycles concern the testing of passenger vehicles, light-duty (commercial) vehicles, heavy-duty vehicles, heavy-duty engines, and non-road mobile engines. Passenger cars and light-duty vehicles usually undergo a vehicle speed vs. time Test Cycle on a chassis dynamometer, and the results are expressed in g/km.

There are various operating conditions experienced by (diesel) engines that can be classified as *transient*; these may last from a few seconds up to several minutes. In this thesis, the term *transient* will be used to describe any of the following three forced changes:

1. cold or hot starting
2. step change in pedal position (change in load request) resulting in engine speed changes (acceleration) while keeping the same output torque
3. load acceptance (change) at quasi-constant engine speed, mainly experienced in industrial applications, e.g., electrical generation, but also observed in various propulsion applications, e.g., when a vehicle climbs a hill;

These are the most fundamental transient cases, while their combination results in:

- simultaneous speed and load changes

- whole vehicle propulsion schedule, e.g., gear shift
- Transient Cycles, which consist of all of the above mentioned transients

A number of typical transient cases will be described in the next chapters, emphasizing the different responses of the engine as well as the requirements of the engine control system.

1.3 Structure of the thesis

The core subject treated during the research activity, and presented in this thesis, is the development of control strategies for diesel engine transient operation. Independently of the specific transient pattern (see list above), the control scheme derived is characterized by the same structure, shown in Figure 1.4. The "static-based controls" represent the typical outputs from the feedforward control based on the steady-state set-point maps, whereas the "transient-accounting controls" are obtained when applying the transient compensation strategy to the static outputs. As can be noted, the inputs to the steady-state set-point maps are engine speed (N_{eng}) and load or fuel request (m_{fcc}), and they are independent of the specific transient pattern. On the contrary, the transient compensation strategy strongly depends on the transient effect to be accounted for, and so do the "transient-related inputs" likewise. This is the crucial point: how to derive such compensation strategies, along with the best suited control inputs. Depending on the control problem to be solved, different solutions are proposed.

In Chapter 2, the main issue is the performance of a Selective Catalytic Reduction (SCR) system during cold start and warm up phases of a light-duty diesel engine; it is thus focused on the transient condition 1, referring to the list previously presented. Transient conditions 2 and 3 are considered instead in Chapters 3 and 4, which focus mainly on the dynamic optimization of diesel engines.

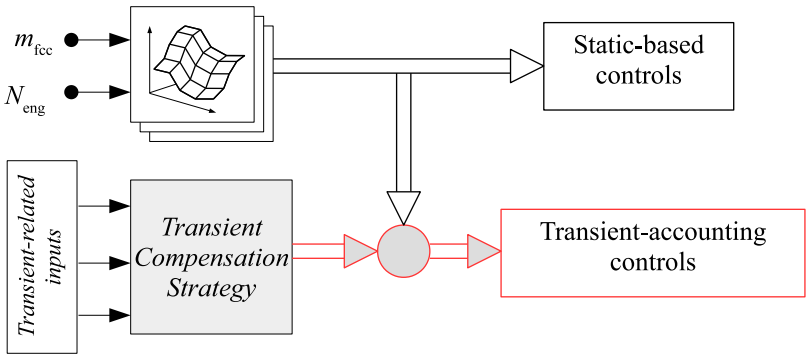


Figure 1.4: Transient control approach presented in the thesis.

Chapter 3 presents the numerical methods and the testbench setup required to perform an iterative dynamic optimization of diesel engines over prescribed driving profiles. Through the synergy of engine modeling and optimal-control theory, control trajectories resulting from the solution of an optimal-control problem are derived. The results achieved are then applied in Chapter 4, where a methodology to derive a transient feedforward control system in an automated way is proposed.

It has to be pointed out that each subject has been tackled with the same approach, which is in logical order:

- analysis of the problem and definition of the main objectives
- experimental and/or numerical methods definition
- development of the methodology
- experimental validation through a systematic and reliable approach
- implementation on a real engine
- final experimental validation and presentation of results

Chapter 2

Transient thermal management of diesel engines

As is well known, Euro 6 (EU6) regulation is about to introduce in Europe a particularly severe reduction of the allowed NO_x emission levels for Diesel passenger cars [9]. Due to the lean composition of diesel engines exhaust gas, Three Way Catalysts (TWCs) prove inefficient to reduce NOx species, and other solutions have been investigated in recent years. Especially in heavy-duty applications, the most widely adopted solution is the installation of Selective Catalyst Reduction (SCR) systems [10, 11, 12]. The use of a SCR system on a diesel engine exhaust line is now becoming a valid solution also for passenger cars, with the aim of drastically reducing NO_x emissions, in an effort to comply with forthcoming EU6 regulations. In [13, 14] a preliminary study about the application of a SCR system on a EU5 B-segment vehicle has been presented, analyzing possible solutions for obtaining high NOx conversion efficiency values from the after-treatment device. Several simulations have been performed, using the New European Driving Cycle (NEDC) as a benchmark test to compare the results. They show that the exhaust gas temperature tends to be too low at SCR inlet to allow for optimal catalyst operation. So, for the considered vehicle (B-segment), the simulation demonstrated the need for a specific heating strategy to rapidly light-off the SCR catalyst, and to subsequently control its operating temperature at optimal values.

This chapter presents the analysis and application of thermal management strategies for the exhaust line, with the constraint of limiting additional fuel consumption [15]. Firstly, a few NEDC cycles have been performed on the roll bench dynamometer, to collect exhaust line temperature dynamics experimental data of the production calibration (EU5), as a reference starting point. Figure 2.1 shows the exhaust line layout selected for this study (Fiat 1.3L Doblò Multi-Jet), with the SCR and its related components installed specifically for this project. The exhaust gas path, from the engine to the

SCR inlet (where urea injection takes place), is particularly long. The challenge is to reach a temperature of approximately 190°C as fast as possible [16], far away from the exhaust valves and without compromising fuel consumption.

Figure 2.2 shows the temperature trend of the exhaust gas at SCR inlet, during NEDC operation (the vehicle speed is reported in the right y-axis). As clearly shown, the target temperature mentioned above is reached almost at the end of the driving cycle (around t=900 s), suggesting the impossibility to effectively take advantage of the selective catalyst reactor. From a more general point of view, during engine warm-up the significant NOx reduction that should be provided by the SCR system seems to be unfeasible, and this consideration is true the farther the SCR system is located from the engine.

To find a possible solution to this issue, an experimental investigation has been performed, focusing the attention on the main engine control parameters that are available to the Engine Control Unit (ECU). The proposal is to analyze the effects on the combustion and the resulting exhaust gas temperature (and composition), by changing the parameters that mostly affect the engine efficiency, reported in the following list (see also Figure 2.3):

- *Injection phasing*, by controlling the Start Of Injection (SOI) position of the pre-selected injection pattern.
- *Turbine efficiency*, and engine back-pressure, by controlling the Variable Geometry Turbine (VGT) actuation.
- *Exhaust Gas Recirculation* (EGR) rate, for its impact on NO_x production and pumping losses, by controlling the engine EGR valve position.
- *Intake pressure and composition*, by controlling the engine Throttle Valve Actuation (TVA) system.

After this first analysis, performed on a test bench with the same engine layout set on the vehicle, several calibration strategies have been developed and validated. Afterwards, calibration profiles derived from the validation process have been applied on the vehicle while executing other NEDC tests. Final results show the improvement reached with the heating strategy methodology, enhancing a faster SCR activation.

2.1 Control parameters sensitivity analysis

In the first phase of the work, different approaches on how to rapidly warm-up the exhaust line have been evaluated, taking into consideration two potential strategies [17, 18, 19]:

Table 2.1: Engine characteristics.

engine type:	**FIAT 1.3 Multi-Jet EU5**
displacement:	1248 cm^3
compression ratio:	16.8
EGR:	high pressure, cooled
turbocharger:	VGT, charge-air cooler
injection system:	common rail DDI
valves/cyl.:	4
max. torque:	200 Nm (1500-3500 rpm)
max. power:	66 kW (4000 rpm)

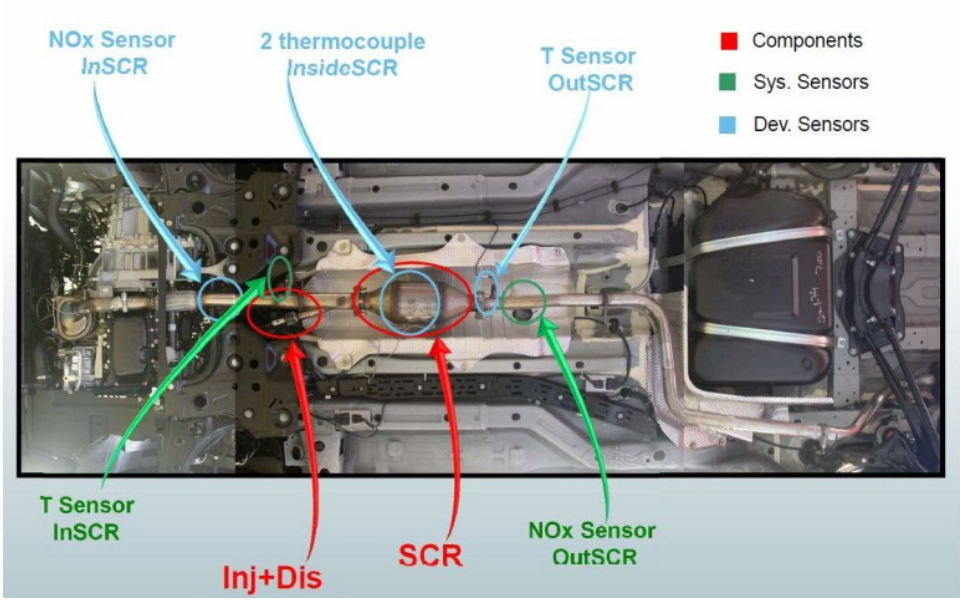


Figure 2.1: Exhaust line layout.

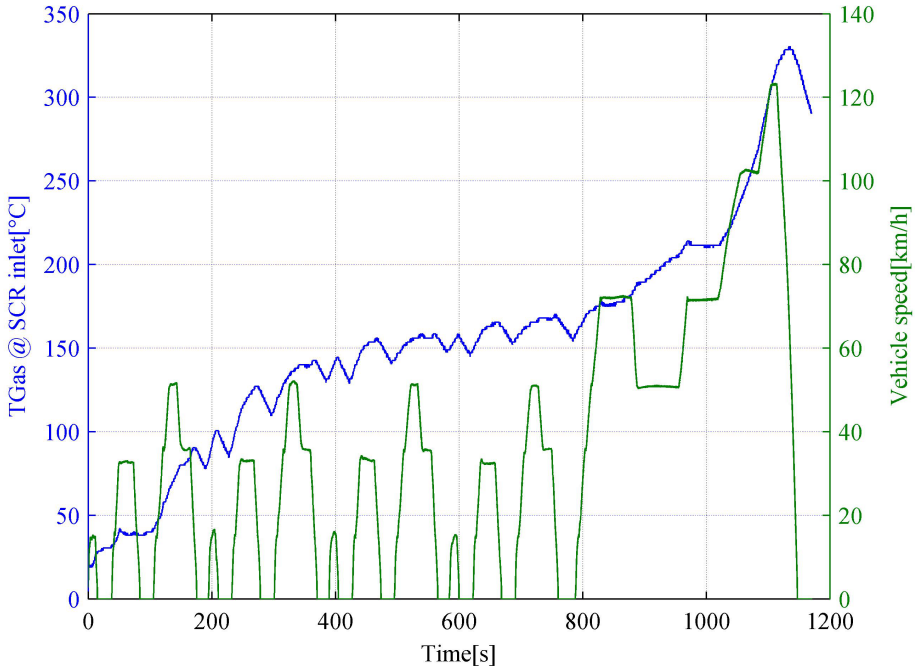


Figure 2.2: Temperature trend at SCR inlet during NEDC cycle without heating strategy (EU5 calibration).

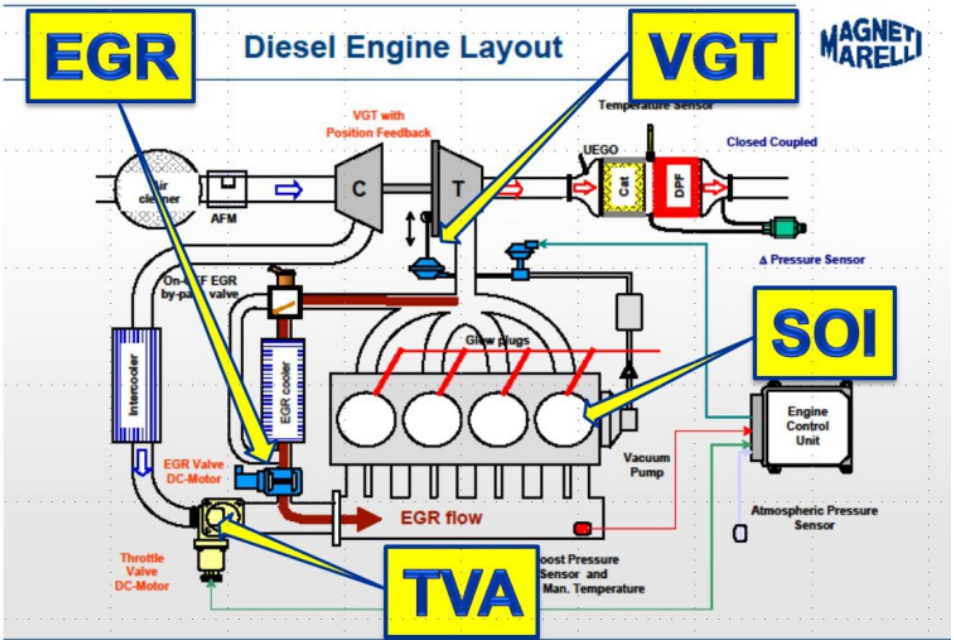


Figure 2.3: Layout of the diesel engine used for the experiments.

- *"Hard" strategy*: late or post injections to promote oxidation processes in the exhaust line (after Diesel Oxidation Catalyst DOC activation). It implies the need to increase the amount of injected fuel, with a consequent predictable fuel penalty.

- *"Soft" strategy*: identify control strategies and calibration parameters, with particular focus on the ones that mainly influence combustion efficiency.

The second choice has been preferred, since it could have the desired effect of satisfying both the exhaust heating and the fuel consumption constraint. As already mentioned, and with reference to Figure 2.3, the main control parameters are SOI, EGR, TVA, and VGT, while the main engine characteristics are presented in Table 2.1. The final effect, in terms of exhaust temperature, must be evaluated downstream of the turbine and upstream of after-treatment devices, so that the measured temperature is not affected by the operating condition of DOC and DPF systems.

The standard ECU calibration (EU5) is the result of a trade-off between engine performance (maximum torque and optimum drivability) and minimum fuel consumption and emissions, so any modification should in principle result in a worse engine behavior. Nevertheless, the addition of a SCR system downstream of DOC and DPF systems introduces new perspectives. It is now possible to identify new calibration parameter combinations that optimize the overall engine (and aftertreatment system) behavior. For instance, the NO_x reduction task can be completely handled by the SCR system, once it has been thermally activated. Therefore, starting from this consideration and given the previous simulation results, the effects of SOI, VGT and TVA have been evaluated on a test bench, performing several steady-state measurements.

The engine operating points have been extrapolated from the ECE driving cycle. Figure 2.4 shows the four operating points (red triangles) in which the engine behavior can be approximately considered stationary; it follows that each point is characterized by a speed-bmep combination, as shown in Table 2.2 (bmep represents the engine normalized output torque). In order to take into consideration also engine transient operation, a fifth operating point has been identified as the average condition in terms of engine speed and bmep. The operating conditions chosen for the sensitivity analysis have been n. 1, n. 3, and n. 5, with the aim of minimizing the number of tests while guaranteeing a substantial representation of the engine operating range.

The EGR valve actuation has not been directly considered, leaving to the ECU the priority to handle it, dependently from the air flow resulting from VGT and TVA positions. The EGR rate has therefore not been externally controlled during the experimental tests, but it depends on the other devices involved in the air-path controller.

Table 2.2: Engine operating points extrapolated from the ECE driving cycle. Only n. 1, n. 3 and n. 5 have been used for the sensitivity analysis.

#	Speed [rpm]	bmep [bar]
1	2200	1.81
2	2400	1.51
3	2500	2.52
4	1700	2.11
5	1800	5.03

2.1.1 Start of Injection (SOI)

Considering a determined steady state operating point, in a maximum efficiency condition, apparently the only way to increase the exhaust temperature is to shift the combustion toward the exhaust phase. This action reduces combustion efficiency, therefore a greater amount of injected fuel is needed to guarantee the same engine torque. The main control parameter of the combustion phase in a diesel engine is the SOI position (expressed in crankshaft degrees Before Top Dead Center BTDC). If not differently specified the SOI refers to the main injection in a multiple injection pattern.

The figures below show the exhaust gas temperature and fuel consumption responses to SOI sweeps performed in different steady state conditions. ΔSOI is defined as the difference between reference (EU5) SOI value and actual SOI value, while ΔTcatIN and ΔFuelConsumption respectively represent the exhaust gas temperature difference between actual and ref. value (with ref. SOI), and the percentage difference between actual and ref. fuel consumption values. Positive values in the x-axis of Figures 2.5 through 2.7 therefore represent retarded injection phases with respect to reference values. Figure 2.5 confirms that the higher is the combustion delay, the higher is the exhaust gas temperature. As previously shown in Table 2.2, tests have been performed while controlling speed and bmep at constant values, so that the efficiency reduction that comes with the combustion delay needs to be compensated by increasing fuel consumption, as clearly visible in Figure 2.6. Furthermore, the same figure confirms that in the $\Delta SOI \approx 0$ range the fuel consumption is minimum, proving that the SOI_{ref} value had been correctly identified.

It is also true that the applied SOI corrections reduce the combustion chamber temperatures, thus decreasing the thermal NO_x production. This explains why sometimes the SOI_{ref} may in other cases prove to be sub-optimal, in terms of efficiency, to limit such type of emissions. Finally, Figure 2.7 shows the Indicated Mean Effective Pressure (IMEP) Coefficient

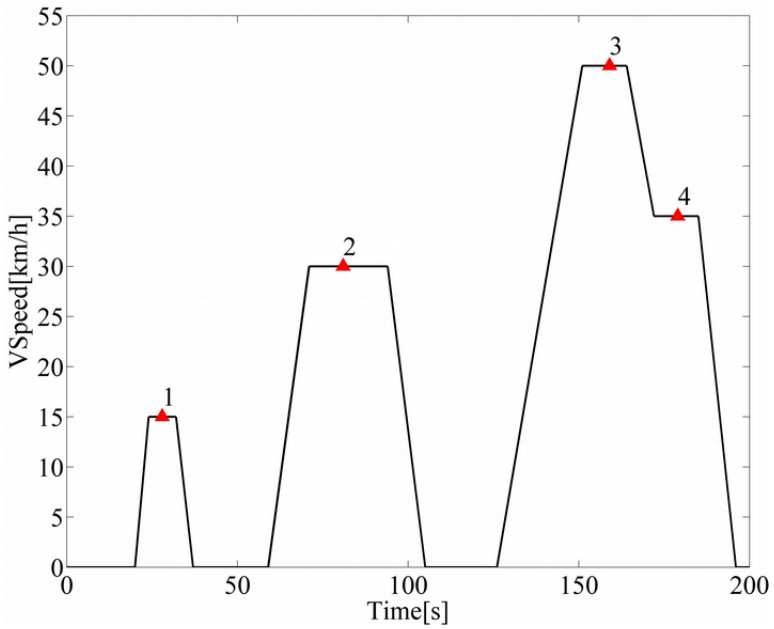


Figure 2.4: Vehicle speed profile in the ECE cycle and operating points investigated.

of Variation, CoV(IMEP), for cylinders 2 and 3, to prove that up to a SOI delay of 12 CA, the combustion stability is always guaranteed. CoV(IMEP) is defined as the ratio between IMEP standard deviation and its mean value, expressed as a percentage value, and it is widely recognized as an indirect but accurate measure of combustion stability [20].

2.1.2 Variable Geometry Turbine (VGT)

As it has been previously mentioned, another important control parameter within this context is the VGT position. In the given control system, a VGT position=100% corresponds to a fully closed condition, and vice versa the distributor is fully open with a VGT position=0%. In other words, when the VGT is fully closed (100%) it means that the turbine, under steady state flow conditions, generates the highest backpressure, correspondent to that mass flow, and therefore the pressure drop through the turbine is maximum. In this condition, the turbine produces its maximum power, increasing the turbocharger speed, with the consequent increase of the intake manifold pressure.

The tests performed have the aim to increase the distributor opening with respect to reference (EU5) calibration, in order to reduce the enthalpy drop and to increase the temperature downstream of the turbine.

As shown in Figure 2.8, where negative ΔVGT values correspond to a

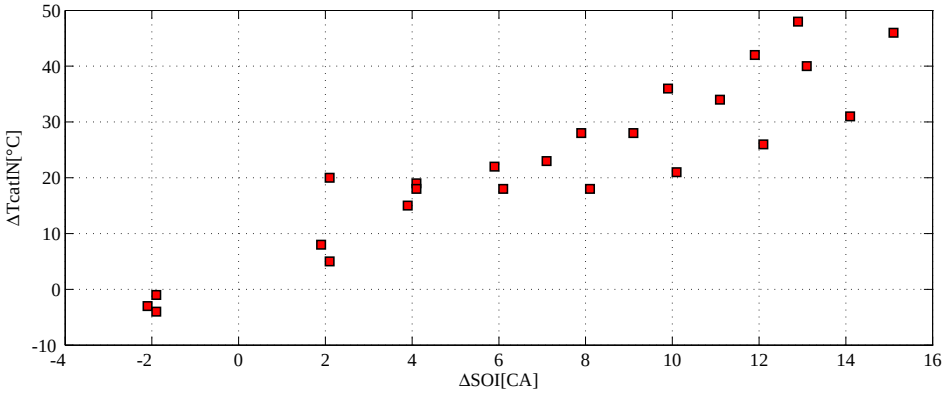


Figure 2.5: Effect of SOI variation on TcatIN.

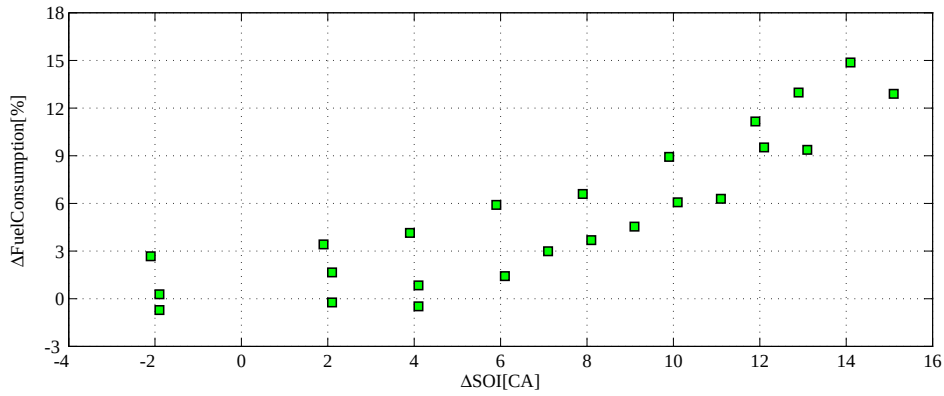


Figure 2.6: Effect of SOI variation on Fuel consumption.

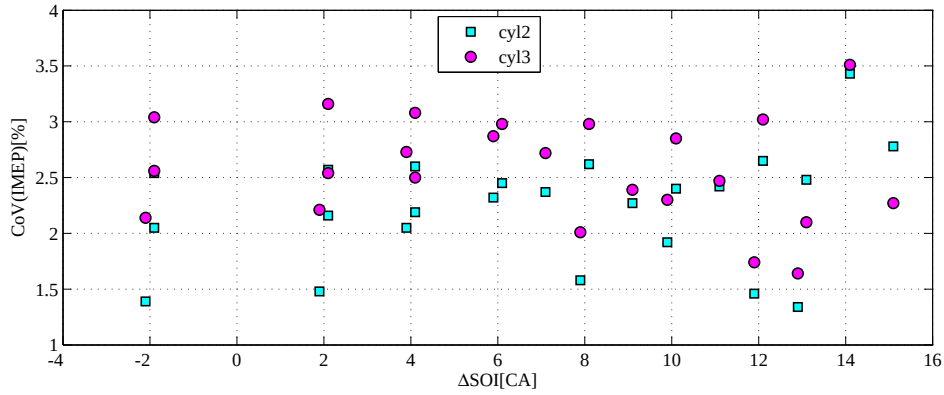


Figure 2.7: Effect of SOI on CoV(IMEP), cylinders 2 and 3.

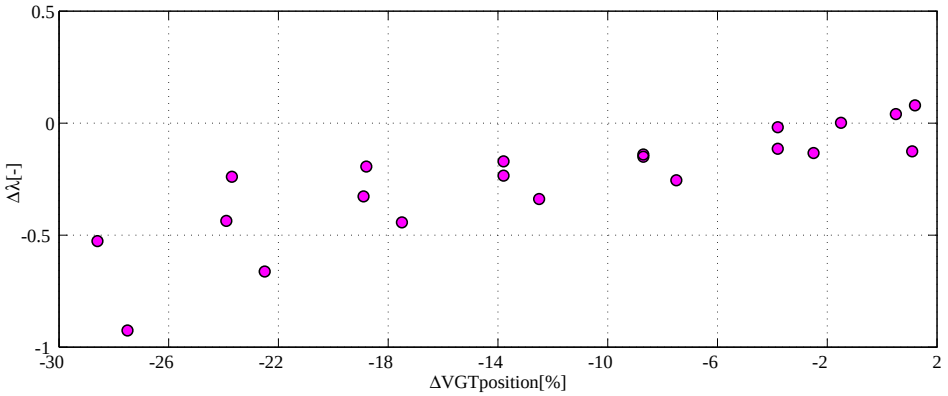


Figure 2.8: Effect of VGT on λ.

wider opening position with respect to reference values, by reducing the VGT position the normalized air-to-fuel ratio (λ) value slightly decreases. This effect is quite predictable, because the fuel mass flow does not change (constant load and essentially constant engine efficiency) while the air mass flow decreases due to intake pressure reduction. During the tests, a lower saturation has been imposed on λ, to guarantee acceptable smoke levels. In this case there is a positive effect both on exhaust gas temperature and fuel consumption (Figure 2.9, Figure 2.10). The first effect is due to the turbine energy conversion reduction, and to a lower λ value (richer combustion, due to reduced excess air mass and therefore reduced thermal capacity of the charge), while the fuel consumption reduction is mainly due to the reduced engine backpressure (higher pumping efficiency). As shown in Figure 2.11, the CoV(IMEP) is rather unaffected by the VGT variation, except for relatively large opening positions, implying lowest boost pressure and lowest lambda. Generally, large deviations from nominal turbocharger operating conditions are not allowed, to avoid negative effects on turbo lag response and vehicle drivability.

2.1.3 Throttle Valve (TVA)

The throttle valve is placed at the intake manifold inlet, before the EGR path outlet (Figure 2.3), and during engine operation the valve is typically fully opened. The same valve is actuated only when the pressure drop across the EGR valve is not sufficient to guarantee the requested exhaust gas recirculation amount. By forcing the throttle valve closing, while controlling EGR and VGT at constant position, the intake manifold pressure may be significantly reduced. Figures 2.12, 2.13 and 2.14 highlight the effects of intake pressure reduction (with respect to the levels achieved with reference TVA calibration) on the most significant engine operating

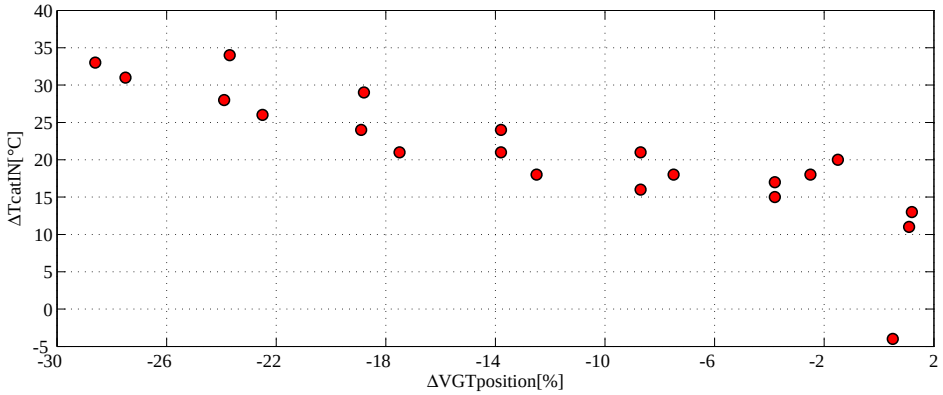


Figure 2.9: Effect of VGT on TcatIN.

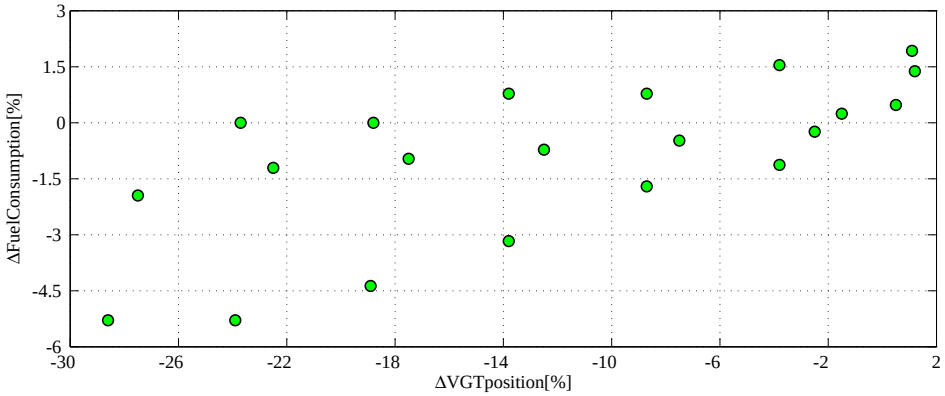


Figure 2.10: Effect of VGT on fuel consumption.

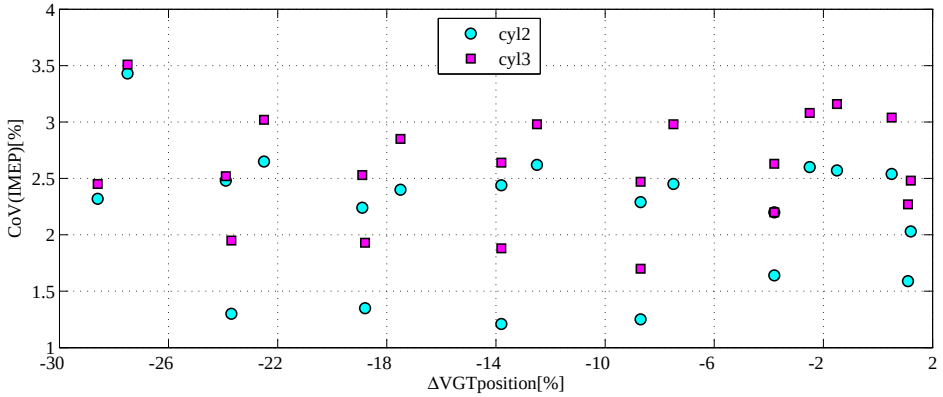


Figure 2.11: Effect of VGT on CoV(IMEP), cylinders 2 and 3.

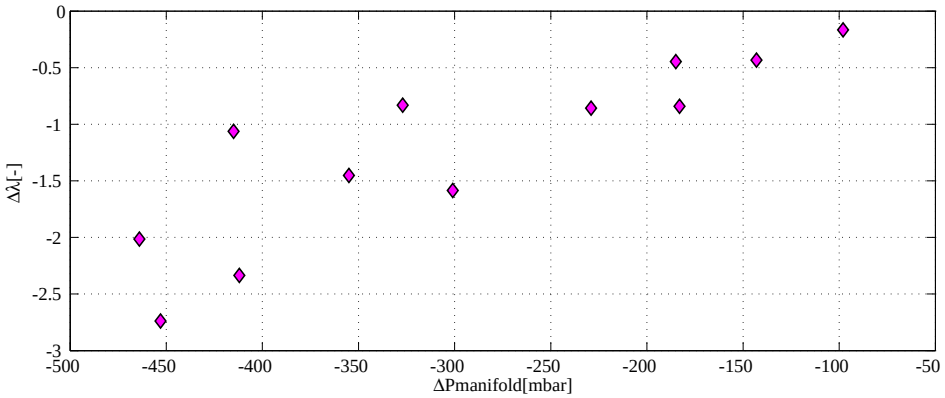


Figure 2.12: Effect of TVA on λ.

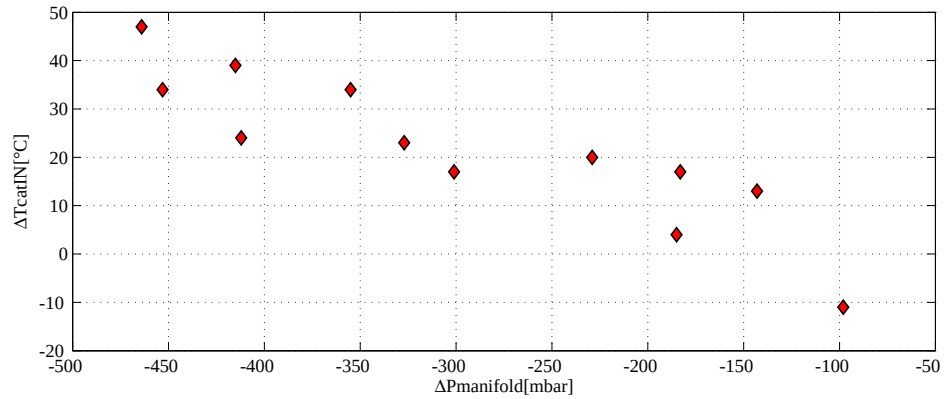


Figure 2.13: Effect of TVA on TcatIN.

parameters. ΔPmanifold represents the difference between reference intake manifold pressure values and the ones measured during the tests. The main effect is a large λ value reduction, which causes the exhaust gas temperature to increase. On the other hand, the specific fuel consumption responds negatively, especially when the intake manifold pressure is lower than the ambient pressure (see Figure 2.14). Even if there is a positive effect on the temperature, this kind of strategy has been discarded, because of the large fuel penalty introduced by pumping losses related to throttle opening reduction.

2.1.4 SOI and VGT combined effects

Summarizing what presented so far, Figures 2.15 and 2.16 show the effects of SOI and VGT variations on exhaust gas temperatures and fuel consump-

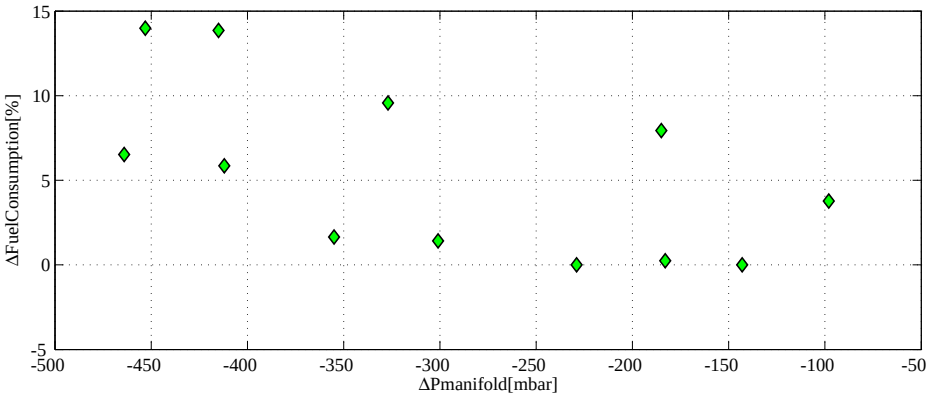


Figure 2.14: Effect of TVA on fuel consumption.

tion, respectively. It is evident that SOI positive variations are essential for achieving a significant increase of the exhaust temperature, while VGT negative variations can compensate for the fuel penalty while at the same time contributing to make the temperature higher.

Given the considerations presented so far, and considering the knowledge gained by simulating the whole system, a possible solution for heating the exhaust line while limiting the fuel consumption penalty can be found by considering the combined effects of SOI and VGT. A more detailed analysis has been carried out in order to define potential heating strategies based on this concept. In Figures 2.17, 2.18 and 2.19, catalyst intake temperature (TcatIN), Fuel Consumption (FC), and NO_x trends are shown, in relation to SOI and VGT, referring to one engine operating point (number 5 in Table 2.2). The main result is that SOI and VGT carry out a synchronous action on TcatIN, increasing the exhaust gas temperature up to about 50 °C with respect to reference conditions, if both control parameters are forced towards their allowed limits. For what concerns the fuel consumption, it is interesting to notice that in a limited region of SOI-VGT combinations, the fuel penalty is almost insignificant, if not even negative (higher fuel efficiency). To complete this scenario, it has also been reported the NO_x trend (Figure 2.19), showing that, as expected, the higher the TcatIN temperature, the lower the combustion chamber temperature, and therefore the NO_x production decreases.

2.2 Heating strategy

Summarizing what has been presented so far, and adding some other considerations, it comes out that if it was possible to separately optimize one parameter at a time, it would be possible to define the following list of

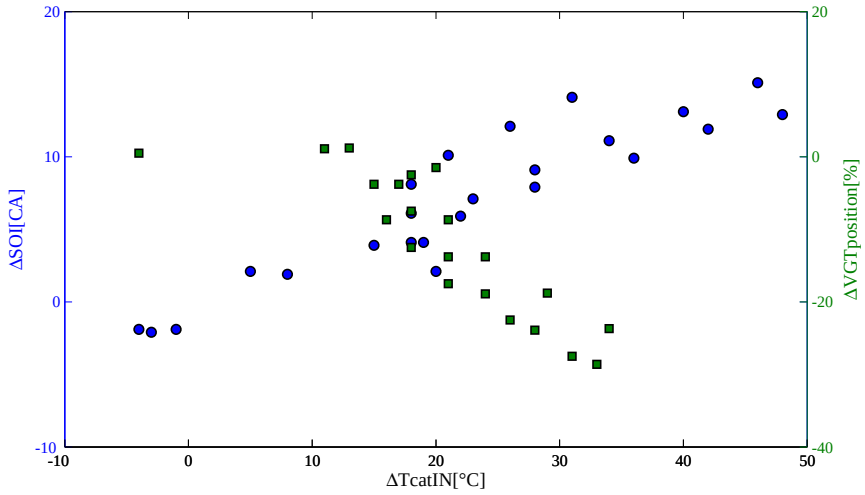


Figure 2.15: Effects of SOI and VGT on Exhaust Temperature Variations.

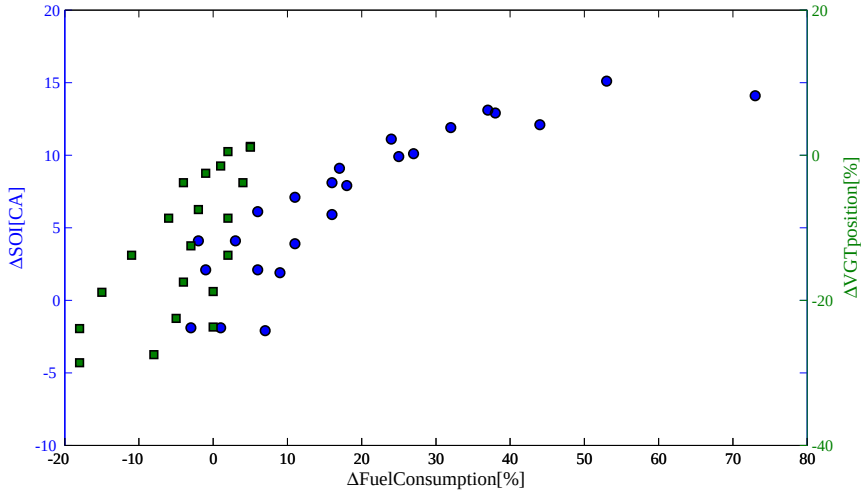


Figure 2.16: Effects of SOI and VGT on Fuel Consumption Variations.

control priorities:

- $max\Delta TcatIN$: to reach the highest admissible exhaust temperature, using both SOI and VGT at their limits, as shown in Figure 2.17.
- $min\Delta FC$: to minimize the fuel consumption without considering exhaust temperature, by trying to work in the small region shown in Figure 2.18.
- $\Delta FC = 0$: rather similar to the previous target, with the slight difference to try to work in a zero fuel penalty region, without searching for its minimum.
- $\Delta Pboost(100)$: by analyzing the effect of VGT on TcatIN (see Figure 2.9) it has been verified that the higher is the VGT opening, the lower is the boost pressure and the higher is TcatIN. It follows that the boost pressure can be reduced, in order to increase TcatIN, but being careful to not exceed smoke limits and drivability constraints (turbo lag). So this control piority refers to a maximum boost pressure reduction of about 100 mbar.

Figure 2.20 shows all these control objectives, represented as SOI-VGT combinations in a diagram prepared as a guideline of a possible control strategy calibration. This synthetic representation makes it quite easy to have a general idea of the engine behavior, when forced to operate under different control (and calibration) conditions. Depending on the priority, a particular region of the plane can be chosen, and not all targets can be achieved at the same time. This graphical representation is a starting point to define optimal heating calibration strategies, to be tested on the vehicle during driving cycle operation.

2.2.1 The concept

The main idea is to apply the results of the previous analysis, while taking into consideration other aspects concerning the vehicle and the driving cycle. For example, all tests performed on the test bench have been executed with a warmed up engine, while the NEDC presents a cold start phase that is quite relevant for what concerns fuel consumption and emissions production. Depending on the engine (and aftertreatment systems) thermal state, the strategy should be able to smoothly adapt SOI and VGT position.

A typical feedforward control strategy based on steady-state actuator setpoint maps is implemented in the ECU; it is thus the starting control structure, as shown in the upper part of Figure 2.21. In order to account for transient operation (thermal transient in this case), the control structure needs to be adapted. The steady-state control maps are spanned by engine

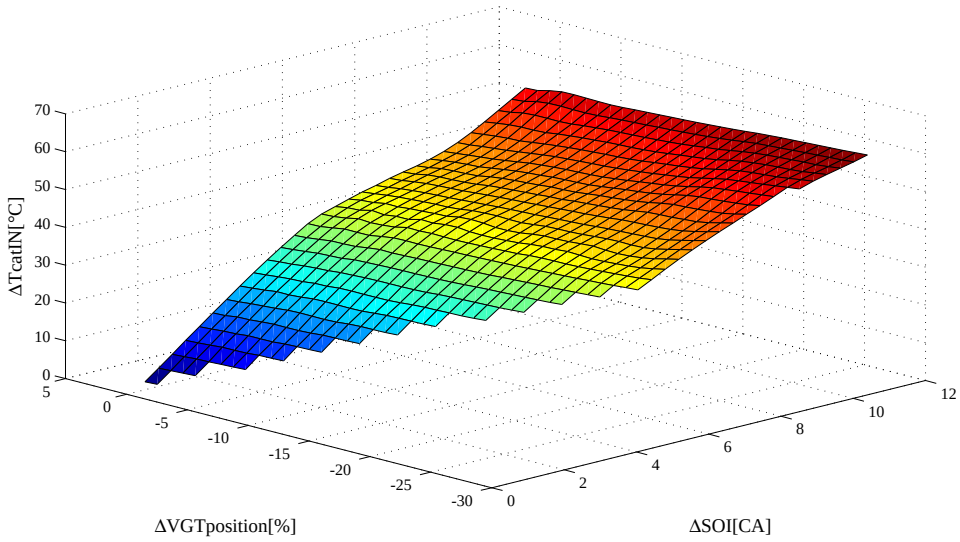


Figure 2.17: Combined effect of SOI and VGT on TcatIN (1800 rpm, 5.03 bmep).

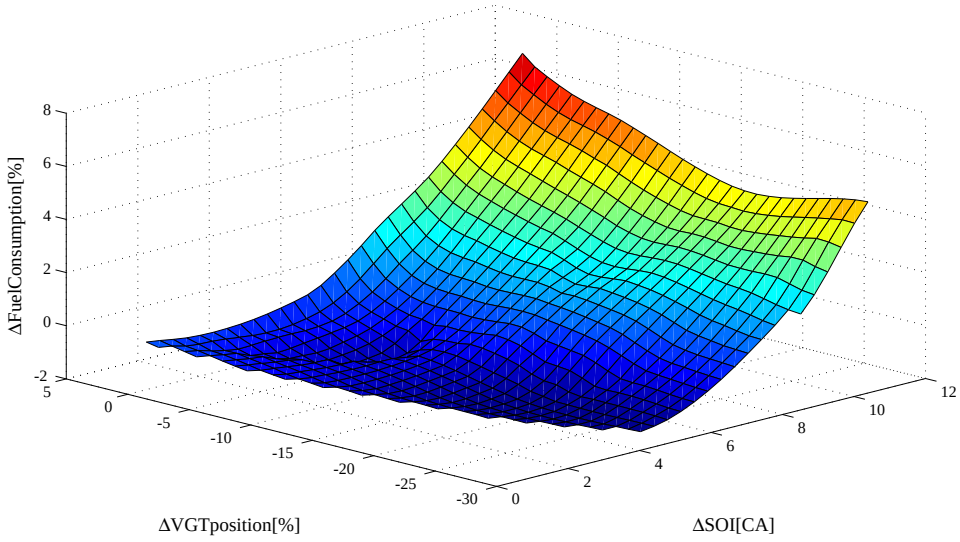


Figure 2.18: Combined effect of SOI and VGT on fuel consumption (1800 rpm, 5.03 bmep).

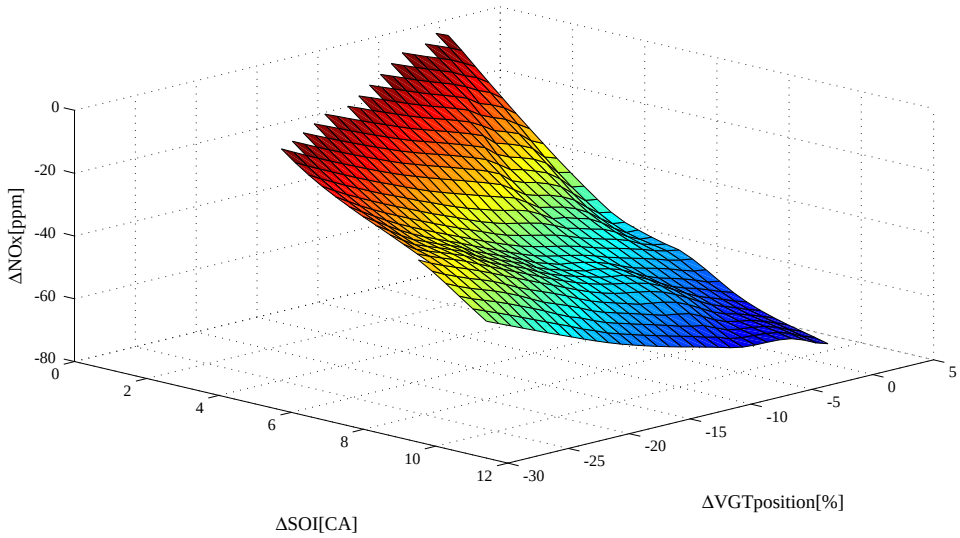


Figure 2.19: Combined effect of SOI and VGT on NO_x (1800 rpm, 5.03 bmep).

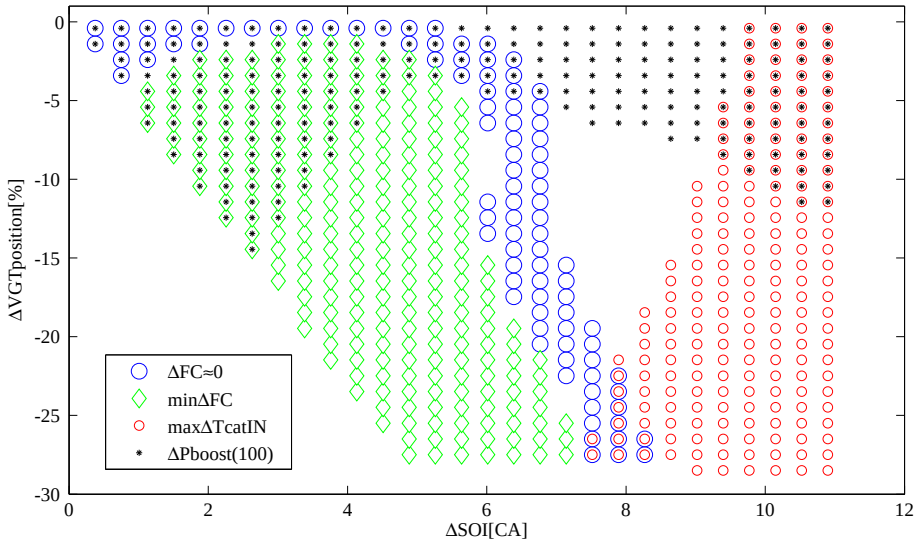


Figure 2.20: Representation of combined VGT and SOI position effects on engine behavior.

speed (N_{eng}) and load (m_{fcc}). The latter variable has not been introduced yet: m_{fcc} is the fuel quantity injected per cycle in each cylinder, and it can be intended as the load request (bmep). If no other corrections were applied, the two control outputs would be $\mathrm{SOI_{ref}}$ and $\mathrm{VGT_{ref}}$, which are solely set-point dependent, regardless of the thermal condition.

On one hand it can be argued that such a control structure is not suitable to account for dynamic phenomena, on the other hand it represents a standard architecture in automotive ECUs. As a consequence, enhancing its control efficacy while keeping the main structure intact may be achieved by adding transient compensation maps. The sensitivity analysis presented in preceding sections may be seen as a tool to derive the compensation maps. More precisely, referring to Figure 2.21, correction maps (spanned by N_{eng} and m_{fcc}) of SOI and VGT have been extracted from the information provided by the calibration table in Figure 2.20. The corrective factors $\mathrm{SOI_{corr}}$ and $\mathrm{VGT_{corr}}$ rely on N_{eng} and m_{fcc} at the same way as the reference values, therefore the thermal-related compensation is given by the third map shown in the figure. Depending on the engine coolant temperature (T_{eng}), the modulation factors $\mathrm{T_{SOI,corr}}$ and $\mathrm{T_{VGT,corr}}$ act on the steady-state correction values, by keeping them unchanged ($\mathrm{T_{i,corr}} = 1$) or eliminating their effect ($\mathrm{T_{i,corr}} = 0$).

Figure 2.22 gives an overview of the ideal control action to be satisfied. A NEDC cycle is considered as a driving profile; as is well known the urban part of the cycle (ECE) is repeated four times, therefore the engine operates in the same operating points (N_{eng} and m_{fcc}). Since the correction maps are spanned by the same input variables, their outputs remain unvaried for each ECE repetition. Supposing to define the target correction with a rectangular area in Figure 2.22, one might want to shift this rectangle over the calibration table, depending on the desired effect (maximize TcatIN versus minimize FC) along the driving profile. This action cannot be accomplished with a single map, therefore the temperature-dependent map helps to fulfil such request.

The bottom plot of Figure 2.22 shows a typical engine coolant temperature trend during the NEDC considered. It can be argued that the results presented so far suit only this specific driving profile, invalidating this approach when considering a different cycle. The global validity is instead guaranteed, since the engine coolant temperature is the key factor used for compensating the set-point related correction factors. Finally, whenever a cold start takes place the heating strategy is activated, speeding up the aftertreatment systems' light-off.

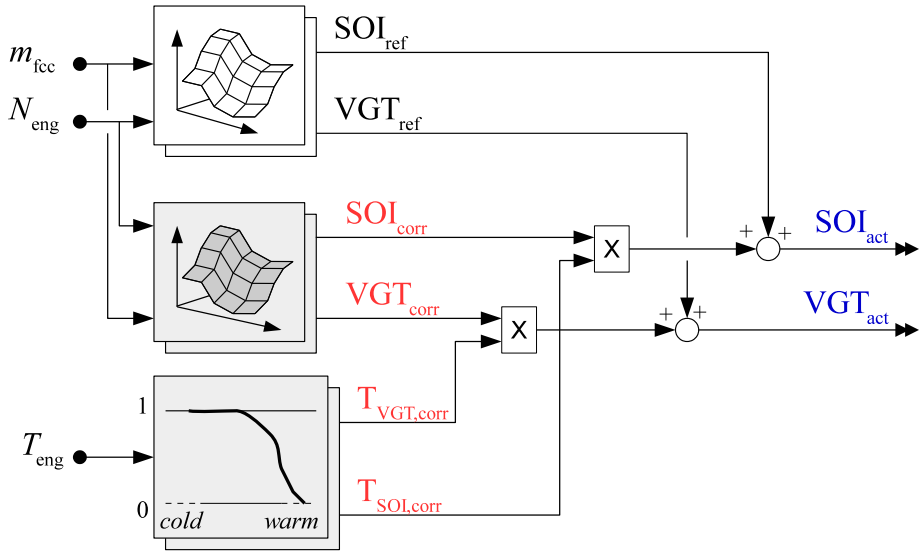


Figure 2.21: Control structure derived for managing SOI and VGT during the engine warm-up.

2.3 Results

Figure 2.23 shows the results achieved by implementing different heating strategies on the vehicle, during roll bench dynamometer NEDC tests. The blue line refers to the original temperature profile without any heating strategy (EU5 reference calibration, also shown in Figure 2.2), while the other ones refer to the various temperature profiles that have been recorded with different control strategies and calibration datasets. In all test cases the target temperature is reached about 600 s earlier than the original behavior, thus highly improving the chances to quickly start catalytic reactions on SCR, enhancing NO_x reduction efficiency. The temperature profiles (and the corresponding heating strategies) differ one from each other because of different SOI and VGT calibrations, confirming the ability to control their effect on exhaust temperature and engine behavior, according to the specific request.

Table 2.3 presents a synthetic comparison in terms of fuel consumption, highlighting that the heating strategy corresponding to the red line ("C") introduces a very limited increase in fuel consumption (+1.47 %). This result is of particular interest, also because "hard" heating strategies based on late or post-injection events typically imply much greater fuel penalties. Heating strategy "C" is based on the implementation of a specific SOI control law, which will be briefly described below.

Some specific considerations need to be expressed, as far as the SOI is concerned. The optimal SOI is typically calibrated on the test bench, as a

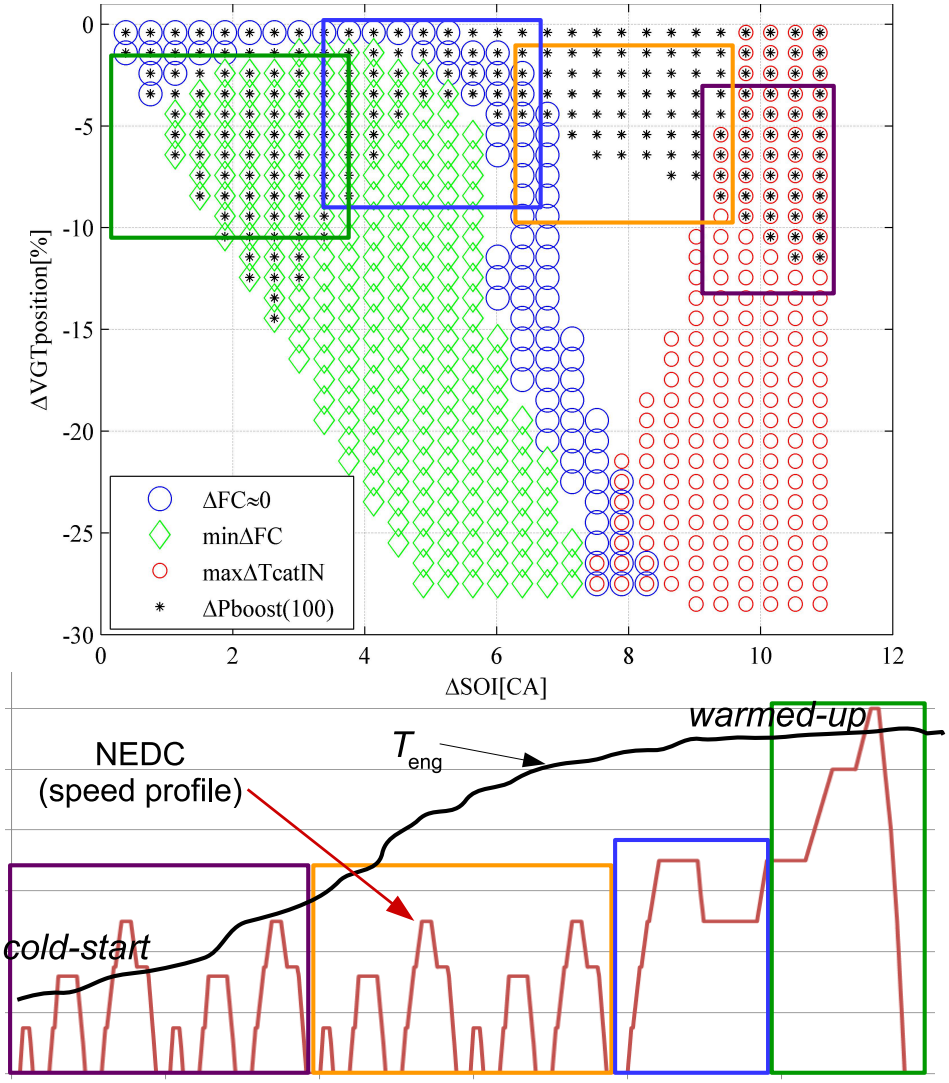


Figure 2.22: Ideal application of different calibration strategies during the engine warm-up. The NEDC cycle has been taken into consideration as exemplary cycle. The calibrations derived as well as the final results achieved can be extended to any driving profile.

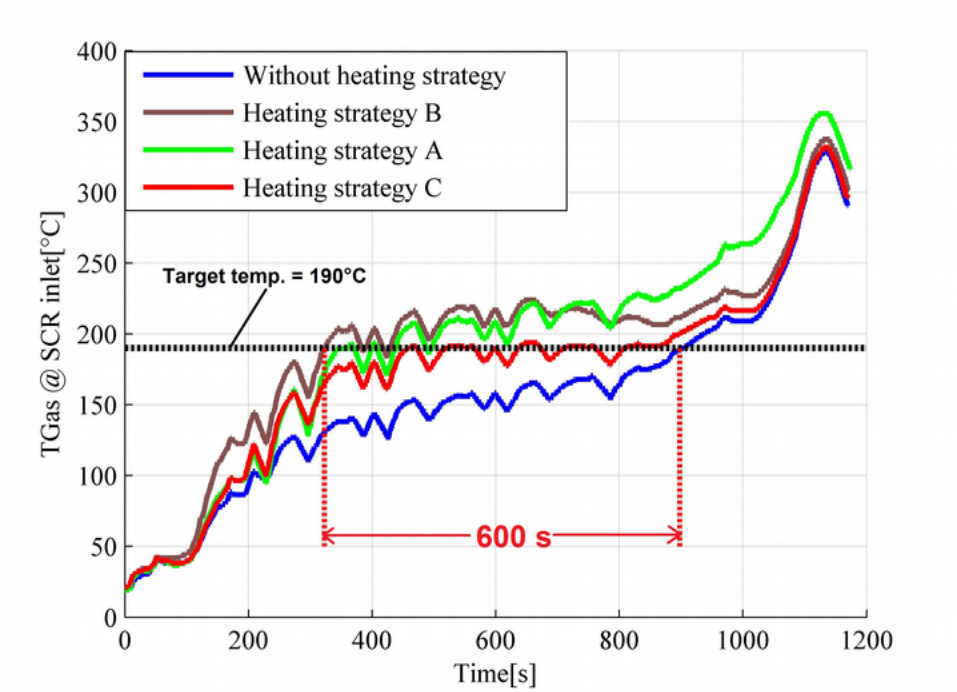


Figure 2.23: SCR inlet temperatures, for NEDC tests with different exhaust line heating strategies.

trade-off between maximum combustion efficiency (minimum fuel consumption) and combustion chamber temperature containment, in order to limit NO_x production. Therefore the ECU cannot always apply maximum efficiency SOI, being forced to delay the combustion process to reduce its temperature. In the system under exam, this dual functionality can be realized by two different control actions, thanks to the SCR system availability, which can reduce NO_x emissions with a higher efficiency if compared to SOI delay action. Referring to the red line (strategy "C"), as soon as the engine temperature approaches the warmed up condition, SOI and VGT corrections gradually decrease. When the final pattern of the driving cycle starts, the engine is substantially warm and the heating strategy adapts again to the new conditions. This part of the cycle is characterized by an higher load, and higher exhaust gas temperatures: the heating strategy is not necessary anymore, and the maximum efficiency SOI is selected, together with a slight increase of VGT opening position, to minimize fuel consumption.

2.3.1 Conclusion

In this chapter, the development of exhaust line heating strategies, based on experimental investigations and previously developed simulation analysis,

Table 2.3: Fuel consumption variation with respect to the reference temperature profile (no heating strategy).

heating strategy:	A	B	C
ΔFuelConsumption[%]*:*	+ 6.76	+ 7.39	+ 1.47

has been presented. The case study is represented by the SCR installation in a small diesel engine exhaust line, to enhance NO_x reduction in an effort to comply with upcoming EU6 regulations. The challenge was to reach a pre-specified temperature (of approximately 190°C) as fast as possible, far away from the exhaust valves and without compromising fuel consumption.

The experimental investigation, supported by a modeling activity that had previously been carried out, allowed the identification of the main effects of the available control parameters on exhaust gas temperature, by considering at the same time eventual negative effects on fuel consumption. The control strategy (and the correspondent calibration dataset) has been based on such approach, and different solutions have been tested on a roll dynamometer, while performing homologation driving cycles (NEDC). The interesting result is that a substantial SCR light-off time reduction (around 600 s) may be achieved with minor fuel penalties, and this may be obtained by implementing a control strategy that is designed to respect different priorities depending on the SCR thermal state. The proposed approach consists of starting with low NO_x production and higher exhaust gas temperature targets until the SCR has been thermally activated, and switching to maximum efficiency priority (with SCR active thermal state constraint) afterwards.

Possible further improvement could be achieved by exploring the effect of the EGR valve partial closing, which can be possible thanks to the increase in NO_x reduction due to the higher efficiency of the SCR after-treatment system. A fuel consumption reduction is foreseen because of the double effect of a better combustion efficiency and a further reduction of VGT closing at constant boost pressure (due to the higher enthalpy available at the turbocharger inlet).

Chapter 3

Dynamic optimization of diesel engines

Optimal control of Diesel engines is becoming increasingly important. More stringent emission regulations [7, 8] require the exploitation of the remaining potential to further reduce the pollutant emissions, especially during transient operation [21, 22]. At the same time, the fuel consumption has to be minimized for economical and environmental reasons. Nowadays, deriving an engine calibration is a highly time demanding task, owing to the many degrees of freedom provided by current diesel engines and to the requirement to consider transient operation. Optimal control is a tool that can support the calibration engineer. The optimal control trajectories indicate directions for the calibration, disclose shortcomings of a chosen control structure and serve as a benchmark.

The common approach to dynamic optimization is to use a large amount of stationary and/or transient measurement data to identify a globally valid model. Relying on this model, the optimal-control problem (OCP) is solved at once. This approach assumes that all influences on the relevant outputs are captured by the model and the selected set of control inputs.

Similar approaches were already used in the early days of automated engine calibration [23, 24]. However, due to limitations in computer hardware and software, a predefined structure of the control system had to be assumed. Furthermore, the transient measurements were used to directly construct gradient information for the optimization, since the vast amount of model evaluations required to numerically solve a nonlinear OCP would have been prohibitive at that time. Nowadays, such computing power is readily available.

This thesis presents the development of an alternative approach to dynamic optimization. It involves an iteration between the dynamic optimization, restricted to the current model-validity range, and the refinement of the model using specific transient measurements [25]. This approach avoids

measurements in irrelevant regions. Furthermore, refining the model using transient measurements on the driving cycle at hand allows accounting for the effect of all influences that are missing in the model. This compensation mechanism allows for the use of a simple mean-value model for the air path as well as empirical setpoint-relative emission models. It is thus not required to develop sophisticated combustion models that have to provide reliable far-field extrapolation. The accuracy of the air-path model is increased by applying a "generalized Kalman filter". All OCPs occurring within the methodology are solved numerically by transcribing the continuous-time problem into a finite-dimensional nonlinear program (NLP).

The drawback of this method is that multiple runs of the transient cycle have to be performed during each iteration. The procedure is thus best suited for compact test cycles containing all relevant transient patterns in a condensed form.

In this chapter, all models and numerical methods required to implement the iterative procedure are described and analyzed. Furthermore, the experimental setup and the testbench-software interfaces are presented. Finally, one iteration of the procedure is executed on an engine testbench and the results are analyzed and discussed.

3.1 Experimental setup

The experimental activity has been carried out on a light-duty engine that is used mainly in EU4 C-segment passenger car. It is a 3.0 l V6 diesel engine, equipped with a high-pressure injection system, a high-pressure cooled EGR (Exhaust Gas Recirculation) circuit and a VGT (Variable Geometry Turbine) turbocharger (Table 3.1).

The engine testbench is equipped with a dynamic brake which allows not only steady-state performance but also transient operation. This characteristic is of crucial importance whenever driving cycles are needed to be carried out, in order to reproduce as closely as possible the real operating condition of the vehicle. The brake controller is remotely connected and driven by a dSPACE rapid prototyping system, in which all algorithms needed to control the desired engine speed and load torque are implemented.

Before focusing on the engine control unit (ECU), it is important to describe the layout of sensors installed on the engine. Generally speaking, every testbench setup follows some common rules or guidelines, which can vary accordingly to the particular research activity conducted. What makes the testbench such an important environment for engine testing, is the possibility of installing several sensors, mainly "laboratory sensors" which are not implemented on the production vehicle, because of their cost, short-term reliability and duration.

Figure 3.1 shows the layout of all sensors that were available when this activity has been carried out. It can be noticed that only a minority of sensors are related to the ECU, which typically are in the minimum number to properly control the engine. In recent control systems there is a growing tendency to use virtual sensors (models) rather than physical ones, since the main objective is to reduce the overall cost of the engine system, while keeping the same amount of information provided by the sensors. Deriving reliable models often relies on the type of sensors installed on the engine during the experiments, for instance the availability of cylinder-pressure signals provides a huge amount of information about the combustion phenomena. Furthermore, several temperature and pressure sensors are usually installed all over the engine, to acquire as much information as possible, which can be used both to identify engine models or to enhance the supervision of the engine. During the calibration activity performed on the testbench, fuel consumption and pollutant emissions are typically measured by means of sophisticated and expensive measurement devices, not usually available in on-board applications. Referring to Figure 3.1, the following sensors and devices have been used during the research activity at hand:

- **pressure sensors based on piezoresistive effect**: widely used for measuring the pressure of fluids (both air and liquid) in various parts of the engine, like in the air-path or in the fuel line. The piezoresistive effect describes change in the electrical resistivity of a semiconductor or metal when mechanical strain is applied. These sensors are robust, reliable and quite inexpensive, therefore they are typically standard sensors (ECU sensors).

- **pressure sensors based on piezoelectric effect**: more sophisticated, accurate and expensive than piezoresistive sensors. Particularly spread in the automotive field for the cylinder-pressure measurement. In this specific case the engine is equipped with six cylinder-pressure sensors, to monitor the combustion process in each cylinder.

- **temperature**: measured by means of thermoresistances and thermocouples. The former are used for on-board application (ECU sensors) while the latter exclusively on the testbench or, more generally, for development applications. Thermocouples are more accurate than thermoresistances, they cover a wider range of temperatures (also high temperatures) and their dynamics may be faster, depending on their size.

- **air-flow meters**: there are several types, based on different principles and dedicated for various applications. Here the standard AFM (Airflow Meter) sensors have been used, one for each engine bank. They are based on the hot-film anemometer measuring principle.

Table 3.1: Main data of the engine used for this research activity.

engine type:	Diesel, V6
displacement:	2987 cm^3
EGR:	high pressure, cooled
compression ratio:	15.5
turbocharger:	VGT, charge-air cooler
bore/stroke:	83/92 mm
injection:	CDI 4, max. 1600 bar
valves/cyl.:	4
max. torque:	400 Nm (1400–3800 rpm)
rated power:	165 kW (3800 rpm)

- **turbocharger speed**: there are various kinds of possible solutions, based on different measuring principles. In this case, it is an eddy current turbocharger speed measurement system (micro-epsilon turboSPEED DZ135).
- **Nitrogen oxide (NOx) emissions**: the NO_x emissions are measured by means of two different types of devices, namely a Cambustion fNOx 400 Fast CLD and a UniNOx sensor. The accurate but slow signal from the latter has been used to calibrate the fast but drifting signal of the fNOx 400.
- **Soot emissions**: the soot emissions are recorded by means of an AVL Micro Soot Sensor. Further details about this measurement device are presented in (4.1.3).

A quite comprehensive list of sensors typically employed in automotive, along with some fundamental measuring principles, are presented in Appendix A.

Electronic Control Unit (ECU) The engine is controlled by a development ECU, which slightly differs from the end of line version implemented on the production vehicle. The Ethernet-based ETK interface by ETAS provides direct access to the control variables and parameters of the ECU via the parallel data and address bus, or via a serial microcontroller testing or debugging interface. Due to its extremely compact design, the ETK can be accommodated inside the housing of the production ECU. Moreover, the ETK interface is real-time capable, and provides a universal ECU interface for sophisticated applications in the development and calibration of engine ECUs. The interface is supported across the board by ETAS hardware modules, the INCA calibration tool (Figure 3.2) as well as the INTECRIO and ASCET development tools. The high standards in terms of performance, safety, responsiveness, drivability, fuel savings and emissions met by today's

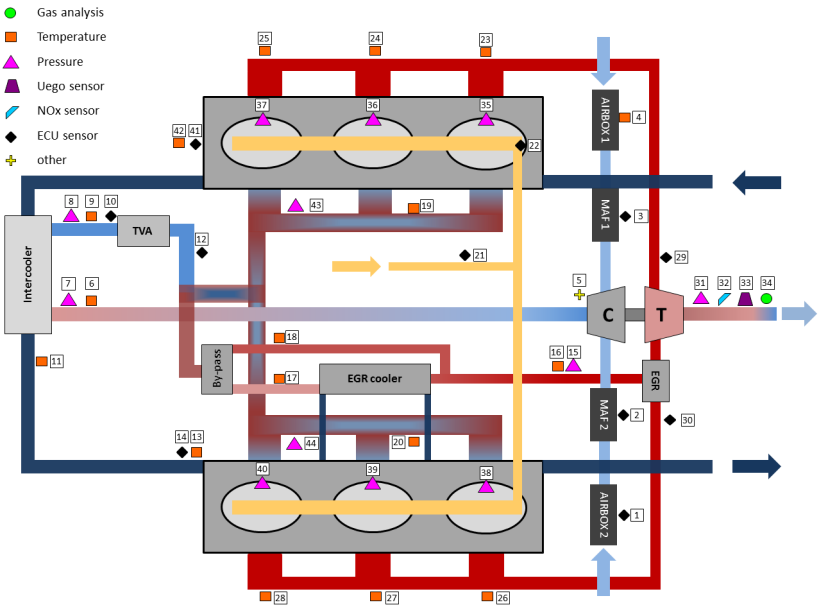


Figure 3.1: Layout of the diesel engine used in this work, showing the various types of sensors installed.

vehicles would not be attainable without the deployment of ECUs featuring a multitude of sophisticated functions. Powerful measuring and calibration access constitutes an essential prerequisite for developing these functions and calibrating the function parameters. With INCA, ETAS offers a flexible family of software products for calibration, diagnostics, and validation. The need to develop, add and/or bypass some control structures implemented on the ECU has been fulfilled by the utilization of a Rapid Prototyping Module (ES910), in combination with the software package INCA-EIP (Experimental Target Integration Package). INCA-EIP enables real-time function developers to use the measurement and calibration functionality of INCA while the ECU software functions are executed on the ES910. The ES910 prototyping and interface module combines high computing performance with all common ECU interfaces in a compact and robust housing (Figure 3.3). A schematic representation of its interaction with the ECU and the Host PC is shown in Figure 3.4. CAN and LIN interfaces provide the connection of the ES910 module to the ECU bus and to the external CAN-modules (Input/Output operations). This setup provides complete control over all relevant control inputs of the engine.

The prototyping tool INTECRIO is used to set up the rapid-prototyping models, previously designed in Simulink (MathWorks), on the ES910. During the INTECRIO experiment the user has runtime access to the model

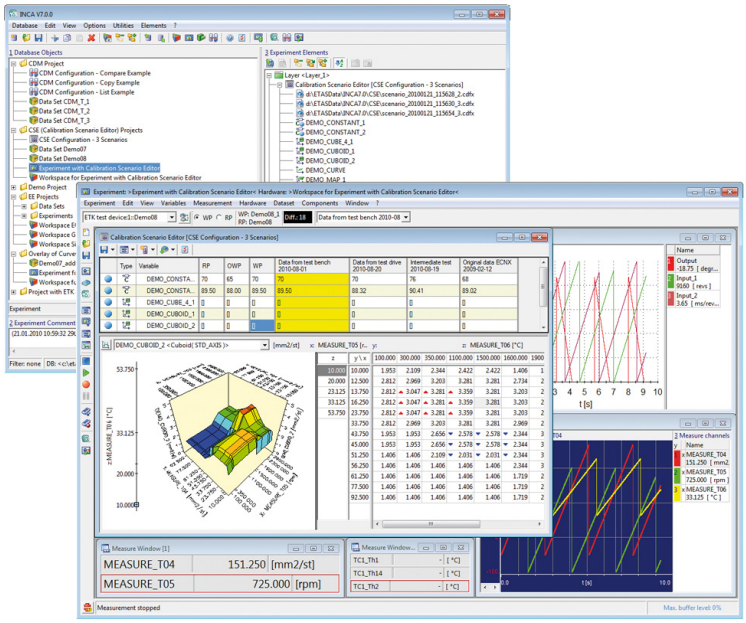


Figure 3.2: Screen shots of the calibration tool INCA.

executed by the ES910 module. With such an experimental setup any ECU function can be bypassed, allowing for the development and the implementation of custom control algorithms, needed to perform the dynamic optimization procedure that is presented in this chapter.

3.2 Engine model

From a modeling point of view, the engine treated in this work can be seen as a nonlinear dynamic system with two dynamic feedback loops coupled to each other, namely the EGR and the turbocharger. Referring to Figure 3.1, the fresh air path (light blue and blue) and the exhaust gas path (red)

Figure 3.3: ETAS ES910 prototyping and interface module.

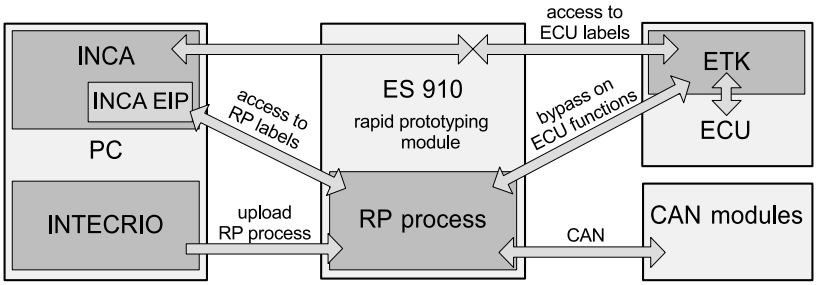


Figure 3.4: Hardware, software, and interfaces of the testbench setup.

are not separate, but connected in different ways. The first coupling is due to the mechanical connection between the compressor and the turbine: the enthalpy in the exhaust gas is transformed by the turbine into mechanical energy (torque) to speed up the turbocharger, and then this mechanical energy is transferred to the intake air by the compressor. The EGR system extracts part of the exhaust gas flow (upstream of the turbine for high-pressure EGR) and drives it to the intake manifold (downstream of the compressor). So, on the one hand the enthalpy related to this EGR mass flow is not available anymore to the turbine to drive the turbocharger, but on the other hand it increases the pressure in the intake manifold and thus also the mass flow into the cylinders (and the composition).

Due to this coupling, the effects of the two actuators associated with the turbocharger and the EGR depend on each other. The EGR valve controls the EGR mass flow, while the turbocharger is equipped with a variable-geometry turbine (VGT). Thanks to the adjustable vanes position, the opening area of the turbine can be modified in order to change the restriction of the flow. While the EGR valve positioning has an intuitive convention, as EGR=0% means fully closed and no gas recirculation, a clarification for the VGT position might be useful for the reader. In the given control system, a VGT position=100% corresponds to a fully closed condition, and vice versa the distributor is fully open with a VGT position=0%. In other words, when the VGT is fully closed (100%) it means that the turbine, under steady state flow conditions, generates the highest backpressure correspondent to that mass flow rate, and therefore the pressure drop through the turbine is at the maximum. In this condition, the turbine produces its maximum power, increasing the turbocharger speed, with the consequent increase of the intake manifold pressure.

To recap, capturing and modeling all the physical processes in a modern diesel engine, equipped with VGT and EGR, is a demanding task. The final purpose has to be clear from the beginning, when deriving an engine model, since its structure and characteristics are strongly related to its intended use. A model able to catch the waves effect in the intake and exhaust manifolds,

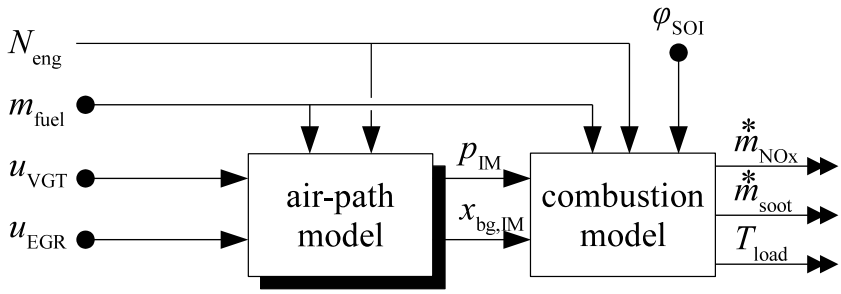


Figure 3.5: Structure of the engine model.

or the emission species formation into the combustion chamber, is hugely more complex than a linearized, control-oriented model that runs online on the ECU. The model derived in this work is close to the latter category, but it may be defined as an optimization-oriented rather than a control-oriented model. The intended use of the model is to be part of a dynamic optimization framework, in which an iterative optimization routine is run by using the engine model as plant. For the results of the optimization to be of any relevance, the model on which the optimization relies has to be sufficiently accurate and must capture all influences of each control input. The vast amount of model evaluations during the iterative optimization requires the model to be fast, which is primarily achieved by keeping the structure simple. Finally, the model has to be smooth since an efficient optimization relies on (first and even second) derivatives. The combustion processes are especially hard to describe by models that combine all those requirements.

In Figure 3.5 the schematic of the engine model derived is shown. The air-path model is presented in details in Sec. 3.2.1, and the combustion model in Sec. 3.2.4. The model can be subdivided into a dynamic and a static part. The subscripts denote the intake and exhaust manifolds (IM/EM), the volume between the turbine and the exhaust-gas aftertreatment system (ATS), and the turbocharger (TC). The positions of the variable-geometry turbine (VGT) and the exhaust-gas recirculation valve (EGR) are pulse-width modulation signals with a range of [0,1]. The SOI is specified in degrees before top dead center. Finally, the burnt-gas fraction in the intake manifold is denoted by $x_{\mathrm{bg,IM}}$.

All dynamics are induced by the air path, i.e. the turbocharger inertia and the volumes. The full model is represented in state-space form as:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{3.1a}$$
$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{3.1b}$$

where $n_{\mathrm{x}} = 6$ state variables, $n_{\mathrm{u}} = 4$ control inputs, and the state, control

and output vectors are:

$$\boldsymbol{x} = \left(p_{\mathrm{IM}}, \vartheta_{\mathrm{IM}}, x_{\mathrm{bg,IM}}, p_{\mathrm{EM}}, p_{\mathrm{ATS}}, \omega_{\mathrm{TC}}\right)^T,$$
$$\boldsymbol{u} = \left(u_{\mathrm{VGT}}, u_{\mathrm{EGR}}, \varphi_{\mathrm{SOI}}, m_{\mathrm{fuel}}\right)^T, \quad \boldsymbol{y} = \left(\overset{*}{m}_{\mathrm{NOx}}, \overset{*}{m}_{\mathrm{soot}}, T_{\mathrm{load}}\right)^T.$$

The engine speed (N_{eng}) is treated as time-varying parameter.

3.2.1 Mean-value model of the air-path

The models for the air path are based on many related works. Most of the modeling components are based on [7], while other publications are referenced along the text.

In the following paragraphs, all the air-path components needed to build the engine model are presented. Their mathematical formulation is described, highlighting the set of parameters that have to be determined for each model. The numerical methods that have been applied to identify such parameters are detailed later, while the set of experimental measurements, needed to run the identification routines as well as the validation runs, is now described. In addition, Figure 3.6 shows a schematic representation of the air-path model that has been derived. It may be helpful to go through the nomenclature used when describing the models.

Measurement data Figure 3.7 shows an overview of the measuring points, spanned by engine speed and load torque. Once the engine is warmed-up, the first measuring point is reached by setting the engine speed, on the test-bench management system, and the torque request on INCA. Afterwards, the engine is considered to be in a stationary condition when the variations of the main physical variables (intake/exhaust temperature and pressure, turbocharger speed, etc.) are not relevant. Finally the first measurement can start. The recording time is at least 20 seconds, so that meaningful average values can be calculated during the post-processing phase. In order to discard meaningless recorded data, resulting for instance from a not real steady-state condition, the post-processing code calculates, besides the average value, the CoV (Coefficient of Variation), and highlights those measurements with an over threshold CoV (defined by the user). All the scripts are written in MatLab 2012b.

The EGR and VGT positions are controlled manually, bypassing the ECU control functions. In doing so, sweeps of both actuators can be performed. A set of measurements is performed without exhaust-gas recirculation, by keeping the EGR valve closed, while VGT sweeps are applied. For each measuring point, approximately ten VGT positions are swept and, obviously, recorded. The following models are identified and validated by using this set of data: *compressor mass-flow, compressor enthalpy, turbine mass-flow, turbine efficiency, restriction mass-flow, cylinder mass-flow, engine-out temperature*. As a general rule, half of the measurement data are used

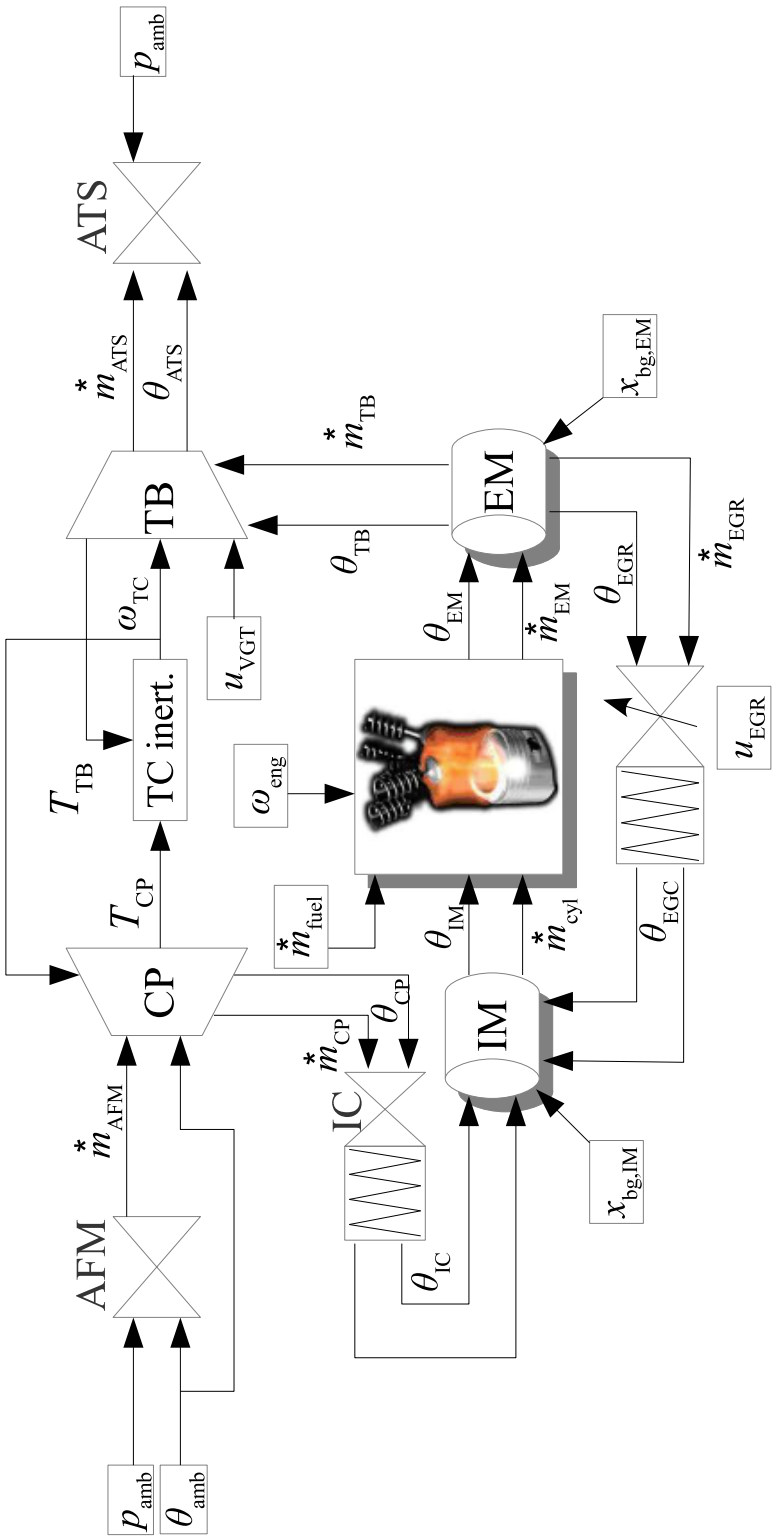


Figure 3.6: Schematic of the air-path model described in this work.

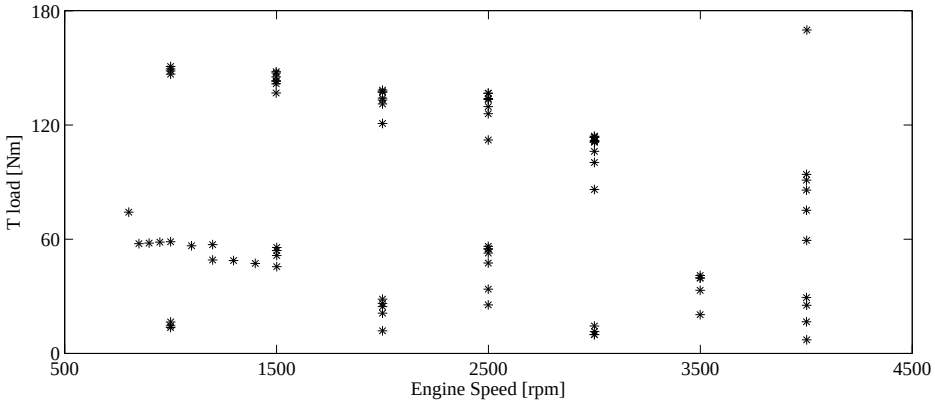


Figure 3.7: Grid of operating points measured on the testbench for identification and validation of the models. For each measuring point (overall roughly 90), the engine is brought in steady-state condition, and then VGT and EGR sweeps are performed.

for the identification, and the other half for the validation. This approach ensures that the prediction quality of the model, assessed by the validation data, is unrelated to the measurements carried out. Once these models are identified, only the *EGR mass-flow* and the *EGR cooler* remain to be analyzed. Similar to the previous case, EGR sweeps are performed, from the lowest mass-flow through the EGR valve to the highest (but tolerable by the engine) one.

Looking back to Figure 3.7 it can be noticed that, while the engine speed range spanned is coherent with its speed limits (1000-4000 rpm), the load torque is significantly below the reachable limit. This apparently not spread experimental plane does not invalidate the results obtained, because the operating region is defined by the levels of pressure and temperatures reached in the intake and exhaust manifolds. In other words, by sweeping the VGT position the extreme conditions for the turbocharger speed can be reached, bringing the intake manifold pressure up to its limit (approx. 2.7 bar absolute). Once the border conditions are reached, there is no need to further increase the output torque (by increasing the fuel injected), especially considering that the measured torque is not accounted for when deriving the air-path models.

Compressor mass-flow The model is based on [26]. According to the original literature, the ellipse equation should be solved for the pressure ratio, if a surge model is included. The engine is operated strictly out of the surge region of the compressor, thanks to constraints imposed by the optimal control problem. Hence, surge has not been considered in this case,

resulting in the solution of the ellipse equation for the mass flow instead of the pressure ratio. The speed lines need to have a positive slope towards lower mass flows. The equations of the original model are presented below:

$$\Pi_\mathrm{s} = 1 + k_{\pi,1} \cdot n_\mathrm{TC}^{k_{\pi,\mathrm{e}}} \tag{3.2a}$$

$$\overset{*}{m}_\mathrm{s} = 0 + k_{\mathrm{m},1} \cdot n_\mathrm{TC}^{k_{\mathrm{m,e}}} \tag{3.2b}$$

$$\overset{*}{m}_\mathrm{max} = k_{\mathrm{max},0} + k_{\mathrm{max},1} \cdot n_\mathrm{TC} \tag{3.2c}$$

$$\overset{*}{m}_\mathrm{CP} = \overset{*}{m}_\mathrm{s} + (\overset{*}{m}_\mathrm{max} - \overset{*}{m}_\mathrm{s}) \cdot \left[1 - \left(\frac{\Pi_\mathrm{CP}}{\Pi_\mathrm{s}}\right)^{c_2}\right]^{\frac{1}{c_1}} \tag{3.2d}$$

where the index s denotes the surge limit. For every turbocharger speed n_TC, these equations define the ellipse-like curve:

$$\left(\frac{\Pi_\mathrm{CP}}{\Pi_\mathrm{s}}\right)^{c_2} + \left(\frac{\overset{*}{m}_\mathrm{CP} - \overset{*}{m}_\mathrm{s}}{\overset{*}{m}_\mathrm{max} - \overset{*}{m}_\mathrm{s}}\right)^{c_1} = 1. \tag{3.3}$$

where the denominators are the semi-axes (for an ellipse $c_1 = c_2 = 2$).

By adopting this kind of model formulation, two potential problems can arise. The first one has to do with the surge line, because when it is approached, a high sensitivity from pressure ratio to mass flow can easily result. Moreover, when the surge line is crossed, the mass flow is not even defined. The solution found has been to shift the ellipses, in order to obtain a constant slope at the surge line, and then to extend it to the abscissa $\overset{*}{m}_\mathrm{CP} = 0$.

The second problem concerns the extrapolation of the speed lines. For the model identification, an appropriate set of experimental data has been used, in order to achieve good interpolation behavior of the model. Outside this experimental region, the model needs to extrapolate, leading to unstable behavior. For instance the mass flow could increase with an increasing pressure ratio. Given these considerations, the model (shown in Figure 3.8 a)) has been modified by replacing the surge line with a vertical line $\overset{*}{m}_\mathrm{s} = \text{const.}(< 0)$, which is achieved by modifying equations (3.2a) and (3.2b) to:

$$\Pi_\mathrm{s} = k_{\pi,0} + k_{\pi,1} \cdot n_\mathrm{TC}^{k_{\pi,\mathrm{e}}} \tag{3.4a}$$

$$\overset{*}{m}_\mathrm{s} = k_{\mathrm{m},0}(< 0) \tag{3.4b}$$

To provide sufficient degrees of freedom to the model, the shape of the ellipses may change as a function of the turbocharger speed:

$$c_2 = c_{2,0} + c_{2,1} \cdot n_\mathrm{TC} + c_{2,2} \cdot n_\mathrm{TC}^2. \tag{3.5}$$

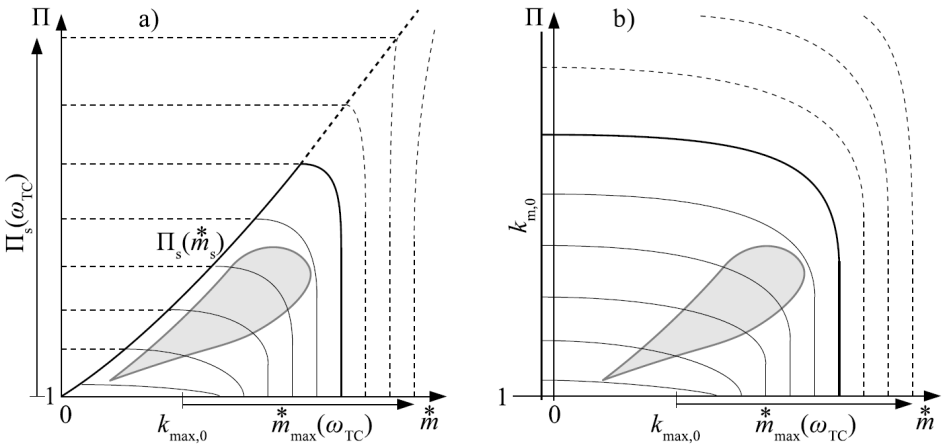


Figure 3.8: a)Original compressor mass-flow model, critical extrapolation of speed lines (dashed speed-lines) and usual operating region on the engine (shaded area). b) Modified compressor mass-flow model.

The modified version of the model, shown in Figure 3.8 b), has been found to be reliably identified by an automated fitting procedure, and exhibits good accuracy for all mass-flow ranges and sufficient flexibility with a small number of parameters.

For the identification, the compressor map, usually provided by the manufacturer, was not available, therefore only experimental data have been used. Stationary measurements have been performed as explained above. During these experiments the EGR valve has been kept closed, while VGT sweeps have been carried out with the aim of exploring all the operating region of the compressor. A compromise between the absolute and the relative squared error has been used as the objective function. The weighting between the two has been tuned to obtain a balanced fit, which showed a satisfying accuracy at low as well as at high mass-flows. The simplex method in MATLAB's `fminsearch` revealed itself to be successful in finding a suitable set of parameters. Multiple iterative runs have been performed, each one initialized with the optimized parameter set from the previous iteration. Figure 3.9 shows the results of the identification, while the parameters identified are listed in the final Table 3.2.

Compressor enthalpy The specific enthalpy increase over the compressor can be expressed by the isentropic enthalpy change, and the isentropic

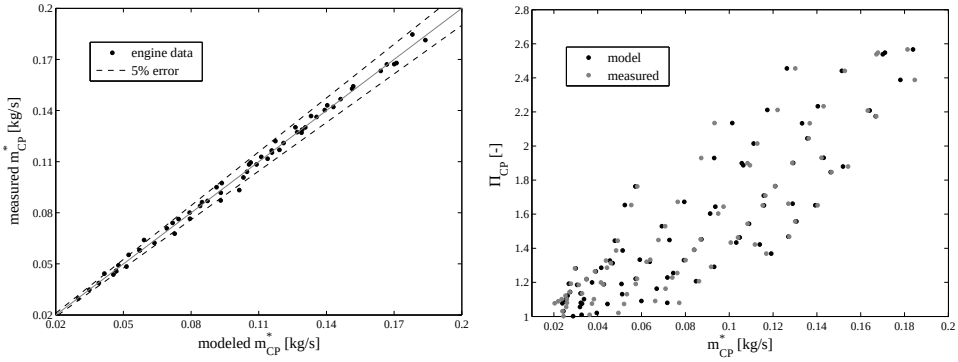


Figure 3.9: Results of the identification of the compressor mass-flow model.

efficiency:

$$\Delta h_{\mathrm{CP}} = c_{\mathrm{p}} \cdot (\vartheta_{\mathrm{CP,out}} - \vartheta_{\mathrm{CP,in}}) \tag{3.6a}$$

$$\Delta h_{\mathrm{CP,is}} = c_{\mathrm{p}} \cdot \vartheta_{\mathrm{CP,in}} \cdot \left(\Pi_{\mathrm{CP}}^{\frac{\kappa-1}{\kappa}} - 1 \right) \tag{3.6b}$$

$$\eta_{\mathrm{is}} = \frac{\Delta h_{\mathrm{CP,is}}}{\Delta h_{\mathrm{CP}}} \tag{3.6c}$$

Instead of this common formulation, the specific enthalpy transferred to the gas can alternatively be calculated through a momentum balance [27], resulting in the Euler equation:

$$\Delta h_{\mathrm{CP}} = \mathrm{U}_2 \mathrm{C}_{\vartheta 2} - \mathrm{U}_1 \mathrm{C}_{\vartheta 1} \tag{3.7}$$

where U_2 is the speed of the blade tip. The tangential speed of the gas at the inlet $\mathrm{C}_{\vartheta 1}$ is assumed to be zero (axial flow). The tangential speed of the gas at the outlet $\mathrm{C}_{\vartheta 2}$ may be described by the radial component, and the back-sweep angle (BSA) of the blade (assuming no slip):

$$\mathrm{C}_{\vartheta 2} = \mathrm{U}_2 - \mathrm{C}_{\mathrm{r}2} cot(\beta_2) \tag{3.8}$$

$\mathrm{C}_{\mathrm{r}2}$ is proportional to the mass-flow divided by the density at the impeller outlet. Research in [28] shows how the assumption $\rho_2 = \rho_1$ may be reasonable, so if (3.8) is inserted in (3.7), all constant parameters are merged, and a constant slip factor is introduced, the specific enthalpy increase over the compressor becomes:

$$\Delta h_{\mathrm{CP}} = k_{\mathrm{slip}} \cdot \left[\mathrm{U}_2^2 - k_{\mathrm{BSA}} \cdot \mathrm{U}_2 \cdot \frac{\dot{m}}{\rho_1} \right] \tag{3.9}$$

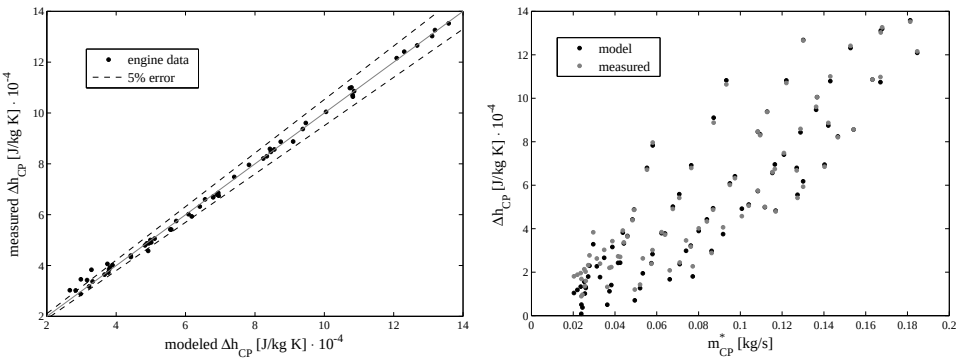


Figure 3.10: Results of the identification of the compressor enthalpy model.

Considering that for single-stage turbocharging the gas density at the inlet ρ_1 can reasonably be assumed constant, and rearranging the equations, the following compressor enthalpy model has been derived:

$$\Delta h_{\mathrm{CP}} = k_{\mathrm{h},1} \cdot n_{\mathrm{TC}}^2 - k_{\mathrm{h},2} \cdot \overset{*}{m} \cdot n_{\mathrm{TC}} \tag{3.10}$$

Once Δh_{CP} is known, the temperature after the compressor can be calculated by means of (3.6a) as:

$$\vartheta_{\mathrm{CP,out}} = \vartheta_{\mathrm{amb}} + \frac{\Delta h_{\mathrm{CP}}}{c_{\mathrm{p}}}. \tag{3.11}$$

The two parameters of the model have been identified by a linear least-squares (LSQ) regression, since (3.10) is linear in the parameters $k_{\mathrm{h,i}}$. Since c_{p} appears in both formulations, it is only used to yield a physical quantity (enthalpy). Moreover its value has been chosen to be constant at $c_{\mathrm{p}} = 1000\ \mathrm{J/(kg \cdot K)}$, since the inlet temperature hardly changes. The identified parameters are shown in Table 3.2. The fit of the model is shown in Figure (3.10).

Intercooler The mass flow through the intercooler has not been modeled explicitly. Therefore, the pressure drop due to the intercooler can be taken into account directly in the compressor mass-flow model, by using the pressure value in the intake manifold when defining Π_{CP}. In other words, the pressure ratio in (3.2d) is $\Pi_{\mathrm{CP}} = p_{\mathrm{IM}}/p_{\mathrm{amb}}$. By doing so, one state variable (pressure between compressor and intercooler) and an orifice function can be eliminated. The intercooler model is based on a simple stationary energy balance:

$$k_{\mathrm{IC}} \cdot \left(\frac{\vartheta_{\mathrm{CP}} + \vartheta_{\mathrm{IC}}}{2} - \vartheta_{\mathrm{cool}} \right) = \overset{*}{m}_{\mathrm{CP}} \cdot (\vartheta_{\mathrm{IC}} - \theta_{\mathrm{CP}}) \tag{3.12}$$

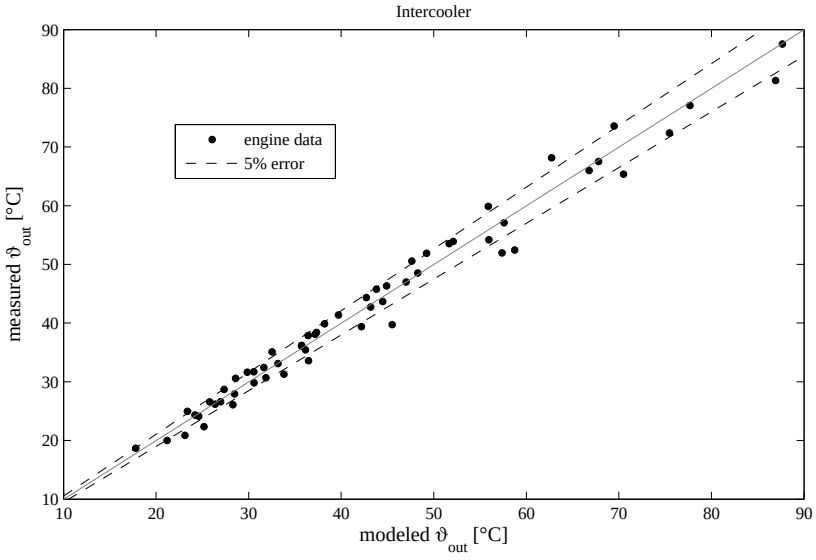


Figure 3.11: Comparison between modeled and measured temperature after the intercooler (ϑ_{IC}).

As representative temperature, the average between the inflow and outflow temperatures has been used. The efficiency parameter k_{IC} depends on the mass flow, so the model has been derived as:

$$k_{\mathrm{IC}} = k_{\mathrm{IC},0} + k_{\mathrm{IC},1} \cdot \overset{*}{m}_{\mathrm{CP}} + k_{\mathrm{IC},2} \cdot \overset{*}{m}{}^{2}_{\mathrm{CP}}. \tag{3.13}$$

The value of k_{IC} has been calculated from the measurement data and an LSQ regression has been applied to identify the model. Figure 3.11 shows the comparison between measured and modeled outflow temperature.

Turbine mass-flow The turbine mass flow can be treated as an orifice, with a variable area in the case of variable nozzle turbines and a blade speed dependent discharge coefficient c_{d}. While in the compressor the mass flow is driven by its rotation, in the case of the turbine the difference between the upstream and downstream pressure drives the fluid in a nonlinear way through the restriction. A detailed description of fluid dynamics formulations used for deriving the turbine mass-flow model can be found in [7]. Only the main definitions are summarized below, starting from the dependent variables of the model chosen:

1. $\Pi_{\mathrm{TB}} = p_{\mathrm{TB,in}}/p_{\mathrm{TB,out}}$: the pressure drop across the turbine is the main factor responsible for the flow generation.

2. u_{VGT}: in the case of variable nozzle turbines, the orifice area varies accordingly to the VGT position.

3. n_{TC}: the turbine rotation restricts the flow due to centrifugal acceleration, therefore the discharge coefficient c_{d}, which appears in the orifice equation, is blade speed dependent.

The variable Π_{TB} (1) is included in the flow function term $\Psi(\cdot)$, which in this case has been considered in its simplified form

$$\Psi_{\mathrm{TB}} = \sqrt{\frac{2}{\Pi_{\mathrm{TB}}} \cdot \left(1 - \frac{1}{\Pi_{\mathrm{TB}}}\right)} \tag{3.14}$$

Concerning the mass flow $\overset{*}{m}_{\mathrm{TB}}$, usually the normalized mass flow

$$\overset{*}{\mu}_{\mathrm{TB}} = \overset{*}{m}_{\mathrm{TB}} \cdot \frac{\sqrt{\vartheta_{\mathrm{TB,in}}}}{p_{\mathrm{TB,in}}} \tag{3.15}$$

is used instead, to make the measurement data obtained on the engine comparable to the manufacturer-supplied turbine map. In order to apply an LSQ regression, the following selection of regressors has been found suitable to fit the experimental data:

$$\overset{*}{\mu}_{\mathrm{TB}} = (k_{\mathrm{TB},0} + k_{\mathrm{TB},1} \cdot u_{\mathrm{VGT}} + k_{\mathrm{TB},2} \cdot {u_{\mathrm{VGT}}}^2) \cdot \Psi_{\mathrm{TB}} + k_{\mathrm{TB},3} \cdot n_{\mathrm{TC}} \tag{3.16}$$

Since the EGR mass-flow is still null (EGR valve closed), the entire mass of engine-out gas flows through the turbine, therefore the exhaust mass flow is simply the sum of the fresh air and fuel mass-flows: $\overset{*}{m}_{\mathrm{TB}} = \overset{*}{m}_{\mathrm{CP}} + \overset{*}{m}_{\mathrm{fuel}}$. Results of the identification are shown in the left-hand plot of Figure 3.12.

Turbine efficiency The turbine efficiency can be expressed as:

$$\eta_{\mathrm{TB}} = \frac{\Delta h_{\mathrm{TB}}}{\Delta h_{\mathrm{TB,is}}} = \frac{c_{\mathrm{p}} \cdot (\vartheta_{\mathrm{TB,in}} - \vartheta_{\mathrm{TB,out}})}{c_{\mathrm{p}} \cdot \vartheta_{\mathrm{TB,in}} \cdot \left(1 - \Pi_{\mathrm{TB}}^{\frac{1-\kappa}{\kappa}}\right)} \tag{3.17}$$

It could be easily calculated from experimental data, since each term is measured and a constant value for κ is assumed. However, the enthalpy change over the turbine Δh_{TB} may be overestimated when using the exhaust temperatures directly (upstream and downstream of the turbine), because of the heat losses in the exhaust manifold. In other words, the temperature decrease due to heat losses is attributed to the enthalpy that the turbine extracts from the gas, leading to an overestimated turbine power:

$$P_{\mathrm{TB}} = w_{TC} \cdot T_{\mathrm{TB}} = \Delta H_{\mathrm{TB}} = \overset{*}{m}_{\mathrm{TB}} \cdot \Delta h_{\mathrm{TB}} \tag{3.18}$$

This effect is ever more emphasized the higher is the turbine case temperature, therefore another approach has been adopted. Considering the energy balance across the turbine

$$\Theta_{\mathrm{TC}} \cdot \frac{dw_{\mathrm{TC}}}{dt} = T_{\mathrm{TB}} - T_{\mathrm{CP}} - T_{\mathrm{fric}} = \frac{\Delta H_{\mathrm{TB}} - \Delta H_{\mathrm{CP}} - \Delta H_{\mathrm{fric}}}{w_{TC}} \tag{3.19}$$

in a steady-state condition (i.e. $\frac{dw_{\mathrm{TC}}}{dt} = 0$) it becomes:

$$\Delta H_{\mathrm{TB}} = \Delta H_{\mathrm{CP}} + \Delta H_{\mathrm{fric}} \tag{3.20}$$

Sometimes, a more general mechanical efficiency is used instead of a friction torque. However, supposing the neglection of the friction contribution, it can be imposed that the enthalpy change over the turbine matches the enthalpy that the compressor transfers to the air, resulting in:

$$\Delta h_{\mathrm{TB}} = \frac{\overset{*}{m}_{\mathrm{CP}} \cdot \Delta h_{\mathrm{CP}}}{\overset{*}{m}_{\mathrm{TB}}} \tag{3.21}$$

This matching ensures an accurate prediction of the turbocharger speed and eliminates the influence of stationary heat losses in the exhaust manifold. Given the considerations so far, the following LIP (linear in parameters) model has shown a good fit (right-hand plot of Figure 3.12) with the measurement data:

$$\eta_{\mathrm{TB}} = h_{\mathrm{TB},0} + h_{\mathrm{TB},1} \cdot u_{\mathrm{VGT}} + h_{\mathrm{TB},2} \cdot u_{\mathrm{VGT}}^2 + h_{\mathrm{TB},3} \cdot \Pi_{\mathrm{TB}} \tag{3.22}$$

Once the efficiency has been determined, the temperature after the turbine and before the aftertreatment system (ATS) can be calculated by applying the following equations:

$$\Delta h_{\mathrm{TB}} = \Delta h_{\mathrm{TB,is}} \cdot \eta_{\mathrm{TB}} \tag{3.23a}$$

$$\Delta \vartheta_{\mathrm{TB}} = \frac{\Delta h_{\mathrm{TB}}}{c_{\mathrm{p}}} \tag{3.23b}$$

$$\vartheta_{\mathrm{ATS}} = \vartheta_{\mathrm{EM}} - \Delta \vartheta_{\mathrm{TB}} \tag{3.23c}$$

where ϑ_{EM} is the temperature in the exhaust manifold, which is later defined in Eq. (3.29).

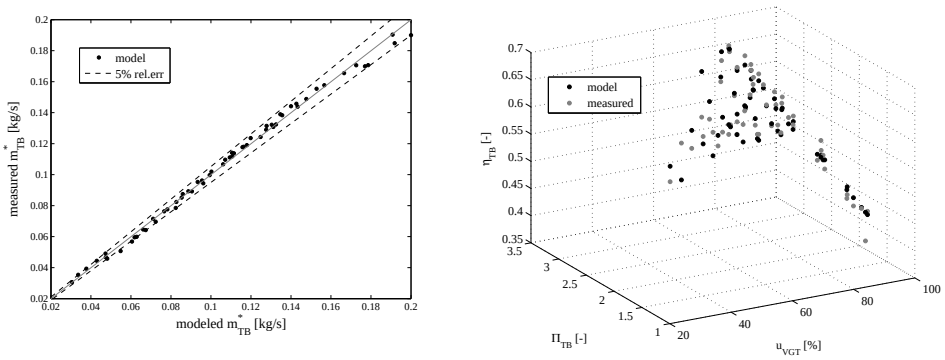


Figure 3.12: Results of the identification of the mass-flow and enthalpy models for the turbine.

Restriction mass-flow The aftertreatment system has been modeled as simple flow restriction, by adopting the approximation of the flow function for compressible flows, already used for the turbine in Eq. (3.14). Since the pressure ratio never reaches critical conditions, no case distinction is necessary. The final model is expressed as

$$\dot{m}^*_{\mathrm{ATS}} = k_{\mathrm{ATS},0} + k_{\mathrm{ATS},1} \cdot \frac{p_{\mathrm{ATS}}}{\sqrt{\vartheta_{ATS}}} \cdot \sqrt{\frac{2}{\Pi_{\mathrm{ATS}}} \cdot \left(1 - \frac{1}{\Pi_{\mathrm{ATS}}}\right)} \tag{3.24}$$

with $\Pi_{\mathrm{ATS}} = p_{\mathrm{ATS}}/p_{\mathrm{amb}}$. Figure (3.13) shows that the usual model formulation (with $k_{\mathrm{ATS},0} = 0$) exhibits a systematic error. For this reason the model has been extended by adding the parameter $k_{\mathrm{ATS},0}$.

Cylinder mass-flow Under normal engine operating conditions, the total mass flow $\dot{m}^*_{\mathrm{cyl}}$ entering the cylinders, is composed by a mixture of fresh air, driven from the compressor ($\dot{m}^*_{\mathrm{CP}}$), and exhaust gas, recirculated from the exhaust manifold ($\dot{m}^*_{\mathrm{EGR}}$) through the EGR valve. Since the EGR valve has been kept closed for the air-path model identification, there is no recirculation of exhaust gas ($\dot{m}^*_{\mathrm{EGR}} = 0$). As a result, the total mass flow equals the compressor mass flow ($\dot{m}^*_{\mathrm{cyl}} = \dot{m}^*_{\mathrm{CP}}$).

Regarding the mass flow entering the cylinders, the engine itself can be approximated as a volumetric pump, i.e. a device that enforces a volume flow approximately proportional to its speed. A typical formulation for such a model is:

$$\dot{m}^*_{\mathrm{cyl}} = \rho_{\mathrm{IM}} \cdot \dot{V}^*_{\mathrm{cyl}} = \eta_{\mathrm{vol,rel}} \cdot \frac{p_{\mathrm{IM}}}{R \cdot \vartheta_{\mathrm{IM}}} \cdot \frac{Vd}{2} \cdot \frac{N_{\mathrm{eng}}}{60} \tag{3.25}$$

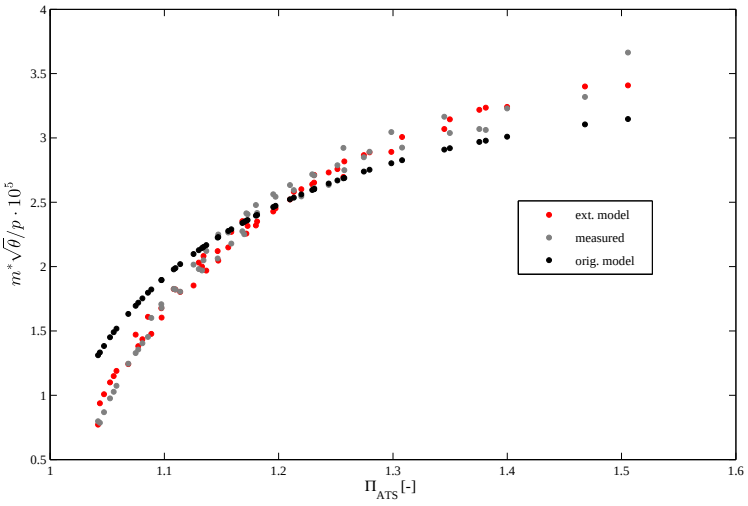


Figure 3.13: Model for the flow restriction representing the aftertreatment system (ATS). The original model formulation suffers a systematic error.

The variable ρ_{IM} is the density of the gas at the engine's intake, related to the intake pressure and temperature by the ideal gas law. The parameter $\eta_{\mathrm{vol,rel}}$ is the relative volumetric efficiency, which describes how far the engine differs from a perfect volumetric device ($\eta_{\mathrm{vol,rel}} = 1$). Several effects degrade the ideal volumetric efficiency to a typical value of 0.7 - 0.95. For instance a volume of residual gas, larger than the volume of the combustion chamber at top dead centre (TDC), may be captured in the cylinder, or flow restrictions causing a reduction of the pressure gradient may cause such an effect. These effects are influenced mainly by the valve timing, which is invariable for the given engine. Finally, the model formulation remains strict to the concept of volumetric efficiency introduced in Eq. (3.25), so the mass of fresh mixture aspirated in each cylinder is calculated as

$$m_{\mathrm{cyl}} = \frac{p_{\mathrm{IM}}}{\vartheta_{\mathrm{IM}}} \cdot \left(k_{\mathrm{cyl},0} + k_{\mathrm{cyl},1} \cdot N_{\mathrm{eng}} + k_{\mathrm{cyl},2} \cdot N_{\mathrm{eng}}^2\right) + k_{\mathrm{cyl},3} \cdot N_{\mathrm{eng}} \tag{3.26}$$

and consequently the total mass flow into the engine is expressed as

$$\overset{*}{m}_{\mathrm{cyl}} = m_{\mathrm{cyl}} \cdot \frac{N_{\mathrm{eng}} \cdot n_{\mathrm{cyl}}}{120} \tag{3.27}$$

The model reproduces the mass flow in the cylinders with a relative error within 5%, as shown in Figure 3.14.

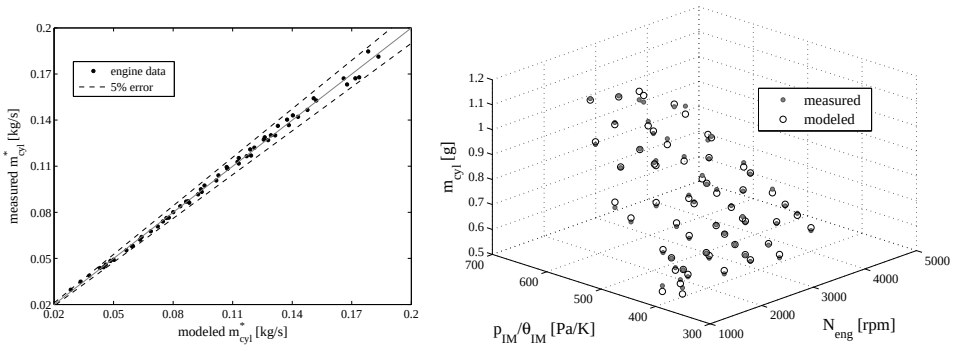


Figure 3.14: Model for the cylinder mass-flow. In right-hand plot, the measured mass entering each cylinder is plotted over the engine speed and the density at engine's intake. The model accurately reproduces all measurement data.

Engine-out temperature The engine-out temperature is calculated by an ideal-gas dual cycle (also known as the Seiliger cycle), splitting the combustion process into an isochoric and an isobaric part. To be precise, a distinction between the cylinder-outlet and the turbine-inlet temperature would be needed, because of the occurring heat losses from the exhaust manifold walls. In this engine at hand, the exhaust pipe is rather short, placing the turbine very close to the cylinders-outlet. For this reason the exhaust gas temperature along the connecting pipe between engine and turbine is assumed to be constant.

The temperature of the aspirated gas is adjusted to compensate for residual gases, heat transfers, turbulence and friction phenomena. The aspirated charge is heated from the hot valves, piston and cylinder walls. In addition, turbulence and flow friction also increase the charge temperature. However, the highest contribution is due to the hot residual gas from the previous combustion cycle. Rather than calculating its temperature (and estimating its fraction) iteratively, correction factors are introduced which also account for the other effects. A linear dependence in engine speed and intake manifold temperature has been found to be sufficient. Therefore, the initial temperature $\vartheta_{\text{cylInit}}$ of the cylinder charge is given by:

$$\vartheta_{\text{cylInit}} = h_{\text{cylInit},1} \cdot \vartheta_{\text{IM}} + h_{\text{cylInit},2} \cdot N_{eng} \tag{3.28a}$$

After defining the pressure ratio over the cylinder block as

$$\Pi_{\text{eng}} = \frac{p_{\text{exhaust}}}{p_{\text{intake}}} = \frac{p_3}{p_2} \tag{3.28b}$$

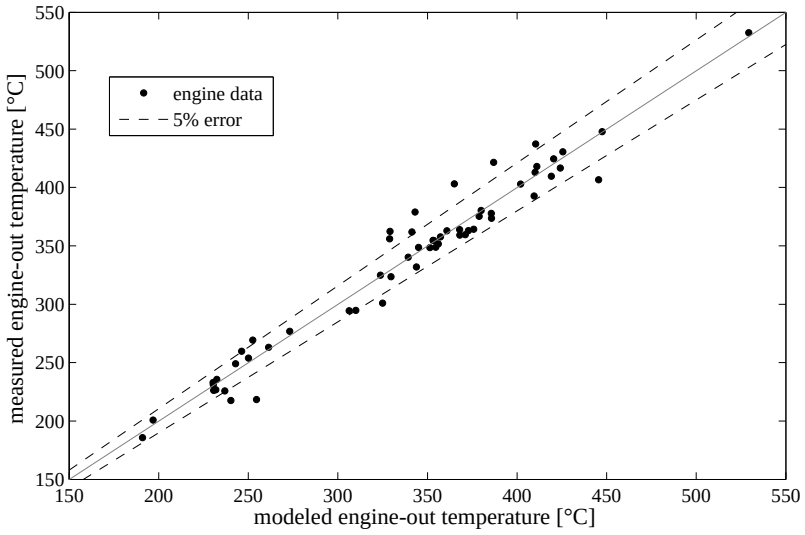


Figure 3.15: Model for the engine-out temperature.

and the "energy density" of the combustion process

$$\tilde{q} = \frac{\overset{*}{m}_{\text{fuel}} \cdot H_l}{\overset{*}{m}_{\text{in}}}, \tag{3.28c}$$

the ideal-gas Seiliger cycle (diesel formulation in this case) has been used to calculate the engine-out temperature as follows:

$$\vartheta_{\text{EM}} = \eta_{\text{comb}} \cdot \Pi_{\text{eng}}^{1-\frac{1}{\kappa}} \cdot r^{1-\kappa} \cdot \left(1 + \frac{\tilde{q}}{c_{\text{v}} \cdot r^{\kappa-1} \cdot \vartheta_{\text{cylInit}}}\right)^{\frac{1}{\kappa}-1} \cdot \left[\frac{\tilde{q}}{c_{\text{v}}} + \vartheta_{\text{cylInit}} \cdot r^{\kappa-1}\right] \tag{3.29}$$

The parameter η_{comb} is a constant correction factor to compensate additional heat losses and incomplete combustion, while r is the engine compression ratio. The MATLAB's function `fminsearch` has been adopted to find the best set of parameters, which are listed in Table 3.2. The prediction accuracy of the model is shown in Figure 3.15.

Dynamic elements In order to present the dynamics of the air-path, the concept of *receiver* needs to be introduced [7]. A *receiver* is a fixed volume in which the thermodynamic states (pressures, temperatures, composition, etc.) are assumed to be the same over the entire volume (lumped parameter system). As shown in Figure 3.16 , the inputs and outputs of the receiver are the mass and enthalpy flows, the storage quantities are the mass and the internal energy, and the state variables are the pressure and the temperature. Assuming no heat or mass transfer through the walls, and no substantial

changes in potential or kinetic energy in the flow, the following two coupled differential equations describe the receiver dynamics:

$$\frac{d}{dt}m(t) = \overset{*}{m}_{\text{in}}(t) - \overset{*}{m}_{\text{out}}(t) \tag{3.30a}$$

$$\frac{d}{dt}U(t) = \overset{*}{H}_{\text{in}}(t) - \overset{*}{H}_{\text{out}}(t) \tag{3.30b}$$

The coupling between these two equations, under the assumption that the fluids can be modeled as ideal gases, is given by the following relations:

$$p(t) \cdot V = m(t) \cdot R \cdot \vartheta(t) \tag{3.31a}$$

$$U(t) = c_{\text{v}} \cdot \vartheta(t) \cdot m(t) = \frac{1}{\kappa - 1} \cdot p(t) \cdot V \tag{3.31b}$$

$$\overset{*}{H}_{\text{in}}(t) = c_{\text{p}} \cdot \vartheta_{\text{in}}(t) \cdot \overset{*}{m}_{\text{in}} \tag{3.31c}$$

$$\overset{*}{H}_{\text{out}}(t) = c_{\text{p}} \cdot \vartheta(t) \cdot \overset{*}{m}_{\text{out}} \tag{3.31d}$$

Note that the temperature $\vartheta_{\text{out}}(t)$ of the out-flowing gas is assumed to be the same as the temperature $\vartheta(t)$ of the gas in the receiver (lumped parameter approach).

Substituting (3.31) into (3.30), the following two differential equations for the level variables pressure and temperature are obtained after some algebraic manipulations (time dependencies omitted):

$$\frac{d}{dt}p = \frac{R}{V} \cdot \kappa \cdot [\overset{*}{m}_{\text{in}} \cdot \vartheta_{\text{in}} - \overset{*}{m}_{\text{out}}(t) \cdot \vartheta] \tag{3.32a}$$

$$\frac{d}{dt}\vartheta = \frac{\vartheta \cdot R}{p \cdot V \cdot c_{\text{v}}} \cdot [c_{\text{p}}(\overset{*}{m}_{\text{in}} \cdot \vartheta_{\text{in}} - \overset{*}{m}_{\text{out}} \cdot \vartheta) - c_{\text{v}} \cdot \vartheta \cdot (\overset{*}{m}_{\text{in}} - \overset{*}{m}_{\text{out}})] \tag{3.32b}$$

Moving the attention towards the air-path model described so far, there are two receivers, namely the intake and exhaust volumes. Therefore, pressure and temperature in the intake and exhaust manifolds would be easily determined through equations (3.32), if there wasn't a connection between these two receivers. Indeed, they are connected through a pipe of small diameter, in which the flow is determined by the pressure difference and by a valve opening area, which serves as a regulation device. Since a higher level of pressure prevails in the exhaust side than in the intake one, the flow is driven in a non linear way through such valve, and as a consequence, part of the exhaust gas is recirculated into the intake manifold. This system, also known as EGR (Exhaust Gas Recirculation), connects the intake and exhaust receivers, therefore it has to be taken into account when defining the balance equations (3.30).

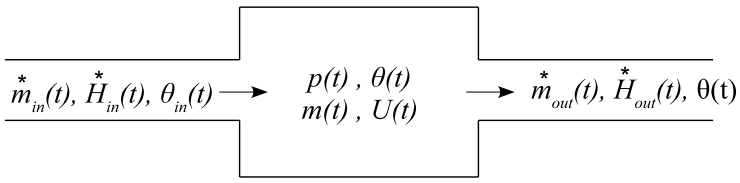


Figure 3.16: Inputs, outputs and level variables of a receiver.

3.2.2 Exhaust Gas Recirculation

The EGR valve has been modeled as a modified form of the simplified isentropic orifice [7] with a variable effective cross-sectional area. An approximate flow function has been used, i.e.

$$\tilde{\Psi}(\Pi_{\mathrm{EGR}}) = \sqrt{\frac{1}{\Pi_{\mathrm{EGR}}} \cdot \left(1 - \frac{1}{\Pi_{\mathrm{EGR}}}\right)} \tag{3.33a}$$

with $\Pi_{\mathrm{EGR}} = p_{\mathrm{EM}}/p_{\mathrm{IM}}$. The mass flow through the EGR may be expressed as:

$$\dot{m}_{\mathrm{EGR}} = A_{\mathrm{EGR}} \cdot p_{\mathrm{EM}} \cdot \tilde{\Psi}(\Pi_{\mathrm{EGR}}) \tag{3.33b}$$

and it is influenced by the term A_{EGR}, which is related to the geometrical opening area of the valve, reduced by a discharge coefficient. Since the opening area is proportional to the EGR valve command u_{EGR}, a relation with this parameter has been identified and the fitting model is:

$$A_{\mathrm{EGR}} = \frac{\dot{m}_{\mathrm{EGR}}}{p_{\mathrm{EM}} \cdot \tilde{\Psi}} = (k_{\mathrm{EGR},0} \cdot u_{\mathrm{EGR}} + k_{\mathrm{EGR},1} \cdot u_{\mathrm{EGR}}^2 + k_{\mathrm{EGR},2} \cdot u_{\mathrm{EGR}}^3) \tag{3.34}$$

Figure 3.17 shows such relation, while the final results for the EGR mass-flow model are shown in the left-hand plot of Figure 3.18.

It can be noticed that the prediction accuracy is slightly poorer compared to the models previously presented. Most likely this result is due to the EGR mass-flow estimation, which has been calculated as the difference between the cylinder and compressor mass-flows. While the compressor mass-flow is measured by the AFM, the cylinder mass-flow has been estimated by using the model itself (3.26), thereby the quantity $\dot{\tilde{m}}_{cyl}$ has been predicted over the entire experimental plan.

$$\dot{m}_{\mathrm{EGR}} = \dot{\tilde{m}}_{\mathrm{cyl}} - \dot{m}_{\mathrm{CP}}. \tag{3.35}$$

Consequently, the error in estimating the cylinder mass-flow (Figure 3.14), may lead to a potential propagation of the prediction error when calculating the EGR mass-flow.

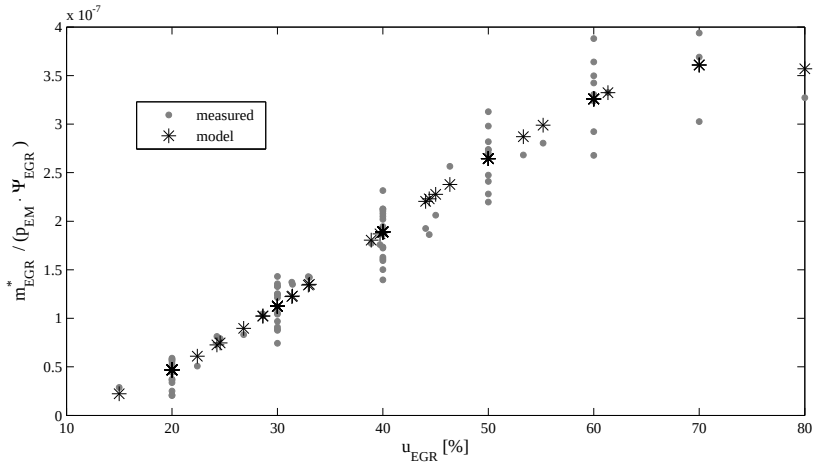


Figure 3.17: Model for the EGR valve. Fitting of the variable effective cross-sectional area.

EGR cooler The recirculated exhaust gas passes through a heat exchanger, called EGR cooler (EGC), with the aim of cooling the gas before mixing inside the intake manifold with the compressed air, which comes from the compressor. For the temperature difference across the EGC, a simple stationary energy balance is evaluated. The average of the inflow and outflow temperatures is used as the representative temperature, which yields:

$$k_{\mathrm{EGC}} \cdot \left(\frac{\vartheta_{\mathrm{EM}} + \vartheta_{\mathrm{EGR}}}{2} - \vartheta_{clf} \right) = \overset{*}{m}_{\mathrm{EGR}} \cdot (\vartheta_{\mathrm{EM}} - \vartheta_{\mathrm{EGR}}) \tag{3.36a}$$

The temperature of the cooling-fluid is assumed to be constant, i.e. $\vartheta_{clf} = 360\mathrm{K}$, and the efficiency parameter k_{EGC} depends on the mass flow:

$$k_{\mathrm{EGC}} = k_{\mathrm{EGC},1} \cdot \overset{*}{m}_{\mathrm{EGR}} + k_{\mathrm{EGC},2} \cdot \overset{*}{m}{}^{2}_{\mathrm{EGR}} \tag{3.36b}$$

The right-hand plot of Figure 3.18 shows the results of the model. As already explained above, the inaccurate estimation of the EGR mass-flow may be the main factor responsible for the poor prediction accuracy achieved.

Dynamic elements with EGR Now that the models related to the EGR system have been introduced, the final balance equations can be defined and the air-path dynamics summarized. Since the objective was to derive a simple air-path model in the first place, the EGR dynamics need to be introduced in accordance with the following assumptions:

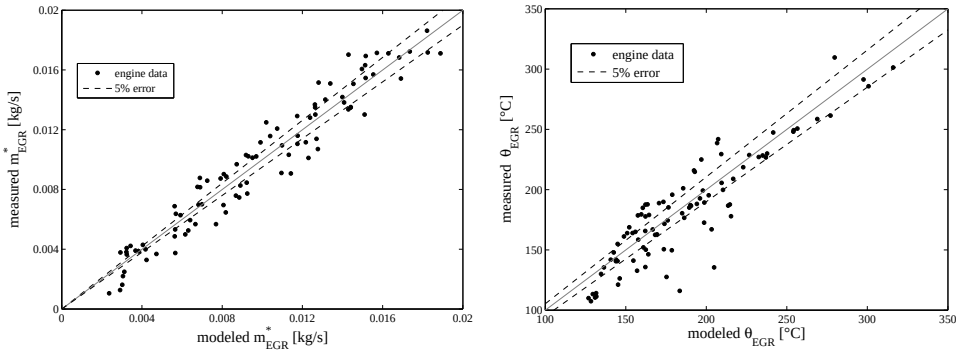


Figure 3.18: EGR mass-flow in left-hand plot. EGR cooler in right-hand plot. For both cases the relative error is above 5 %, this inaccuracy might be due to the approximation made when estimating the realEGR mass-flow.

- No transport delay of the mass through the EGR system. In a plug-flow, the flow through the EGR valve instantly pushes the same flow into the intake manifold.
- No gas dynamics from the exhaust ports to the EGR valve. In accordance with the plug-flow assumption, no mixing or diffusion takes place.
- No transport delay of the gas composition through the EGR system. The flow entering the intake manifold is assumed to have the composition of the exhaust gas leaving the exhaust ports at the same time.
- Perfectly mixed gas in the intake manifold. This assumption is common to all mean-value modelling approaches (lumped parameters).
- The EGR cooler is installed close enough to the valve that the two parts can be represented by a single flow restriction with a variable opening area.
- The mixture is stoichiometric or lean (but not rich). For diesel engines, this is a normal operating condition.

Some of these assumptions may appear quite strong but, as long as the EGR path is short, they can be accepted without significantly degrading the quality of the model.

The differential equations (3.32) are now applied to the intake receiver, where there are two input and one output flows, as shown in Figure (3.6). The fresh air coming from the compressor mixes with the exhaust gas, consequently the gas in the intake manifold can be subdivided into two species, namely burnt gas and air. Burnt gas denotes the fraction of mixture that actually burned during the combustion, accordingly to its stoichiometric

factor, thereby no oxygen is left in it. Hence, the burnt-gas fraction in the exhaust manifold (bg,EM) is

$$x_{\mathrm{bg,EM}} = \frac{x_{\mathrm{bg,IM}} \cdot \overset{*}{m}_{\mathrm{cyl}} + (1+\sigma_0) \cdot \overset{*}{m}_{\mathrm{fuel}}}{\overset{*}{m}_{\mathrm{cyl}} + \overset{*}{m}_{\mathrm{fuel}}} \tag{3.37}$$

where σ_0 is the stoichiometric air-to-fuel ratio. Due to the distinction of the two species in the intake volume, two mass-balance equations result:

$$\begin{aligned} \frac{d}{dt} m_{\mathrm{IM,air}} &= \overset{*}{m}_{\mathrm{CP,air}} + \overset{*}{m}_{\mathrm{EGR,air}} - \overset{*}{m}_{\mathrm{cyl,air}} \\ &= \overset{*}{m}_{\mathrm{CP}} + \overset{*}{m}_{\mathrm{EGR}} \cdot (1 - x_{\mathrm{bg,EM}}) - \overset{*}{m}_{\mathrm{cyl}} \cdot (1 - x_{\mathrm{bg,EM}}) \end{aligned} \tag{3.38a}$$

$$\frac{d}{dt} m_{\mathrm{IM,bg}} = \overset{*}{m}_{\mathrm{EGR}} \cdot x_{\mathrm{bg,EM}} - \overset{*}{m}_{\mathrm{cyl}} \cdot x_{\mathrm{bg,IM}} \tag{3.38b}$$

Concerning the energy balance (3.30b), constant and uniform specific heats are assumed:

$$\begin{aligned} \frac{d}{dt} \left(c_{\mathrm{v}} \cdot m_{\mathrm{IM}} \cdot \vartheta_{\mathrm{IM}} \right) &= c_{\mathrm{v}} \cdot \left(m_{\mathrm{IM}} \cdot \frac{d\vartheta_{\mathrm{IM}}}{dt} + \vartheta_{\mathrm{IM}} \cdot \frac{dm_{\mathrm{IM}}}{dt} \right) \\ &= c_{\mathrm{p}} \cdot \left(\overset{*}{m}_{\mathrm{CP}} \cdot \vartheta_{\mathrm{IC}} + \overset{*}{m}_{\mathrm{EGR}} \cdot \vartheta_{\mathrm{EGC}} - \overset{*}{m}_{\mathrm{cyl}} \cdot \vartheta_{\mathrm{IM}} \right). \end{aligned} \tag{3.38c}$$

where $m_{\mathrm{IM}} = m_{\mathrm{IM,air}} + m_{\mathrm{IM,bg}}$. The solution of the equation system for the pressure, the temperature and the mass fraction is given by applying the ideal gas law (3.31a):

$$f_1 = \frac{d}{dt} p_{\mathrm{IM}} = \frac{R}{V_{\mathrm{IM}}} \kappa \left(\overset{*}{m}_{\mathrm{CP}} \vartheta_{\mathrm{IC}} + \overset{*}{m}_{\mathrm{EGR}} \vartheta_{\mathrm{EGC}} - \overset{*}{m}_{\mathrm{cyl}} \vartheta_{\mathrm{IM}} \right), \tag{3.39a}$$

$$\begin{aligned} f_2 = \frac{d}{dt} \vartheta_{\mathrm{IM}} = \frac{R\vartheta_{\mathrm{IM}}}{p_{\mathrm{IM}} V_{\mathrm{IM}} c_v} \Big[& c_p \left(\overset{*}{m}_{\mathrm{CP}} \vartheta_{\mathrm{IC}} + \overset{*}{m}_{\mathrm{EGR}} \vartheta_{\mathrm{EGC}} - \overset{*}{m}_{\mathrm{cyl}} \vartheta_{\mathrm{IM}} \right. \\ & \left. - c_v \vartheta_{\mathrm{IM}} \left(\overset{*}{m}_{\mathrm{CP}} + \overset{*}{m}_{\mathrm{EGR}} - \overset{*}{m}_{\mathrm{cyl}} \right) \Big], \end{aligned} \tag{3.39b}$$

$$f_3 = \frac{d}{dt} x_{\mathrm{bg,IM}} = \frac{R\vartheta_{\mathrm{IM}}}{p_{\mathrm{IM}} V_{\mathrm{IM}}} \left(\overset{*}{m}_{\mathrm{EGR}} \cdot (x_{\mathrm{bg,EM}} - x_{\mathrm{bg,IM}}) - \overset{*}{m}_{\mathrm{CP}} x_{\mathrm{bg,IM}} \right). \tag{3.39c}$$

An alternative option would have been to consider the exhaust-gas fraction rather than the burnt-gas fraction. In that case, the fraction of exhaust gas would be equal to 1 in the exhaust manifold, but the derivation of the balance equations for the intake manifold would not change. The choice of using the burnt-gas fraction has been guided by the fact that the burnt-gas composition depends only on the air composition. During transient

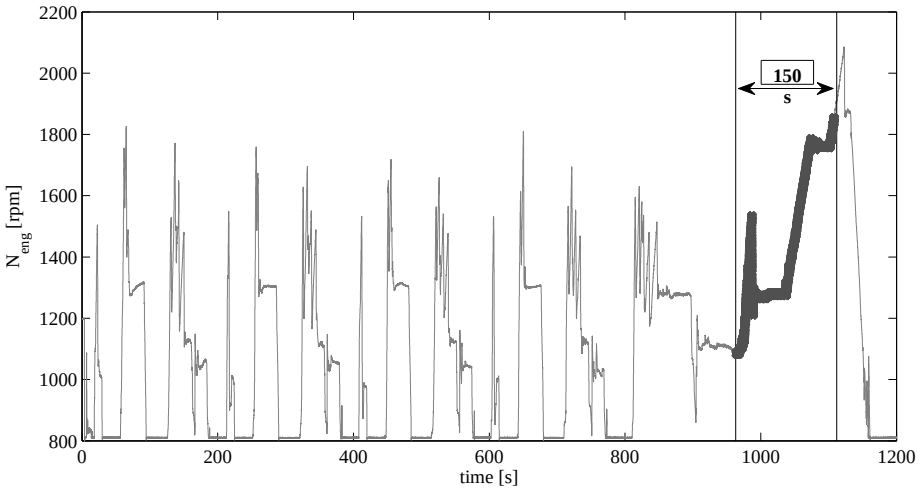


Figure 3.19: Transient validation of the air-path model. A sub-part of the NEDC cycle (highlighted in dark grey) has been used for the validation.

operation a more accurate prediction of the mixture composition in the intake manifold can be accomplished, since it can be influenced only by the ambient conditions, which can be assumed to be constant or, at least, to change significantly more slowly than the EGR dynamics.

To complete the set of differential equations describing the air-path model dynamics, the following equations for the turbocharger speed, the exhaust manifold and the ATS are here listed:

$$f_4 = \frac{d}{dt}p_{\mathrm{EM}} = \frac{R \cdot \vartheta_{\mathrm{EM}}}{V_{\mathrm{EM}}}\kappa\left(\overset{*}{m}_{\mathrm{EM}} - \overset{*}{m}_{\mathrm{TB}} - \overset{*}{m}_{\mathrm{EGR}}\right), \tag{3.39d}$$

$$f_5 = \frac{d}{dt}p_{\mathrm{ATS}} = \frac{R \cdot \vartheta_{\mathrm{ATS}}}{V_{\mathrm{ATS}}}\kappa\left(\overset{*}{m}_{\mathrm{TB}} - \overset{*}{m}_{\mathrm{ATS}}\right), \tag{3.39e}$$

$$f_6 = \Theta_{\mathrm{TC}} \cdot \frac{dw_{\mathrm{TC}}}{dt} = T_{\mathrm{TB}} - T_{\mathrm{CP}} = \frac{\overset{*}{m}_{\mathrm{TB}}\Delta h_{\mathrm{TB}} - \overset{*}{m}_{\mathrm{CP}}\Delta h_{\mathrm{CP}}}{w_{\mathrm{TC}}}. \tag{3.39f}$$

3.2.3 Transient validation

In the previous sections, the prediction accuracy of each sub-model has been shown in the respective figures. It results that every model predicts the measurement data with an error below 5%, except for the EGR models which fall slightly short of expectations. All parameters identified are listed in Table 3.2.

Finally, a transient driving cycle has been performed in order to assess the quality of the entire air-path model. Hence, a NEDC (New European

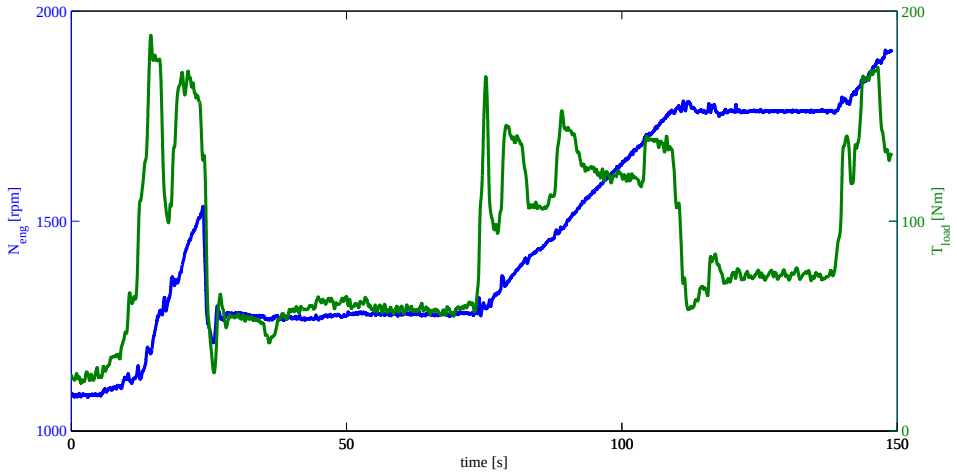


Figure 3.20: Engine speed (blue) and load torque (green) requested during the validation driving cycle.

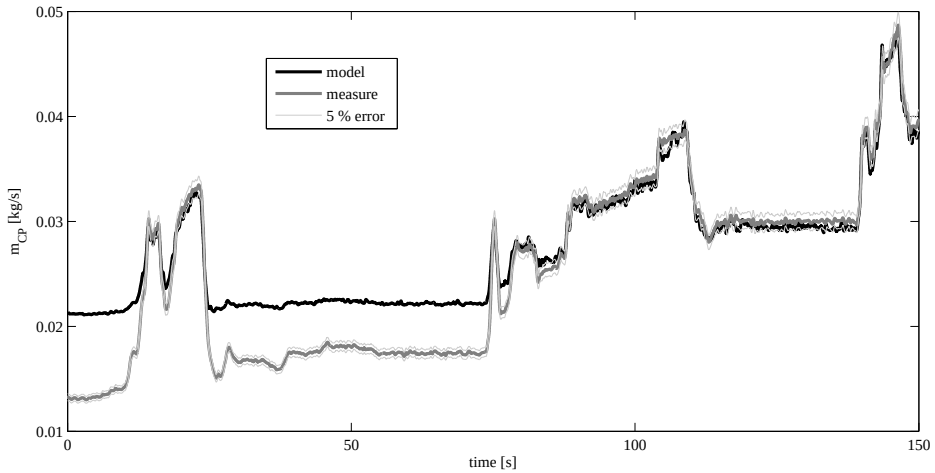


Figure 3.21: Comparison between the modeled and measured compressor mass-flows. The light grey lines delimit the region characterized by a relative error of 5%.

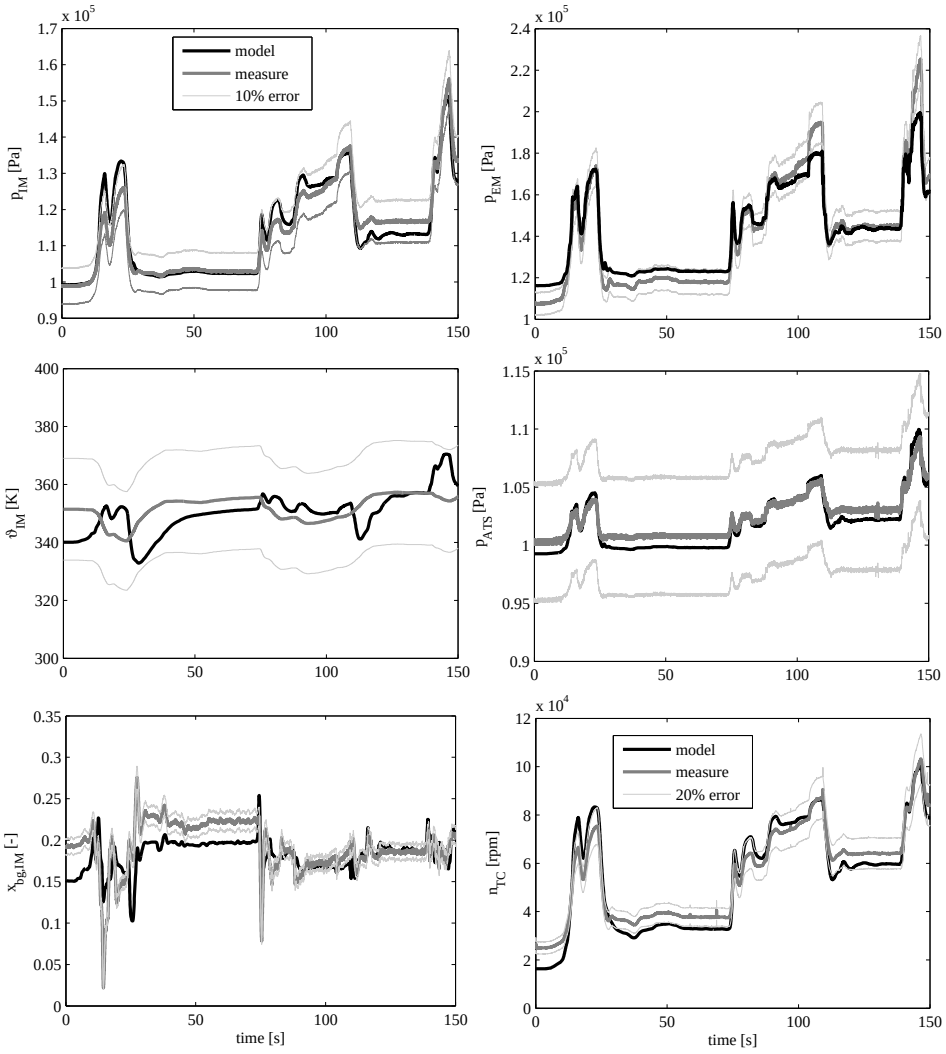


Figure 3.22: Results of the air-path model including the dynamic part. The results related to the six equations expressed in (3.39) are plotted, in comparison with the recorded data. If not differently specified, the error region showed in the plots above is 10%.

Table 3.2: Identified parameters for the engine model.

model	**parameter**	**value**	**equation**
compressor mass-flow	$k_{\pi,0}$	8.5224	(3.4a)
	$k_{\pi,1}$	$7.2326 \cdot 10^{-15}$	(3.4a)
	$k_{\pi,\mathrm{e}}$	2.7310	(3.4a)
	$k_{\mathrm{m},0}$	-0.2784	(3.4b)
	$k_{\mathrm{max},0}$	0.0097	(3.2c)
	$k_{\mathrm{max},1}$	$1.1097 \cdot 10^{-6}$	(3.2c)
	c_1	$1.5796 \cdot 10^{-11}$	(3.2d)
	$c_{2,0}$	12.2056	(3.2d), (3.5)
	$c_{2,1}$	$4.1484 \cdot 10^{-11}$	(3.5)
	$c_{2,2}$	$3.0921 \cdot 10^{-10}$	(3.5)
compressor enthalpy	$k_{\mathrm{h},1}$	$6.99 \cdot 10^{-5}$	(3.10)
	$k_{\mathrm{h},2}$	2.1380	(3.10)
intercooler	$k_{\mathrm{IC},0}$	0.0215	(3.13)
	$k_{\mathrm{IC},1}$	1.3144	(3.13)
	$k_{\mathrm{IC},2}$	0.3853	(3.13)
turbine mass-flow	$k_{\mathrm{TB},0}$	$3.1697 \cdot 10^{-5}$	(3.16)
	$k_{\mathrm{TB},1}$	$-2.1797 \cdot 10^{-7}$	(3.16)
	$k_{\mathrm{TB},2}$	$-7.2556 \cdot 10^{-10}$	(3.16)
	$k_{\mathrm{TB},3}$	$2.7540 \cdot 10^{-11}$	(3.16)
turbine efficiency	$h_{\mathrm{TB},0}$	0.2584	(3.22)
	$h_{\mathrm{TB},1}$	0.0113	(3.22)
	$h_{\mathrm{TB},2}$	$-1.1033 \cdot 10^{-4}$	(3.22)
	$h_{\mathrm{TB},3}$	0.0634	(3.22)
restriction mass-flow	$k_{\mathrm{ATS},0}$	0.0300	(3.24)
	$k_{\mathrm{ATS},1}$	$5.2651 \cdot 10^{-4}$	(3.24)
cylinder mass-flow	$k_{\mathrm{cyl},0}$	$1.2120 \cdot 10^{-6}$	(3.26),(3.27)
	$k_{\mathrm{cyl},1}$	$2.5939 \cdot 10^{-10}$	(3.26),(3.27)
	$k_{\mathrm{cyl},2}$	$-6.0851 \cdot 10^{-14}$	(3.26),(3.27)
	$k_{\mathrm{cyl},3}$	$2.1706 \cdot 10^{-8}$	(3.26),(3.27)
engine-out temperature	$h_{\mathrm{cylInit},1}$	1.0393	(3.28a)
	$h_{\mathrm{cylInit},2}$	0.0320	(3.28a)
	κ	1.2513	(3.29)
	η_{comb}	0.8169	(3.29)
EGR mass-flow	$k_{\mathrm{EGR},0}$	$-1.7193 \cdot 10^{-9}$	(3.34)
	$k_{\mathrm{EGR},1}$	$2.4500 \cdot 10^{-10}$	(3.34)
	$k_{\mathrm{EGR},2}$	$-2.0968 \cdot 10^{-12}$	(3.34)
EGR cooler	$\mathrm{k}_{\mathrm{EGC},1}$	1.2512	(3.36b)
	$\mathrm{k}_{\mathrm{EGC},2}$	-33.1900	(3.36b)

Driving Cycle) has been carried out on the testbench, and a short sub-part of 150 seconds has been taken into consideration. Looking at Figure 3.19, the driving cycle considered is highlighted in dark grey, and corresponds to the extra-urban cycle. This choice has been made since a sufficiently large region of the engine operating range can be investigated. Furthermore, the load profile, shown in Figure 3.20, is quite demanding as to allow for an actual transient operation of the engine.

The first result presented is the compressor mass-flow model, in Figure 3.21. There are two visible regions, corresponding both to a low-speed and low-load condition, in which the model is clearly far from the measure. These operating conditions can be considered difficult to be accurately modeled when deriving the compressor mass-flow model, since the mass flow and the turbocharger speed are rather low. In Figure 3.22, where the intake manifold pressure is plotted, it can be noted how low the pressure value is in the region under discussion, confirming that the turbocharging effect is almost absent. Such conditions are considered critical for the turbocharger to work properly. Except for this unstable region, the prediction capability is significantly accurate, being the model enclosed within the 5% error region.

Concerning the dynamic equations (3.39), the results obtained are plotted in Figure 3.22. As explained at the beginning of this chapter, the primary goal was to design a simple model, fast to be executed but still able to capture the main dynamic phenomena concerning the air-path. The attention has to be especially focused on its ability to catch the trend of the physical variables during transient operation. Since no thermal dynamics have been introduced, it should not be surprising that the modeled intake manifold temperature ϑ_{IM} shows different dynamics w.r.t the measured value.

The overall result is satisfying, therefore the next step involves deriving the combustion model.

3.2.4 Time-resolved combustion model

Time-variable quadratic setpoint-deviation models are used for the emissions and for the combustion efficiency. The cross-terms of the full second-order Taylor expansion are omitted since the $n_{\mathrm{u}}(n_{\mathrm{u}}-1)/2$ cross variations would also need to be performed to reliably identify the corresponding model coefficients. The model coefficients are identified around the current reference trajectory during each iteration. The vector $\boldsymbol{w} = (p_{\mathrm{IM}}, x_{\mathrm{bg,IM}}, \varphi_{\mathrm{SOI}}, m_{\mathrm{fuel}})^T$ denotes all inputs to the combustion model. The model for each output thus reads:

$$y_i(t) = y_{\mathrm{ref},i}(t) + \boldsymbol{k}_{\mathrm{lin}}(t)^T \cdot \Delta\boldsymbol{w}(t) + \Delta\boldsymbol{w}(t)^T \cdot \boldsymbol{K}_{\mathrm{quad}}(t) \cdot \Delta\boldsymbol{w}(t), \quad (3.40a)$$

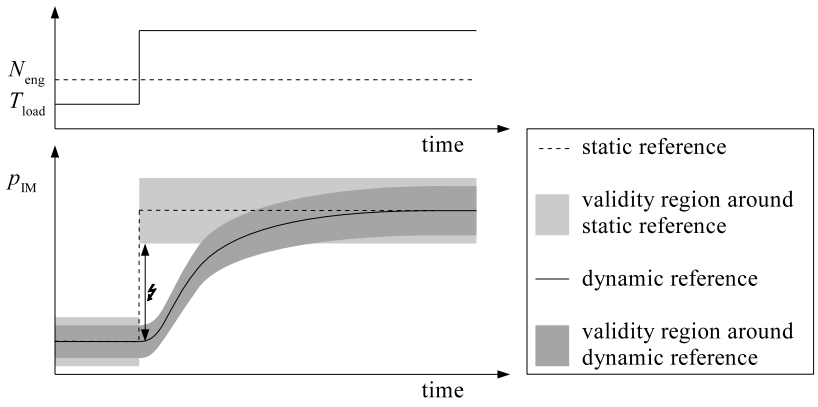


Figure 3.23: Exemplary static and dynamic references over a load step, and the validity regions of the corresponding setpoint-relative models.

with

$$\boldsymbol{K}_{\text{quad}}(t) = \text{diag}\big(k_{\text{quad},1}(t), \ldots, k_{\text{quad},4}(t)\big), \tag{3.40b}$$

$$\Delta\boldsymbol{w}(t) = \boldsymbol{w}(t) - \boldsymbol{w}_{\text{ref}}(t). \tag{3.40c}$$

Such setpoint-relative models are often used for control and optimisation applications [29, 30]. However, usually the references for the inputs $\boldsymbol{w}$ are stored as a static lookup map over engine speed and load. During transient operation such as a load step, the actual values of the dynamic inputs (p_{IM} and $x_{\text{bg,IM}}$ in the case at hand) can be far from the steady-state reference values. By using time-resolved reference values and correction factors, the validity range of the model is relocated to the actual relevant region, which is illustrated in Fig. 3.23.

The outputs of the air-path model, i.e. the dynamic inputs to the combustion model, are denoted by $\boldsymbol{v} := (p_{\text{IM}}, x_{\text{bg,IM}})^T$, with

$$\boldsymbol{v}(t) = \boldsymbol{g}_{\text{AP}}(\boldsymbol{x}(t), \boldsymbol{u}(t)) = \boldsymbol{C}_{\text{AP}} \cdot \boldsymbol{x}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \cdot \boldsymbol{x}(t). \tag{3.41}$$

Identification The models for the emissions and the torque production are re-identified at each iteration. In the final version of the algorithm, all data collected so far should be used during the identification to successively expand the model validity region over the iterations. It is subject to debate whether the model structure should be adapted, i.e. its complexity increased according to the data available. Alternatively, the measurements could be weighted by their distance to the current $\boldsymbol{w}(t)$ to yield a tradeoff between the representation of local and far-field trends. For now, only the variations

around the initial trajectory are considered, and the quadratic setpoint-relative model described above is used.

A fast measurement of the instantaneous emissions is required. The NO_x emissions are measured by a Cambustion fNOx 400 FastCLD. The soot emissions are recorded by means of an AVL Micro Soot Sensor. The dynamic response of this device is identified by performing several stepwise variations of the injected fuel mass at different operating points. A constant delay of 0.95s and a first-order element with a time constant of 0.25s are found to closely compensate the sensor dynamics.

A slight averaging w.r.t. time is performed to suppress measurement noise and to provide a smooth model, which is advantageous in the context of optimization. A windowed Gauss curve is used to weight the preceding and consecutive samples for the identification of the model at each time instant,

$$\theta(\Delta t) = \begin{cases} \exp\left(-\frac{\Delta t^2}{\sigma^2}\right), & \text{if } |\Delta t| < \sqrt{-\sigma^2 \cdot \log_e(\theta_{\text{cut}})}, \\ 0, & \text{else.} \end{cases} \tag{3.42}$$

The parameter θ_{cut} defines the value for θ at which the Gauss curve is cut. Figure 3.24 shows the curve for the parameter values $\sigma = 0.4s$ and $\theta_{\text{cut}} = 0.5\%$, which are found to be a reasonable choice.

For each output y_i, a weighted linear least-squares regression is used to identify the coefficients at each sampling point t_k. The $N_{\text{m}} - 1$ variations are used, which are again denoted by the superscript index in round brackets. The equation system

$$(\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X}) \cdot \boldsymbol{p} = \boldsymbol{X}^T \boldsymbol{W} \Delta \boldsymbol{y}_i, \tag{3.43a}$$

is solved for $\boldsymbol{p}$, where

$$\boldsymbol{p} = \big(k_{\text{lin},1}(t_k), \ldots, k_{\text{lin},4}(t_k), k_{\text{quad},1}(t_k), \ldots, k_{\text{quad},4}(t_k)\big)^T, \tag{3.43b}$$

$$\Delta \boldsymbol{y}_i = \begin{pmatrix} \boldsymbol{y}_i(t_{-l}) - \boldsymbol{y}_{\text{ref},i}(t_{-l}) \\ \vdots \\ \boldsymbol{y}_i(t_l) - \boldsymbol{y}_{\text{ref},i}(t_l) \end{pmatrix}, \tag{3.43c}$$

$$\boldsymbol{X} = \begin{bmatrix} \Delta \boldsymbol{w}_1(t_{-l}) & \cdots & \Delta \boldsymbol{w}_4(t_{-l}) & \Delta \boldsymbol{w}_1(t_{-l})^2 & \cdots & \Delta \boldsymbol{w}_4(t_{-l})^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ \Delta \boldsymbol{w}_1(t_l) & \cdots & \Delta \boldsymbol{w}_4(t_l) & \Delta \boldsymbol{w}_1(t_l)^2 & \cdots & \Delta \boldsymbol{w}_4(t_l)^2 \end{bmatrix}, \tag{3.43d}$$

$$\boldsymbol{W} = \operatorname{diag}\big(\theta(t_{-l} - t_k) \cdot \boldsymbol{I}_{N_m - 1}, \ldots, \theta(t_l - t_k) \cdot \boldsymbol{I}_{N_m - 1}\big). \tag{3.43e}$$

Here, l is the number of samples inside the window in both directions, and $\boldsymbol{I}_{N_{\text{m}}-1}$ is the identity matrix. Each vector $\Delta \boldsymbol{w}_j := \big(w_j^{(1)} - w_{\text{ref},j}, \ldots, w_j^{(N_{\text{m}}-1)} - w_{\text{ref},j}\big)^T$ stacks all variations (similarly for $\boldsymbol{y}_i$). Here, $\Delta \boldsymbol{w}_j^2$ denotes element-wise squaring.

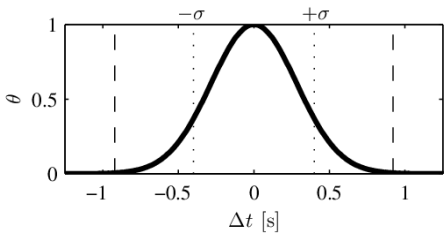


Figure 3.24: Windowed Gauss function used for the time-averaging of the combustion model. The dashed lines delineate the window.

3.3 Numerical optimal control

A continuous-time optimal control problem (OCP) consists of an objective function to be minimized (or maximized), and a set of ordinary differential equations (ODE) that need to be satisfied, yielding dynamic constraints:

$$\min_{\boldsymbol{x}(\cdot),\boldsymbol{u}(\cdot)} \int_0^T L\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\pi}(t)\big)\ dt \tag{3.44a}$$

$$\text{s.t.}\quad \dot{\boldsymbol{x}}(t) - \boldsymbol{f}\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\pi}(t)\big) = 0, \quad t \in [0,T] \tag{3.44b}$$

The integral cost $L\big(\boldsymbol{x}(t),\boldsymbol{u}(t)\big)$ is called the Lagrange term, and the end cost $E\big(\boldsymbol{x}(T)\big)$ is known as Mayer term. The combination of the two is called a Bolza objective. In the remainder of this work, only a Lagrange term is considered. Every (differentiable) Mayer term can be replaced by an equivalent Lagrange term, but the Lagrange formulation is preferable from a numerical point of view [31].

A common extension of the unconstrained OCP (3.44) is to prescribe an initial state $x(0) = x_0$, or to enforce some conditions on the end state $x(T)$. Often, "simple bounds" are imposed directly on the state variables x and on the control inputs u. These limits stem from the ranges of the physical actuators represented by the control inputs, or mechanical limits on certain state variables such as rotational speed or temperature. More general "path constraints" may be imposed, which are nonlinear functions of the state variables and the control inputs.

Integral equalities or inequalities are another type of constraint that often arise in the formulation of engineering problems. These constraints may also represent a Mayer term that is rewritten in Lagrange form. Since the problem to be tackled in this thesis has to be formulated using time-variable parameters $\boldsymbol{\pi}(t)$, this special case is explicitly included in the formulation of the general OCP, which becomes:

$$\min_{\boldsymbol{x}(\cdot),\boldsymbol{u}(\cdot)} \int_0^T L\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\pi}(t)\big)\ dt \tag{3.45a}$$

$$\text{s.t.} \quad \dot{\boldsymbol{x}}(t) - \boldsymbol{f}\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\pi}(t)\big) = 0, \quad t \in [0,T] \tag{3.45b}$$

$$\int_0^T \boldsymbol{g}\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\boldsymbol{\pi}(t)\big)\ dt - \hat{\boldsymbol{g}} \leq 0 \tag{3.45c}$$

$$\underline{\boldsymbol{u}}(t) \leq \boldsymbol{u}(t) \leq \overline{\boldsymbol{u}}(t), \quad t \in [0,T] \tag{3.45d}$$

$$\underline{\boldsymbol{x}}(t) \leq \boldsymbol{x}(t) \leq \overline{\boldsymbol{x}}(t), \quad t \in [0,T] \tag{3.45e}$$

A detailed description about optimal control theory applied to diesel engines can be found in [32].

3.3.1 The Diesel-engine problem

Since the main objective is to minimize the fuel consumption while maintaining the same emission levels, the Diesel-engine problem can be cast in the form of the following OCP:

$$\min_{\boldsymbol{x}(\cdot),\boldsymbol{u}(\cdot)} \left\{ m_{\text{fuel}} = \int_0^{t_\text{f}} \overset{*}{m}_{\text{fuel}}\big(\boldsymbol{x}(t),\boldsymbol{u}(t),\tilde{n}_{\text{eng}}(t)\big)\ dt \right\} \tag{3.46a}$$

$$\text{s.t.} \quad \dot{\boldsymbol{x}}(t) - \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),\tilde{n}_{\text{eng}}(t)) = 0 \tag{3.46b}$$

$$\int_0^{t_\text{f}} \overset{*}{\boldsymbol{m}}_{\text{em}}(\boldsymbol{x}(t),\boldsymbol{u}(t),\tilde{n}_{\text{eng}}(t))\ dt - \hat{\boldsymbol{m}}_{\text{em}} \leq 0 \tag{3.46c}$$

$$\tilde{T}(t) - T_{\text{load}}(\boldsymbol{x}(t),\boldsymbol{u}(t),\tilde{n}_{\text{eng}}(t)) \leq 0 \tag{3.46d}$$

$$\underline{\boldsymbol{u}}(t) \leq \boldsymbol{u}(t) \leq \overline{\boldsymbol{u}}(t), \quad \underline{\boldsymbol{x}}(t) \leq \boldsymbol{x}(t) \leq \overline{\boldsymbol{x}}(t) \tag{3.46e}$$

The instantaneous fuel consumption $\overset{*}{m}_{\text{fuel}} = m_{\text{fcc}} \cdot \frac{\tilde{n}_{\text{eng}}}{120} \cdot N_{\text{cyl}}$ is integrated to yield the cumulative fuel consumption m_{fuel}. The dynamic constraints (3.46b) enforce the model equations and (3.46c) limit the cumulative emissions. The engine speed prescribed by the driving cycle, $\tilde{n}_{\text{eng}}(t)$, is considered a time-varying parameter. The desired load torque $\tilde{T}(t)$ is imposed as a time-local inequality constraint (3.46d).

3.3.2 Direct transcription

The continuous-time OCP (3.46) is transformed into a finite-dimensional mathematical program by direct transcription. An integration scheme is applied to approximate all continuous signals. Due to the stiffness of the system, the method of choice is the family of Radau collocation schemes [33]. These stiffly accurate schemes provide stiff decay and algebraic stability [34]. Their implicit nature is irrelevant when used in the context of direct transcription. The first-order representative, the well-known Euler-backward

discretization, is used for the solution of the OCPs in this study. It is found to be very robust and it provides the solution on a uniform discretization grid, which can be consistently implemented on the testbench.

The set of continuous ordinary differential equations (3.46b) is transformed to

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + h \cdot \boldsymbol{f}(\boldsymbol{u}_{k+1}, \boldsymbol{x}_{k+1}, \tilde{n}_{\mathrm{eng},k+1}) \tag{3.47}$$

adopting the notation $\boldsymbol{x}_k := \boldsymbol{x}(t_k)$. A uniform step size of $h = 0.1s$ is used, yielding a grid of N points $0 = t_0 < t_1 < \ldots < t_{N-1} = t_\mathrm{f}$. Accordingly, all integrals are approximated as sums

$$\int_0^{t_\mathrm{f}} \overset{*}{m}_\bullet(\boldsymbol{u}(t), \boldsymbol{x}(t), \tilde{n}_\mathrm{eng}(t))\, dt \approx h \cdot \sum_{k=0}^{N-2} \overset{*}{m}_\bullet(\boldsymbol{u}_{k+1}, \boldsymbol{x}_{k+1}, \tilde{n}_{\mathrm{eng},k+1}). \tag{3.48}$$

The bounds on the controls and on the state variables are imposed at the grid points.

This discretization of the problem yields a sparse NLP. Sparsity signifies that only a few entries in the Jacobian matrix, which contains the first partial derivatives of the constraints, are non-zero. In fact, (3.47) reveals that they are only related in neighboring sets of two. Furthermore, the control inputs and the state variables appear in nonlinear form only in $\boldsymbol{f}(\boldsymbol{u}_{k+1}, \boldsymbol{x}_{k+1}, \tilde{n}_{\mathrm{eng},k+1})$ and in the output function $\boldsymbol{g}$. Thus, only the partial derivatives of the model functions $\boldsymbol{f}$ and $\boldsymbol{g}$ w.r.t. the control inputs and the state variables at each discretization point have to be calculated to construct the Jacobian of the NLP. Assembling the first-derivative information of the NLP from the model derivatives corresponds to a perfect exploitation of the problem sparsity [35].

The derivatives of the model equations are calculated by forward finite differences. Therefore, one additional model evaluation is required for each partial derivative. The solver SNOPT 7.2 [36] is used to solve the sparse NLP. It approximates the second partial derivatives by iterative updates using the first-derivative information along the solution steps.

3.3.3 Regularization

Singular arcs are time intervals during which the Hamiltonian (the combination of the objective and the appropriately weighted constraint violations) becomes affine in the controls. During such intervals, the second derivatives vanish, which results in spurious oscillations when applying direct transcription to solve the OCP. A more detailed analysis of this phenomenon is provided in [37].

In the aforementioned thesis, a regularization based on the "piecewise derivative variation of the control" is proposed as a countermeasure. The regularization term, for a scalar control input u, is

$$L_{\mathrm{reg}}(u_\bullet) := c_N \cdot \dot{\mathrm{Var}}^2_{t_N}(u_\bullet) = \frac{c_{\mathrm{reg}}}{(N-3)(N-1)^2} \cdot \frac{1}{2} \sum_{l=3}^{N-1} |s_{l+1} - s_l|^2, \quad (3.49)$$

where $s_l = (u_l - u_{l-1})/(t_l - t_{l-1})$ is the slope of the control in each discretization interval l. The factor $(N-3)$ in c_N accounts for the number of summation terms, whereas $(N-1)^2$ is an approximation of the average step size. This formulation scales the regularization term according to the chosen resolution of the approximation, such that the effect of the user-specified parameter c_{reg} is invariant. The regularization term (3.49) is summed over all control inputs and added to the discretized form of the objective (3.46a).

3.4 Transient model refinement

This section describes the methods to refine the dynamic part of the model. The interaction between the air-path and combustion models is indicated at the appropriate locations in the text.

3.4.1 Refinement of the dynamic air-path model

A generalization of a Kalman filter is applied to the dynamic part of the model [38]. In order to correct for systematic model errors, the model equations are augmented by the dynamic and the static corrective variables $\tilde{\boldsymbol{x}}_{\mathrm{f}}$ and $\tilde{\boldsymbol{x}}_{\mathrm{g}}$,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)) + \boldsymbol{K}_{\mathrm{f}} \cdot \tilde{\boldsymbol{x}}_{\mathrm{f}}(t), \quad (3.50a)$$

$$\boldsymbol{x}_{\mathrm{corr}}(t) = \boldsymbol{x}(t) + \boldsymbol{K}_{\mathrm{g}} \cdot \tilde{\boldsymbol{x}}_{\mathrm{g}}(t). \quad (3.50b)$$

The matrices $\boldsymbol{K}_{\mathrm{f}}$ and $\boldsymbol{K}_{\mathrm{g}}$ define which dynamics are adjusted and for which state variables the absolute values are corrected.

The optimal-control framework described in Sec. 3.3.1 is used to derive the trajectories of the corrective variables. As objective, the minimization of the integrated squared error in the corrected state variables is used, yielding a least-squares fit of these trajectories. The original state variables remain in the OCP, but the control inputs become time-varying parameters. Instead, the corrective variables $\tilde{\boldsymbol{x}}(t) = (\tilde{\boldsymbol{x}}_{\mathrm{f}}^T(t), \tilde{\boldsymbol{x}}_{\mathrm{g}}^T(t))^T$ are optimised in the OCP,

$$\min_{\boldsymbol{x}(t), \tilde{\boldsymbol{x}}(t)} \int_0^{t_{\mathrm{f}}} \boldsymbol{b}^T \cdot \boldsymbol{K}_{\mathrm{g}}^T \cdot \Big(\underbrace{\boldsymbol{x}(t) + \boldsymbol{K}_{\mathrm{g}} \cdot \tilde{\boldsymbol{x}}_{\mathrm{g}}(t)}_{\boldsymbol{x}_{\mathrm{corr}}(t)} - \hat{\boldsymbol{x}}(t) \Big)^2 \, dt \quad (3.51a)$$

$$\text{s.t.} \qquad \dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \hat{\boldsymbol{u}}(t)) + \boldsymbol{K}_\mathrm{f} \cdot \tilde{\boldsymbol{x}}_\mathrm{f}(t). \tag{3.51b}$$

Quantities with a hat, e.g. $\hat{\boldsymbol{x}}$, denote measured signals. The vector $\boldsymbol{b}$ contains the weights to put more emphasis on the accuracy of some of the corrected state variables. The regularization term introduced in Sec. 3.3.3 can be used to penalize fast changes of corrective variables. In fact, smooth trajectories are desirable since the model errors are assumed to be of a systematic nature.

If the data from a single measurement is used to identify the corrective variables by solving (3.51), a perfect match of the corrected state trajectories is obtained by adjusting $\tilde{\boldsymbol{x}}_\mathrm{g}$ only. Therefore, multiple measurements need to be considered simultaneously. In the optimal-control framework, N_m instances of the model are stacked to yield a new system with $N_\mathrm{m} \cdot n_\mathrm{x}$ state variables. The error in the relevant outputs is cumulated over all measurements.

$$\min_{\boldsymbol{x}^{(1)}(t),\ldots,\boldsymbol{x}^{N_\mathrm{m}}(t),\tilde{\boldsymbol{x}}(t)} \sum_{k=1}^{N_\mathrm{m}} \int_0^{t_\mathrm{f}} \boldsymbol{b}^T \cdot \boldsymbol{K}_\mathrm{g}^T \cdot \Big(\underbrace{\boldsymbol{x}^{(k)}(t) + \boldsymbol{K}_\mathrm{g} \cdot \tilde{\boldsymbol{x}}_\mathrm{g}(t)}_{\boldsymbol{x}_\mathrm{corr}^{(k)}(t)} - \hat{\boldsymbol{x}}^{(k)}(t) \Big)^2 \, dt \tag{3.52a}$$

$$\text{s.t.} \qquad \begin{pmatrix} \dot{\boldsymbol{x}}^{(1)}(t) \\ \vdots \\ \dot{\boldsymbol{x}}^{(N_\mathrm{m})}(t) \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}(\boldsymbol{x}^{(1)}(t), \hat{\boldsymbol{u}}^{(1)}(t)) + \boldsymbol{K}_\mathrm{f} \cdot \tilde{\boldsymbol{x}}_\mathrm{f}(t) \\ \vdots \\ \boldsymbol{f}(\boldsymbol{x}^{(N_\mathrm{m})}(t), \hat{\boldsymbol{u}}^{(N_\mathrm{m})}(t)) + \boldsymbol{K}_\mathrm{f} \cdot \tilde{\boldsymbol{x}}_\mathrm{f}(t) \end{pmatrix}. \tag{3.52b}$$

3.5 Iterative procedure

As introduced at the beginning of this chapter, the main purpose was to set up a framework for the iterative dynamic optimization of diesel engines. Building up such a framework has been possible thanks to the great flexibility provided by the experimental setup.

The iterative procedure relies on two modes of operating the engine:

A) The ECU with its standard calibration controls the engine. The test-bench controller is used to follow the desired profiles of the engine speed (by controlling the brake torque) and the load torque (by controlling the fuelling). This mode is used for the initialization of the iterative procedure. The resulting trajectories of the controls, including the injected fuel mass, are recorded.

B) Time-resolved trajectories are prescribed for all control inputs using the rapid-prototyping module and the bypasses of the ECU. The test-bench controller is only used to follow the engine-speed profile. This mode is used for validation runs as well as for all variations.

The inputs to the procedure are a transient driving profile (3.25), consisting of engine speed and load torque trajectories, and any calibration of the ECU that is able to operate the engine along this profile.

1. Initialization: drive the profile in mode A. Save the resulting trajectories of the control inputs, set them as the "current controls" $\boldsymbol{u}(t)$. Set the corrective variables $\tilde{\boldsymbol{x}}(t)$ to constant zero.

2. Perform $1 + 2 \cdot n_{\mathrm{u}}$ testbench runs in mode B. Thereby, apply

   (a) the current controls $\boldsymbol{u}(t)$, and

   (b) isolated perturbations of all controls in both directions, i.e. $u_i(t) \leftarrow u_i(t) \pm \Delta u_i$, for $i = 1, \ldots, n_{\mathrm{u}}$.

3. Use the measurement data from step 2 to identify the corrective variables $\tilde{\boldsymbol{x}}(t)$ by solving (3.52). Run simulations of the refined air-path model for all variations, and save the resulting corrected state trajectories $\boldsymbol{x}_{\mathrm{corr}}(t)$.

4. Use the state trajectories of the air-path model from step 3 to identify the torque and emission models around the current references by solving (3.43).

5. Solve the control and state constrained OCP (3.46) to derive the improved control trajectories $\boldsymbol{u}^*(t)$. Thereby, set

$$\underline{\boldsymbol{u}}(t) = \boldsymbol{u}(t) - k_{\mathrm{u}} \cdot \Delta\boldsymbol{u}, \quad \overline{\boldsymbol{u}}(t) = \boldsymbol{u}(t) + k_{\mathrm{u}} \cdot \Delta\boldsymbol{u}, \tag{3.53a}$$

$$\underline{\boldsymbol{x}}(t) = k_{\mathrm{x}} \cdot \min\{\boldsymbol{x}(t)^{(k)} |\ k = 1, \ldots, 1 + 2 \cdot n_{\mathrm{u}}\}, \tag{3.53b}$$

$$\overline{\boldsymbol{x}}(t) = k_{\mathrm{x}} \cdot \max\{\boldsymbol{x}(t)^{(k)} |\ k = 1, \ldots, 1 + 2 \cdot n_{\mathrm{u}}\}. \tag{3.53c}$$

6. Set $\boldsymbol{u}(t) \leftarrow \boldsymbol{u}^*(t)$, repeat steps 2.-5. until the change in the controls is small.

Using the simulated state trajectories during the identification of the combustion model in step 4 ensures a consistent prediction of the emissions and of the torque inside the validity region. This fact is important since the optimization in step 5 is also restricted to this region. This "trust region" can be slightly expanded by the factors k_{u} and k_{x}.

Since the identification of the models for the air path and for the combustion are identified separately, the physical causality is preserved. More

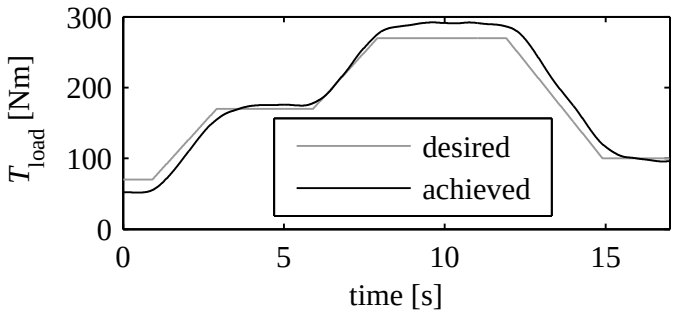


Figure 3.25: The desired load-torque profile and the trajectory achieved by the testbench controller.

precisely, there is no way that the combustion model corrects errors in the air-path model or vice-versa. Although a combined identification could yield a slightly higher accuracy, it would introduce unphysical cross corrections. Furthermore, the identification procedure would become more complex and non-convex.

3.6 Results

A short test cycle is used to evaluate the methods presented in the preceding sections. The engine is operated at a constant speed of 2500 rpm, and the desired load-torque profile is shown in Fig. 3.25. As variations, an additive offset of ±5% PWM is applied to the VGT and EGR positions, and ±2° to the SOI. A multiplicative factor of 1.05 defines the variation of the fuel mass. These variations are referred to as the *small* variations in the remainder of the text. To assess the accuracy of interpolation and extrapolation, *large* variations with offsets of ±8% for the VGT, ±10% for the EGR, ±4° for the SOI and a factor of 1.1 for the fuel mass are performed. Since the VGT and the EGR both dynamically affect the burnt-gas fraction and the pressure in the intake manifold, the four *cross variations* of these two actuators are additionally recorded.

For the identification of the combustion model, it would be desirable to have a constant offset in $\boldsymbol{w}$. However, the EGR-VGT controller cannot perfectly follow reference trajectories for the burnt-gas fraction and the boost pressure. After initial tests, it has been found that applying constant offsets directly to the two dynamic controls is the best choice.

3.6.1 Transient air-path model refinement

The turbocharger speed and the pressure in the exhaust manifold represent the relevant dynamics in a turbocharged engine system with EGR. The boost

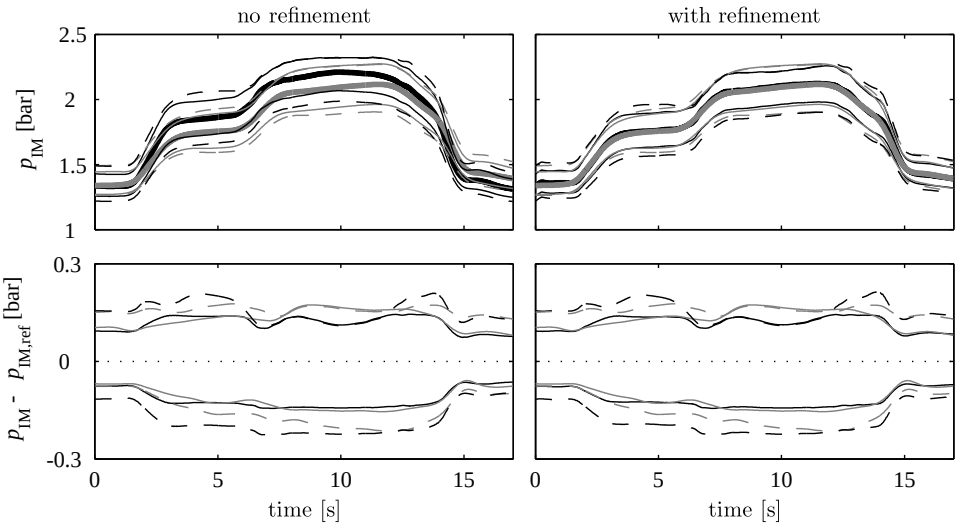


Figure 3.26: Effects of the dynamic model refinement, VGT variations. Measurement data (grey) versus model outputs (black). Line styles: reference (bold), small variations (solid), large variations (dashed).

pressure closely follows the turbocharger speed and the EGR mass-flow is defined, aside from the position of the EGR valve, by the pressure ratio between exhaust and intake manifolds. Therefore, it is sufficient to correct these two dynamics.

Since for the combustion model only the *relative* accuracy is of interest, it is not necessary to require $\boldsymbol{v}(t)$ to be accurate in terms of absolute values. Rather, the state variables that are critical for a safe engine operation should be matched to the measured trajectories to enable an accurate limitation of these quantities in the OCP. Since no limits on any state variables are included in the OCP as of yet, the pressures in the exhaust and intake manifolds have been chosen to illustrate the methodology. The choice of the dynamics and the state variables to be corrected are represented by

$$\boldsymbol{K}_{\mathrm{f}}^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{K}_{\mathrm{g}}^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{3.54}$$

The errors in the two pressures are equally weighted, i.e. $\boldsymbol{b}^T = (1\ 1)$.

The top plot in Fig. 3.26 shows the measured trajectories of the intake pressure along with the model output before and after the refinement. The refinement is performed using only the small variations. The resulting good match for the large variations indicates that the model errors in fact are systematic. For example, the too fast speedup of the turbocharger predicted by the model might be caused by the omission of the thermal models. The

Table 3.3: Static combustion model: average magnitude of the relative error in % for the instantaneous NO_x and soot emissions, and the torque T_{load}

ident. data:	**small variations**			**small & cross vars.**			**large variations**		
	NO_x	soot	T_{load}	NO_x	soot	T_{load}	NO_x	soot	T_{load}
small vars.:	0.14	0.89	0.09	0.70	4.45	0.25	1.14	7.70	0.75
cross vars.:	2.85	15.84	0.82	1.60	12.41	0.27	2.60	18.92	0.73
large vars.:	3.31	20.56	2.36	3.46	21.17	2.11	0.19	1.10	0.09

heat losses to the manifold walls are neglected and consequently, the enthalpy available to the turbine is overestimated.

The bottom plot shows the difference of the variations to the reference trajectory. Obviously, the refinement has no influence on the predicted difference. Therefore, the refinement of the air-path model is not critical for the combustion model but rather a tool that allows an accurate limitation of any state variable, e.g. maximum turbocharger speed, exhaust-manifold pressure and temperature, etc.

In the context of the generalized Kalman filter, the regularization described in Sec. 3.3.3 may be used to enforce smooth trajectories of the corrective variables $\tilde{\boldsymbol{x}}$. The demand for smooth trajectories is justified by assuming the model errors to be of a systematic nature and thus not to exhibit a stochastic or arbitrarily fast changing behavior. For the results presented here, a value of $c_{reg} = 50$ is used.

3.6.2 Static combustion model

The results presented here are all derived using the time averaging introduced in Sec. 3.2.4. Interpolation and extrapolation are not influenced by the time averaging. Table 3.3 shows the errors of the combustion model when identified using three different data sets, namely the small variations, the small and the cross variations, and the large variations. The figures indicate that the model identified using the small variations is able to accurately predict the cross variations, except for the soot emissions. Furthermore, interpolation is significantly more reliable than extrapolation. This fact encourages the use of rather large variations and small factors k_u and k_x instead of the reliance on small variations and extrapolation.

3.6.3 Optimal control

For the OCP, the regularization was set to $c_{reg} = 100$. This choice successfully suppresses oscillating solutions while not affecting the parts that exhibit smooth trajectories anyway. Especially for the SOI, which has no influence on the air-path dynamics, fast oscillations result when no regu-

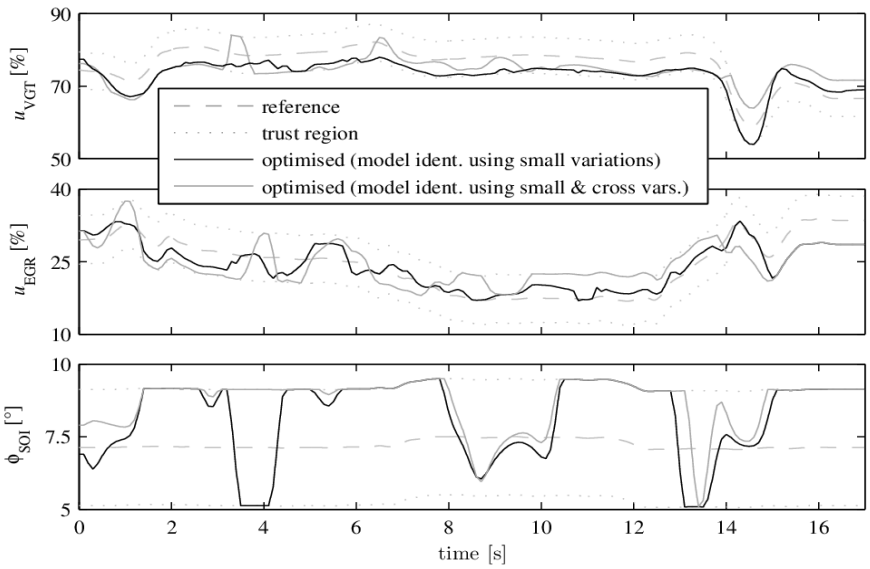


Figure 3.27: Controls resulting from the solution of the constrained OCP.

larization is applied. The time averaging during the identification of the combustion model has a similar effect to the regularization. Determining which approach provides the most plausible solutions and the fastest convergence of the iterative procedure, is left to be analyzed in more detail.

Figure 3.27 shows the optimal control trajectories obtained when using the model identified by the small variations only, or by the small and the cross variations. In both cases, no extrapolation is performed, i.e. $k_{\mathrm{u}} = 1$ and $k_{\mathrm{x}} = 1$. The limits for the cumulative emissions, $\hat{\boldsymbol{m}}_{\mathrm{em}}$, are chosen such that the brake-specific emissions of the reference are maintained. The torque profile resulting from the reference control trajectories is imposed by (3.46d).

The optimal control trajectories are experimentally validated on the engine testbench. To account for possible deviations in the resulting torque, the emissions and the fuel consumption are related to the integrated engine power. Figure 3.28 summarizes the measured cumulative emissions and the fuel consumption.

Five runs are recorded for each set of trajectories. The average as well as the minimum and maximum values are shown in the plots. The emissions remain within the measurement uncertainty, while the fuel consumption is reduced. The soot measurement exhibits a large variability, which is also present in the identification data. Therefore, the corresponding model is not reliable and possibly hinders a more effective reduction of fuel consumption. Furthermore, due to this inconsistency, the model quality seems not to improve when including the cross variations in the identification data.

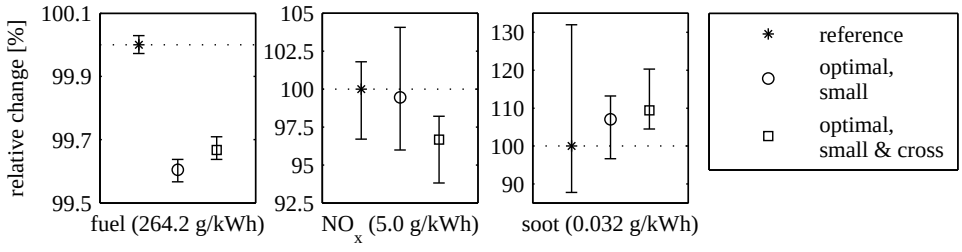


Figure 3.28: Experimental results of the first iteration. The error bars indicate the range (minimal to maximal values) of the five measurements.

3.6.4 Conclusion

In conclusion, this chapter has presented the numerical methods and the testbench setup, required to perform an iterative dynamic optimization of diesel engines over prescribed driving profiles. One exemplary iteration has been performed and experimentally validated, in order to assess the validity of the methodology. A recognizable progress towards lower fuel consumption while maintaining the emission levels has been observed.

Further development of the methodology focused firstly on the replacement of the fast but manually calibrated, failure-prone and high-maintenance emission-measurement devices by standard instrumentation, and secondly on a fully automated implementation of the iterative procedure.

These mentioned further developments, along with an application of the presented methodology exclusively oriented to engine control during transient operation, will be the subject of next chapter.

Chapter 4

Transient control of diesel engines

Due to ever more challenging goals, in terms of fuel economy improvement and reduction of exhaust emissions, noise and vibration, complexity of engine systems has proportionally increased. New technologies help to fulfill such demanding requests, but they come at a price. The increased complexity and the exploitation of the entire hardware potential introduce two important subjects, which are optimal engine calibration and transient control. The aim of the present thesis is to merge these two requirements, or rather to apply the optimal control theory to achieve an incisive transient control.

Static optimization of the control inputs of an engine is the first step towards the calibration of an engine control unit. This optimization can be carried out directly on the real engine or by utilizing models to represent the engine [39, 40]. The number of actuations influencing the engine performance is increasing, and, as a consequence, the effort in the calibration of control parameters can be very costly and demanding. Often, with the aim of speeding up the calibration process several methodologies can be applied, for e.g. statistical tools [41] or automatic calibration procedures carried out along engine transients [42, 43]. The resulting steady-state maps usually constitute the first step to parameterize the engine control unit in form of a feedforward control structure. Without any further manipulation, this kind of control system cannot capture dynamic phenomena, leading to high emission peaks and fuel inefficiency during transient operation. An alternative approach could take into account engine dynamics through mathematical models, in order to implement a model-based rapid transient calibration optimization process [44].

The following chapter is still focused on transient operation of diesel engines, but it shifts the attention towards a control-oriented approach [45]. In order to derive a transient controller, engine modelling for control purposes

is the first step. Hence, control oriented models (COM) of engines have been presented in several works [7, 46, 47, 48]. Although many COMs have been developed for the purpose of transient control of different types of engines, deriving accurate physics-based models, which cover the whole engine operating region, is a difficult and not always successfully achievable task [49]. For this reason, in this context an alternative approach is proposed; it involves a custom-tailored model that is valid in a region around the actual state and input trajectories.

With the difficulties in estimating engine responses, transient control of engines can be realized by using feedforward controls, based on steady-state actuator set-point maps, and transient compensation maps. However, generating compensation maps suitable for transient operation is of crucial importance. This role often relies on engineering experience more than on engine physics, resulting in a highly iterative calibration process, which is not systematically repeatable and provides only suboptimal results.

Chapter 3 focused on the numerical methods needed to derive optimal control trajectories over a predetermined driving cycle. Besides the validity and the usefulness of the methodology in itself, it might be utilized as a tool to derive optimal transient compensation maps, implementable in a feedforward control structure, as it will be shown in next paragraphs. From the optimal solutions, the relevant information may in fact be extracted and stored in maps spanned by the engine speed and the torque gradient. These maps complement the static control maps by accounting for the dynamic behavior of the engine. The procedure is implemented on a real engine and experimental results are presented along with the development of the methodology. The experimental setup is the one already presented in Sec. 3.1.

4.1 Methods

Although most methods used in this context have been presented in details in Chapter 3, in some sections they need to be reminded for completeness, and their application will be clarified whenever it differs from the one previously illustrated.

4.1.1 The concept

The top left corner of Figure 4.1 shows a typical feedforward control structure based on steady-state control maps, spanned by engine speed (N_{eng}) and load (m_{fcc}). In a diesel engine, if neither torque limitation nor smoke limitation are active, the load request directly corresponds to a specific fuel quantity injected per cycle in each cylinder (m_{fcc}). To investigate the typical dynamic behavior of the engine, load transients at constant engine speed are performed by operating the engine over load ramps, and the results are

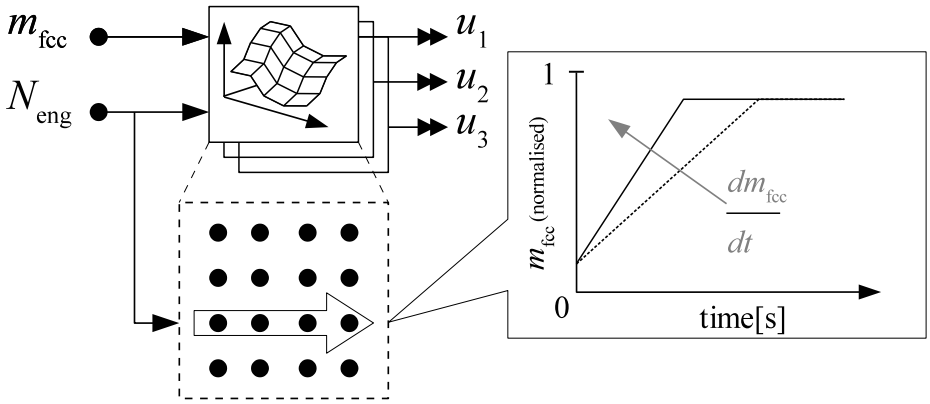


Figure 4.1: Feedforward (FF) control based on the steady-state actuator set-point maps

analyzed in terms of fuel consumption and emissions. With the aim of exploring the entire load range, fuel ramps from 10 to 85 percent of maximum load and with different durations, have been carried out on the testbench (see Figure 4.1).

Looking at Figure 4.2, the slowest ramp (30 s duration) can be considered to represent the static case, i.e. to consist of a sequence of steady-state conditions. The results in Table 4.1 show cumulative fuel consumption, NO_x and soot emissions variations with respect to this stationary-like profile. From now on, the values of the three quantities just mentioned are always normalized with respect to the integrated engine power, to allow for meaningful comparisons. It turns out that the higher is the load gradient, the higher is the specific fuel consumption. Despite this statement could be quite predictable, figuring out the responsible physical effect is of sure interest. The third graph in Figure 4.2 shows that the increasing pumping effect corresponding to increasing gradients, is probably the main reason for the fuel penalty. In other words, the faster is the transient the higher is the negative backpressure effect caused by the turbine. This effect disappears as soon as quasi-static state conditions are reached.

This is the crucial point: a stationary lookup map by nature is not able to account for dynamic effects, so it sets the actuator set-point to the value that will be appropriate when the transient is over. This yields, for instance, a VGT position that increases the engine backpressure with negative effect on efficiency. When the VGT is closed in order to increase the intake manifold pressure (p_{IM}), the exhaust pressure (p_{EM}) increases too, but with a faster dynamic, resulting in a growing effect of the backpressure. This is proven by the third graph, which shows the ratio between p_{EM} and p_{IM}. Moreover, it is also true that the air handling system has a slower dynamic than the

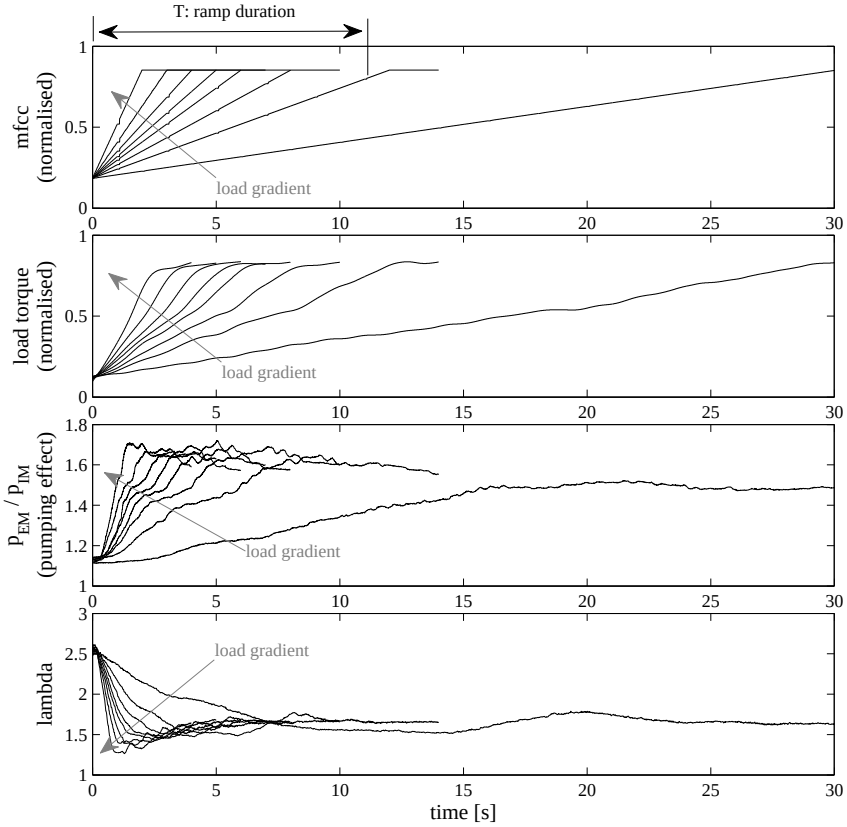


Figure 4.2: Ramp profiles performed at 1950 rpm.

Table 4.1: Effect of increasing load gradients: fuel, NOx and soot related to the integrated engine power (normalized quantities).

ramp duration [s]	$\frac{d}{dt}m_{\text{fcc}}$ $[\frac{mm^3}{s}]$	**fuel**	**NO$_x$**	**soot**
30	*1*	0%(ref)	0%(ref)	0%(ref)
12	*2.5*	2.77%	21.44%	-15.37%
8	*3.75*	4.34%	30.91%	-15.53%
6	*5*	4.42%	41.46%	-18.81%
5	*6*	5.31%	44.42%	-19.24%
4	*7.5*	5.79%	51.58%	-10.42%
3	*10*	6.73%	57.55%	-3.79%
2	*15*	8.30%	65.48%	31.83%

fuelling system, so an air reservoir is needed to prevent an excessive air-to-fuel ratio (AFR) reduction during a positive load gradient. The fourth graph proves this statement, showing the drop of the lambda value, ever more accentuated as the load gradient increases. The phenomenon just explained does not depend only on the turbocharger, but also on the EGR rate, so that the combined effect of these two control levers has to be taken into account.

Obviously all the considerations presented so far have an influence also on exhaust emissions, as shown in Table 4.1. Concerning NO_x emissions, there is a combined effect of the reduced amount of air entering the cylinders and the delayed effect of the EGR, which together lead to a higher combustion temperature. The soot trend is not as monotone as the NO_x one, and the reason can be related to several aspects. First of all, the measurement device is not fast enough to capture the transient peaks, resulting in a potentially misleading cumulative value. It is also reminded that raw emissions have been measured, while the traditional use of the engine at hand sees a DPF installed on the exhaust line, resulting in a drastic reduction of tailpipe soot emissions. Moreover, this aspect does not invalidate the presented results, since, as remarked later, soot emissions are considered in this context as a qualitative quantity not to be widely exceeded.

Another control lever which strongly effects fuel efficiency and NO_x emissions is the crank angle corresponding to the start of injection (SOI). This control input directly influences the combustion process, therefore it does not involve any dynamics. However, it may be used to compensate for the transient effects described above, which is also proposed in [50]. For this reason, the SOI is included in the dynamic optimization procedure proposed below.

Once that the indisputable effect of transients has been remarked, it comes out that finding the perfect combination of multiple cross-coupled control inputs is a difficult engineering problem, which cannot be solved with just a calibration-based approach. Before describing the new control structures derived, all methods concerning engine models and optimal control are outlined in the next sections.

4.1.2 Engine model

The nomenclature is consistent with the one previously introduced, therefore the reader may refer to Figure 3.6. The model structure (Figure 4.3) is basically the same, except for the number of inputs. In this case, not only the engine speed (N_{eng}) but also the fuel mass injected (m_{fcc}) is treated as a time-varying parameter, moreover both are fixed during the optimization procedure.

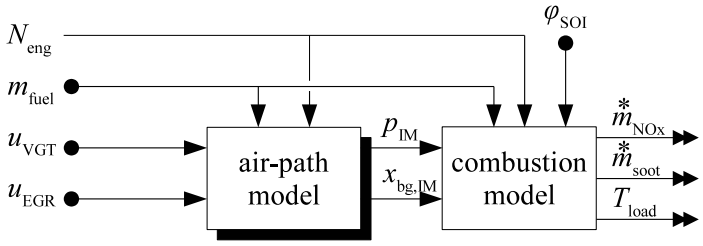


Figure 4.3: Structure of the engine model.

The model is represented in state-space form as:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{4.1a}$$
$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \tag{4.1b}$$

where the $n_\mathrm{x} = 6$ state variables, the $n_\mathrm{u} = 3$ control inputs, and the outputs are:

$$\boldsymbol{x} = \left(p_\mathrm{IM}, \vartheta_\mathrm{IM}, x_\mathrm{bg,IM}, p_\mathrm{EM}, p_\mathrm{aTB}, \omega_\mathrm{TC}\right)^T,$$
$$\boldsymbol{u} = \left(u_\mathrm{VGT}, u_\mathrm{EGR}, \varphi_\mathrm{SOI}\right)^T, \quad \boldsymbol{y} = \left(\overset{*}{m}_\mathrm{NOx}, \overset{*}{m}_\mathrm{soot}, T_\mathrm{load}\right)^T.$$

The air-path model and all related sub-models remain unvaried (see Sec. 3.2.1).

4.1.3 Time-resolved combustion model

Time-variable quadratic setpoint-deviation models are used for the emissions and for the combustion efficiency. The vector $\boldsymbol{w} = (p_\mathrm{IM}, x_\mathrm{bg}, \varphi_\mathrm{SOI})^T$ denotes all inputs to the combustion model. The model for each output thus reads

$$y_i(t) = y_{\mathrm{ref},i}(t) + \boldsymbol{k}_\mathrm{lin}(t)^T \cdot \Delta\boldsymbol{w}(t) + \Delta\boldsymbol{w}(t)^T \cdot \boldsymbol{K}_\mathrm{quad}(t) \cdot \Delta\boldsymbol{w}(t), \tag{4.2a}$$

with

$$\boldsymbol{K}_\mathrm{quad}(t) = \mathrm{diag}\big(k_\mathrm{quad,1}(t), \ldots, k_\mathrm{quad,3}(t)\big), \tag{4.2b}$$
$$\Delta\boldsymbol{w}(t) = \boldsymbol{w}(t) - \boldsymbol{w}_\mathrm{ref}(t). \tag{4.2c}$$

As it can be noted, the input to the combustion model is not exactly the vector $\boldsymbol{w}$, but rather its variation with respect to a reference vector. A detailed description of the identification procedure is presented in 4.1.5, however the reference condition is nothing other than the starting ramp profile, while the real input $\boldsymbol{\Delta w}$ is generated by performing several variations of the control inputs along the reference trajectory.

Sensor dynamics of the NO_x and soot sensors A further critical aspect regarding empirical modelling of engine-out emissions, is the phase shift between transient engine events and transient emission measurements. In the case at hand the NO_x emissions are measured by a Continental UniNOx sensor, while the soot emissions are recorded by means of an AVL Micro Soot Sensor.

Accounting for the transient transport delays and sensor lags, which together constitute the phase shift, is very important for the quality of the model prediction. For this reason, the dynamic characteristic of the sensor signals has been modelled as a first-order lag element, with a time delay resulting from the transport of the gas to the sensor. The signal $\xi_{s,i}$ of emission sensors, can thus be expressed by the differential equation (4.3), where i corresponds to the emission species NO_x or soot:

$$\frac{d}{dt}\xi_{s,i}(t) = -\frac{1}{\tau_{s,i}} \cdot \left[\xi_{s,i}(t) - \psi_i(t - \Delta t_{s,i})\right]. \tag{4.3}$$

The variable ψ_i refers to the corresponding engine-out emission species that reaches the sensor after a delay of $\Delta t_{s,i}$. In order to influence the engine-out emissions separately, measurements with stepwise changes in the SOI and in the EGR valve have been carried out, for several engine speeds and loads. The recorded emission signals have then been used to identify the sensor dynamics of the NO_x and soot sensors, respectively. Figure 4.4 shows the results of the identification. Interesting correlations with the NO_x mass flow (md_{NOx} in figure) and the total exhaust mass flow (md_{EM} in figure) have been observed, respectively for the NO_x and soot sensors. For this reason, the fitted curves shown in red dots have been used to account for the sensors dynamics within the combustion model.

Regarding the soot measurement, the dynamic characteristics of the Micro Soot Sensor are significantly influenced by the dilution ratio. In the event of low ambient temperatures or insufficiently diluted exhaust gas there is the risk of the formation of condensate in the measuring chamber. In order to avoid condensate formation, sufficient dilution of the exhaust gas should be provided for. On the other hand, the delay and the time constant are smaller for low dilution ratios. As a compromise a dilution ratio of five (in a range of 2 to 20) has been used for the experiments.

4.1.4 The optimal control problem

The optimal control problem (OCP) needs a different formulation with respect to the previous one (Sec. 3.3.1), thus it reads:

$$\max_{\boldsymbol{u}(\cdot),\boldsymbol{x}(\cdot)} \left\{ E_{\text{load}} = \int_0^{t_f} T_{\text{load}}\big(\boldsymbol{x}(t), \boldsymbol{u}(t), m_{\text{fcc}}(t)\big) \cdot N_{\text{eng}}\, dt \right\} \tag{4.4a}$$

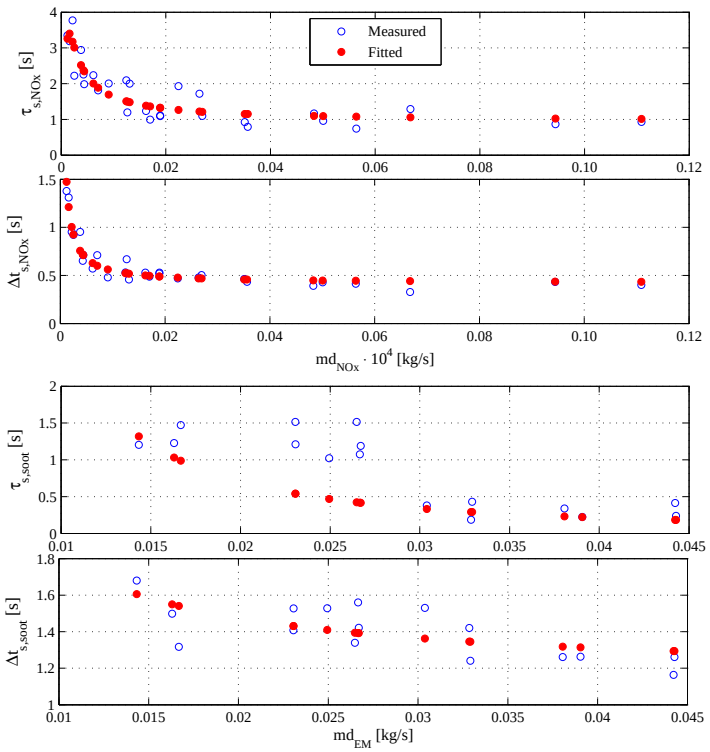


Figure 4.4: Trend of the time constant $\tau_{s,i}$ and the time delay $\Delta t_{s,i}$ identified for the NO_x and soot sensors respectively.

$$\text{s.t.} \qquad \dot{\boldsymbol{x}}(t) - \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), m_{\text{fcc}}(t), N_{\text{eng}}) = \boldsymbol{0}, \tag{4.4b}$$

$$\int_0^{t_f} \overset{*}{\boldsymbol{m}}_{\text{em}}(\boldsymbol{x}(t), \boldsymbol{u}(t), m_{\text{fcc}}(t), N_{\text{eng}})\ dt - \hat{\boldsymbol{m}}_{\text{em}} \leq \boldsymbol{0}, \tag{4.4c}$$

$$\underline{\boldsymbol{u}}(t) \leq \boldsymbol{u}(t) \leq \overline{\boldsymbol{u}}(t), \quad \underline{\boldsymbol{x}}(t) \leq \boldsymbol{x}(t) \leq \overline{\boldsymbol{x}}(t). \tag{4.4d}$$

with $m_{\text{em}} = \left(m_{\text{NO}_x}, m_{\text{soot}}\right)^T$.

The cost function is the integrated power over the load ramp profile, called E_{load} in (4.4a). The instantaneous fuel mass flow $\overset{*}{m}_{\text{fuel}} = m_{\text{fcc}} \cdot \frac{N_{\text{eng}}}{120} \cdot n_{\text{cyl}}$ is integrated to yield the cumulative fuel consumption m_{fuel}. The true objective would be the minimization of the specific fuel consumption, which can be expressed as $:= \frac{m_{\text{fuel}}}{E_{\text{load}}}[\frac{g}{kWh}]$, but the fuel quantity is assigned for each ramp, so it never changes. Consequently, the target becomes the maximization of the energy released, in other words the net engine power (or T_{load} being N_{eng} constant). The dynamic constraints (4.4b) enforce the model equations and equation (4.4c) limits the cumulative emissions. Considering that the soot measurements exhibit a large variability, which is also present in the identification data, the corresponding model is not as reliable as the NO_x model. Despite that, a cumulative but more permissive constraint for soot emissions is set during the optimization anyway; otherwise the optimization

could be negatively effected, resulting in a suboptimal solution for the main objective.

4.1.5 Optimization procedure

Differently from the preceding case (Sec. 3.5), the dynamic optimization procedure is not meant to be iterative. Similarly to the iterative procedure, it relies on two modes of operating the engine, but it has been adapted to the different problem at hand:

A) The ECU with its standard calibration controls the engine, in particular it prescribes the fuel profile related to each ramp. The testbench controller has to guarantee the desired engine speed (constant in this case). This mode is used for the initialization of the optimization procedure. The resulting trajectories of the controls (VGT, EGR and SOI) are recorded.

B) Time-resolved trajectories are prescribed for all control inputs using the rapid-prototyping module and the bypasses of the ECU. The testbench controller has exactly the same task as in mode A). This mode is used for all variations runs as well as for validation.

The inputs to the procedure are the transient profiles of Figure 4.2, performed one at a time, by using a standard engine calibration that is able to operate the engine along these profiles. The following list describes the individual steps of the dynamic optimization:

1. Initialization: drive the profile in mode A. Save the resulting trajectories of the control inputs $\boldsymbol{u}(t)$.

2. Perform $1 + 4 \cdot n_{\mathrm{u}}$ testbench runs in mode B. Thereby, apply

   (a) the reference controls $\boldsymbol{u}(t)$, and

   (b) two perturbations of all controls towards each direction, i.e. $\mathrm{u}_i(t) \leftarrow u_i(t) \pm \{0.5 \cdot \Delta u_i, \Delta u_i\}$, for $i = 1, \ldots, n_{\mathrm{u}}$.

   Run simulations of the air-path model for all variations.

3. Use the state trajectories of the air-path model and the measured emission and torque trajectories from step 2, to identify the torque and emission models around the current references.

4. Solve the control and state constrained OCP (4.4) to derive the improved control trajectories $\boldsymbol{u}^*(t)$. Thereby, set

$$\underline{\boldsymbol{u}}(t) = \boldsymbol{u}(t) - \Delta\boldsymbol{u}, \quad \overline{\boldsymbol{u}}(t) = \boldsymbol{u}(t) + \Delta\boldsymbol{u}. \tag{4.5}$$

In order to automate the optimization procedure, it has been necessary to remotely manage some basic functionalities provided by INCA, such as setting calibrations, reading signals and recording measurements. This additional functionality can be achieved through several methods; in this particular case a COM communication between Matlab and INCA has been established, by means of an in-house Matlab script.

4.1.6 From time-based to map-based optimal control

The optimal control input set, achieved with the optimization procedure, is non causal and valid only for the specific profile used for the optimization, as opposed to a conventional table or map-based control structure. Moreover, referring to Figure 4.1, there is not a unique optimal trajectory for a given speed, but several, corresponding to each ramp performed.

Finding a way to transform these control trajectories to an operating point based representation, while preserving the information about the optimal solution, is thus of particular interest. In view of the fact that keeping intact the content of the time-resolved control vector is unfeasible, its information could be stored in a scalar value of a new lookup table, namely the transient compensation map. Since this map has to be activated only when transients occur, a direct dependency with the load gradient is needed. Therefore, the second quantity, besides the engine speed, that spans the compensation maps is the derivative of the fuel quantity, while the fuel quantity itself is used for the steady-state maps.

Figure 4.5 shows the two solutions proposed, both based on compensation maps that contain the information derived from the solution of the OCP. What distinguishes the two control structures is the way the cylinder charge is managed, while the SOI is treated identically. Controlling the cylinder charge is equivalent to controlling the intake manifold pressure p_{IM} and the burnt-gas fraction $x_{\mathrm{bg,IM}}$, by means of EGR and VGT, which represent the relevant actuators. In the first case, shown in Figure 4.5a, the physical conditions in the intake manifold are not considered at all, resulting in a feedforward (FF) control based on the steady-state actuator set-point maps. Because of its simplicity, this control structure is still used as basic layer in engine control systems. Unfortunately, as soon as a transient occurs all its limits come to light, suggesting that a pure steady-state approach is not satisfactory anymore. Enhancing its performance, by adding information related to transient operation, can be an interesting solution, especially if such a simple structure can be preserved. The compensation maps are highlighted, and their outputs are the dynamic compensation (DC) factors to be applied to VGT and EGR. If the load request gradient is below a certain threshold, set within the compensation map, there is no alteration of the stationary actuator set-point.

The second case, shown in Figure 4.5b, is characterized by an alternative

and more complex kind of cylinder charge control, called air-path control. In this case the actuation of VGT and EGR are feedback (FB) controlled, therefore the set-point variables are p_{IM} and $x_{\mathrm{bg,IM}}$. The implementation of this advanced control structure has been possible thanks to the RCP system presented in Sec. 3.1. All details about the air-path controller implemented and used during the experimental validation, can be found in [51]. Differently from the previous case, the lookup tables store the steady-state set-points of the two state variables mentioned above. At the same time, the compensation maps act directly on these references for the FB controller. If the correction was applied directly to the actuators ($u_{\mathrm{VGT,FB}}$ and $u_{\mathrm{EGR,FB}}$), it would act as a disturbance that the feedback controller would try to compensate. Regardless of the control structure, the dynamic compensation maps have the goal to reproduce as strictly as possible the optimal trajectories. Figure 4.6 helps explaining how they have been derived, for one exemplary ramp profile. The first row shows the comparison between the reference trajectory, obtained by using a standard engine calibration with the ECU running in mode A, and the optimized trajectory (Sec. 4.1.5). VGT and EGR profiles are the real optimal control inputs derived from the optimization procedure, while p_{IM} and $x_{\mathrm{bg,IM}}$ profiles are the corresponding measurements. Since one single ramp profile is considered, which means a single value of the fuel derivative, a single scalar value has to capture the information coming out from this comparison.

The second row shows how the scalar value is calculated. For VGT and EGR, it is the mean value of the difference between optimal and reference trajectory, starting 0.5 seconds after the actual ramp starts (dash-dot line), since the initial conditions of reference and optimal controls coincide. For p_{IM} and $x_{\mathrm{bg,IM}}$, the final value of, respectively the ratio and the difference between the trajectories is considered.

The third and final row highlights the differences between optimal solutions and dynamically compensated reference control inputs. Coherently with the distinction made between the two control structures, VGT and EGR refer to case (a) while p_{IM} and $x_{\mathrm{bg,IM}}$ refer to case (b). For what concerns the SOI, its compensation is identical, no matter the control structure, and likewise to the VGT and EGR cases, the average difference is used to calculate the correction factor (last column).

4.2 Results and discussion

Firstly, the results obtained from the dynamic optimization are presented, by focusing on the capabilities of the optimization tool. Afterwards, the potential of this novel approach to dynamic feedforward control is assessed by the validation procedure. Since the goal was to stimulate the dynamic response of the engine, a transient demanding cycle has been employed.

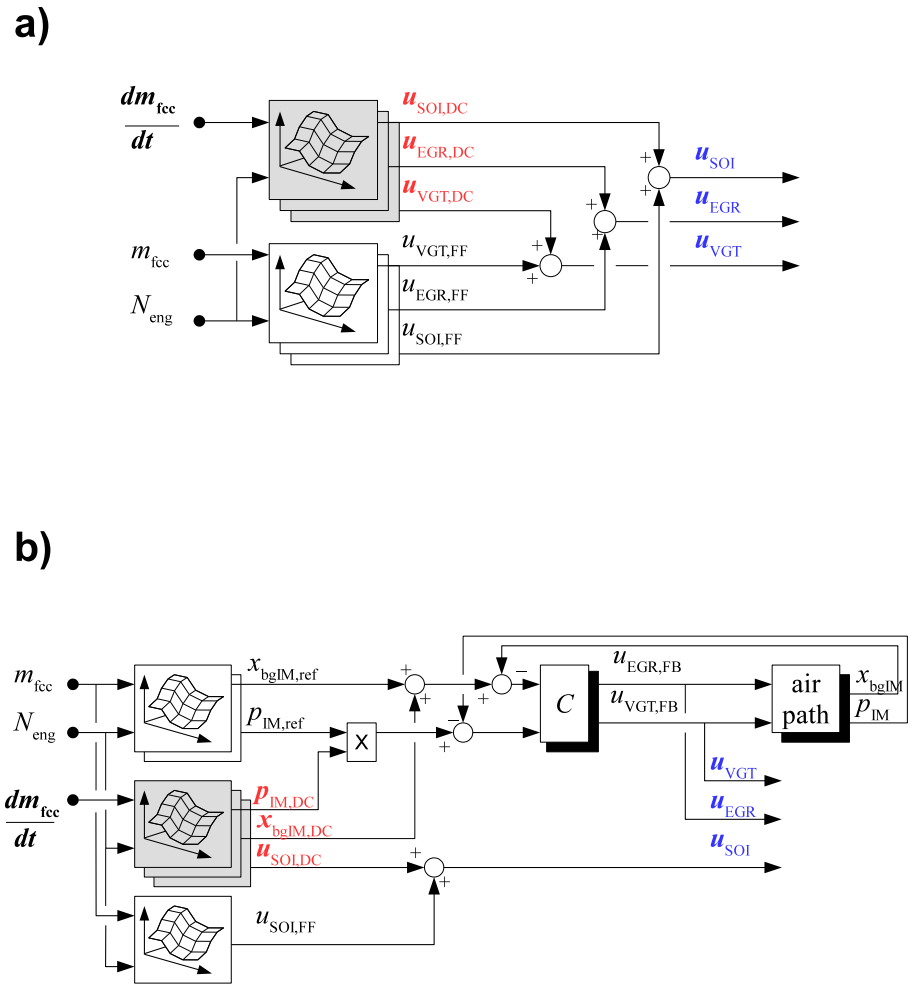


Figure 4.5: a) Feedforward (FF) control based on the steady-state actuator set-point maps: compensation of actuator set-point values. b) Feedback (FB) control of the state variable set-point maps: compensation of state variable set-point values.

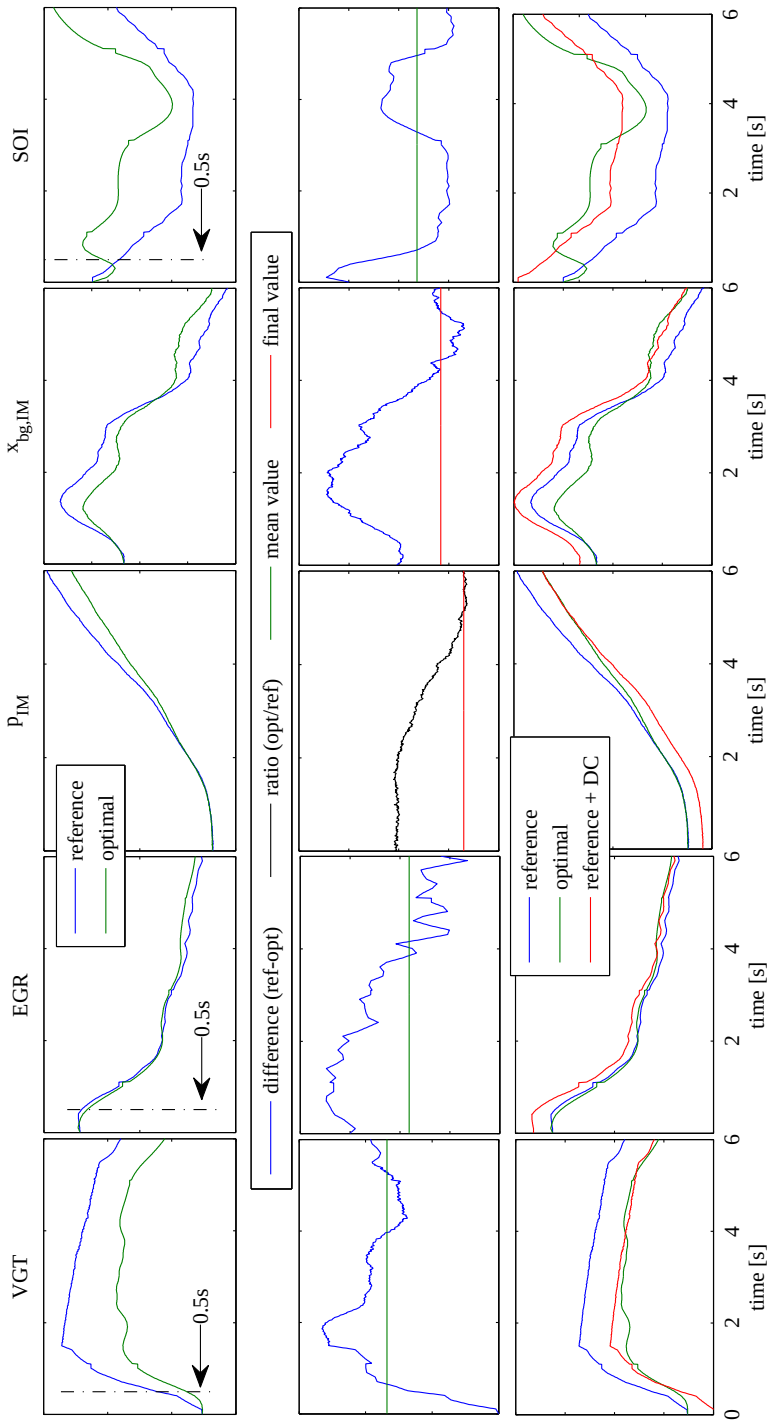


Figure 4.6: Example of calculation of the dynamic compensation factor, for the 6s ramp. First row: comparison between reference and optimal trajectories. Second row: difference (blue line) and ratio (black line) between optimal and reference trajectories. Compensation factors are derived by using either the average (green line) or the final value (red line) of the respective compensation vectors. Third row: comparison between optimal solutions and dynamically compensated reference control inputs.

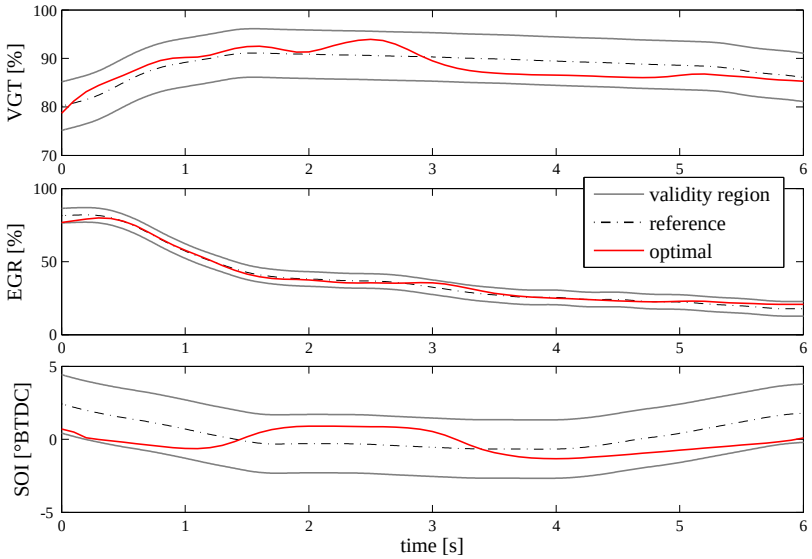


Figure 4.7: Optimal control-input trajectories for 6s duration ramp.

4.2.1 Dynamic optimization

The outputs resulting from the solution of the OCP laid inside the validity region (4.4d) without being constrained. This fact implies that the variations of the inputs chosen, namely $\Delta u_{\mathrm{VGT}} = 0.05, \Delta u_{\mathrm{EGR}} = 0.05$ and $\Delta u_{\mathrm{SOI}} = 2$, yield a good compromise between model quality and broadness of the validity region. If such outputs were limited by the validity limits, a second dynamic optimization would be necessary, shifting the validity region in order to leave the optimizer work properly.

Figure 4.7 shows an exemplary solution coming out from the optimization framework. Table 4.2 shows the relative changes between the optimal and the reference time-based control input set, in terms of fuel consumption and emissions, related to the total energy released. In addition, Figure 4.8 assesses the repeatability of the measurements, performed five times for each ramp, both for reference and for optimal conditions. The objective was to minimize fuel consumption, without penalizing NO_x emissions (4.4c) and keeping soot emissions below an acceptable level. Regarding the former, a significant improvement of fuel efficiency has been accomplished, except for the slowest ramp (30 s duration). The result is plausible since that ramp is sufficiently slow to represent stationary operation. This result is also consistent with the assertion made in Sec. 4.1.1, where this ramp has been considered as the reference condition, in order to assess the effect of transient operation on fuel consumption.

Another proof of the effectiveness of the optimization is the reduction of

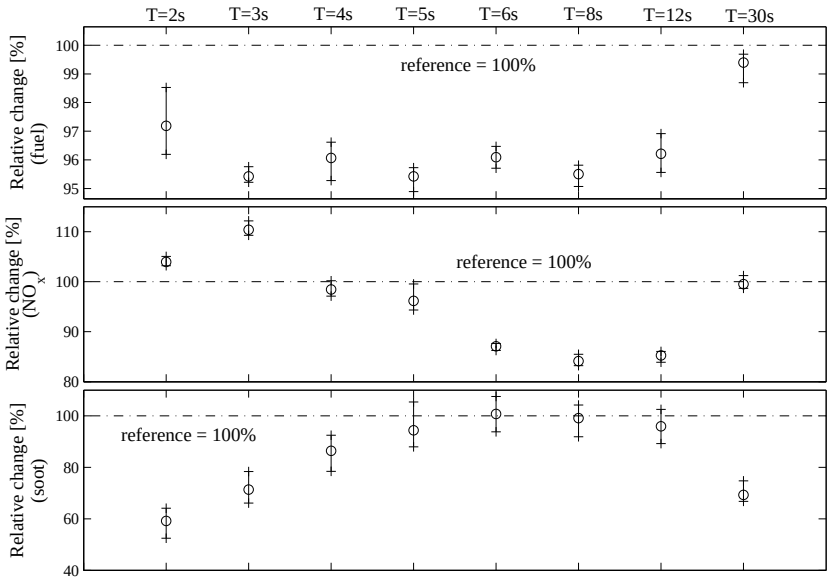


Figure 4.8: Relative change of fuel consumption, NO_x and soot emissions after the optimization. Average value and error bars calculated by repeating the measurement five time for each ramp.

the pumping effect, as compared to suboptimal control of VGT and EGR performed by the production ECU. Figure 4.9 shows a comparison between the reference and the optimal solutions. The negative backpressure effect has been reduced, consistently with the observed fuel consumption reduction, without compromising the air-to-fuel ratio. Likewise to EGR and VGT, the SOI plays a fundamental role, due to its strong effect on fuel efficiency and NO_x emissions.

The usefulness of the optimization framework can be evaluated also from an other perspective. For example, by supposing that the controller is asked to fulfill higher or lower NO_x emissions. In the first case, for instance, an even lower fuel consumption can be achieved, while accepting the increase of NO_x emissions (that could afterwards be reduced by means of an after-treatment system). This result can be obtained only if the engine model correctly predicts the NO_x-fuel tradeoff, and the optimization tool is able to cope with changing constraints. In order to assess this ability, NO_x constraint targets have been altered in positive or negative direction, for two cases representing fast and slow ramps. The results are listed in Table 4.3, and demonstrate that the optimization framework is sensitive to the different requests of NO_x target, showing that lower NO_x emissions allow for higher improvement of fuel efficiency, and vice versa.

Table 4.2: Results of the dynamic optimization for each ramp. Relative changes between optimal and reference condition (average value over 5 repetitions).

ramp duration [s]	fuel,measured	NO_x, measured	soot,measured
2	-2.81%	3.96%	-40.84%
3	-4.58%	10.4%	-28.66%
4	-3.93%	-1.53%	-13.58%
5	-4.58%	-3.85%	-5.63%
6	-3.91%	-12.92%	0.75%
8	-4.50%	-15.87%	-0.92%
12	-3.79%	-14.73%	-4.09%
30	-0.6%	-0.49%	-30.70%

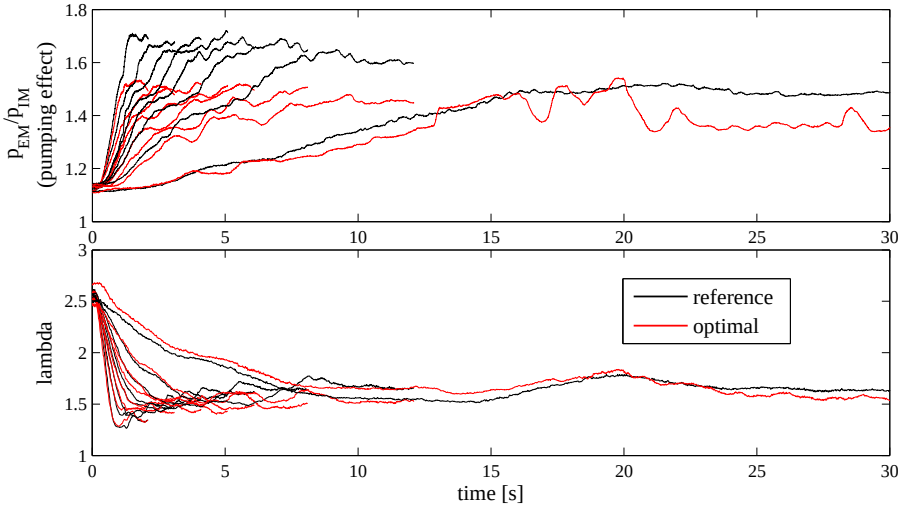


Figure 4.9: Comparison of the pumping effect and the air-to-fuel ratio between reference and optimal ramp profiles.

Table 4.3: Sensitivity analysis of the optimization framework, analyzed for two cases representing fast and slow ramp. Arrows indicate the NO_x target in qualitative form.

ramp	*NO_x target*	*NO_x measured*	*fuel measured*
5s	↓	-16.86%	0.15%
30s	↑	10.37%	-0.57%

Table 4.4: Effect of the implementation of the compensation maps, for three cases representing very fast, fast and medium ramps. Comparison to optimal time-based control, in terms of relative change with respect to reference trajectories.

ramp duration[s]	*fuel*	NO_x	*soot*
2, optimal	-2.81%	3.96%	-40.84%
ref + compens	-0.68%	-8.39%	31.67%
5, optimal	-4.58%	-3.85%	-5.63%
ref + compens	-2.92%	4.33%	26.63%
8, optimal	-4.5%	-15.87%	-0.92%
ref + compens	-3.00%	-1.89%	16.86%

4.2.2 Optimal control versus compensation maps

As explained in Sec. 4.1.6, the time-based optimal control trajectories need to be transformed into a causal form, in order to be used for online control. Moreover, a simple map-based structure is chosen so that the transient compensation strategy may be directly implemented in a standard production ECU (Figure 4.5a). Since the impossibility to exactly copy out the optimal trajectories, an analysis of the effects when going from the non-causal to the causal and implementable structure has been conducted.

To perform this analysis, a specific control structure has been used (Figure 4.10). The ramp is performed and the air-path controller runs on the ECU, without any transient compensation. The control signals are recorded and then set as reference trajectories for the validation run. These control signals are referred as TB (time-based), and they can be applied only when using the ECU in B mode. The dynamic compensation maps act directly on the reference trajectories. This case, from now on called (TB + DC), is not implementable outside dedicated testbench setup, but it is the closest one to the optimal time-based control. Therefore, this procedure highlights the effects of switching from the non-causal optimal solution to the map-based representation, isolated from other influences. Table 4.4 displays the relative changes with respect to the reference trajectories. As expected, a deterioration of the fuel efficiency is unavoidable, but the positive effect produced by the dynamic compensation is still recognizable.

4.2.3 Validation

Figure 4.11 shows the validation cycle. It is important to remind that all the methodology described has been applied for one engine speed (1950 rpm), therefore that speed has been used also for the validation cycle. Three cases,

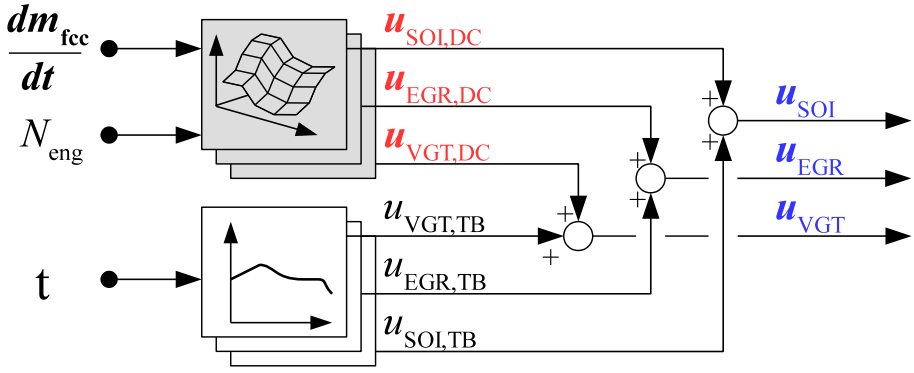


Figure 4.10: time-based (TB) control of actuator set-points: compensation of the time-resolved reference values.

corresponding respectively to three different control structures, have been compared:

1. ***TB + DC*** The non-implementable solution presented in Sec. 4.2.2 (Figure 4.10).

2. ***FF + DC***. The feedforward control based on the steady-state actuator set-point maps runs on the ECU (Figure 4.5a). This is the easiest implementation possible, the compensation maps act on the feedforward actuator outputs.

3. ***FB + DC***. The air-path control introduced in Figure 4.5b runs on the ECU. Only the reference values for the air-path controller are changed, except for the SOI, which is the same as in case 2. Although this implementation is as straightforward as the previous case, the basic control structure, to which the compensation maps are added, is more complex.

In order to analyze the effect of the transient compensation in more detail, two different types of result are presented. Beside the overall effect of the dynamic compensation, the sub-part of the cycle where the compensation is active can be considered, i.e. where the fuel derivative is non negative (Figure 4.12). Table 4.5 summarizes both results for the three cases under discussion.

It turns out that the best result is achieved by implementing case 1, as could be expected. It serves as a benchmark or a guideline to design alternative near-optimal control structures. The case 3 is very promising, especially if considered the fact that it can be applied easily, once an air-path control is available. The case 2 falls slightly short of expectations, because even

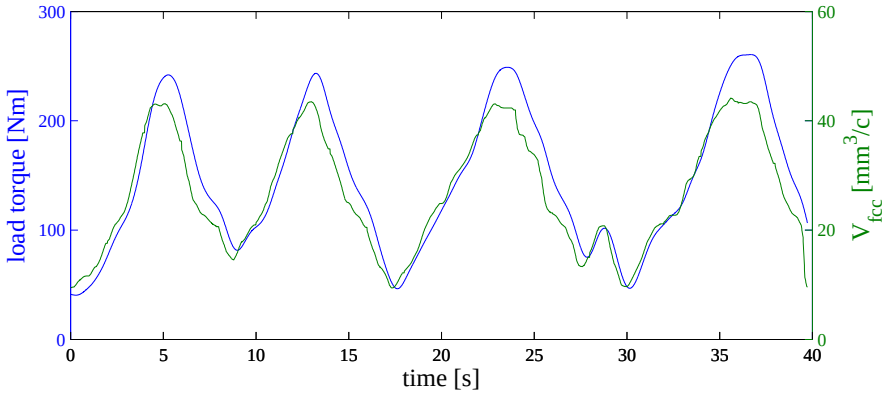


Figure 4.11: Validation cycle at constant engine speed of 1950 rpm. V_{fcc} is the volume of fuel injected per cycle per cylinder.

Table 4.5: Performance of the compensation-map based control systems on the validation cycle, in terms of relative change with respect to reference trajectories.

implementation case	*1)TB+DC*	*2)FF+DC*	*3)FB+DC*
fuel, overall	-2.08%	-1.68%	-1.22%
sub-part	-2.62%	-2.00%	-1.68%
NO_x, overall	0.28%	17.11%	-4.96%
sub-part	-0.19%	15.07%	-1.65%
soot, overall	-0.67%	-30.37%	18.05%
sub-part	-0.55%	-36.80%	25.96%

if there is a significant fuel reduction, it substantially increases the NO_x emissions. Most likely NO_x emissions might be reduced to the detriment of fuel efficiency, according to the well known trade-off between these two quantities. However, these results can be further explained by comparing the different control trajectories, as shown in Figure 4.13. It turns out that the higher is the deviation of the control signals from the "optimal" solution (TB+DC), the worse is the result. In fact, there is a discrepancy between the three cases in terms of approximation of the optimal trajectories, especially for the EGR valve position in case 2. Regarding the latter, the control inputs generated by using the ECU maps are rather far away from the optimal solution, and the transient maps cannot fully compensate such a deviation, resulting in a suboptimal result. However, the simplest case 2, as opposed to the near-optimal but not implementable case 3 which serves as a benchmark, has been employed to assess the effectiveness of the overall methodology.

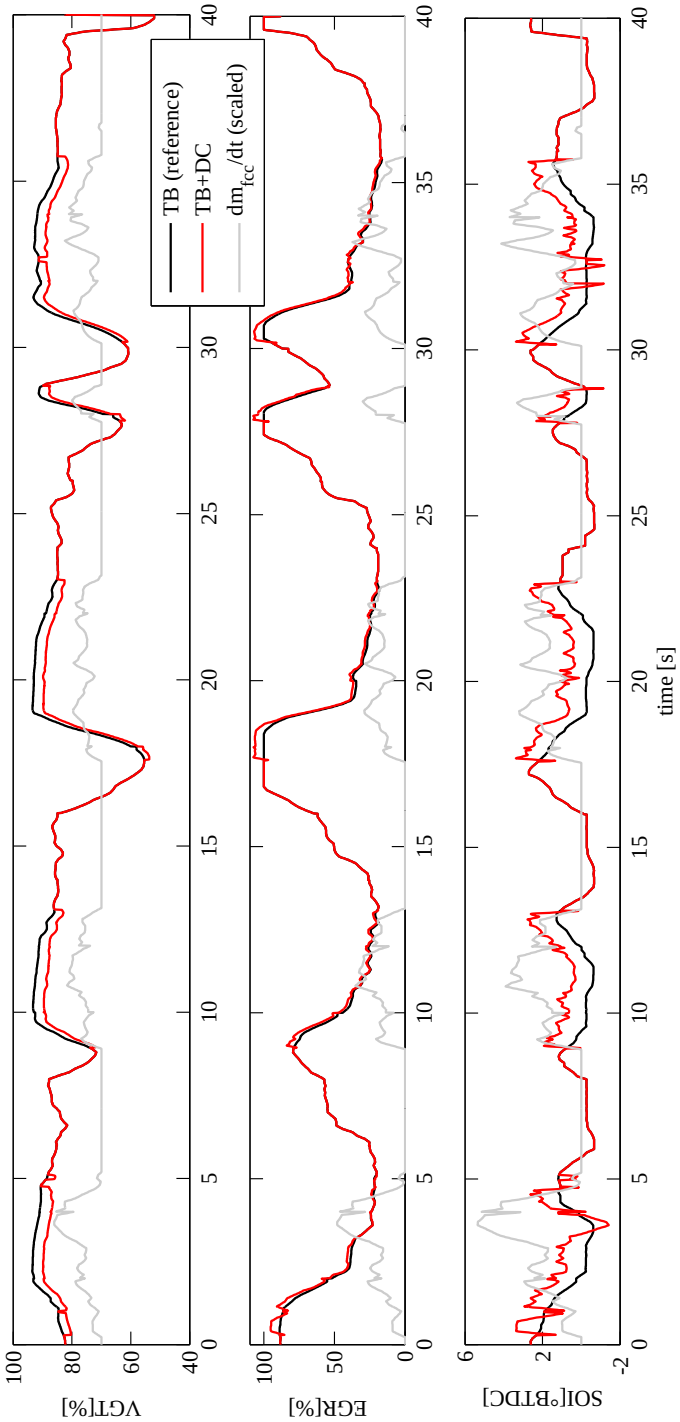


Figure 4.12: Control signals for case 1. In order to highlight when the compensation occurs, the fuel derivative trend is superimposed on the graph.

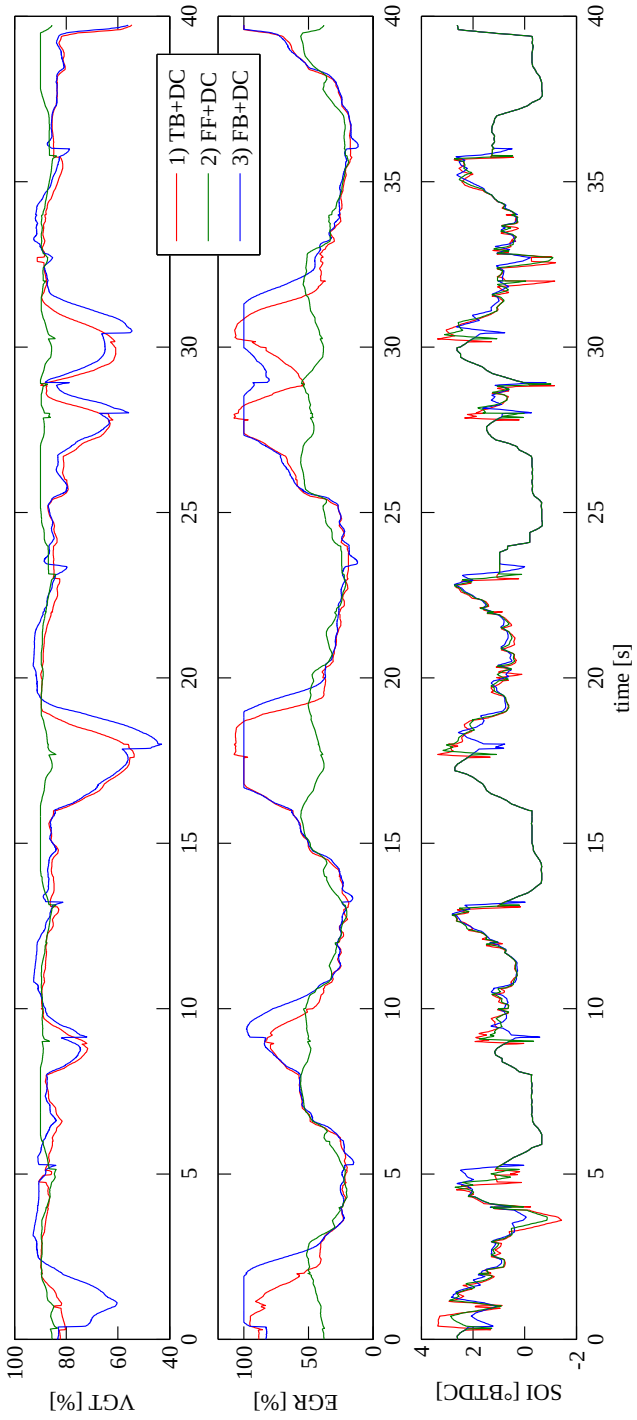


Figure 4.13: Comparison between the three cases tested.

4.2.4 Conclusion

A methodology to derive compensation maps suitable for the transient control of a diesel engine has been presented and experimentally validated. These maps are obtained by means of a dynamic optimization process, of which a beginning-to-end implementation and validation has been detailed.

A significant improvement of fuel efficiency, without compromising emission levels, can be achieved by applying the methodology developed during the doctorate. Moreover, the optimization framework shows a good sensitivity to the NOx-fuel tradeoff when changing the NOx emission constraints. As a consequence, the optimization tool can provide different levels of fuel reduction correspondingly to different tunings of the emission level targets.

Since the main goal was to test the overall validity of the methodology, only a single engine speed has been used. In order to cover the full operating range, the same procedure has to be repeated for different engine speeds. Thanks to the automated optimization procedure developed, the total time needed to perform a full calibration of the transient maps is directly proportional to the number of engine speeds spanned.

As concluded in Sec. 4.2.3, the transient maps cannot fully recover large deviations, typical of static control inputs significantly far away from the optimal solution. Therefore, further development of the methodology will focus on different approaches to extend the performance of the compensation maps to that case. An improvement could be achieved by deriving multi-dimensional maps, which should be able to reproduce the optimal solution with more accuracy.

Chapter 5

Summary and outlook

The research activity carried out during the doctorate, and presented in this thesis, focused on the development of control strategies for diesel engine transient operation.

In Chapter 2, the development of exhaust line heating strategies, based on experimental investigations and previously developed simulation analysis, has been presented. The case study is represented by the SCR installation in a small diesel engine exhaust line, to enhance NO_x reduction in an effort to comply with upcoming EU6 regulations. The challenge was to reach a pre-specified temperature (of approximately 190°C) as fast as possible, far away from the exhaust valves and without compromising fuel consumption. The interesting result is that a substantial SCR light-off time reduction (around 600 s) may be achieved with minor fuel penalties, and this may be obtained by implementing a control strategy that is designed to respect different priorities depending on the SCR thermal state. Possible further improvement could be achieved by exploring the effect of the EGR valve partial closing, which can be possible thanks to the increase in NO_x reduction due to the higher efficiency of the SCR after-treatment system. A fuel consumption reduction is foreseen because of the double effect of a better combustion efficiency and a further reduction of VGT closing at constant boost pressure (due to the higher enthalpy available at the turbocharger inlet).

Chapter 3 has presented the numerical methods and the testbench setup, required to perform an iterative dynamic optimization of diesel engines over prescribed driving profiles. One exemplary iteration has been performed and experimentally validated, in order to assess the validity of the methodology. A recognizable progress towards lower fuel consumption while maintaining the emission levels has been observed.

In Chapter 4, a methodology to derive compensation maps suitable for the transient control of a diesel engine has been presented and experimentally validated. These maps are obtained by means of a dynamic optimization

process, of which a beginning-to-end implementation and validation has been detailed. A significant improvement of fuel efficiency, without compromising emission levels, can be achieved by applying the methodology developed during the doctorate. Moreover, the optimization framework shows a good sensitivity to the NOx-fuel tradeoff when changing the NOx emission constraints. As a consequence, the optimization tool can provide different levels of fuel reduction correspondingly to different tunings of the emission level targets.

As concluded in Sec. 4.2.3, the transient maps cannot fully recover large deviations, typical of static control inputs significantly far away from the optimal solution. Therefore, further development of the methodology will focus on different approaches to extend the performance of the compensation maps to that case. An improvement could be achieved by deriving multi-dimensional maps, which should be able to reproduce the optimal solution with more accuracy.

Appendix A

Sensors

As far as performing measurement is concerned, sensors play a primary role. In this Appendix several kinds of sensors are analyzed, both automotive sensors and "laboratory" sensors, being the latter pretty much utilized in a testbench environment. The list presented below does not fully cover all measurements potentially achievable in a test cell, but focuses on the most widespread sensors in automotive field while highlighting their measuring principle.

A.1 Temperature Sensors

Temperatures are measured by means of thermoresistances and thermocouples. The main distinction to be made lays in their different applicability, being the former type more suitable for production application whereas the second type is exclusively relegated to the testbench environment.

A.1.1 Thermoresistances

Most temperature measurements in the automotive utilize the temperature sensitivity of electric resistance materials with negative temperature coefficient (NTC). The strong nonlinearity enables a large temperature range to be covered (Figure A.1).

For applications with very high temperatures (exhaust gas temperatures up to 1000°C), platinum sensors are employed. The change in resistance is converted into an analog voltage by a voltage grading circuit with an optional parallel resistance to the linearization.

A.1.2 Thermocouples

A thermocouple is a sensor for measuring temperature. It consists of two dissimilar metals, joined together at one end. When the junction of the two metals is heated or cooled a voltage is produced that can be correlated back

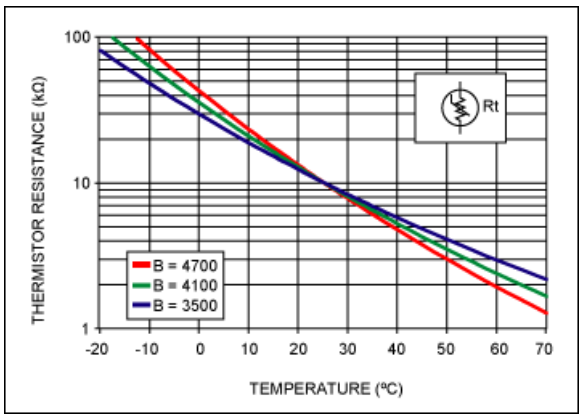


Figure A.1: Typical characteristics of temperature sensors (NTC).

to the temperature. The thermocouple alloys are commonly available as wire.

A thermocouple is available in different combinations of metals or calibrations. The four most common calibrations are J, K, T and E. There are high temperature calibrations R, S, C and GB. Each calibration has a different temperature range and environment, although the maximum temperature varies with the diameter of the wire used in the thermocouple. Although the thermocouple calibration dictates the temperature range, the maximum range is also limited by the diameter of the thermocouple wire. That is, a very thin thermocouple may not reach the full temperature range.

Since a thermocouple measures in wide temperature ranges and can be relatively rugged, thermocouples are very often used in industry. The following criteria are used in selecting a thermocouple:

- Temperature range
- Chemical resistance of the thermocouple or sheath material
- Abrasion and vibration resistance
- Installation requirements (may need to be compatible with existing equipment; existing holes may determine probe diameter)

The response time is an important characteristic to be considered when choosing a thermocouple. A time constant has been defined as the time required by a sensor to reach 63.2% of a step change in temperature under a specified set of conditions. Five time constants are required for the sensor to approach 100% of the step change value. An exposed junction thermocouple is the fastest responding. Also, the smaller the probe sheath diameter, the faster the response, but the maximum temperature may be lower. It has to be noticed that sometimes the probe sheath cannot withstand the full temperature range of the thermocouple type.

A.2 Piezoresistive effect

Piezoresistivity is a common sensing principle for micromachined sensors. First discovered by Lord Kelvin in 1856, the piezoresistive effect is a widely used sensor principle. In a few words, an electrical resistor may change its resistance when it experiences a strain and deformation. This effect provides an easy and direct energy/signal transduction mechanism between the mechanical and the electrical domains. Today, it is used in the MEMS (Micro Electro-Mechanical Systems) field for a wide variety of sensing applications, including accelerometers, pressure sensors, gyro rotation rate sensors, tactile sensors, flow sensors, sensors for monitoring structural integrity of mechanical elements, and chemical/biological sensors.

The resistance value of a resistor with the length l and the cross-sectional area A is given by

$$R = \rho \frac{l}{A} \tag{A.1}$$

The resistance value is determined by both the bulk resistivity ρ and the dimensions. Consequently, there are two important ways by which the resistance value can change with applied strain. First, the dimensions, including the length and cross section, will change with strain. This is easy to understand, though the relative change in dimensions is generally small. Note that transverse strains may be developed in response to longitudinal loading. For example, if the length of a resistor is increased, the cross section will likely decrease under finite Poissons ratios. Secondly, the resistivity of certain materials may change as a function of strain. The magnitude of resistance change stemming from this principle is much greater than what is achievable from the first one. By strict definition, piezoresistors refer to resistors whose resistivity changes with applied strain. Metal resistors change their resistance in response to strain mainly due to the shape deformation mechanism. Such resistors are technically called strain gauges. The resistivity of semiconductor silicon changes as a function of strain. Silicon is therefore a true piezoresistor.

Focusing now on the macroscopic description of the behavior of a piezoresistor under a normal strain, the change in resistance is linearly related to the applied strain, according to

$$\frac{\Delta R}{R} = G \cdot \frac{\Delta L}{L} \tag{A.2}$$

The proportional constant G in the above equation is called the *gauge factor* of a piezoresistor. We can rearrange the terms in this equation to arrive at an explicit expression for G:

$$G = \frac{\frac{\Delta R}{R}}{\frac{\Delta l}{l}} = \frac{\Delta R}{\varepsilon R} \tag{A.3}$$

The resistance of a resistor is customarily measured along its longitudinal axis. Externally applied strain, however, may contain three primary vector components: one along the longitudinal axis of a resistor and two arranged 90° to the longitudinal axis and each other. A piezoresistive element behaves differently towards longitudinal and transverse strain components.

The change of measured resistance under the longitudinal stress component is called longitudinal piezoresistivity. The relative change of measured resistance to the longitudinal strain is called the *longitudinal gauge factor*. On the other hand, the change of resistance under transverse strain components is called transverse piezoresistivity. The relative change of measured resistance to the transverse strain is called the *transverse gauge factor*. For any given piezoresistive material, the longitudinal and transverse gauge factors are different.

It is important to realize that longitudinal and transverse strains are often present at the same time though one of them may play a clearly dominating role. The total resistance change is the summation of changes under longitudinal and transverse stress components, namely

$$\frac{\Delta R}{R} = \left(\frac{\Delta R}{R}\right)_{longit} + \left(\frac{\Delta R}{R}\right)_{transv} = G_{longit} \cdot s_{longit} + G_{transv} \cdot s_{transv} \quad \text{(A.4)}$$

Resistance changes are often read using the Wheatstone bridge circuit configuration. A basic Wheatstone bridge consists of four resistors connected in a loop. An input voltage is applied across two junctions that are separated by two resistors. Voltage drop across the other two junctions forms the output. One or more resistors in the loop may be sensing resistors, whose resistances change with the intended variables. In the bridge shown in Figure A.2a, one resistor (R_1) is variable by strain. The other resistors (R_2, R_3 and R_4) are made insensitive to strains by being located in regions where mechanical strain is zero, such as on rigid substrates.

The output voltage is related to the input voltage according to the following relationship,

$$V_{out} = \left(\frac{R_2}{R_1 + R_2} - \frac{R_4}{R_3 + R_4}\right) \cdot V_{in} \quad \text{(A.5)}$$

In many practical applications, all four resistors share an identical nominal resistance value. A representative case is shown in Figure A.2b. In this case, the resistance of the variable resistor (sensor) is represented as

$$R_s = R + \Delta R \quad \text{(A.6)}$$

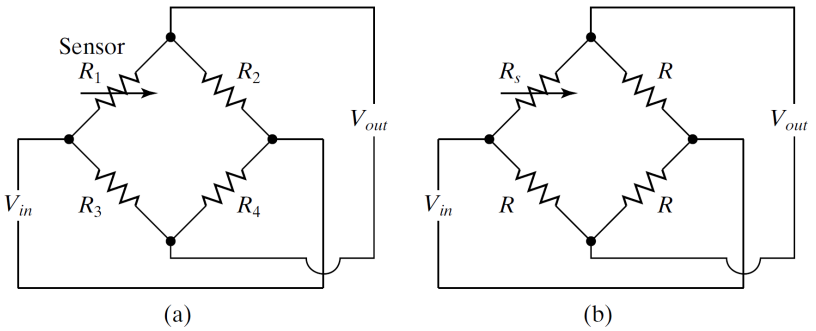


Figure A.2: Wheatstone bridge circuits.

whereas the nominal resistance values of other three resistors are denoted R. The output voltage is linearly proportional to the input voltage according to

$$V_{out} = \left(\frac{-\Delta R}{2R + \Delta R} \right) \cdot V_{in} \tag{A.7}$$

Most piezoresistors are temperature sensitive. For the purpose of eliminating the effect of changing environmental temperature on the output, the Wheatstone bridge is particularly effective. Variation of environmental temperature would cause changes to all resistances in the bridge with the same percentage. Hence, the temperature variation would cause the numerator and the denominator of the right-hand terms of (A.7) to be scaled by an identical factor. The temperature effect is therefore cancelled out.

A.3 Piezoelectric effect

The piezoelectric effect produces an opposed accumulation of charged particles on the crystal. This charge is proportional to applied force or stress. A force applied to a quartz crystal lattice structure alters alignment of positive and negative ions, which results in an accumulation of these charged ions on opposed surfaces. These charged ions accumulate on an electrode that is ultimately conditioned by transistor microelectronics.

There are two types of piezoelectric material that are used for PCB accelerometers: quartz and polycrystalline ceramics. Quartz is a natural crystal, while ceramics are man-made. Each material offers certain benefits, and material choice depends on the particular performance features desired of the accelerometer.

Quartz is widely known for its ability to perform accurate measurement tasks and contributes heavily in everyday applications for time and frequency measurements. Examples include everything from wrist watches and radios

to computers and home appliances. Accelerometers benefit from several unique properties of quartz. Since quartz is naturally piezoelectric, it has no tendency to relax to an alternative state and is considered the most stable of all piezoelectric materials. This important feature provides quartz accelerometers with long-term stability and repeatability. Also, quartz has virtually no pyroelectric effect (output due to temperature change), which provides stability in thermally active environments. Because quartz has a low capacitance value, the voltage sensitivity is relatively high compared to most ceramic materials, making it ideal for use in voltage-amplified systems. Conversely, the charge sensitivity of quartz is low, limiting its usefulness in charge-amplified systems, where low noise is an inherent feature. The useful temperature range of quartz is limited to approximately 600 °F (315 °C).

A variety of ceramic materials are used for accelerometers, depending on the requirements of the particular application. All ceramic materials are man-made and are forced to become piezoelectric by a polarization process. This process, known as "poling," exposes the material to a high-intensity electric field. This process aligns the electric dipoles, causing the material to become piezoelectric. Unfortunately, this process tends to reverse itself over time until it exponentially reaches a steady state. If ceramic is exposed to temperatures exceeding its range or electric fields approaching the poling voltage, the piezoelectric properties may be drastically altered or destroyed. Accumulation of high levels of static charge also can have this effect on the piezoelectric output. PCB uses three classifications of ceramics. First, there are high-voltage-sensitivity ceramics that are used for accelerometers with built-in, voltage-amplified circuits. There are high-charge-sensitivity ceramics that are used for charge mode sensors with temperature ranges to 400 °F (205 °C). This same type of crystal is used in accelerometers that use built-in charge-amplified circuits to achieve high output signals and high resolution. Finally, there are high-temperature ceramics that are used for charge mode accelerometers with temperature ranges to 600 °F (316 °C) for monitoring of engine manifolds and superheated turbines.

A.4 Pressure Sensors

Different sensor types are employed in order to meet the various demands of the pressures to be measured.

Piezoresistive sensors The micromachined pressure sensor was one of the earliest demonstrations of micromachining technology. It is commercially very successful because of several important traits, including high sensitivity and uniformity. Bulk microfabricated pressure sensors with thin deformable diaphragms made of singlecrystal silicon are the earliest products and still dominate the market today. One example is shown in Figure

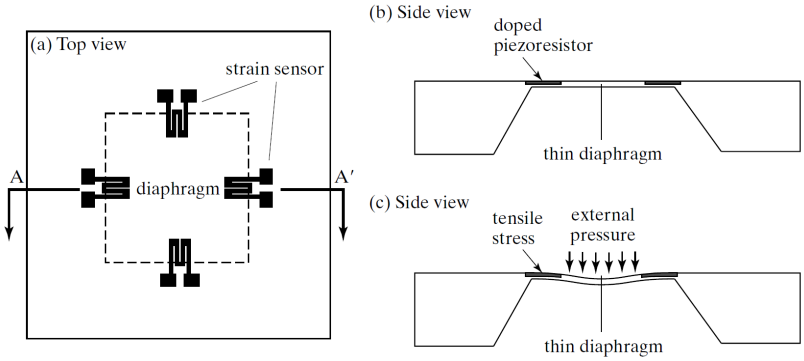


Figure A.3: Piezoresistive pressure sensor.

A.3. Piezoresistors are located in the center of four edges. The location of these piezoresistors corresponds to regions of maximum tensile stress when the diaphragm is bent by a uniformly applied pressure difference across the diaphragm. Four resistors are connected in a full Wheatstone bridge configuration. A fully functional on-chip signal-processing unit consists of two stage amplifiers, compensation circuitry, and two forms of output (frequency and voltage). In the Wheatstone bridge configuration, the temperature sensitivity of the piezoresistors cancels each other. The diaphragm with embedded piezoresistors is made by using silicon bulk micromachining steps. Piezoresistors are made by selectively doping the silicon diaphragm.

Using microfabrication, the diaphragm thickness can be controlled precisely (at approximately $25\mu m$ or below). The sensor chip provides a sensitivity of 4 mV/mm Hg, with the nonlinearity lower than 0.4% over the full scale. The temperature coefficient of the sensitivity is less than 0.06%/°C in the temperature range of to -20 to 110°C.

Piezoelectric sensors Piezoelectric pressure sensors measure dynamic pressures. They are generally not suited for static pressure measurements. Dynamic pressure measurements including turbulence and engine combustion under varying conditions may require sensors with special capabilities. Fast response, ruggedness, high stiffness, extended ranges, and the ability to also measure quasi-static pressures are standard features associated with these quartz pressure sensors.

Piezoelectric pressure sensors are available in various shapes and thread configurations to allow suitable mounting for various types of pressure measurements. Quartz crystals are used in most sensors to ensure stable, repeatable operation. The quartz crystals are usually preloaded in the housings to ensure good linearity. When the crystal is stressed, a charge is generated. This high-impedance output must be routed through a special low-noise cable to an impedance-converting amplifier, such as a laboratory charge

amplifier or source follower. High insulation resistance must be maintained in the cables and connections. The primary function of the charge or voltage amplifier is to convert the high-impedance output to a usable low-impedance voltage signal for recording purposes. Laboratory charge amplifiers provide added versatility for signal normalization, ranging, and filtering. Miniature in-line amplifiers are generally of fixed range and frequency.

Typical applications in automotive field are in-cylinder pressure measurements, but also intake/exhaust manifold pressure measurements whenever the dynamic effects need to be captured.

A.5 Acceleremoters

Acceleration is the rate of change of velocity. Measurement units for acceleration include m/s^2, ft/s^2, and g.

An accelerometer is a sensor, or transducer, which is designed to generate an electrical signal in response to acceleration (or deceleration) that is applied along (parallel with) its sensitive axis.

The applied, or experienced acceleration can fall into one or more of the following categories:

- **Constant Acceleration** - acceleration that does not change during an event. Examples include the acceleration due to earth's gravity or the centrifugal acceleration of a merry-go-round at constant rotational speed.
- **Transient Acceleration** - acceleration that varies over the duration of the event, but is not repetitive. Examples include the deceleration that an automobile undergoes during braking or the acceleration effects experienced by a roller coaster as it negotiates its track. Transient acceleration is the result of discontinuous motion.
- **Periodic Acceleration** - acceleration that continuously varies over the duration of the event, and is quite repetitive. Examples include the vibration of rotating machinery such as motors and bearings or the acceleration experienced by a free-swinging pendulum. Periodic acceleration is the result of continuous motion.

A.5.1 Piezoelectric accelerometers

Piezoelectric accelerometers employ either natural quartz crystals or man-made polycrystalline ceramics as their sensing elements. A proof mass is mated with the crystal and output is generated when a force is imposed upon the crystal during acceleration. This force causes stress in the crystal, which then generates an electrical charge that is relative to the applied force

- the piezoelectric effect. The amount of force is proportional to applied acceleration as governed by Newton's law of motion $F = m \cdot a$. Piezoelectric accelerometers cannot measure constant acceleration because they are inherently AC coupled, however, they are typically the most versatile and economic choice for measuring fast transient and periodic acceleration. In an accelerometer, the stress on the crystals occurs as a result of the seismic mass imposing a force on the crystal. Over its specified frequency range, this structure approximately obeys Newton's law of motion, F=ma. Therefore, the total amount of accumulated charge is proportional to the applied force, and the applied force is proportional to acceleration. Electrodes collect and wires transmit the charge to a signal conditioner that may be remote or built into the accelerometer. Sensors containing built-in signal conditioners are classified as Integrated Electronics Piezoelectric (IEPE) or voltage mode; charge mode sensors require external or remote signal conditioning. Once the charge is conditioned by the signal conditioning electronics, the signal is available for display, recording, analysis, or control.

A very widespread accelerometer, as far as Spark Ignition engines are concerned, is the so called "knock sensor". The name obviously originates from the abnormal combustion of the same name (knock). The functional principle of a knock sensor is typically based on a piezoceramic ring that converts the engine vibrations into electrically processable signals using a superimposed (seismic) mass.

Sensor sensitivity is expressed in mV/g or pC/g and is practically constant over a wide frequency range. The transmission behavior of the knock sensor can be adapted by the choice of the seismic mass.

A.5.2 Piezoresistive accelerometers

Piezoresistive accelerometers may be fabricated from metal strain gauges, piezoresistive silicon, or as a MEMS device. In such designs resistive material is typically bonded to a cantilever beam that undergoes bending under the influence of acceleration. This bending causes deformation of the resistor, leading to a change in its resistance. The resistors are normally configured into a Wheatstone bridge circuit, which provides a change in output voltage that is proportional to acceleration. Piezoresistive accelerometers are capable of measuring constant, transient, and periodic acceleration.

A.5.3 Capacitive accelerometers

Capacitive accelerometers utilize the properties of an opposed plate capacitor for which the distance between the plates varies proportionally to applied acceleration thus altering capacitance. This variable is used in a circuit to ultimately deliver a voltage signal that is proportional to acceleration. Capacitive accelerometers are capable of measuring constant as well as slow

transient and periodic acceleration.

A.6 Microphones

When an object vibrates in the presence of air, the air molecules at the surface will begin to vibrate, which in turn vibrates the adjacent molecules next to them. This vibration will travel through the air as oscillating pressure at frequencies and amplitudes determined by the original sound source. The human eardrum transfers these pressure oscillations, or sound, into electrical signals that are interpreted by our brains as music, speech, noise, etc. Microphones are designed, like the human ear, to transform pressure oscillations into electrical signals, which can be recorded and analyzed to tell us information about the original source of vibration or the nature of the path the sound took from the source to the microphone. The typical audible range of a healthy human ear is 20 to 20000 Hz. Like the human ear, microphones are designed to measure a very large range of amplitudes, typically measured in decibels (dB) and frequencies in hertz (Hz).

In order to convert acoustical energy into electrical energy, microphones are used. There are a few different designs for microphones. The more common designs are Carbon Microphones, Externally Polarized Condenser Microphones, Prepolarized Electret Condenser Microphones, Magnetic Microphones, and Piezoelectric Microphones.

The carbon microphone design is a value-oriented design. This design is a very low quality acoustic transducer type. An enclosure is built. This enclosure houses lightly packed carbon granules. At opposite ends of the enclosure, electrical contacts are placed, which have a measured resistance. When the pressure from an acoustical signal is exerted on the microphone, it forces the granules closer together. This force presses the granules together, which decreases the resistance. This change in resistance is measured and output.

A condenser microphone operates on a capacitive design. The cartridge from the condenser microphone utilizes basic transduction principles and will transform the sound pressure to capacitance variations, which are then converted to an electrical voltage. This is accomplished by taking a small thin diaphragm and stretching it a small distance away from a stationary metal plate, called a *back plate*. A voltage is applied to the back plate to form a capacitor. In the presence of oscillating pressure, the diaphragm will move which changes the gap between the diaphragm and the back plate. This produces an oscillating voltage from the capacitor, proportional to the original pressure oscillation. The back plate voltage can be generated by two different methods. The first is an externally polarized microphone design where an external power supply is used. The power source on this traditional design is 200 volts. The second or newer design is called a prepolarized

microphone design. This modern design utilizes an electret layer placed on the backplane, which contains charged particles that supply the polarization.

A magnetic microphone is a dynamic microphone. The moving coil design is based on the principal of magnetic induction. This design can be simply achieved by attaching a coil of wire to a light diaphragm. Upon seeing the acoustical pressure, the coil will move. When the wire is subjected to the magnetic field, the movement of the coil in the magnetic field creates a voltage, which is proportional to the pressure exerted on it.

A Piezoelectric microphone uses a quartz or man-made ceramic crystal structure, which is similar to electrets in that they exhibit a permanent polarization and can be coupled with an IEPE design. Although these sensor type microphones have very low sensitivity levels, they are very durable and are able to measure very high amplitude (decibels) pressure ranges. Conversely, the floor noise level on this type of microphone is generally very high. This design is suitable for shock and blast pressure measurement applications.

The most popular test and measurement microphones are the capacitor condenser designs.

When choosing the optimum microphone, the parameters to look at include the type of response field, dynamic response, frequency response, polarization type, sensitivity required, and temperature range. There are also a variety of specialty type microphones for specific applications. In order to select and specify a microphone, the first criteria that needs to be looked at is the application and what the sound and environment represent.

Microphones Field Types There are three common application fields for precision condenser microphones. The first and most common is the *free-field type*. The free-field microphone is most accurate when measuring sound pressure levels that radiate from a single direction and source, which is pointed directly (0° incidence angle) at the microphone diaphragm, and operated in an area that minimizes sound reflections.

The second type is called a *Pressure Field*. A Pressure Field microphone is designed to measure the sound pressure that exists in front of the diaphragm. It is described to have the same magnitude and phase at any position in the field. It is usually found in an enclosure, or cavity, which is small when compared to wavelength. The microphone will include the measurement changes in the sound field caused by the presence of the microphone. The sound being measured is typically coming from a single source.

The third type is called a Random Incident Microphone. This is also referred to as a *Diffuse Field Type*. The Random Incident type of microphone is designed to be omni-directional and measure sound pressure coming from multiple directions, multiple sources and multiple reflections. The Random Incident type microphone will have typical correction curves for different

angles of incidence. The random incidence microphone will compensate for its own presence in the field. An average of the net effect of all the calibrated incidence angles will be taken into account, in order to come up with a net zero correction factor.

Dynamic Response The main criteria to describe sound, is based upon the amplitude of the sound pressure fluctuations. The lowest amplitude that a healthy human ear can detect is 20 millionths of a Pascal (20mPa). Since the pressure numbers represented by Pascal's are generally very low and not easily managed, another scale was developed and is more commonly used, called the Decibel (dB). The decibel scale is logarithmic and more closely matches the response reactions of the human ear to the pressure fluctuations. Manufacturers specify the maximum decibel level based on the design and physical characteristics of the microphone.

In order to calculate the maximum output for a microphone, using a specific preamplifier and its corresponding peak voltage, you first need to calculate the pressure in Pascals that the microphone can accept. The amount of pressure can be calculated by using the following formula:

$$P = \frac{Voltage(mV)}{Sensitivity\left(\frac{mV}{Pa}\right)}$$

where P = Pascal's (Pa) and $Voltage$ is the preamps output peak voltage. Once the maximum pressure level that the microphone can sense at its peak voltage is determined, this can then be converted to decibels (dB), using the following logarithmic scale:

$$dB = 20 \cdot log\left(\frac{P}{P_0}\right)$$

where P is the pressure in Pascal's and P_0 is the reference value (0.00002 Pa). The above formula will provide the maximum rating that a microphone (when combined with a specific preamplifier) can be capable of measuring.

Frequency Response Once the type of microphone field response and dynamic range has been taken into consideration, the frequency range (Hz) of interest, for the test requirement should be reviewed. Upon inspecting the microphones specification sheet you will find the usable frequency range of the specific microphone. Smaller diameter microphones will usually have a higher upper frequency level capability. Conversely, larger diameter microphones will be able to detect lower frequencies, generally better.

Polarization Type As explained previously, test and measurement microphones can be broken down into two categories, traditional Externally Polarized microphones and modern Prepolarized microphones. For most

applications either type will work well. The prepolarized tend to be more consistent in humid applications. They are recommended when changes of temperature may cause condensation on the internal components. This may short-out externally polarized microphones. Conversely, at high temperatures, between 120-150° C, externally polarized microphones are a better choice, since the sensitivity level is more consistent in this temperature range.

Temperature Range Temperature has an effect on the microphones performance. Sensitivity levels can be directly affected by extreme environmental conditions. As the temperature approaches the maximum specifications of the microphone, its sensitivity specification will decrease. The owner needs to be aware of not only the operating temperature, but also the storage temperature of the microphones. If operated and/or stored in extreme conditions, the microphone can be adversely affected and also require to be calibrated more often.

A.7 Turbocharger speed

Turbocharger speed mainly depends on the exhaust and the conditions under which the engine is running. Under certain circumstances the speed can reach critical values for the safety and lifetime of the turbocharger. As the need for higher pressure values and hence higher nominal turbo speeds has constantly increased over recent years, it has become more and more important to know the instantaneous speed of the turbo. This is valid for nearly all sizes of turbochargers from the very small and fast rotating passenger car turbochargers up to the large units used in 2-stroke ship engines.

There are three main concepts of sensors used in this application: VRS (Variable Reluctance Sensors), optical sensors and eddy current sensors.

A.7.1 Variable Reluctance measurement system

A variable reluctance sensor (VRS) is used to measure position and speed of moving metal components. This sensor consists of a permanent magnet, a ferromagnetic pole piece, a magnetic pickup, and a rotating toothed wheel. As the teeth of the rotating wheel (or other target features) pass by the face of the magnet, the amount of magnetic flux passing through the magnet and consequently the coil varies. When the gear tooth is close to the sensor, the flux is at a maximum. When the tooth is further away, the flux drops off. The moving target results in a time-varying flux that induces a proportional voltage in the coil. Subsequent electronics are then used to process this signal to get a digital waveform that can be more readily counted and timed.

Regarding turbocharger applications, a strong magnet and a coil system build the core of these sensors. The target is normally a disc which is located

in the bearing section of the turbocharger. That disc has notches or holes so that the air-gap alters at these sections when the disc rotates with the turbo. The induced voltage is the desired output signal. Many different sizes and housings are available to adapt the sensors to the properties and specific mechanical needs of the application.

A.7.2 Optical measurement system

Specially designed for non contact measurement of turbochargers running at speed up to 200000 rpm, for purposes of measurement in stationary and transient engine-mode. A Laser-beam is sent to the compressor wheel, scattered back by means of a reflecting mark once each revolution. The scattered light is detected and converted into a periodical sequence of voltage signals that is available for further processing. The large optical range (usually more than 300mm) enables measurement without changes in design.

A.7.3 Eddy Current measurement system

Eddy currents (also called Foucault currents) are electric currents induced within conductors by a changing magnetic field in the conductor. These circulating eddies of current have inductance and thus induce magnetic fields. These fields can cause repulsion, attraction, propulsion, drag, and heating effects. The stronger the applied magnetic field, the greater the electrical conductivity of the conductor, and the faster the field changes, the greater the currents that are developed and the greater the fields produced.

As far as turbocharger speed measurement is concerned, a coil is potted in the sensor case and is energized by a high frequency alternating current. The electromagnetic field from the coil generates eddy currents in the turbocharger blade. Every blade generates a pulse. The controller identifies the speed (analog 0...10V) by considering the number of blades.

A potential problematic of this kind of sensor is related to EMC (Electromagnetic compatibility) emissions. Particularly where multi-test cells are in use, very high levels of EMC emissions are causing effect on test cell instrumentation. New generations of controller offer a new electronic circuit which 'boosts' signal levels from the sensor and also dramatically improves circuit shielding. The eddy current measurement technique is also immune to the effects of oil, dirt, carbon particles that can be found in the engine, which can affect the measurement output quality of other measurement principles, particularly optical measurement technologies.

A.8 TDC measurement

In engine indicating measurements, special attention should be given to the correct identification of the TDC (Top Dead Center) position, as small errors

in this measurement lead to significant errors in the evaluation of indicated work as well as combustion heat release rate. For instance, concerning the imep, a deviation of only 0.1 ° crank angle from the true top dead location already results in an error of several percentage points in the imep value. In order to achieve adequate precision in determining the TDC position it is recommended to perform a dynamic measurement, running the engine motored and unfired or, alternatively, preventing combustion only in the cylinder where the measurement is going on, while the other cylinders function fired to keep the engine running. When performing dynamic measurement, the inaccuracies that would be generated by the bearing clearances are eliminated. Such measurement is usually carried out with a capacitive proximity sensor, or in the absence of such a sensor, the TDC position can be inferred from the motored-engine pressure data.

The **capacitive proximity sensor** uses two conductive objects separated by a dielectric material. A voltage difference applied to the conductive objects generates an imbalance of electrical charges between them, originating an electric field in the dielectric material. When this voltage is alternated the electrical charges move continuously, going from one of the conductive objects to the other and generating an alternating electric current, which is the output signal of the sensor. The amount of current flow is determined by the capacitance, and the capacitance depends on the proximity of the conductive objects. Closer objects cause greater current than more distant ones.

In the capacitive sensors used to determine the TDC position, one of the conductive objects is the sensor probe itself, while the piston plays the role of the second conductive object (Figure A.4). The sensor is mounted in the head in such a way that when the engine is running the piston will move closer or away from the sensor, but without actually touching it. Thus, the sensor will produce a signal with amplitude, which is inversely proportional to the distance between the TDC sensor tip and the piston top. The exact TDC position will correspond to the maximum amplitude of the TDC sensor signal, which can be determined with great accuracy because of the high degree of symmetry of the signal.

When using the **motored-engine pressure data** for identifying the TDC position, the main arising difficulty is that the peak pressure precedes the actual TDC position, which corresponds to the minimum volume. This occurs due to heat transfer and mass losses, and the angle interval between these events is named “loss angle” (Figure A.5). Several enough accurate methods have been proposed for determining the loss angle, and usually manufacturers of indicating equipment include in their manual recommendations to estimate loss angle values, which depend on the kind of the engine (spark ignition or diesel) and compression ratio.

The advantage of direct determination of TDC, compared with determining the position of pressure maximum from the motored-engine pressure

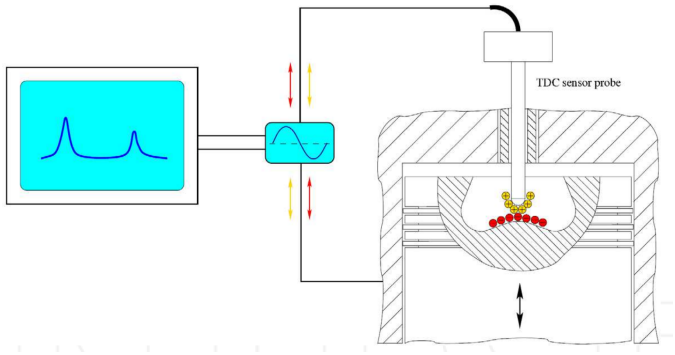


Figure A.4: The capacitive TDC sensor.

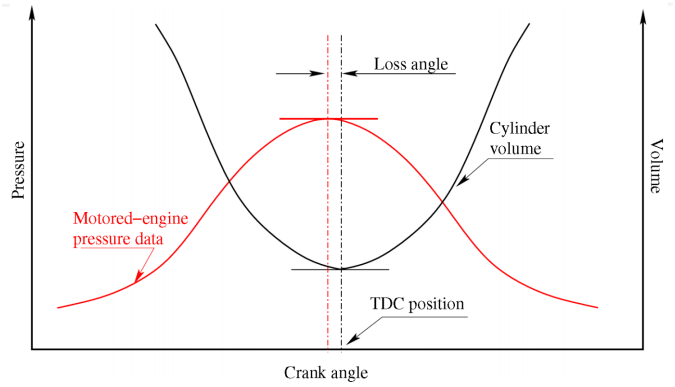


Figure A.5: Definition of loss angle.

curve, is that there is then no need for a correction involving the degree of the thermodynamic loss angle.

A.9 Air Mass Sensors

Hot-film Anemometer Practically all air mass sensors employed in the automobile today follow this principle. A heated element dissipates energy to the surrounding air. The dissipated heat energy is dependent on the air flow and can be used as a measurement parameter.

The HFM consists of a tubular housing with flow straightener (honeycomb-lattice combination), sensor guard, and the sensor module. The tube diameter is adapted to the air mass range required in each case. Sensors, electronics, connecting elements, and plug are integrated into the sensor module.

Two temperature-dependent metallic-film resistors on a glass substrate (R_S and R_T) are positioned inside the tube in the direct intake air stream. These two resistors, in combination with R_1 and R_2, are linked in a Wheatstone bridge circuit, Figure A.6. The voltage at R_2 is a measure of the air mass flow rate. Depending on the intaken air mass flow, R_S is cooled more or less strongly. The electronics control the necessary heating current through R_S so that there is always a constant temperature difference (e.g. 130 K) at R_S to the air temperature measured at R_T. The heating current is transformed into a voltage signal at resistor R_2.

The resistors R_S and R_T are matched to one another in such a way that the map is independent of the air temperature. The map also exhibits (thanks to the physics) an advantageous nonlinear characteristic permitting a practically constant proportional resolution.

Thanks to the use of materials specially adapted to the conditions in the automobile engine compartment, flow control, circuiting technology, and the mechanical configuration, the HFM signal is more or less independent of the temperature, pressure, and soiling.

Internal combustion engines with four or fewer cylinders create extreme pulsations in the intake manifold at wide-open throttle or in the case of no throttle valve (e.g. in diesel engines). At certain engine speeds, at the resonance point, a pulsating return flow occurs that with conventional HFM results in a positive measurement error as the air passes over the sensor three times.

The air mass flow (Q) is calculated as a function of the degree of modulation (m) to

$$Q = Q_{mean} \cdot [1 + m \cdot sin(wt)]$$
$$m = \frac{\hat{Q}}{Q_{mean}}$$

This effect can be compensated with an additional heating resistor R_H (booster). Returning air is heated by the booster and passes over the heating

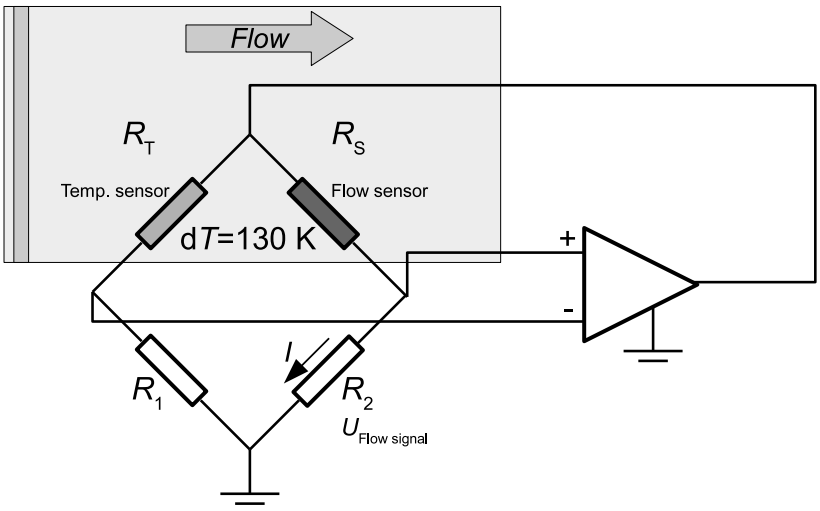


Figure A.6: Principle of a hot-film air mass meter.

resistor R_S. This prevents R_S from being cooled again by the returning air. Overheating the returning air produces an overcompensation that ensures that the air flowing toward the engine again is not measured a second time.

The return-flow compensation is independent of the resonance frequencies, temperature, and air pressure.

In many applications the temperature sensor (NTC resistor) for determining the intake air temperature is also integrated into the HFM.

Ultrasonic transit-time differential method With the ultrasonic runtime method, the time taken for a sound wave to travel from the transmitter to the receiver is measured. This parameter enables the flow velocity to be measured and, in combination with air density and temperature, the mass flow to be determined.

The measuring principle is based on the ultrasonic transit-time differential method. In this method, two ultrasonic pulses are sent simultaneously from Transmitter 1 (T1) and Transmitter 2 (T2) right through the flowing medium. One pulse is propagating into and the other one against flow direction. The interaction between the speed of sound c and the velocity of flow v accelerates the pulse on one of the paths and decelerates the pulse on the other path. This effective propagation velocity results in different transit times through the medium: the signal at Receiver 2 (R2) arrives faster than the signal arriving at Receiver 1 (R1) (Figure A.7). The device measures the speed of sound traveling either way, corresponding to $t_{1\to2}$ and $t_{2\to1}$.

- T - Transmitter
- R - Receiver

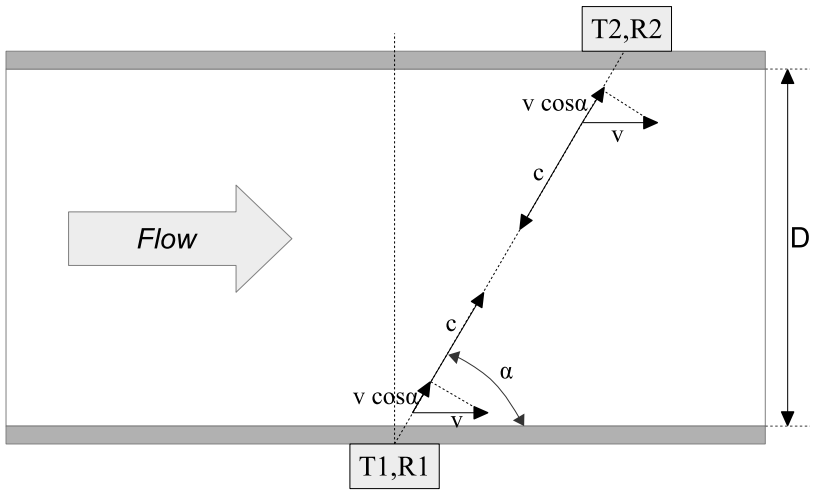


Figure A.7: Principle of ultrasonic flow measurement according to the transit-time differential method.

- c - speed of sound
- v - flow Media velocity
- α - inclination angle

A pulse travelling *with* the current from T1 to R2 needs a transit time of:

$$t_{1\to2} = \frac{D}{sin\alpha} \cdot \frac{1}{(c + v \cdot cos\alpha)}$$

A pulse travelling *against* the current from T2 to R1 needs a transit time of:

$$t_{2\to1} = \frac{D}{sin\alpha} \cdot \frac{1}{(c - v \cdot cos\alpha)}$$

The time difference of both pulses comes to:

$$\Delta t = t_{2\to1} - t_{1\to2} = v \cdot \frac{t_{2\to1} \cdot t_{1\to2} \cdot sin(2\alpha)}{D}$$

$$v = \frac{D}{sin(2\alpha)} \cdot \frac{t_{2\to1} - t_{1\to2}}{t_{2\to1} \cdot t_{1\to2}}$$

The flow rate Q is determined from the mean flow velocity. In a pipeline with circular cross-section the following applies:

$$Q = v \cdot A = v \cdot \pi \cdot \frac{D^2}{4}$$

$$Q = \frac{\pi \cdot D^3}{4 \cdot sin(2\alpha)} \cdot \frac{t_{2\to1} - t_{1\to2}}{t_{2\to1} \cdot t_{1\to2}}$$

The transit time difference is therefore a precise linear measure of the mean flow velocity v along the measuring path (ultrasonic beam).

Additionally, the sound velocity c can be determined on-line from the sum total of transit times $t_{1\to2}, t_{2\to1}$:

$$\sum t = t_{2\to1} + t_{1\to2} = \frac{1}{c} \cdot \frac{2D}{sin\alpha}$$

$$c = \frac{2D}{sin\alpha} \cdot \frac{1}{t_{2\to1} + t_{1\to2}}$$

Bibliography

[1] Lino Guzzella and Alois Amstutz. Control of diesel engines. *Control Systems, IEEE*, 18(5):53–71, 1998.

[2] Alois Amstutz and Luigi R Del Re. Ego sensor based robust output control of egr in diesel engines. *Control Systems Technology, IEEE Transactions on*, 3(1):39–48, 1995.

[3] Mrdjan Jankovic and I Kolmanovsky. Robust nonlinear controller for turbocharged diesel engines. In *American Control Conference, 1998. Proceedings of the 1998*, volume 3, pages 1389–1394. IEEE, 1998.

[4] MJ Van Nieuwstadt, IV Kolmanovsky, PE Moraal, A Stefanopoulou, and M Jankovic. Egr-vgt control schemes: experimental comparison for a high-speed diesel engine. *Control Systems, IEEE*, 20(3):63–79, 2000.

[5] Merten Jung and Keith Glover. Calibratable linear parameter-varying control of a turbocharged diesel engine. *Control Systems Technology, IEEE Transactions on*, 14(1):45–62, 2006.

[6] Johan Wahlstrom, Lars Eriksson, and Lars Nielsen. Egr-vgt control and tuning for pumping work minimization and emission control. *Control Systems Technology, IEEE Transactions on*, 18(4):993–1003, 2010.

[7] L. Guzzella and C. H. Onder. *Introduction to Modeling and Control of Internal Combustion Engine Systems*. Springer Berlin Heidelberg, 2nd edition, 2010.

[8] C. Rakopoulos and E. Giakoumis. *Diesel Engine Transient Operation*. Springer London, 2009.

[9] Regulation (EC) No 715/2007 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information. *Eur-lex.europa.eu. Retrieved*, 2011-02-02.

[10] R. Conway, S. Chatterjee, A. Beavan, and M. Lavenius et al. Combined SCR and DPF technology for heavy duty diesel retrofit. *SAE Technical Paper*, 2005-01-1862, 2005.

[11] C. Schär, C. Onder, H. Geering, and M. Elsener. Control of a Urea SCR catalytic converter system for a mobile heavy duty diesel engine. *SAE Technical Paper*, 2003-01-0776, 2003.

[12] S. Moon, B. Choi, H. Lee, and G. Hyun et al. Development of DPF/SCR system for heavy duty diesel engine. *SAE Technical Paper*, 2012-01-0369, 2012.

[13] F. Covassin, M. Preziuso, M. De Cesare, and G. Serra. A mean value model of the exhaust system with SCR for an automotive diesel engine. *SAE Technical Paper*, 2009-24-0131, 2009.

[14] M. De Cesare and F. Covassin. Neural network based models for virtual NO_x sensing of compression ignition engines. *SAE Technical Paper*, 2011-24-0157, 2011.

[15] Nicolo Cavina, Giorgio Mancini, Enrico Corti, Davide Moro, Matteo De Cesare, and Federico Stola. Thermal management strategies for SCR after treatment systems. *SAE Technical Paper*, 2013-24-0153, 09 2013.

[16] Johnson Matthey. Focus on SCR technology. *Global Emissions Management*, February 2012.

[17] C. Ericson, B. Westerberg, I. Odenbrand, and R. Egnell. Characterisation and model based optimization of a complete diesel engine/SCR system. *SAE Technical Paper*, 2009-01-0896, 2009.

[18] G. Neely, J. Sarlashkar, and D. Mehta. Diesel cold-start emission control research for 2015-2025 LEV III Emissions. *SAE International Journal of Engines*, 1009-1020, doi: 10.4271/2013-01-1301., 2013.

[19] S. Honardar, H. Busch, T. Schnorbus, and C. Severin et al. Exhaust temperature management for diesel engines. assessment of engine concepts and calibration strategies with regard to fuel penalty. *SAE Technical Paper*, 2011-24-0176, 2011.

[20] E. Krengel. Development of IMEP analysis for transient data. *SAE Technical Paper*, 2004-01-1463, 2004.

[21] M. Lindgren and P.-A. Hansson. Effects of transient conditions on exhaust emissions from two non-road diesel engines. *Biosyst Eng*, 87:57–66, 2004.

[22] Jonathan R. Hagena, Zoran S. Filipi, and Dennis N. Assanis. Transient diesel emissions: Analysis of engine operation during a tip-in. *SAE technical paper*, 2006-01-1151, 2006.

[23] John F. Cassidy. A computerized on-line approach to calculating optimum engine calibrations. *SAE technical paper*, 770078, 1977.

[24] Alan R. Dohner. Optimal control solution of the automotive emission-constrained minimum fuel problem. *Automatica*, 17:441–458, 1981.

[25] J. Asprion, G. Mancini, S. Zentner, Ch. H. Onder, N. Cavina, and L. Guzzella. A framework for the iterative dynamic optimisation of diesel engines: Numerical methods, experimental setup, and first results. *Accepted for publication in WIT Transactions on Ecology and the Environment*, 190, 2014.

[26] Oskar Leufven and Lars Eriksson. Surge and choke capable compressor model. *Preprints of the 18th IFACWorld Congress, Milano*, 2011.

[27] N. Watson and M.S. Janota. Turbocharging the internal combustion engine. *The Macmillan Press Ltd, London and Basingstoke*, 1982.

[28] Martin Müller, Elbert Hendricks, and Spencer C. Sorenson. Mean value modelling of turbocharged spark ignition engines. *SAE Technical Paper*, 980784, 1998.

[29] Indranil Brahma, Mike C. Sharp, and Tim R. Frazier. Empirical modeling of transient emissions and transient response for transient optimization. *SAE technical paper*, 2009-01-1508, 2009.

[30] M. Benz, M. Hehn, C.H. Onder, and L. Guzzella. Model-based actuator trajectories optimization for a diesel engine using a direct method. *J Eng Gas Turbines Power*, 133(3):032806–1–11, 2010.

[31] John T. Betts. *Practical methods for optimal control and estimation using nonlinear programming.* 2010.

[32] J. Asprion, O. Chinellato, and L. Guzzella. Optimal control of diesel engines: Numerical methods, applications, and experimental validation. *Accepted for publication in Mathematical Problems in Engineering*, 2014.

[33] Ernst Hairer and Gerhard Wanner. Stiff differential equations solved by Radau methods. *J Comput Appl Math*, 111(12):93–111, 1999.

[34] J.C. Butcher. *Numerical Methods for Ordinary Differential Equations.* John Wiley & Sons, Winchester, 2003.

[35] Michael A. Patterson and Anil V. Rao. Exploiting sparsity in direct collocation pseudospectral methods for solving optimal control problems. *J Spacecraft Rockets*, 49:364–77, 2012.

[36] Philip E. Gill, Walter Murray, and Michael A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Rev*, 47:99–131, 2005.

[37] Adam Lowell Schwartz. *Theory and Implementation of Numerical Methods Based on Runge-Kutta Integration for Solving Optimal Control Problems*. PhD thesis, University of California, Berkeley, 1996.

[38] Boris Houska, Filip Logist, Moritz Diehl, and Jan Van Impe. A tutorial on numerical methods for state and parameter estimation in nonlinear dynamic systems. In Daniel Alberer, Hakan Hjalmarsson, and Luigi del Re, editors, *Identification for Automotive Systems*. Springer London, 2012.

[39] M. Hafner and R. Isermann. Multiobjective optimization of feedforward control maps in engine management systems towards low consumption and low emissions. *T I Meas Control*, 25:57–74, 2003.

[40] E. Corti, G. Mancini, D. Moro, and C. Forte. Automatic combustion phase calibration with extremum seeking approach. *Proceedings of American Society of Mechanical Engineers*, ICEF2013-19132, 2013.

[41] M. Guerrier and P. Cawsey. The development of model based methodologies for gasoline IC engine calibration. *SAE technical paper*, 2004-01-1466, 2004.

[42] E. Corti, A. Cerofolini, N. Cavina, C. Forte, G. Mancini, D. Moro, F. Ponti, and V. Ravaglioli. Automatic calibration of control parameters based on merit function spectral analysis. *Energy Procedia*, 45(0):919 – 928, 2014. {ATI} 2013 - 68th Conference of the Italian Thermal Machines Engineering Association.

[43] Enrico Corti, Nicolò Cavina, Alberto Cerofolini, Claudio Forte, Giorgio Mancini, Davide Moro, Fabrizio Ponti, and Vittorio Ravaglioli. Transient spark advance calibration approach. *Energy Procedia*, 45(0):967 – 976, 2014. {ATI} 2013 - 68th Conference of the Italian Thermal Machines Engineering Association.

[44] M. Allain C. Atkinson and H. Zhang. Using model-based rapid transient calibration to reduce fuel consumption and emissions in diesel engines. *SAE technical paper*, 2008-01-1365, 2008.

[45] Giorgio Mancini, Jonas Asprion, N. Cavina, C.H. Onder, and Lino Guzzella. Dynamic feedforward control of a diesel engine based on optimal transient compensation maps. *Submitted to Advances in Mechanical Engineering - Hindawi*, 2014.

[46] Elbert Hendricks and Spencer C. Sorenson. Mean value modeling of spark ignition engines. *SAE technical paper*, 900616, 1990.

[47] A. G. Stefanopoulou, J. A. Cook, J. S.Freudenberg, and J. W. Grizzle. Control-oriented model of a dual equal variable cam timing spark ignition engine. *ASME Journal of Dynamic Systems, Measurement, and Control*, 120:257–266, 1998.

[48] D. Blomqvist, S. Byttner, U. Holmberg, and T. S. Rognvaldsson. Different strategies for transient control of the air-fuel ratio in a SI engine. *SAE technical paper*, 2001-01-2835, 2000.

[49] Jonas Asprion, Oscar Chinellato, and Lino Guzzella. Optimisation-oriented modelling of the NO_x emissions of a Diesel engine. *Energ Convers Manage*, 75:61–73, 2013.

[50] Frédéric Tschanz, Alois Amstutz, C. H. Onder, and Lino Guzzella. Feedback control of particulate matter and nitrogen oxide emissions in diesel engines. *Control Engineering Practice*, 21:1809–1820, 2013.

[51] Stephan Zentner, Erika Schäfer, Gerald Fast, C. H. Onder, and Lino Guzzella. A cascaded control structure for air-path control of diesel engines. *Accepted for publication in Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 2013.